

STUDYING DISORDERED DYNAMICAL SYSTEMS USING COMBINATORIAL AND ALGEBRAIC TOPOLOGY

**Thesis submitted to the Degree of Doctor of Philosophy (Science)
in Physics**

**by
Anamika Roy**

**Post Graduate and Research Department
of
Physics**

St. Xavier's College (Autonomous), Kolkata

Affiliated to the University of Calcutta

2024

DECLARATION

CERTIFICATE

I certify that the thesis entitled “**STUDYING DISORDERED DYNAMICAL SYSTEMS USING COMBINATORIAL AND ALGEBRAIC TOPOLOGY**” submitted by **Anamika Roy** for the degree of Doctor of Philosophy (Ph.D.) in Physics in the area of Condensed Matter Physics is the record of research work carried out by her during the period from 29.05.2018 to 23.09.2024 under my guidance and supervision and that this work has not formed the basis for the award of any Degree, Diploma, Associateship, Fellowship, Titles in this University or any other University or other similar institution of Higher learning.

OVERALL SIMILARITY PERCENTAGE: 2%

Signature of Principal Supervisor:

Signature of Co-Supervisor: *Amitabh . 13 September 2024*

Signature of PhD Coordinator:

I dedicate my dissertation to my mother, Dipti.

“All is flux”
- Heraclitus

Acknowledgements

When I embarked on my Ph.D. journey, following a long hiatus after completing my masters degree, I had little hope of balancing the demands of academic research with a full-time job and other challenges. I often joked that while the probability of success was small, it wasn't zero and hence I continued. Now, six years later, I can say that it is the people who have been part of my Ph.D. journey have played an invaluable role in shaping the entire process. Their support made all the difference.

First and foremost, I am deeply grateful to my supervisor, Dr. Tapati Dutta. She has been the kindest and the most supportive supervisor one could ask for, patiently teaching me the process of doing research. I want to thank her for giving me the opportunity for all the academic endeavours I had with her and especially for giving me the space, and the flexibility to work and study on my way. Apart from academic mentoring, she has always been there positively assuring me in difficult times. Her zeal and conscientious attitude in doing science have inspired me the most.

This dissertation would not have taken the shape it has today without the guidance of my co-supervisor, Dr. Atish J. Mitra. I was introduced to him through Dr. Dutta during a lab discussion, and that meeting led to an interdisciplinary collaboration that reshaped the course of my research. Dr. Mitra introduced me to Topology in early 2020, and together, we explored the fusion of dynamical systems and topology. His mathematical insight, perfectionism, and constructive feedback were invaluable in shaping this dissertation. Despite the geographical distance, our weekly online group discussions kept the momentum going, even in the face of COVID and other obstacles. I am grateful for his encouragement, which gave me the confidence to pursue this interdisciplinary work.

Besides my supervisors, I am indebted to Prof. Sujata Tarafdar from CM-PRC, JU, who has been a part of my Research Advisory Committee and has provided insightful comments and scientific contributions to my work. I am extremely grateful for the opportunity to learn in her company, which has stimulated both my scientific and creative pursuits, making learning enjoyable and grounding. I also extend my gratitude to Prof. Parongama Sen for her valuable suggestions and assistance with the thesis, from the very beginning to the end.

In addition to my mentors, I would like to express my heartfelt thanks to my colleague and friend, Ruhul Amin Ibne Haque. Over the past few years, we have grown and learned together, and it was Ruhul who first introduced me to Python in 2018, encouraging me to integrate it into my research. Our academic and personal conversations have been instrumental in shaping my work. I also want to acknowledge my colleagues Samiul Haque, Moutushi Dutta Choudhury, and Ankita Paul for their support, resources, and guidance, which were vital for the experiments conducted in this dissertation.

This research would not have been possible without the opportunity to work as a Ph.D. scholar in the Postgraduate and Research Department at St. Xaviers College, Kolkata. I extend my deepest gratitude to Father Principal, Dr. Dominic Savio S.J., and all the faculty members for providing a platform and a nurturing environment for research. I owe special thanks to Dr. Samrat Roy, Ph.D. coordinator, for his consistent guidance and support throughout this process. I would also like to express my gratitude to my workplace, Charuchandra College, Kolkata, for providing me with the opportunity to balance my research with my professional responsibilities, and for their cooperation whenever it was required.

Then there are my everyday people – my family and friends – who have been patiently there for me with their love and affection. I want to thank every one of them: my close friends Doyel, Arunita, Shamayeeta, and Pampi; my close colleagues at work, Susmita di, Mita di, Munmun di, and Sudip da; and my close acquaintances, Poulami and Amrita di. I am deeply grateful for their warmth, companionship, and constant emotional support over the past few years.

The biggest and most persistent support system throughout my journey has been my family. Maa, Didibhai, Bhaiada, Dadamani, Borda – everyone in my family has been there for me, encouraging and supporting me since my early kindergarten days. Baba, though physically not with us, left a profound impact on me, guiding and

shaping the path I chose to follow. A special mention goes to my nephew, Wrick, who is seven years younger than me and a constant source of positivity, always pushing me to step out of my comfort zone. My niece, Raya, is the newest addition to our family whose presence makes me smile the widest, even on my dullest days.

Lastly, I am grateful for the random interactions, reflections, and moments of resonance that have brought unexpected glimmers to my journey, letting me grow in my pursuit.

Abstract

This dissertation investigates various dynamical systems through the frameworks of combinatorial and algebraic topology and geometry, with the goal of understanding how physical dynamical processes imprint themselves on the topology of the system. By developing a structured approach for analyzing topological information, this research applies these methods to both simulated and empirical dynamical data. The study encompasses a wide range of physical processes, beginning with an exploration of crack mosaics formed by drying solutions/suspensions of different materials. These are characterized by relatively simple patterns, and the study progressively extends to more intricate problems, such as flow patterns in fluid dynamical systems.

A key contribution of this work is the construction of topological and geometric tools that capture the essential features of dynamical systems, enabling their classification based on topological and geometric similarities and differences. For crack mosaics, a four-parameter tuple space, denoted as (n, v, D, λ) , was introduced. This combinatorial and geometric framework successfully investigates and distinguishes crack pattern characteristics according to material properties of the drying system. In the analysis of spatio-temporal evolution in fluid dynamics, which is more involved, a novel topological tool, termed the Euler Characteristic Surface (ECS), was developed. The ECS encapsulates the topological signature of a dynamical system and provides a unique descriptor for it.

To quantify the similarity and dissimilarity between ECS representations, a new metric referred to as the Euler Metric was introduced, facilitating the comparison of dynamical systems through their topological characteristics. The effectiveness of these topological constructs has been thoroughly analyzed, with results that validate their ability to capture and distinguish dynamical behavior.

The dissertation concludes by establishing the mathematical stability of the proposed topological constructs and by drawing connections to widely-used techniques in Topological Data Analysis (TDA).

Contents

1	Introduction:	13
1.1	Characteristics of frameworks suitable to study complex dynamical systems:	14
1.2	Topology and Dynamical Systems:	16
1.2.1	Studying topological invariants:	16
1.2.2	Topological data analysis:	18
1.3	Outline	18
2	Studying combinatorial topology of crack networks: A 4-parameter space	20
2.1	Background on planar mosaics	22
2.1.1	The topological combinatorics (n,v) and $n-v$ plane :	22
2.1.2	Introduction of geometrical measures	24
2.1.3	The parameter space of quadruples $(\mathbf{n}, \mathbf{v}, \mathbf{D}, \boldsymbol{\lambda})$	26
2.2	Analysis work-flow	26
2.2.1	Extracting crack skeleton from images	27
2.2.2	Identification of nodes: calculation of angles at nodes	28
2.2.3	Simulated crack mosaics	31
2.3	Observation from the analysis	33
2.3.1	Mapping the topology of the crack mosaic on the (\mathbf{n}, \mathbf{v}) plane	33
2.3.2	Refining by introducing geometric measures - the $(\mathbf{n}, \mathbf{v}, \mathbf{D}, \boldsymbol{\lambda})$ space	34
2.3.3	Including Hausdorff Measure for comparison	36
2.4	Comparing real and simulated crack mosaics	38
2.5	Evolution of crack pattern in $(\mathbf{n}, \mathbf{v}, \mathbf{D}, \boldsymbol{\lambda})$ space:	39
2.6	Discussion	41

2.7	Conclusion	42
3	Studying the topological picture in fluid dynamical systems: Introducing the Euler Characteristic Surface	50
3.1	Introduction:	50
3.2	The physical system	51
3.3	Image Processing:	53
3.3.1	Thresholding for grey scaling	54
3.3.2	Gaussian Filter	56
3.3.3	Image and tessellation:	59
3.3.4	Euler Characteristic on triangular lattice:	60
3.3.5	Extension to hexagonal tessellation:	62
3.4	Evolving topological features of agglomeration patterns and the dynamical system	63
3.4.1	Time evolution of the Euler Characteristic through images	64
3.5	Introducing the approach for multiscale study of topological features:	67
3.6	Introduction to the Euler Characteristic Surface (ECS) :	69
3.6.1	Euler Characteristic Level Curves	70
3.6.2	Proposing the Euler Metric for dynamical systems	71
3.7	Discussion	72
3.7.1	Particle agglomeration features	73
3.7.2	Flow stability features	75
3.8	Conclusion	76
4	Characterizing fluid dynamical systems using Euler Characteristic Surface and Euler Metric	78
4.1	Introduction	78
4.2	Material and Methods	79
4.2.1	Simulated dynamical fluid flow system	79
4.2.2	Stochastic flow systems: Flow patterns in drying droplets	82
4.3	Results and Discussion	84
4.3.1	Analysis of simulated fluid flow	84
4.3.2	Analysis on stochastic dynamical fluid system	88
4.4	Conclusion	93
5	The mathematical framework and stability of the Euler Characteristic Surface	94
5.1	Introduction	94
5.2	Mathematical Framework	96

5.2.1	Cell Complexes and Euler Characteristic Surfaces	96
5.2.2	Homology, Betti Surfaces and Persistence Diagrams	101
5.2.3	Comparison with Persistent Homology	105
5.3	Application on simulated data	112
5.3.1	Eggbeater Flow: Dependence on parameters, Stability of ECS	112
5.3.2	Vicsek Model: A case study	117
5.4	Conclusion	123
6	Conclusion and future directions	126
6.0.1	Future directions:	129

List of Tables

2.1	Measures of real crack mosaics	37
3.1	Summary of topological tools studied	73
4.1	Measures of Euler Metric($d\chi_k, \chi_{k+dk}$)	89
4.2	Measures of Euler Metric	90
5.1	Measures between point clouds for $\eta = 0.0$ and $\eta = 0.90$	120

List of Figures

1.1	The objects in the figure have topologically equivalent shapes, that can be transformed from one to other through continuous deformations[1].	17
2.1	Rectangular and Triangular lattices with - (a) all “regular nodes”, (b) with all “irregular nodes”, (c) with both regular and irregular nodes present, (d) A, B, and C are polygons in a schematic mosaic, nodes marked by numbers 1 - 9; tabulation of nodes and vertices of each polygon; tabulation of regular and irregular nodes of the mosaic. . . .	23
2.2	(a) (n,v) phase diagram. The range (2,4) to (4,4) i.e. bricklayer type to square-tiled type patterns, cluster around ‘Platonic attractor’. Hexagonal tiling pattern with (3,6), cluster around the ‘Voronoi attractor’. (b) θ_i for a 4-sided polygon and a 5-sided polygon. Corresponding regular polygons are shown in dotted lines.	25
2.3	Sequential steps towards determination of (n, v, D, λ) data. (a) A desiccated layer of bentonite clay. (b) Binary image after suitable filtration of noise with a schematic of crack skeleton extraction with respect to mid-points of crack width. (c) Skeletonized image. (d) All nodes (junction points) detected on the skeleton. (e) All polygons identified and labelled. (f) Vertices v determined and marked on each polygon. (g) Nodes n determined and marked. (h) Superposition of Voronoi mosaic constructed (red edges) from centroids of real crack polygons.	28
2.4	(a) Convex hull constructed on some polygons of the mesh. Red stars indicate non-convex polygons determined from the construction. (b) Schematic guide to the determination of internal angles of non-convex polygons. Image has been adapted from [2]	31

2.5	(a) Simulated Gilbert mosaics with random uniform and random normal seed distribution, (b) Simulated Voronoi mosaics with random uniform and random normal seed distribution, (c) Simulated mosaics with iterative cell divisions at the end of 1 st , 4 th and 8 th iterations respectively.	32
2.6	(n, v) data of real and simulated crack mosaics. The legend indicates the source of the experimental crack mesh. The iso-p lines at $p = 0.3$ and $p = 0.5$ are drawn as a guide to the eye.	35
2.7	(a1)(n, v, D) space of real and simulated crack mosaics.(a2) Zoomed in view of (n, v, D) data. (c1) (n, v, λ) space of real and simulated crack mosaics. (c2) Zoomed in view of (n, v, λ) data for experimental and simulated crack mosaics. Description of data points provided in the legend.(b) legends for (a1), (a2) and (d) legends for (c1) and (c2).	44
2.8	(a) Image of crack mosaic on Laponite clay of pH 6.73, (b) Image of crack mosaic on Laponite clay of pH 13.05, (c) Hausdorff distance $d_{H(sv)}$ estimated between the real crack mosaic (labelled with blue line) and its corresponding Voronoi tessellation (labelled with black broken line) for (a), (d) Hausdorff distance $d_{H(vs)}$ estimated between the real crack mosaic (labelled with blue line) and its corresponding Voronoi tessellation (labelled with black broken line) for (b).	45
2.9	Histogram of isoperimetric ratio λ for real crack mosaics of :(a) TiO ₂ , (b) PDMS, (c) Laponite (pH 6.73), (d)Laponite (pH 13.45), (e) Tempered Glass, (f) Resin, (g) Corn starch, (h) Mud crack.	46
2.10	Histogram of isoperimetric ratio λ for simulated crack mosaics: Gilbert tessellations with slopes(a) random, (b) 0 and 90 degrees, (c) 0, 45, 90 degrees, (d) 0 ± 10 , and 90 ± 10 ; Voronoi tessellation for (e) uniform and (f) normal seed distribution.	47
2.11	(a)Columnar joints [3], (b) Cracks in dried river bed [4].	47
2.12	(a) Convection flow cooling, indicated by white arrows, takes place in crack joints; conduction cooling, indicated by black arrows, takes place across column interior. The dashed lines are the isotherms. Figure adapted with permission from [5]. Copyright 1994 Elsevier. New crack deviates in the direction of the wider column, following the higher thermal gradient. (b) Schematic for algorithm to estimate the shift in crack nodes in each time step of the simulation at different horizontal sections of adjacent columns meeting at a point. Black circles represent centroids of the neighboring polygons of the node(for which the shift is estimated for the next time step) marked by red circle.	48

2.13	Trajectory of the crack in the (n, v, D, λ) domain: Gilbert cracks with uniform seed distribution and crack orientation - (a) random (b) parallel (c) $\pm 10^\circ$ (d) 45° and 90° ; Gilbert cracks with normal seed distribution and crack orientation (e) random (f) parallel (g) $\pm 10^\circ$ (h) 45° and 90° (i) Voronoi cracks.	49
3.1	Different modes of droplet evaporation from a substrate- (a) Constant contact radius (CCR) with contact angle decreasing with time, (b) Constant contact angle (CCA) with contact radius decreasing with time. Fluid flow patterns in an evaporating droplets on hydrophilic rigid substrate- (c) Arrows represent a schematic of evaporation flux on droplet interface. The radial convective flow pattern towards the triple phase contact line carrying particles is represented by dashed lines. (d) Marangoni flows along a section of a droplet represented by arrows. The flow is induced by the surface tension gradient along the fluid-air interface.	52
3.2	(I) Images of top surface of a droplet containing water and polystyrene beads during evaporation with their binary images at (a) 12.24 min, (b) 16.25 min, (c)20.24 min, (d)22.01 min, (e)25.241 min, (f)28.24 min, (g) 33.24 min, (h)39.24 min. The binary images after applying a Gaussian filter are displayed alongside their original images; (II) Time variation of Euler Characteristic during droplet evaporation. The insets display the corresponding binary maps of the system.	54
3.3	Best possible binarization of images of the droplet at (a)16.25 min and (d)20.24 min- without filter:(b and e); with Gaussian filter:(c and f), respectively.	55
3.4	Intensity distribution of the images. The black graph displays the intensity distribution. The red graph is a fitted Gaussian curve. . . .	56
3.5	Distribution of Gaussian function values.	57
3.6	(a) Images composed of hexagonal cells do not present the connectivity problems, sometimes associated with images composed of squares or triangles. As can be seen from this figure (cases (b) and (c)), cells might appear connected by their corners. (d)A random distribution of a simulated triangular lattice of size 16×16 with $p = 0.48$ with black clusters being marked in the picture. Hexagonal lattice generated as a subset of triangular lattice, marked by red	60

3.7	The variation of Euler Characteristic (χ) for triangular lattice of different sizes. In each cases, the value becomes zero near the percolation threshold, $p_c = 0.5$	61
3.8	The four possible choices to suppress rows and columns of square lattice to turn into a hexagonal grid-like format.	63
3.9	Time variation of the Euler Characteristic (χ) of the evaporating droplet system using a hexagonal grid. The insets display the corresponding binary maps at different time instants.	65
3.10	Size Distribution of clusters at time instants (a) 17.05 min , (b) 22.01 min, (c) 25.24 min, (d) 39.24 min. A zoomed view is shown in the inset.	66
3.11	Top-view of drying droplet image as each black cluster is enlarged by a scale of r-neighborhood. Scale $r = 0$ corresponds to the original binary image at $t=20.24$ min. Scales in the figure vary between $r = 0$ to $r = 7$	68
3.12	(a) Variation of χ with different scales for constant times. (b) Variation of χ with time at different r-neighbourhood scales.	69
3.13	The Euler Characteristic surface in scale (r) and time (t) of the drying droplet. The experimental data points are marked by red circles, the blue surface is the best-fit surface constructed by the $r-t$ phase space. Two orientations of the same surface are displayed in figures(a) and (b).	70
3.14	(a) The Euler Characteristic Surface Level Curves. Curves having the equal value of Euler Characteristic have χ displayed. Positive χ curves are bold lines and negative χ curves are dotted lines. (b) Color plot of the Euler Characteristic Level Curves. The color legend displays different χ values.	71
3.15	Comparison table illustrating a schematic of topological features (column 1) and their reflection in Euler Characteristic Numbers (column 2), compared with images of droplet interface features observed during different time instants of evaporation (column 3).	74
4.1	Patterns generated with $k = 4.1$ at time steps- (a) $t = 500$, (b) $t = 1000$, (c) $t = 2000$, (d) $t = 3000$, (e) $t = 5000$ and (f) $t = 7000$	80
4.2	Final patterns generated after 10000 time steps for (a) $k = 4.00$, (b) $k = 4.20$, (c) $k = 4.40$, (d) $k = 4.60$, (e) $k = 4.80$ and (f) $k = 5.00$	81
4.3	Final pattern after 10,000 time steps for $k = 4.1$ at different scales (a) $r = 1$, (b) $r = 2$, (c) $r = 3$, (d) $r = 4$, (e) $r = 5$, (f) $r = 6$, (g) $r = 7$, (h) $r = 8$	82

4.4	Top-view images of flow patterns at different time instants for 2 droplet drying systems. The binary image is displayed along with every real-time image. (a) Patterns of agglomeration while evaporation of water droplet containing polystyrene micro-beads on PDMS substrate (40:1) at (i) 12.24 min, (ii) 15.24 min, (iii) 18.24 min, (iv) 21.24 min, (v) 26.24 min, (vi) 29.24 min, (vii) 32.24 min, (viii) 36.24 min and (b) Agglomeration patterns for evaporation of water-ethanol droplet containing silica nano-particles on PDMS (40:1) substrate at (i) 3.12 min, (ii) 8.06 min, (iii) 10.23 min, (iv) 12.03 min, (v) 14.15 min, (vi) 15.42 min, (vii) 16.39 min, (viii) 17.35 min.	83
4.5	Plot of Euler Characteristic (χ) computed for final patterns i.e. after 10,000 time steps for different values of k in the simulated particle flow.	85
4.6	(a) Final pattern after 10,000 time steps for $k = 4.1$ at scale 2, (b) Final pattern after 10,000 time steps for $k = 4.0$ at scale 2, (c) Euler Characteristic Surface for $k = 4.1$ and (d) Euler Characteristic Surface for $k = 4.0$	86
4.7	(a) Final pattern after 10,000 time steps for $k = 4.2$ and its corresponding ECS, (b) Final pattern after 10,000 time steps for $k = 4.4$ and its corresponding ECS, (c) Final pattern after 10,000 time steps for $k = 4.6$ and its corresponding ECS and (d) Final pattern after 10,000 time steps for $k = 4.8$ and its corresponding ECS.	87
4.8	(a) Final pattern after 10,000 time steps for $k = 4.1$, (b) Final pattern after 10,000 time steps for $k = 4.0$, (c) Euler Metric : Arrows indicate the difference between corresponding points between the Euler Characteristic Surfaces for $k = 4.1$ and $k = 4.0$	88
4.9	Euler Metric(d) computed for Euler Characteristic Surfaces corresponding to flow patterns with $k + dk$ and k . The insets depict flow patterns corresponding to significant k values at the end of 10,000 time steps. EM is computed at two values of variation dk as indicated in the legend. Point A marks EM between flow pattern with $k = 4.0$ and 4.1, is a point of critical topological changes. The region between C and D are stable flow regions.	89
4.10	Euler Characteristic Surfaces – (a) for evaporating water-ethanol droplets containing silica flakes for Set 1(blue) and Set 2 (orange), (b) for evaporation of water droplet containing polystyrene beads (orange) and water-ethanol droplet containing silica flakes (blue).	90

4.11	Euler Characteristic versus time at scale $r = 1$ for evaporating water-ethanol droplet containing silica flakes and for evaporation of water droplet containing polystyrene beads.	91
5.1	Schematic of Coarse-graining/union of r -neighborhood approach: (a) pixels at scale 0 (no-coarse graining applied) (b) The nearest neighborhood of the black pixels marked with blue color and the next nearest neighbors marked with pink color, (c) coarse-graining applied at scale 1 with diagonal connections attached with 0.5 probability.	98
5.2	(a) The collection of subsets, (b) The corresponding Nerve complex, (c) Schematic to represent the intersection of Voronoi cells and balls of radius r around a point cloud and (d) The corresponding Alpha complex.	100
5.3	Alpha complexes of a point distribution generated using Vicsek model at scale - (a) $r = 0.03$, (b) $r = 0.045$, (c) $r = 0.055$	101
5.4	(a)-(g) represent different possible bijections or matching between two Persistence Diagrams(red and blue), (h) The table showcases the matching distances and the estimated metrics.	104
5.5	a schematic of persistence diagram. The points within the shaded box whose birth at $\leq r$ and death at $> r$ contribute to the simplicial complex at scale r and so to the Betti r curve.	106
5.6	(a) and (b) are the possible orientations for (b_k, d_k) and (b_l, d_l) for a possible optimum matching between a pair of points from two different persistent diagrams. Matching (c) is not optimal.	107
5.7	(a) Alpha Complex of flow pattern of modified egg-beater flow at time-step $t=500$ with $k=4.1$, at scale $r=0.01447$ estimated $\chi=161$, (b)Same point set at same scale($r = 10$) with “coarse-graining” or “union of r -neighbourhood” approach, estimated $\chi =148$	113
5.8	For Flow pattern with $k = 4.1$ -(a) 0-dimensional Betti surface, (b)1-dimensional Betti surface, (c) The Euler Characteristic surface.	114
5.9	(a)Euler characteristic surface constructed through Alpha complexes, (b)Euler characteristic surface constructed through ‘coarse-graining’ or union of r -neighborhood approach on the pixelated grid, (c) The values of estimated Euler Characteristic (χ) at different scales using the two methods.	115
5.10	Euler metric estimated between surfaces with k and $k+dk$ where $dk=0.1$, using both approaches.	116

5.11	Euler metric estimated between surfaces keeping R upto scale 10(black)and upto scale 11 (red).	117
5.12	Slices of Euler Characteristic Surfaces at different time steps for the ECS corresponding to $\eta = 0.9$	121
5.13	The L_1 Euler metric between ECS time-slices with dt distance apart from time slice $t = 39$	122
5.14	(a) L_2 Euler metric estimated between ECSs of noise η and zero noise. (b) The order parameter vs noise (η) is plotted for different interaction radius R	123

Chapter 1

Introduction:

The story of my thesis started in a serendipitous way. Initially when I had joined the lab to pursue my research, there were scholars from the lab who had published work on studying morphological phase transition in real and simulated dynamical systems. One of them being the Island to Mainland transition in square lattice without and with modified diagonal connectivity [6, 7], the other being the development of crystalline phase in drying droplet[8]. Interestingly, in both of the works, there were analysis based on a parameter called “Euler Characteristic”, whose minima and maxima were observed along with significant morphological transition in the system. “Euler Characteristic” is a very popular topological invariant(the details about it are discussed later in this chapter and as well as in Chapter 3).

Reading these works made me interested in exploring the parameter in other systems. My initial endeavour was to compute this “Euler Characteristic” in a simulated system of triangular lattice with the percolation theory approach. The output of the simulation matched with the results reported in [9, 10], the curve of the Euler characteristic reaches zero at percolation threshold $p_c = 0.5$ and changes its slope. Generally, the question that followed after this, whether this parameter is that powerful to capture transitions and dynamical changes in real systems as well. To follow the quest, I gradually developed algorithms and methods that studies the variation of Euler Characteristic in dynamical systems that also helps in understanding the physical processes going on in the dynamical systems. The study then got widened with inclusion of the approaches from combinatorial topology and algebraic topol-

ogy, influenced by the collaboration with my co-supervisor. We delved into analysing different real and simulated dynamical systems such as crack networks, fluid mixing and droplet evaporation, with the topological perspective and carried on improving the robustness of our tools and methods. The intention behind this dissertation were – 1) Building frameworks that analyses the topology of dynamical systems, 2) Whether studying dynamical systems in terms of topological approaches is at all beneficial/insightful, 3) Testing the robustness of the established framework with the already established topological tools and methods. The studies done for this dissertation are presented chapterwise, before that a very brief description of the different standard methods to study dynamical systems follows in the subsequent sections here.

Dynamical systems have been a domain of critical importance for more than centuries. It simply describes the time evolution of system by differential equations where time is a continuous parameter or by iterated maps where time is discrete. The very first foundation of it started with the introduction of Newtonian classical mechanics to describe planetary motion in late 16 th century. Later in 18th century Poincare proposed to look at dynamical systems with geometrical approach. The field gradually got improved with study of complex behaviour with Hamiltonian mechanics by Birkhoff, Kolmogorov, Arnol'd in early 19 th century. The invention of the computer in mid 19th century and therefore running numerical simulation helped advancement to understand dynamical system more vividly. From 1960 to 1980 the emergence of strange attractor and Chaos by Lorenz opened a new side of the studies. Later the introduction of fractal by Mandelbort and the study of non-linear biological oscillator by winfree added new perspectives to the field[11, 12]

1.1 Characteristics of frameworks suitable to study complex dynamical systems:

For real world dynamical systems having a large number of components, the problem becomes complex when the conceptual and analytical framework of the simple dynamical system theory approach often does not yield an effective understanding. A complex dynamical system is one made with many constituents which may interact with one another. Thus studying the constituents in isolation makes no sense. The attributes associated with complex systems are their unpredictable outcomes and the lack of precision in linking them to their causes[13]. E.O. Wilson had mentioned that the challenge in studying a complex dynamical system is to specify an accurate

and complete description through a suitable framework[14].

Complex dynamical systems have been studied from a network approach by different scientists during the last three decades. The focus of studying a dynamical system as a network is to analyze the structure or anatomy of the network and describe how it relates to the functions or behavior of the system. The standard network models used are the following : regular networks (grid, lattice, chain etc.), random graphs, small world networks, the scale free networks etc. Analysis of a network involves study of the following features - structural complexity, network evolution, connection diversity, dynamical complexity, node diversity, meta complications and other such details. Further details are explained in [15]. Real world complex networks, e.g. the World Wide web, Ecological networks, Cellular networks, Power and neural networks etc have been studied to understand their network topology. Average path length, clustering coefficient and degree distribution are some of the parameters that are studied to quantify the network structures[16]. It is to be noted that a complex network may not always be a complex dynamical system.

A statistical physics inspired approach for characterization of complex system dynamics has also been applied by scientists. In a complex system the interactions between the constituents are space and time dependent. The state of the constituents and their interaction co-evolve over time, causing the microstate and macrostate to dynamically update each other. Behavior of complex dynamical systems are often explained using tools of statistical physics such as - universality, scaling, phase transition, diffusion and so on. Concepts and techniques such as Entropy, Random walk and other such tools are often employed[17]. Estimation of the free energy to describe different spatio-temporal states in a complex dynamical system made of N agents acting collectively was studied by Koorehdavoudi et al., [13].

Complex system science, that focuses on how the components within a system are related to one another, provides a structure to study multiscale behavior of a real system which is often complicated. Its basic principles include multiscale analysis, trade off between small scale and large scale complexity, matching the complexity of a system to that of its environment etc. A system is characterized as random, coherent, or correlated according to the interaction between its components. This is similar in approach to statistical mechanics where the macroscopic properties of a system is studied by analyzing all its possible states. Complex system science - studies the space of all possible modes of behavior of a complex system by identifying the complexity profiles. The larger the number of possible ways of behavior for a

system, the greater is its complexity. Complexity of a system is very much a scale dependent phenomenon and it requires multiscale analysis. Clickable maps is one such tool that aids in analyzing multiscale complexity of a system. Although it often may be inadequate to understand critical macroscopic phenomena such as phase transitions in a system[18].

1.2 Topology and Dynamical Systems:

The aim of this work is to look at dynamical systems from topological perspectives, using tools and parameters of topology. Topology, is a branch of mathematics that studies the properties of a geometric object (represented by graphs, simplicial complexes, or other discrete structures) that remains unchanged under continuous transformation. The usual examples of continuous transformation of space are stretching, twisting, bending, crumpling etc and the topology of an object remains unchanged under these processes. Tearing, gluing, self crossing do not preserve topology. Figure(1.1 illustrates how objects with varying sizes, curvatures, and shapes can undergo continuous deformations while maintaining their global topological properties. Despite the transformations, key features such as connectivity, holes, and boundaries remain unchanged, showcasing the robustness of topology in preserving fundamental structures during deformation. In a complex dynamical system the above mentioned physical processes get combined repeatedly leading to a structure. Thus the signature of the mechanics of the dynamical system is hidden in its topology. Studying topology of the geometric structures present in the dynamical system may give information about the physical forces acting in the system. Poincare in 1892 had instigated the question that whether it is possible to identify dynamical systems and its evolution in terms of topology.

1.2.1 Studying topological invariants:

With the advances in computation the theory of algebraic topology has branched into ‘Applied Topology’ where one uses the concept of topology to analyse physical or biological or societal dynamical systems. The application was vastly on understanding patterns or morphology of complex dynamical systems. One of the most popular topological invariant is the Euler Characteristic. It is one of the measures in Minkowski functionals as well. Euler Characteristic is a very efficient marker to study connectivity of a network or system. It has been used to study the structure in porous systems [19] and has been related to permeability of the porous matrix[20]. The large dataset produced by Astronomical observations have been interpreted [21] as well as



Figure 1.1: The objects in the figure have topologically equivalent shapes, that can be transformed from one to other through continuous deformations[1].

the non-trivial connectivity in the complex functional brain network, protein or gene networks have been studied with the help of this topological invariant[22, 23]. The field of Image processing has been vastly benefited with using topological tools in last fifty years. Pattern Recognition, Optimal thresholding, Skeletonizing etc have been improved with algorithms inspired from topology[24].

Chaos and topology:

There exist other topological measures that have been used to understand dynamics in complex disordered systems. In a review Gilmore describes how topological methods can help analysis of dissipative dynamical system operating in chaotic regime. The study highlights how studying topological invariant is beneficial over metric invariant and dynamical invariant. The topological invariants studied were linking number and relative rotation rates related to unstable periodic orbits occurring in strange attractors. The topological invariants are robust under change in parameter for chaotic data whereas the other two variants are not. [25]. The non-autonomous and stochastic dynamical behavior in climate is also being looked at with insights of algebraic topology[26] by studying the branched manifolds of the dynamical systems through cell complex and homology. Branched manifolds are topologically invariant and two dynamical systems with same branched manifolds are considered

as dynamically equivalent.

1.2.2 Topological data analysis:

With recent advanced computation technology, there has been a data avalanche for different kind of dynamical systems. Interpretation of such time series data of dynamical systems to interpret the dynamics is the new challenge scientists are running for. One of the most efficient approach to deal with this has been the development of the subject “Topological Data Analysis” or TDA. The problem with working with large data is to identify the noise and the missing information. Often times these datasets have high dimensional relation that is not possible to visualise with open eyes. Topological data analysis extends the concept of homology to extract features from such complex data sets. A data set is basically considered as a vertex set or point cloud, where the whole set is constructed with geometric structures of different dimensions called simplices. The tools of homology is then applied on these simplicial complexes to extract different dimensional features like connected components, loops, cavity etc. One of the most popular branch of TDA is Persistent Homology, first conceptualised by G. Carlsson in 2004[27, 28]. There have been many successful application of TDA on different dynamical systems. A concise review can be found in [29].

1.3 Outline

In this dissertation I shall first explore static patterns, specifically crack patterns, left by dynamical processes. The patterns shall be classified in terms of combinatorial topology and geometry. This format of pattern classification will help classify systems/materials that share some common intrinsic properties as their combinatorial information will be clustered together. A 4-dimensional domain of topological and geometrical descriptors of the tuples will act as a calibration map for materials. This study has further lead me to explore the possibility of whether there existed dynamical systems whose birth patterns and mature dynamical patterns were located at distinctly different positions in the aforementioned 4-dimensional domain. This led me to the study of the time evolution of columnar joints - gigantic 3-dimensional crack systems that evolve from cooled lava. It is demonstrated that the geometric and topological evolution of these cracks is a journey towards minimization of system energy.

The second part of my dissertation is concerned with changing patterns ob-

served in the dynamical state of systems. The ‘tracers’ that help track the changing patterns with time are particles of length scales much smaller than the length scales of flow lines in these systems. This ensures that the particles do not significantly affect the flow dynamics itself. In my dissertation I demonstrate that flow patterns can be mapped as functions of time and resolution scales of observation on 2-dimensional topological surfaces proposed in this work. Streamlines and vortices leave their signatures on these surfaces leading to the quest of why and how these changes occur.

Once dynamical systems are captured on their ‘characteristic surfaces’, one can distinguish between dynamical systems via a metric measure proposed in this dissertation. Questions of stability of flow systems to small perturbations have also been addressed using the topological maps and measures introduced in this dissertation.

Finally I have discussed the merits and demerits of the topological analysis developed in this dissertation in comparison to other existing techniques often used to characterize disordered dynamical systems.

In the following chapters I shall first discuss the static patterns left behind by disordered dynamical systems that have reached maturation and their characterization in terms of combinatorics of topology and geometry. This will be followed by characterizing flow patterns in disordered dynamical systems using topological descriptors. Lastly we shall build the mathematical stability of our topological methods and conclude its merits and demerits.

Chapter 2

Studying combinatorial topology of crack networks: A 4-parameter space

In nature, various patterns exist generated by physical dynamic processes. Natural crack networks are one such genre where we observe intriguing patterns/ tessellations. For more than decades scientists have been involved in exploring the mechanics of crack formation and its network progression [30, 31, 32, 33, 34, 35, 36, 37]. Studies also include fluid transport through crack networks [38, 39, 40, 41], scaling laws in fracture interfaces [42, 43, 44, 45], and mechanisms of failure [46, 47, 48, 49]. In every case crack morphology plays a non-trivial role and therefore different measures have been used to characterize patterns in crack networks, integral geometry based Minkowski function [50, 51, 52, 53, 54, 55], fractal dimension analysis [56, 57, 58, 59, 60, 61], summary statistics such as distance characteristics, spherical contact distribution functions, and the J-function, as well as second-order characteristics like the two-point correlation function, Ripley's K-function [62], and the L-function, have generally been utilized [63, 64, 65, 66, 67, 68]. Mecke et. al introduced a morphological description of a triplet function that included normalized values of integral-geometric quantities such as area, boundary length, and Euler number of patterns of discs centered on the points of a stationary point distribution [69]. This approach is similar to that of Adler [70] and Worsley [71, 72, 73]. Andresen et al. [74] analyzed the topology of 3-dimensional fractured systems as an abstract map of

nodes and links using network theory tools. Hope et al. [75] studied the topology of 3-D fracture models including a Poissonian discrete fracture model and a mechanical discrete fracture model.

In this work, we look at the crack networks from geometrical and topological aspects and try to find whether the geometry of these crack patterns carries the signature of the physical desiccation process. Here, we are not concerned with a single crack but rather with networks of cracks. These networks may often form distinctive patterns, such as those seen in mud cracks beside a dried riverbed or in a shattered glass pane. A 2-d crack network consists of solid polygonal shapes which are called “peds”, separated by narrow gaps which are the “cracks”. The images of the final pattern of such real crack networks which we consider as a tessellation of the Euclidean plane have been analyzed focusing on two aspects: (i) the geometry, which includes the shapes and relative sizes of the peds and how they depend on various factors that create the crack pattern, and (ii) the topology, which concerns the connectivity of the pattern - how the adjacent peds connect with each other and how the crack network pervades the whole system.

We propose a 4-tuple (n, v, D, λ) to classify planar surface crack networks, including both convex and non-convex polygons, as detailed later in Section 2.1. Briefly, the first two elements of the 4-tuple are the average number of nodes and vertices of the network, forming an (n, v) pair that determines the topological connectivity of the network. This information encodes connectivity similarly to the coordination number in sphere packings [76]. The angular defect D measures average geometric regularity by comparing the deviation in polygonal angles to those of a regular polygon with the same number of sides. Finally, λ , the normalized isoperimetric ratio, quantifies the polygonal shape.

We expect that crack networks in similar materials, or with similar modes of cracking [77], or similar underlying mechanisms, will cluster in this 4-dimensional space. With sufficient data on crack networks, a ‘well-stocked’ 4-parameter phase space could be constructed. Ideally, this would act as a calibration space for any crack network, allowing it to be placed as a point in one of the classified clusters. This type of crack network calibration would provide immediate insights into material information, crack modes, and the responsible mechanisms for an unidentified crack pattern. The time evolution of a crack network would trace a trajectory in this phase space, providing key insights into the underlying principles responsible for such transitions.

In the next Section 2.1, after a brief discussion on cracks and mosaics, we introduce the theory for analyzing the topology and geometry of planar polygonal mosaics. Section 2.2 describes image processing and analysis of real crack systems, followed by comparing real crack and simulated crack networks in section 2.4. Section 2.6 highlights the findings of the study and finally, in Section 2.7, the conclusions of the work have been summarized.

2.1 Background on planar mosaics

As the study is based on the crack patterns on plane only that resemble regular mosaic-like tessellation, a brief description of the classical combinatorics used to study planar mosaics is given at first. The idea of considering not only the combinatorial topology of a mosaic but also its geometry follows after that where two geometrical measures are introduced. Finally, the 4-parameter phase space of topology and geometry is manifested and is used to study a given set of crack patterns of different materials.

2.1.1 The topological combinatorics (n,v) and $n-v$ plane :

There exist comprehensive studies on the mosaics of convex polygons on the Euclidean plane [78]. Two-dimensional mosaics are also called planar tessellations or tilings. In a combinatorial sense, a mosaic is a countable set of compact regions that tessellate the plane such that the regions intersect pairwise only at their boundaries. In the theory, the regions are idealized to be convex polygonal. However, the real cracks, natural or experimental, this restrictive assumption often does not hold in some cases. The classical theory is based on the non-trivial assumption of the polygons being convex only. In standard terminology of combinatorial topology, a zero-dimensional face of a polygonal region is called a vertex, and a one-dimensional face is called an edge. The vertices of the polygons are known as the nodes or junction points of the mosaic. Not every node (junction point) on the boundary of a polygon is necessarily a vertex of that polygon. If a node is a vertex of n cells/polygons, it is said to have degree n . A polygon/cell is assigned a degree of v if it has v vertices. In a mosaic, there can exist two kinds of nodes, (i) Regular nodes, and (ii) Irregular nodes.

A node is called regular if it is the vertex of all the cells on whose boundary it lies. Figure (2.1a) shows tilings where all nodes are regular. A node is not

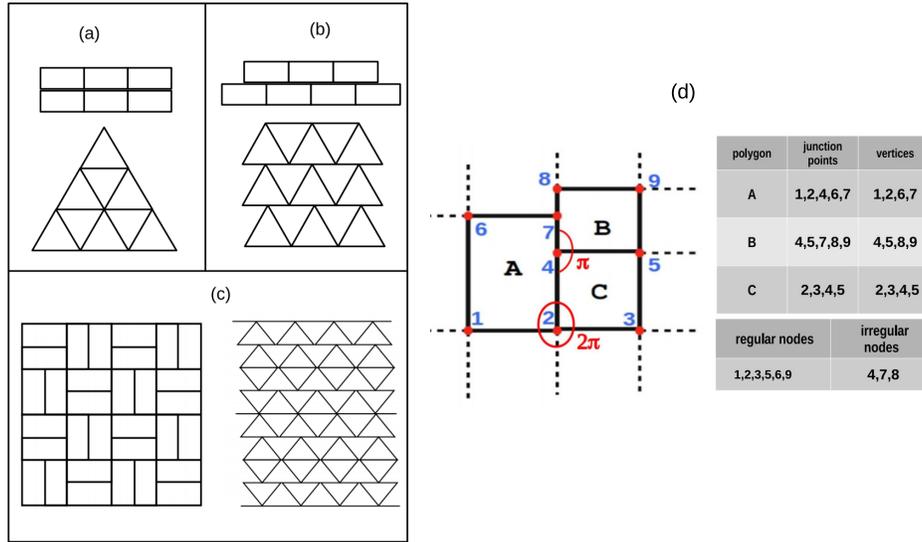


Figure 2.1: Rectangular and Triangular lattices with - (a) all “regular nodes”, (b) with all “irregular nodes”, (c) with both regular and irregular nodes present, (d) A, B, and C are polygons in a schematic mosaic, nodes marked by numbers 1 - 9; tabulation of nodes and vertices of each polygon; tabulation of regular and irregular nodes of the mosaic.

regular when it is not the vertex for any of the cells on whose boundary it lies. It is then called an irregular node of the mosaic, as represented in Figure (2.1b) where all nodes are irregular in the tilings. Figure (2.1c) presents tilings with both regular and irregular nodes. Figure (2.1d) showcases how nodes and vertices are marked, and the distinction between regular and irregular nodes. A schematic of a polygonal mosaic in which 3 polygons A, B, and C are highlighted is shown. The nodes (junction points) are numbered from 1 to 9. Given a planar convex mosaic, we consider the number average values of the degrees n and v , \bar{n} and \bar{v} respectively, over the entire mosaic. If the mosaic is infinite, we consider the limits of the averages of these degrees over planar disks with radius going to infinity. Finally, the regularity index p of the mosaic is measured by the ratio of regular nodes to the total number of nodes in the mosaic.

Any convex polygonal normal mosaic must lie in a compact combinatorial domain in the (n, v) plane.

It can be easily shown that

$$\frac{p+1}{n} + \frac{2}{v} = 1, \quad (2.1)$$

where $v \geq 3$ is the total number of vertices of a polygon. This relation gives a natural combinatorial classification of mosaics by their p -regularity, referred to as iso- p lines.

For the values of $p = 0$ and $p = 1$, the above gives two more natural combinatorial curved boundaries for the (n, v) domain. Moreover, for any regular node $n \geq 3$, and for any irregular node $n \geq 2$, Figure(2.1). Thus the average number of polygons per node n , of the mosaic must also satisfy the minimum number $p \cdot 3 + (1 - p) \cdot 2 = p + 2$ or,

$$n \geq p + 2. \quad (2.2)$$

Combining Eq.(2.1) with the condition Eq. (2.2), the final boundary of $v \leq 2n$ is defined. This defines the 2-dimensional (n, v) plane for all convex polygonal normal mosaics, schematically shown in Figure (2.2a). Domokos et al. [79] carried out extensive fieldwork on fractured materials and concluded that there were primarily two regions on the allowed (n, v) map where patterns generated by the natural fragmentation of rocks and soil tended to cluster.

A group of natural crack mosaics could be classified as having (n, v) ranging from (2, 4) to (4, 4). This implied rectangular patterns, ranging from irregular to regular, as shown in the schematic Figure (2.2a). Domokos et al. considered rectangular or cubic symmetry as the ‘Platonic attractor’. Another group of natural crack mosaics had (n, v) values clustered around (3, 6), implying hexagonal symmetry which is referred to as the ‘Voronoi attractor’.

2.1.2 Introduction of geometrical measures

As the combinatorial measures (n, v) do not differentiate between cracks from different materials, the question was how materialistic individuality can be reflected in the patterns. This led to considering to study the geometry of the crack networks. Two geometrical measures were used - (i) a measure of geometric regularity of a mosaic in the form of its non-dimensionalized “angular defect”, (ii) a measure of regularity in the shape of the polygons in the form of non-dimensionalized and normalized Isoperimetric ratio.

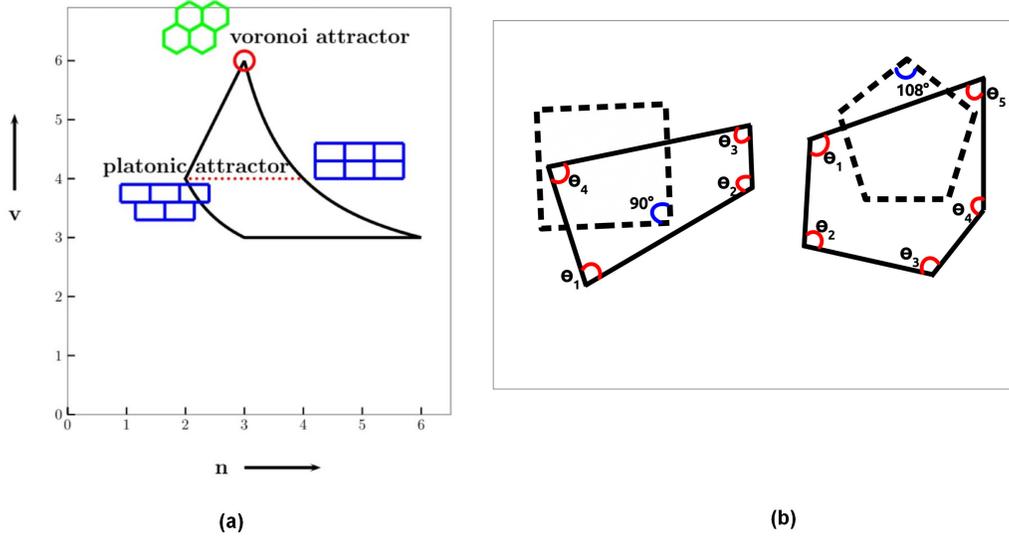


Figure 2.2: (a) (n,v) phase diagram. The range (2,4) to (4,4) i.e. bricklayer type to square-tiled type patterns, cluster around ‘Platonic attractor’. Hexagonal tiling pattern with (3,6), cluster around the ‘Voronoi attractor’. (b) θ_i for a 4-sided polygon and a 5-sided polygon. Corresponding regular polygons are shown in dotted lines.

The angular defect D and the (n, v, D) space: The i th angle of a polygon, θ_i is a function of the angular departure of the i^{th} vertex of a polygon with respect to the internal angle of a regular polygon of the same number of sides, Figure (2.2b). For the N^{th} polygon having v_N number of vertices, this takes the form

$$D_N = \frac{1}{\sum_{i=1}^{v_N} |\theta_i - \frac{(v_N-2)\pi}{v_N}| + 1}, \quad (2.3)$$

Thus, The measure of the geometric regularity in terms of the defect in the internal angle of the entire mosaic is given by

$$D = \frac{\sum_{i=1}^M D_i}{M}. \quad (2.4)$$

where M is the total number of polygons (cells) in the mosaic. By construction, these measures D_N and D are in $[0, 1]$, i.e. normalized. To combine both the combinatorial and geometric qualities, a three-dimensional (n, v, D) space is constructed wherein all our convex planar mosaics must lie.

Normalized Isoperimetric ratio λ and the (n, v, λ) space: The angles alone do not fully specify the polygon. A polygon having the same angles can be constructed with a larger or smaller area and perimeter. It is therefore important to include the diversity in the polygonal shapes. According to the classical isoperimetric inequality [80] for a planar simple closed curve of length L , having enclosed area A , $L^2 \geq 4\pi A$. The shape of the polygon is thus reflected upon its perimeter and area and thus the attempt to quantify the regularity in the shape of the polygon was done with another non-dimensionalized measure, the “normalized isoperimetric ratio” $\lambda = \frac{4\pi A}{L^2}$ (averaged out over all peds present in a mosaic). A circle has $\lambda = 1$.

The isoperimetric ratio for a regular N -gon can also be written as $\lambda = \frac{\frac{\pi}{N}}{\tan(\frac{\pi}{N})}$. As the number of vertices/edges N increases, λ converges to 1. Thus, when λ is close to 1, the shape of polygons is nearly symmetric or round and for polygons having more elongated shapes it is close to 0. Similar to the case of an angular defect, for the N^{th} polygon, λ_N is calculated, and then the values of all λ_N s are averaged out over the entire mosaic.

$$\lambda = \frac{\sum_{i=1}^M \lambda_i}{M} \quad (2.5)$$

where M is the total number of polygons (cells) in the mosaic. The three-dimensional space (n, v, λ) is constructed along with the (n, v, D) space described earlier.

2.1.3 The parameter space of quadruples (n, v, D, λ)

Thus the idea of combinatorial (n, v) was expanded to a space of the quadruples (n, v, D, λ) , which are points in a subset of \mathbb{R}^4 . The purpose behind this construction was to characterize crack mosaics successfully by their topology and geometry. For low dimensional classification and visual representation of the 4-parameter space, the crack mosaics shall be represented as points in two separate 3-dimensional spaces - the space of points (n, v, D) and the space of points (n, v, λ) .

2.2 Analysis work-flow

With the established parameter tuple (n, v, D, λ) , crack mosaics from different physical systems were analyzed to examine how topological and geometrical characteristics of the crack systems influence their classification in the (n, v, D) and (n, v, λ) space. The intention behind the study is to find out if crack mosaics of similar materials, or having similar dynamics of cracking exist in the same place in the (n, v, D, λ) space,

as well as to compare the position of the real crack mosaic with that from simulated crack mosaics in the (n, v, D, λ) . If the objective to distinguish crack mosaics in terms of the combined 4-parameter tuple (n, v, D, λ) gets accomplished, it may aid to identify the material of given crack mosaic from calibration in 3-dimensional (n, v, D) and (n, v, λ) spaces, provided we build a set of enough samples of crack mosaics of different materials with their (n, v, D, λ) values. The analogy between real and simulated cracks may help to understand the crack mechanisms following the geometric algorithms that guide the crack simulation.

2.2.1 Extracting crack skeleton from images

The set of images that were studied were either experimental crack patterns from our lab [81] and from the articles [82, 83, 84, 2, 85] or natural crack mosaics collected from [86, 87]. The challenging part of the study was the computational process of extracting information from images to estimate the n, v, D, λ values. The computational approach is discussed below.

The image analysis starts with proper gray-scale thresholding of the images. It was done through in-house coding in conjunction with suitable Python library functions. For most of the images of 2-dimensional planar real crack networks, a thresholding algorithm was made following Otsu's binarization [88]. After we set up a binary bit-map of the images, the crack mosaic was skeletonized using standard Python libraries [89], Figure (2.3 a-c). The algorithm of crack skeleton extraction is based on the idea of joining the mid-points of crack widths and therefore, a slight change in the shape of the system boundary may be introduced during the process. For a large crack mosaic, one may neglect a boundary layer of polygons to circumvent this problem. The crack skeleton has a thickness of a single pixel and effectively constitutes the mosaic for all analysis henceforth. If a black pixel on the skeleton is shared with more than two white pixels in its δ -neighbourhood, it is identified as a node or junction point, Figure (2.3d). All the nodes or junction points of the skeletonized mosaic are thus identified and tagged. The mosaic polygons (painted white) are identified using the Hoshen-Kopelman algorithm for cluster detection and labelled.¹ [90], Figure (2.3e), the cracks are marked in black.

¹This is a variant of the Union-Find class of algorithms commonly used in computer science.

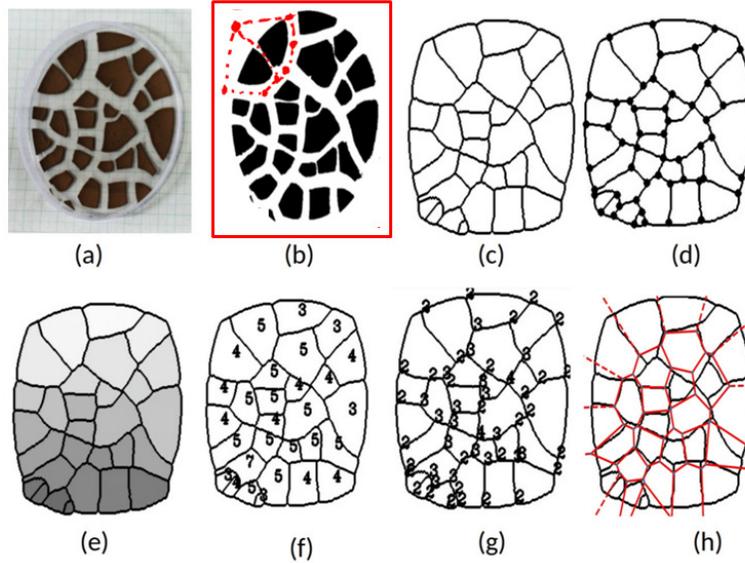


Figure 2.3: Sequential steps towards determination of (n, v, D, λ) data. (a) A desiccated layer of bentonite clay. (b) Binary image after suitable filtration of noise with a schematic of crack skeleton extraction with respect to mid-points of crack width. (c) Skeletonized image. (d) All nodes (junction points) detected on the skeleton. (e) All polygons identified and labelled. (f) Vertices v determined and marked on each polygon. (g) Nodes n determined and marked. (h) Superposition of Voronoi mosaic constructed (red edges) from centroids of real crack polygons.

2.2.2 Identification of nodes: calculation of angles at nodes

As explained earlier, all nodes in a polygon may not be its vertices. A polygon or ped in the crack mosaic may contain regular and irregular nodes. The regular node is certainly its vertex but the irregular node may or may not be its vertex. Correctly identifying the vertices of any polygon gives the values of v . For a convex polygon, the angle made by irregular node which is not its vertex, is always 180° . Therefore, angles formed by three consecutive junction points of a polygon are measured. However, crack mosaics of real systems can contain both convex and non-convex polygons and the algorithm holds only for convex cases. Hence, it is important to check the convexity of the polygon being considered for analysis. For this, using the ‘gift wrapping algorithm’ [88], the convex hull of all nodes of the polygon is drawn. If even one node of the polygon is not contained on the convex hull, the polygon is identified as non-convex. In our samples, a few mosaics contained

non-convex polygons but the numbers were one or two per mosaic. The non-convex polygons were labelled differently and the angles were calculated differently with active supervision from visual estimation. Sequencing nodes in a non-convex polygon by a general algorithm is an open problem to date, that has not been successfully solved yet!

Characterizing convex polygons in a crack mosaic

For convex polygons, the first step is tagging every node to each polygon that contains it in its boundary. The centroid of every polygon is determined. The polygon is then traversed along its edges following a fixed direction, either clockwise or anticlockwise. In this process, the angle that the polygon centroid makes at every junction point is calculated. Our algorithm identifies a node (junction point) not as a vertex of a particular polygon if the angle θ between two consecutive edges of the polygon is $180^\circ \pm \epsilon$, where $\epsilon = 15^\circ$. This process establishes the vertex-edge connectivity of every polygon. If a node (junction point) is not a vertex of the polygon, it is tagged as an irregular node of the mosaic.

Characterizing non-convex polygons in a crack mosaic

The problem of characterizing non-convex polygons in a mosaic is well-known to be messy and an open challenge for a general solution. If the degree of non-convexity is small, most often the problem of non-convex polygons is bypassed by replacing them with the approximately closest convex polygons. In many cases, this problem is avoided by simply neglecting the contribution of non-convex polygons to the statistics, especially if the ratio of the number of non-convex to convex polygons is very small. However, the experimental crack meshes studied for the work are finite-sized. Therefore any approximation or neglect of non-convexity can be expensive and lead to poor statistics. Hence the contribution of non-convex polygons in determining the (n, v, D, λ) point of a crack mosaic was included.

The process that was adopted for identifying non-convex polygons in the mosaic and analyzing them to find (n, v, D, λ) data is described here. The vertex-edge connectivity is based on the idea that if two neighboring polygons share the same nodes or junction points, then those nodes may be connected via an edge. If the number of common nodes between two neighboring polygons is two, then those two nodes are surely connected via an edge but if the number of shared junction points is more than two, the vertex-edge connectivity of those shared points can not be estimated clearly if either one of them or both are non-convex. For example in

Figure (2.4b), polygon 1 is a nonconvex polygon, hence its vertices(edge vectors) can not be sorted in a particular direction(clockwise/anti-clockwise) using the previously used convex hull approach. So, to estimate the angle at each junction point of a non-convex polygon, the process described below was followed. First, its neighbor polygons and sets of all the points it shares with each of its neighbors are identified i.e. (A, B, C) with polygon 2 ; (C, D) with polygon 3 ; (D, E) with polygon 4; (E, F) with polygon 5, (F,G, H) with polygon 6, (H, I) with polygon 7 and (I, A) with polygon 8. These sets are called ‘common points’ here. Now, to calculate the angle at a point, another set is constructed that contains all the points that are likely to be connected with the point considered via an edge, defined by the union of the “common points” sets containing that point as an element. These sets are called ‘connectivity sets’ for the point. For example, in Figure (2.4b) for point A, this set is $(A, B, C) \cup (I, A) = (A, B, C, I)$; for B, it is (A, B, C); for C, it is (A, B, C, D); for D, it is (C, D, E); for E, it is (E, D, F); for F, it is (F, E, G, H); for G, it is (F, G, H); for H, it is (H, G, F, I); for I, it is (I, H, A). If a ‘connectivity set’ contains 3 points, then there is no problem in estimating vertex-edge connectivity and hence in angle calculation which is done from the dot product between the corresponding edge vectors. For example, the angle at D is calculated from the dot product of edge CD and DE (sorted counterclockwise). The same applies to angle calculation at points E, I, B. Whereas, if the connectivity set for a point contains more than 3 elements, sorting points simply by orientation would not yield the right result all the time. For example, to calculate the angle at C, sorting A, B, C, and D counterclockwise would not yield the right vertex-edge connectivity. To tackle these cases, the points of the ‘connectivity set’ are sorted with respect to their distances from the point considered for angle calculation; (here, distance of C from points A, B, and D), and thus BC and CD are considered as edge and angle at C is calculated from $\angle BCD$). In a similar fashion, angles at F and H are calculated by considering their nearest neighbors in their corresponding ‘connectivity sets’ and thus, by angle $\angle EFG$ and $\angle GHI$. again. Surely, this is not a stable fix for defining edge-connectivity in non-convex polygons but it worked for the mosaics used for analysis here.

Angle approximation is a non-trivial step in the algorithm because it determines whether a junction point is also a vertex of a polygon or not. If the angle at a point is $180 \pm 15^\circ$ then it is not considered as a vertex of that polygon. However, the same point can be the vertex of other neighboring polygons where the angle at that point is not $180 \pm 15^\circ$. The point then acts as an irregular node of the network.

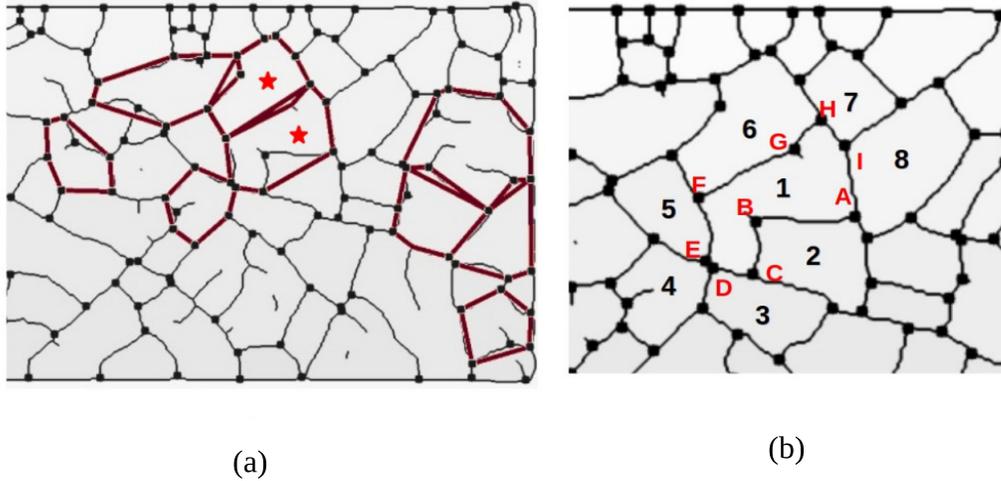


Figure 2.4: (a) Convex hull constructed on some polygons of the mesh. Red stars indicate non-convex polygons determined from the construction. (b) Schematic guide to the determination of internal angles of non-convex polygons. Image has been adapted from [2]

2.2.3 Simulated crack mosaics

Simultaneously with me working in the analysis of real crack images, one of my colleagues also worked on simulating geometric planar tessellations. Broadly, three different kinds of tiling algorithms were used- (a) Gilbert tessellation, (b) Voronoi tessellation, (c) Iterative cell division. Different variations in parameters for each case were studied. Every crack mosaic was examined for determination of the average n, v, D, λ over 50 configurations. The average mosaic characteristic is then plotted as a representative point in the (n, v, D) and (n, v, λ) spaces, and analyzed and compared in terms of their features.

Gilbert tessellation and its variations:

In the Gilbert tessellation, first, a set of points were chosen with a random distribution. These points were the seeds through which crack lines were drawn with randomly chosen angles between a range. The lines were extended on both sides, till they reached the boundary or another line. The variations that were applied - (1) both uniform and normal random distribution functions for seeds, (2) Different sets of angles for the slope of the crack lines. The choices of crack slopes were - (a) random choice between 0 – 180 degrees, (b) slopes parallel to reference axes, i.e. 0 or 90 degrees, (c) slopes chosen randomly from 0, 45 and 90 degrees and (d) slope determined randomly between 0 ± 10 degrees and 90 ± 10 degrees. Figure (2.5a) displays the results for a uniform and normal random distribution of crack seeds respectively for Gilbert tessellation on a 2-dimensional plane.

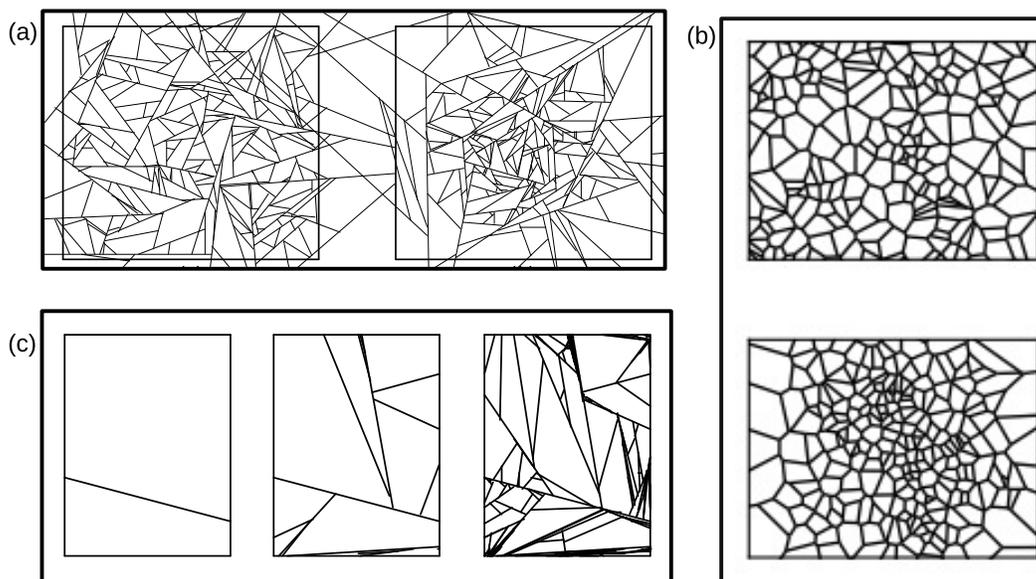


Figure 2.5: (a) Simulated Gilbert mosaics with random uniform and random normal seed distribution, (b) Simulated Voronoi mosaics with random uniform and random normal seed distribution, (c) Simulated mosaics with iterative cell divisions at the end of 1st, 4th and 8th iterations respectively.

Voronoi tessellation

Voronoi tessellation, also known as Voronoi diagram or Voronoi partition, is a method of dividing a plane into regions based on distance to a specific set of points. Given a set of seed points, the plane is divided such that each region contains all the points closer to one seed point than to any other. Each region is called a Voronoi cell. The algorithm starts with creating a set of random seeds, followed by drawing the lines that bisect imaginary lines joining pairs of nearest neighboring seeds. These lines form a set of close-packed convex Voronoi cells making all points in a cell closest to the seed of that particular cell. This enhances the ‘roundness’ of the cells, i.e. makes λ closer to 1 than 0. Simulation of Voronoi tessellation was done following the Python class ‘Voronoi’[91]. Both the uniform and normal random distribution of seeds were chosen, Figure (2.5b), and the average n, v, D, λ were estimated along with the statistics. The details are given in [92].

Iterative Cell Division

Another simple simulation algorithm the Iterative Cell Division (ICD) was also tested. The algorithm starts with outlining a square area. Two points are randomly chosen on any two sides taken randomly and joined by a straight line, producing 2 daughter cells. The process is repeated on each daughter cell to get 4 cells, and so on as long as desired. Figure (2.5 c) displays the crack mosaic development after the 1st, 4th, and 8th iterative steps respectively, the newest crack lines are indicated by broken lines. These type of mosaics in cracks is not very commonly seen.

2.3 Observation from the analysis

After successfully setting up the vertex-edge connectivity in the crack skeleton, the outcome of the analysis was mapped into the topological-geometrical (n, v, D, λ) space. The analysis was sectioned broadly into three parts- (a) the $n - v$ plot of the studied mosaics, (b) portraying the mosaics in n, v, D and n, v, λ space, (c) comparing the results of real mosaics with that of simulated ones.

2.3.1 Mapping the topology of the crack mosaic on the (n, v) plane

Once the nodal values for each node and number of vertices of every polygon were calculated, the average of the nodal values, n , and average number of vertices per

polygon v were determined for each crack mosaic. Figure (2.3 f and g) displays respectively, the degrees v for each polygon and the n -value for each node on the sample crack mosaic of bentonite clay. Figure (2.6) displays the (n, v) points of all real and simulated crack mosaics studied. Following equation 2.1, the curves for a specific p value can be drawn in the $n - v$ plane. In Figure (2.6) for $p = 0.5$ and $p = 0.3$ the iso- p lines are drawn which helps to roughly estimate the p value of the mosaics. Remember, the p value gives the fraction of regular nodes present in these crack mosaics. As the plot displays that most of the (n, v) values of the real crack mosaics cluster between $p = 0.3$ to $p = 0.5$, confirming that natural crack mosaics are commonly irregular. At the scales displayed in the figure, the notable point is that most of the data points cluster around a small region in the $n - v$ plane, matching with the observation made by [79]. The crack mosaics that were studied, included different types of materials - natural mud, natural clay as in Bentonite, synthetic clay as in Laponite, corn starch, resin, glass, and metal oxide films. The (n, v) analysis results in their similar nature in the combinatorial topology domain. Therefore, the (n, v) measures alone are not sufficient to distinguish among cracks of different materials. It is to be noted in the Figure(2.6), that a few points are a little outside the theoretical domain for the mosaics in the $n - v$ plane, because we did consider the non-convex polygons as well if presented in the real crack mosaic. Also, real cracks are inhomogeneous and disordered, and a good measure of n, v requires preferably a large polygonal mesh to diminish the boundary effect.

2.3.2 Refining by introducing geometric measures - the (n, v, D, λ) space

As the purpose behind the study was how the difference in chemical and physical properties of materials manifest in the crack mosaics, the idea to include geometrical characteristics of the crack network came into play. The result of the analysis on the extended (n, v, D, λ) space is presented next. The parameter D , defined in Eq.(2.4), which measures the average angular defect, was estimated following the estimation of angular defect for each node, Eq.2.3. The method is discussed in 2.2.2. The other measure, the normalized iso-perimetric ratio of the mosaic, λ , was also estimated following Eqs. (2.5). Combining all measures, a crack mosaic was then represented as a point in the (n, v, D, λ) space.

Crack mosaics on the (n, v, D) and (n, v, λ) spaces: To pictorially represent the result of the 4-parameter tuple (n, v, D, λ) for each crack mosaic, 3-dimensional (n, v, D) and (n, v, λ) graphs were plotted. Figure (2.7) displays the points for the

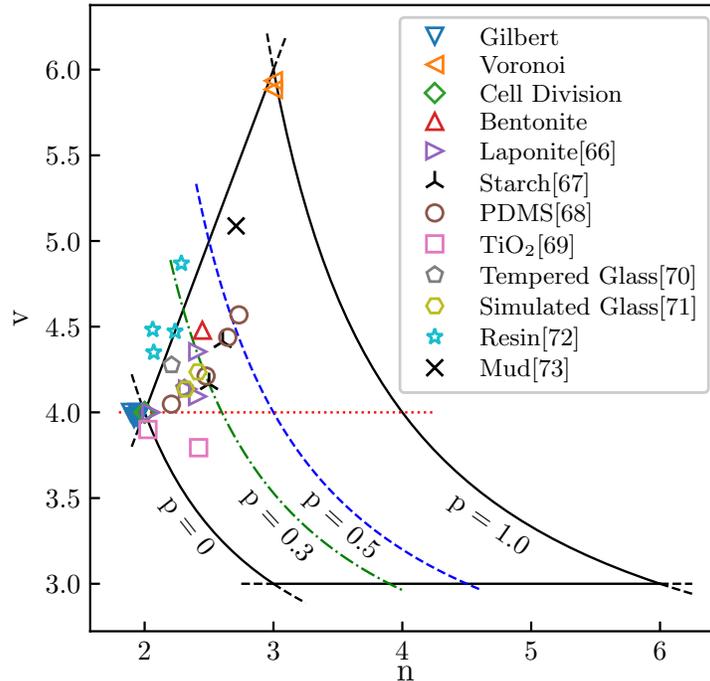


Figure 2.6: (n, v) data of real and simulated crack mosaics. The legend indicates the source of the experimental crack mesh. The iso- p lines at $p = 0.3$ and $p = 0.5$ are drawn as a guide to the eye.

studied crack mosaics, both real and simulated, in the (n, v, D) space. Almost all the data points are contained inside the allowed space defined analytically. However, a few points appear a bit outside the allowed (n, v, D) space, Figure (2.7a). The experimental cracks reported in this paper have not always been measured on large enough systems, thereby bringing in unwarranted boundary effects. While boundary effects cannot be avoided in real systems, it is desirable that the ratio of the number of boundary polygons to inner polygons, be as small as possible, and the number of polygons in the mesh be as large as possible to give robust statistics. The real crack mosaics examined here, throw up points that are clustered around $n = 2$ (a little foe than $n = 2$), v between 3.5 and 4, and D values between 0.38 to 0.46. Figure (2.6) shows that for $n = 2$, only allowed v is 4, whatever be D and λ . For Simulated crack mosaics, points corresponding to the Gilbert and the Iterative cell division are also near $n = 2$ and $v = 4$ with the Gilbert mosaics having higher values of D than the

Cell division ones. For the Voronoi tessellation, the points are positioned near $n = 3$ and $v = 5.9$ with $D \approx 0.3$. The Voronoi and Iterative cell division yields D values ~ 0.2 to 0.3 implying a high degree of irregular polygons which is also consistent with a small value of p . The variants of Gilbert tessellation show a clustering of D around 0.5 .

Similar to the (n, v, D) domain, the points of the respective crack mosaics on the (n, v, λ) space are displayed in Figure (2.7c). As one can see the points are not cluttered at the same region, but have different positions. Crack samples from similar systems have their (n, v, λ) data in clusters, distinguished by their colours. Figures (2.7b and d) are zoomed-in versions of Figures (2.7a and c) respectively, to highlight that crack mosaics of different materials are distinctly differentiable by the different values of their (n, v, D) and (n, v, λ) data in the 4-parameter phase space. To display the zoomed view of Figures (2.7a and c), data points are spaced out along the vertical axis that shows the variation in D and λ values prominently in Figure (2.7b and d) respectively. The data for similar materials have been averaged and represented as a single data point in these figures. Table. (2.1) displays the average geometrical measures for 22 real crack mosaics studied.

Similar analysis on simulated crack mosaics shows that both in the (n, v, D) and (n, v, λ) spaces, points corresponding to the Gilbert and Iterative Cell Division are roughly closed in a region whereas the Voronoi tessellation yields another cluster in the domains.

2.3.3 Including Hausdorff Measure for comparison

There was another metric that was used to compare between original image mosaics and their corresponding Voronoi mosaics and it was the Hausdorff distance d_H between the two network matrices. There are literature suggesting that many of the real crack mosaics in nature follow Voronoi-like tessellation [30, 5]. The Hausdorff distance d_H measures the greatest of all the distances from a point in one set to the closest point in the other set². Here, the two sets are - i) the real crack network and ii) the Voronoi network generated from the seeds that were the centroids of the polygons in the crack mosaic. The Hausdorff distance d_H was calculated using

²For compact subsets $A, B \subset \mathbb{R}^2$, the Hausdorff distance $d_H(A, B)$ is defined as $d_H(A, B) = \max\{\max_{a \in A} d(a, B), \max_{b \in B} d(b, A)\}$, where $d(x, C) = \inf\{\|x - c\| : c \in C\}$.

Table 2.1: Measures of real crack mosaics

Sample	n	v	d_H	D_N	average D_N	λ	average λ
Bentonite	2.442	4.479	30.36	0.463	0.463	0.739	0.739
Laponite (pH10) [81]	2.333	4.136	16.12	0.483	-	0.579	-
Laponite (pH13.05)	2.048	4.0	42.48	0.415	0.446	0.618	0.6179
Laponite (pH13.45)	2.410	4.093	33.61	0.484	-	0.625	-
Laponite (pH6.73)	2.411	4.353	14.86	0.467	-	0.647	-
Corn starch [82]	2.5	4.166	16.40	0.466	0.446	0.674	0.614
Potato starch [82]	2.606	4.411	22.36	0.425	-	0.554	-
PDMS [83]	2.731	4.569	19.41	0.428	-	0.618	-
PDMS [83]	2.644	4.438	20.25	0.445	0.454	0.587	0.609
PDMS [83]	2.477	4.212	21.37	0.453	-	0.612	-
PDMS [83]	2.207	4.047	33.3	0.490	-	0.619	-
TiO_2 (primary) [84]	2.025	3.9	46.87	0.415	-	0.513	-
TiO_2 (secondary) [84]	2.419	3.794	17.88	0.461	0.438	0.497	0.505
Tempered glass [85]	2.208	4.276	31.95	0.416	-	0.547	-
Tempered glass [85]	2.309	4.141	29.15	0.428	0.422	0.548	0.547
Simulated glass [86]	2.318	4.136	31.76	0.414	-	0.456	-
Simulated glass [86]	2.415	4.236	31.78	0.407	0.411	0.498	0.477
Resin [2]	2.284	4.868	29.83	0.410	-	0.789	-
Resin [2]	2.063	4.482	32.52	0.370	0.398	0.657	0.724
Resin [2]	2.236	4.472	38.6	0.393	-	0.675	-
Resin [2]	2.071	4.35	42.72	0.418	-	0.775	-
Mud [87]	2.271	5.088	28.17	0.379	0.379	0.633	0.633

modified Python library functions. Figure (2.8) displays the skeletonized real crack mosaic, and its Voronoi mosaic constructed from the centroids of the polygons. The Hausdorff distance from ‘skeleton to Voronoi’ $d_{H(sv)}$, is the distance between a pair of points that are the farthest from crack skeleton to the Voronoi matrix, indicated in Figure (2.8c). The Hausdorff distance between ‘Voronoi to skeleton’, $d_{H(vs)}$, is the distance between the pair of points that again are the farthest from the Voronoi matrix to the skeleton matrix, indicated in Figure (2.8d). These two distances usually are not equal for inhomogeneous structures. The final measure of d_H is the bigger of the two Hausdorff values.

Table. (2.1), fourth column displays the Hausdorff distance of the crack mosaics studied with respect to their corresponding Voronoi mosaics.

2.4 Comparing real and simulated crack mosaics

Comparing the topological characteristics in the $n - v$ plane suggests that most of the real crack mosaics studied were more towards the Gilbert tilings and the Iterative cell division, $\sim (2, 4)$, except for the mud crack which was close to the Voronoi tessellation, $\sim (3, 6)$. Figures (2.7b and d) display the geometric measures D and λ respectively for all experimental and simulated cracks. Real crack systems (experimental or natural) may thus be compared with the appropriate tessellation. Most of the real cracks are thus more close to the Gilbert mosaic and not to the cell division. Further, comparing the histograms of both the geometric measures, the angular defect D and the isoperimetric ratio λ , the latter seems more efficient in comparing between real and simulated crack mosaics. Figure (2.9) displays the histograms of the isoperimetric ratio λ for a few crack mosaics of real systems studied and Figure (2.10) displays the same for all the variations of the Gilbert and the Voronoi type tessellations. For Gilbert tessellation, the shapes of the polygons and thus the distribution of λ in the mosaic depends largely on the slope that a growing crack makes with respect to a reference line and hence, Figures (2.10a-d), show a broad dispersion even for the same kind of crack model.

Comparison of the distribution of λ values for real crack mosaics in Figure (2.9) with that of simulated cracks in Figure (2.10) suggests that the mosaics of TiO_2 , PDMS and Laponite at $pH = 6.73$ follow with the distribution of the Gilbert tessellation with random slopes as in Figure (2.10a). Whereas, Laponite at $pH = 13.45$ has a λ distribution that is almost similar to the λ distribution of the Gilbert

distribution with crack slopes tilted at 45 and 90 degrees only, Figure (2.10c). The mosaic corresponding to the tempered glass, Figure (2.9e), has a close resemblance to the bimodal distribution of Gilbert tessellation with slopes oriented at 0 ± 10 , and 90 ± 10 degrees, Figure (2.10d). Similarly one may identify the λ distribution of crack mosaics of resin, corn starch, and mud of Figures (2.10f, g, and h) with the distribution observed in Voronoi mosaics of Figures (2.10e and f). This comparison may suggest that the development of these real crack mosaics may have followed a mechanism similar to their simulated counterparts.

The Hausdorff measure d_H between the crack mosaic and its corresponding Voronoi mosaic, given in Table.2.1, can also be used to compare the real crack systems with the Voronoi tessellation. Of all the materials for which crack mosaics were analyzed, corn starch showed the closest match with Voronoi cracks with $d_H = 16$. This supports the finding that in laboratory experiments corn starch shows columnar joint-like cracking with a hexagonal Voronoi-type crack mosaic in the transverse plane [93]. Thus studying crack patterns with the 4-parameter n, v, D, λ not only helps to fulfill the purpose of characterizing different crack mosaics in terms of their materials but also the qualitative comparison of the mosaics with their λ distribution and the quantitative comparison with the Hausdorff distance helps to have an idea about the tiling mechanism that the crack system follows.

2.5 Evolution of crack pattern in (n, v, D, λ) space:

The successful outcome of the (n, v, D, λ) measure further encouraged the idea of observing the time variation of crack mosaics or the dynamical progression of cracks in the (n, v, D) and (n, v, λ) spaces. In nature, there exist examples of crack tilings that change their shapes gradually, finally settling to a stable tiling pattern [94, 30]. Columnar joints observed in lava basalt, and mud crack in river beds are such systems, Figure(2.11). The study to analyze static crack patterns (left out by the dynamical process of desiccation and crack propagation) was then upgraded by examining dynamical crack mosaics with respect to the combined topological and geometrical measures. A colleague in the lab worked on simulating crack formation in columnar joints following the thermal energy and elastic energy conservation principle. In columnar joints the molten volcanic lava cools and cracks. The process of fracturing advances downward making up new interfaces. These interfaces observed along the height of the columnar joints were studied as mosaics following the same n, v, D, λ measure.

The mechanism that guided the evolution of the cracks was primarily based on approaching thermal equilibrium in the system. Both convection and conduction processes were considered for heat transfer in the system. Usually, the vertical boundaries carry the convective flow of gas/vapor while the interior of the columns loses heat through conduction to the crack boundary. The heat flow is higher at the crack interface and irregularity of the column widths causing irregular thermal gradient results in movement of the crack tip. The crack tip moves towards the wider column where the thermal gradient is higher [5]. We studied the gradual evolution of tiling in these crack mosaics at different horizontal cross-sectional interfaces along the column length. These interfaces underwent a metamorphosis starting from a disordered-looking Gilbert tessellation to a well-ordered Voronoi tessellation. The simulation started with the initial mosaic or the topmost interface being a Gilbert mosaic or a Voronoi mosaic. The positions of the nodes in the crack mosaics were changed following the thermal-field gradient. Figure(2.12 a) illustrates the thermal flows in the columnar joints which in turn dictate the shift of crack nodes in each time-step/ horizontal section, Figure (2.12b). Details of the algorithm can be found in [95]. Variants were introduced in the simulation by considering different seed distributions, seed densities, and crack orientations. The time evolution of the crack interfaces finally resulted in a Voronoi tessellation for every kind of variation. Estimation of (n, v, D, λ) values at each time steps were done to map the evolution of these crack mosaics as a trajectory of the 4-parameter tuple in the geometry-topology domain. The analysis revealed that these trajectories in the (n, v, D) and (n, v, λ) spaces depend on the crack seed distribution and crack orientation of the columnar joints, Figure (2.13). An interesting result of the study was that an empirical relation between the energy of the crack system and the shape parameter, iso-perimetric ratio λ of the mosaic was proposed, Eq.(2.6).

$$E = \alpha\lambda^{-\beta} + c. \tag{2.6}$$

α and β are parameters and c is a system constant. The total system energy showed a power-law dependence on λ with the exponent ≈ 0.3 at the final time of crack maturation, regardless of the variation in the initial interface mosaic. The dynamical process of the system moving towards energy minimization was thus manifested in the topological-geometrical changes, in the form of evolving polygonal shapes along the cross-sections of the columnar joints system.

2.6 Discussion

This whole work of studying crack network characteristics in terms of topological and geometrical measures was a relatively new approach, apart from the work by Domokos et al.[79]. Images of real crack mosaics formed on widely different materials - natural mud, clays like Laponite and Bentonite, starch extracts of corn and potato, polymers such as PDMS, metal oxide films, glass, and resins, were analyzed with an attempt to sort them out on the basis of their topological and geometrical properties. The intention was to find out whether crack mosaics can be classified or distinguished according to their composite materials which the classical combinatorial $n - v$ plane could not. For that, the approach that was taken was to combine the system's geometric features along with the combinatorial topological measures. The approach was successful for the 22 mosaics studied. It was intuitive that the system's intrinsic properties, such as the nature of molecular bonding, physical or rheological properties, etc have to affect the geometry of the tiling somehow but quantification of that geometry through the new 4-parameter proposed, the n, v, D, λ was a novel approach of characterizing crack tilings. This 4-parameter tuple (n, v, D, λ) , carries the unique combinatorial-geometric footprint of the crack mosaic. The topological parameters v and n describe the vertex-edge connectivity of the cells of the crack mosaic and their connectivity to neighboring cells. The geometric parameters D and λ describe the regularity of the polygons and their shapes in the mosaic. The four-parameter space was represented graphically as two 3-dimensional spaces - the (n, v, D) space and the (n, v, λ) space. The study by Domokos et al. [79] indicated that fractures and cracks on geological systems mostly belong to the 'Platonic' category with $v = 4$ and some to the Voronoi category with $n = 3, v = 6$. Our study resonated with it. Considering only the combinatorial (n, v) aspect, our studied crack mosaics belonging to different materials, experimental and natural, (besides the geological ones studied in [79]), also fall approximately in either of these categories. Further, these crack mosaics from different material classes were found forming distinguishable clusters in our modified 4-parameter phase space. This implies that *materials having physical and chemical similarities tend to have similar geometry of cracks*. It is hypothesized if a huge amount of data on crack patterns covering a wider array of materials with multiple samples, gets provided in the 4-parameter (n, v, D, λ) phase space, a crack mosaic of unknown material may be identified from its position in the phase space.

Comparison of real crack mosaic with simulated crack mosaics(Gilbert, Voronoi, Cell division) through the shape parameter λ gives hindsight about the geometrical

tiling algorithm and the variation of seed distribution or slope of crack edges that the crack network resembled. Voronoi tessellation was used as a reference to compare with the real crack skeleton. Further, quantification of the deviation of the experimental crack mosaics from their corresponding Voronoi counterpart was done with the measure of the Hausdorff distance d_H between the two.

It may be noted that all data on cracks in real systems, including some experiments done by our group and some collected from other sources (that have been categorically cited in the Table 2.1) were processed from images, that had sometimes noisy or of insufficient resolution. The non-trivial part of this study was to construct suitable in-house codes for accurate data extraction to estimate the measures without errors along with tackling to address the vertex-edge connectivity in non-convex polygons that were present in the studied images.

2.7 Conclusion

The conclusions of the study may be summarised as follows:

1. It is not possible to classify planar mosaics solely on the basis of the topological combinatoric (n, v) , the geometric features also need to be taken into account.
2. The 4-parameter tuple, (n, v, D, λ) serves as a possible marker for a full description of the crack mosaic.
3. Representative points of natural and experimental crack mosaics, of materials with similar physical and/or chemical properties, are shown to form closely spaced clusters in the 3-dimensional spaces (n, v, D) and (n, v, λ) . This may aid in the identification of an unidentified material from its crack pattern.
4. Crack patterns simulated from well-studied models, Gilbert and Iterative cell division can be compared with real crack patterns with respect to the proposed measures to have an idea about the mechanism of crack formation.
5. In-house codes have been developed to estimate the topological and geometric measures of non-convex polygons, which appeared in the crack patterns.
6. This study inspired the idea of tracing the trajectory of time-development of a

crack mosaic in the topology-geometry architecture (n, v, D, λ) , later done in [95].

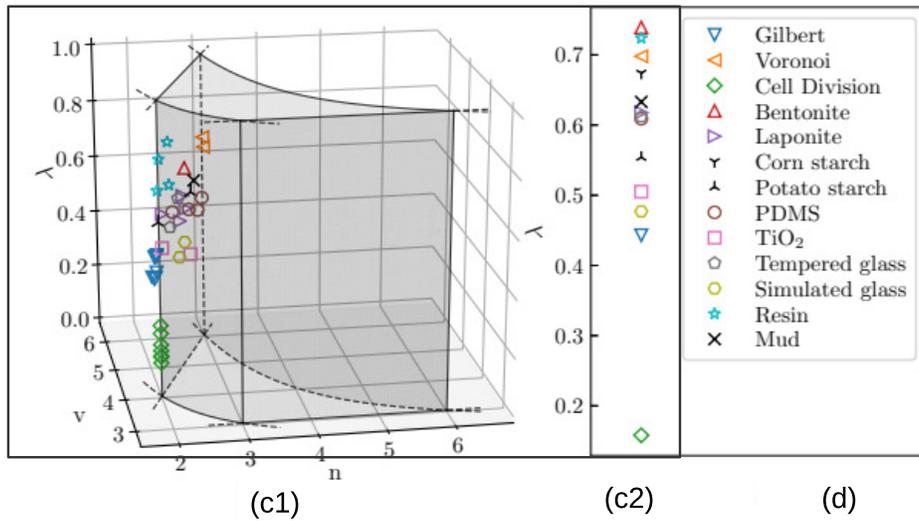
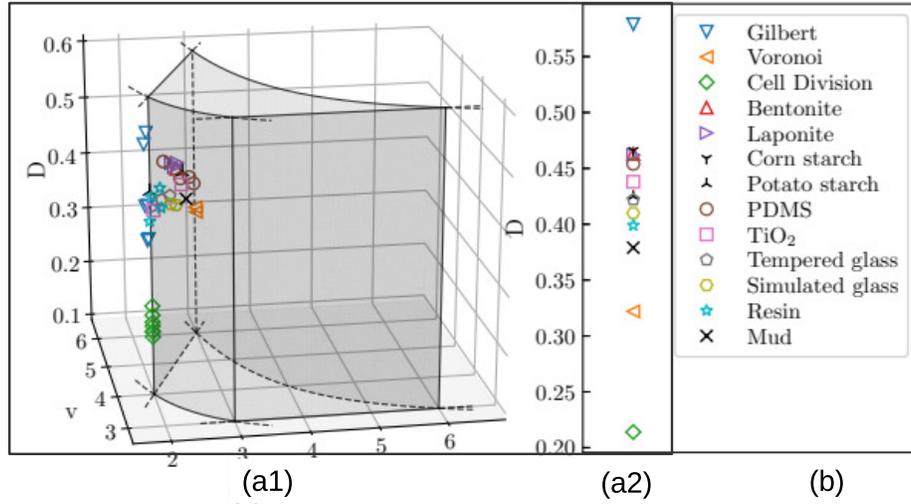


Figure 2.7: (a1) (n, v, D) space of real and simulated crack mosaics. (a2) Zoomed in view of (n, v, D) data. (c1) (n, v, λ) space of real and simulated crack mosaics. (c2) Zoomed in view of (n, v, λ) data for experimental and simulated crack mosaics. Description of data points provided in the legend. (b) legends for (a1), (a2) and (d) legends for (c1) and (c2).

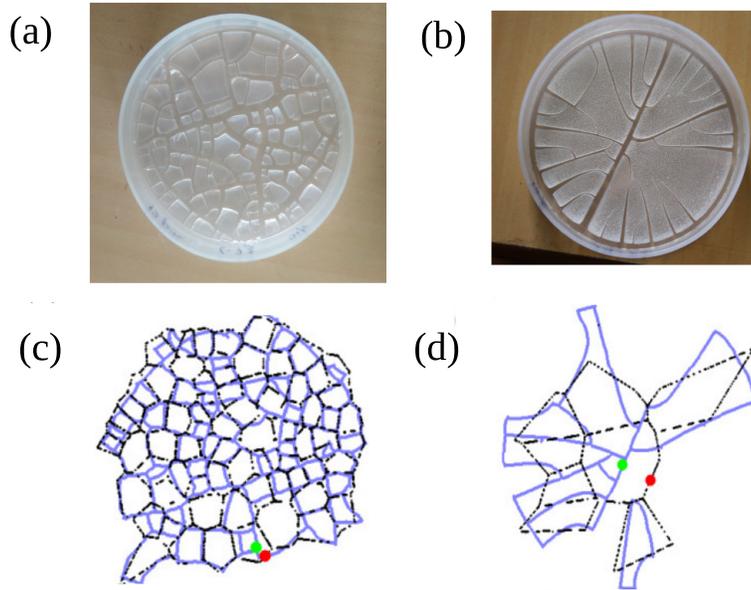


Figure 2.8: (a) Image of crack mosaic on Laponite clay of pH 6.73, (b) Image of crack mosaic on Laponite clay of pH 13.05, (c) Hausdorff distance $d_{H(sv)}$ estimated between the real crack mosaic (labelled with blue line) and its corresponding Voronoi tessellation (labelled with black broken line) for (a), (d) Hausdorff distance $d_{H(vs)}$ estimated between the real crack mosaic (labelled with blue line) and its corresponding Voronoi tessellation (labelled with black broken line) for (b).

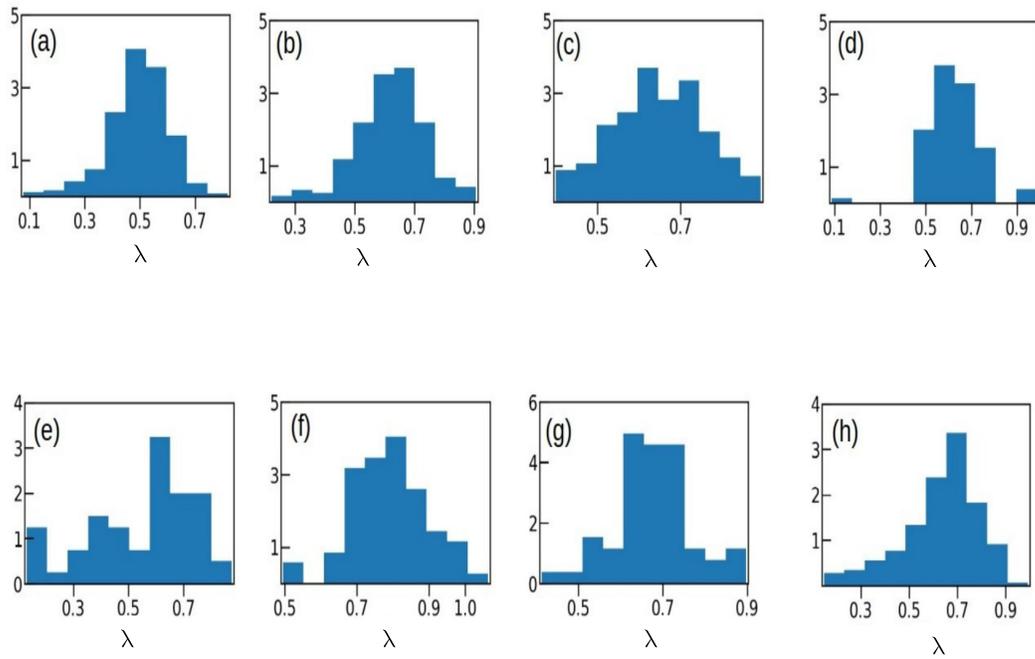


Figure 2.9: Histogram of isoperimetric ratio λ for real crack mosaics of :(a) TiO_2 , (b) PDMS, (c) Laponite (pH 6.73), (d) Laponite (pH 13.45), (e) Tempered Glass, (f) Resin, (g) Corn starch, (h) Mud crack.

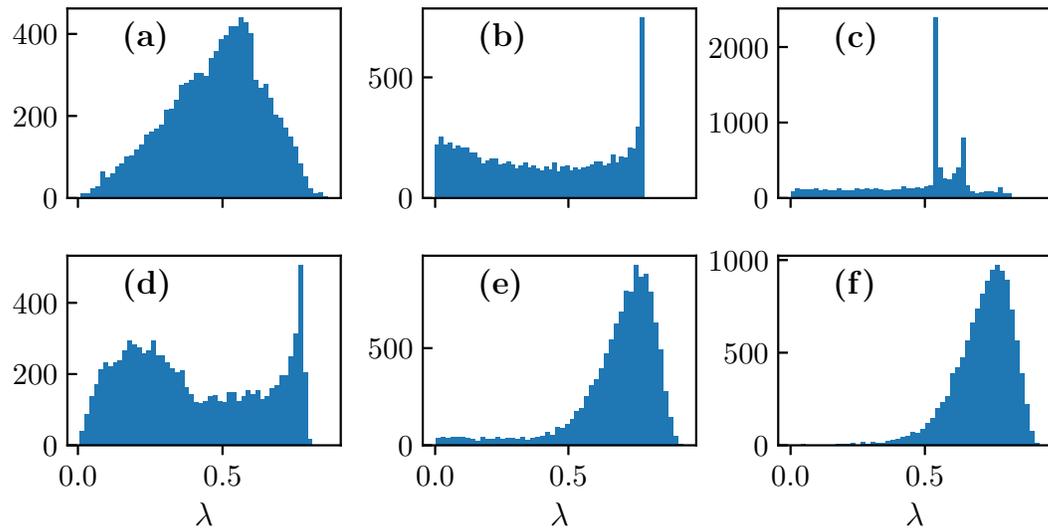


Figure 2.10: Histogram of isoperimetric ratio λ for simulated crack mosaics: Gilbert tessellations with slopes (a) random, (b) 0 and 90 degrees, (c) 0, 45, 90 degrees, (d) 0 ± 10 , and 90 ± 10 ; Voronoi tessellation for (e) uniform and (f) normal seed distribution.

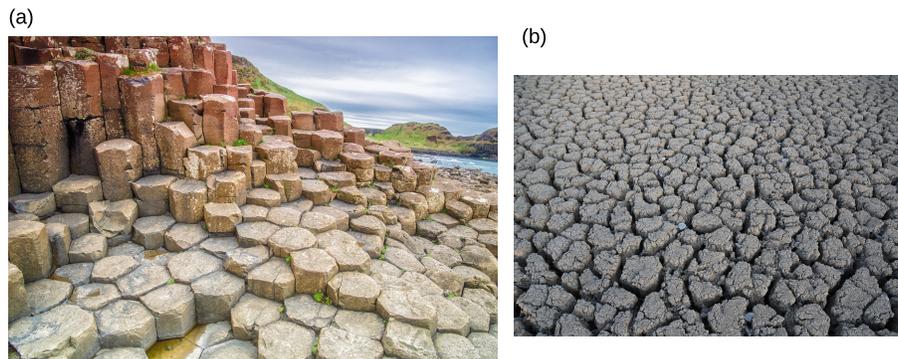


Figure 2.11: (a) Columnar joints [3], (b) Cracks in dried river bed [4].

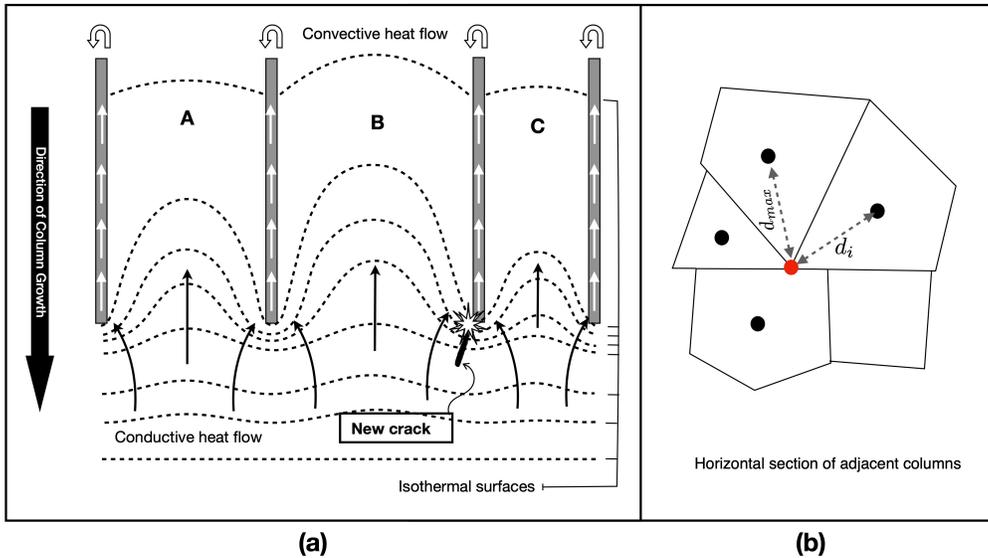


Figure 2.12: (a) Convection flow cooling, indicated by white arrows, takes place in crack joints; conduction cooling, indicated by black arrows, takes place across column interior. The dashed lines are the isotherms. Figure adapted with permission from [5]. Copyright 1994 Elsevier. New crack deviates in the direction of the wider column, following the higher thermal gradient. (b) Schematic for algorithm to estimate the shift in crack nodes in each time step of the simulation at different horizontal sections of adjacent columns meeting at a point. Black circles represent centroids of the neighboring polygons of the node (for which the shift is estimated for the next time step) marked by red circle.

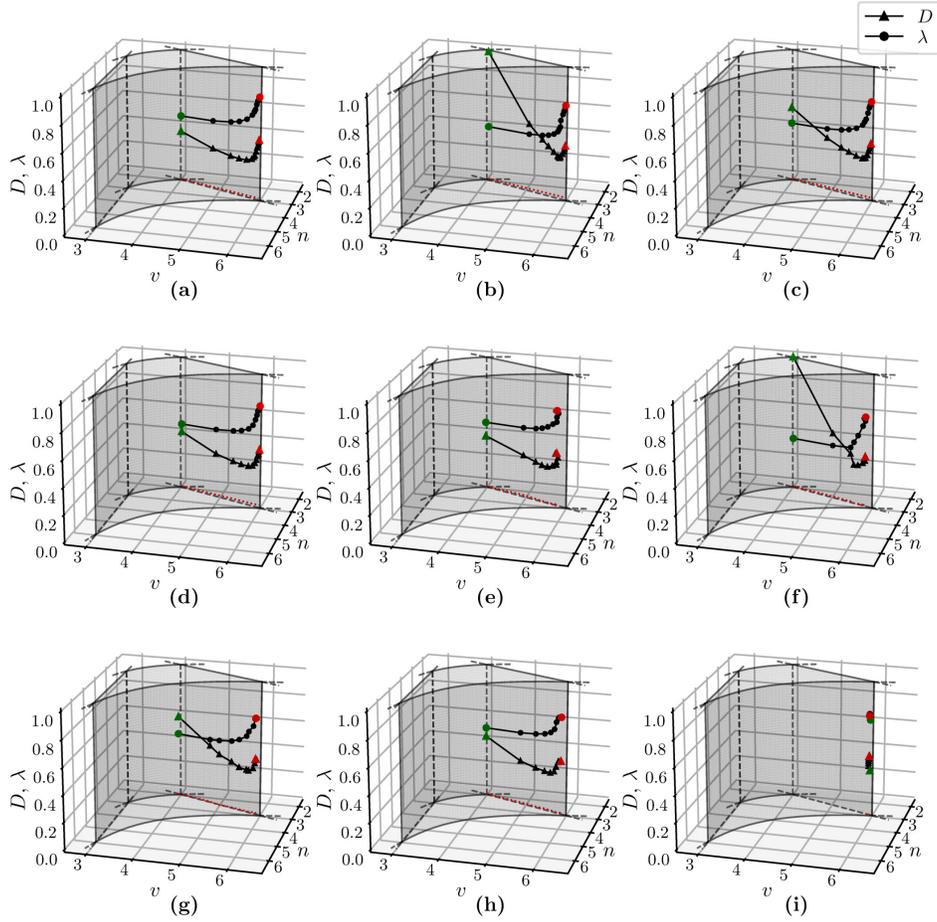


Figure 2.13: Trajectory of the crack in the (n, v, D, λ) domain: Gilbert cracks with uniform seed distribution and crack orientation - (a) random (b) parallel (c) $\pm 10^\circ$ (d) 45° and 90° ; Gilbert cracks with normal seed distribution and crack orientation (e) random (f) parallel (g) $\pm 10^\circ$ (h) 45° and 90° (i) Voronoi cracks.

Chapter 3

Studying the topological picture in fluid dynamical systems: Introducing the Euler Characteristic Surface

3.1 Introduction:

After delving into the topology and geometry of static mosaics, the next step was to investigate complex patterns within real fluid dynamical systems through the lens of topology. This transition from static to dynamic systems opened up a new dimension of complexity and beauty. Fluid dynamics, inherently more intricate and variable than static mosaics, presented an array of evolving patterns that required sophisticated mathematical tools to analyze. The drying droplets can be the perfect examples of a complex non-equilibrium dynamical system that involves interplay between several forces and parameters at the macroscopic level. A decent amount of theoretical and experimental studies have been done to understand the fluid dynamics and surface properties governing the drying process in droplet, [96, 97] give a compact review of the studies on droplet evaporation. In this study, the intention was to look at the dynamical process of droplet drying by studying their dynamical topological features and how the topological picture connects to the physical process. For the study, the topological measure we used is one of the most well-known and sig-

nificant invariants in topology, the ‘Euler characteristic’. Similar to other topological invariants, it remains constant under continuous deformations (stretching, twisting but not tearing) of a topological space. We intended to study a non-equilibrium dynamical system that involves intriguing dynamic pattern formations through this parameter.

The introduction and very first application of Euler Characteristic was done in characterizing polyhedra by Euler [98]. Later, a wide range of theoretical work has been done on Euler Characteristics of stationary point sets, both deterministic and random [99, 100]. Its application became popular in different scientific fields such as random fields [101], cosmology [102, 103, 104], neuro-science [105, 106], material science [107, 108, 20], thermodynamics [109, 110, 111], percolation theory [112], geophysics [20, 113] and complex data analysis [114, 115]. Euler Characteristic of a complex can be defined in various ways. In algebraic topology, the Euler Characteristic is defined as the alternating sum of k-dimensional Betti numbers [116] or equivalently as the alternating sum of the number of k-dimensional simplices (This will be discussed more in detail in Chapter 5). For a planar complex it is simplified to the number of objects/clusters minus the number of holes present in a binary network. [10].

$$\chi = N_b - N_w. \tag{3.1}$$

Where N_b is the number of black cluster(object) and N_w is the number of white cluster(holes) in the binary network/image.

3.2 The physical system

The physical system we studied here is an evaporating droplet composed of a uniform suspension of polystyrene beads in water, situated on a 40:1 polydimethylsiloxane (PDMS) substrate. This substrate is soft, with an elastic modulus of 0.3 MPa. The polystyrene beads, each with a diameter of 0.5 μm , were uniformly mixed with water using a magnetic stirrer. A droplet of this solution, with a volume of 0.5 μl , was carefully deposited onto the substrate, which had been cleaned first with deionized water and then ethanol. The measured contact angle was approximately 120 $^\circ$, indicating that the substrate is both soft and hydrophobic. A vertically positioned camera captured video footage of the droplet’s top view throughout the entire evaporation process. The temperature and relative humidity were maintained at 25 $^\circ\text{C}$

and 55%, respectively. Images were extracted from the video at various time points.

A sessile droplet can evaporate in a constant contact radius (CCR) mode, where the initial contact angle decreases as evaporation progresses, or in a constant contact angle (CCA) mode, where the contact area with the substrate decreases over time. In some cases, a droplet may exhibit both modes at different stages of evaporation. The specific mode of evaporation is determined by the cohesive forces within the droplet and the adhesive forces between the droplet and the surface.

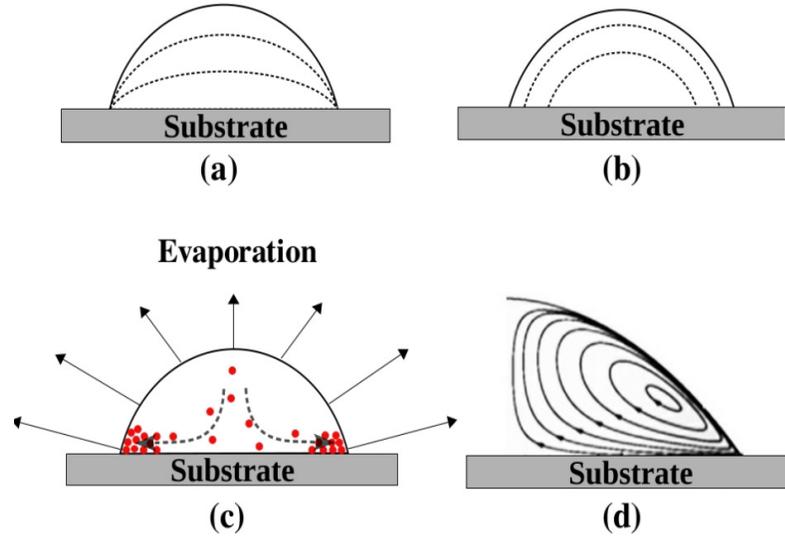


Figure 3.1: Different modes of droplet evaporation from a substrate- (a) Constant contact radius (CCR) with contact angle decreasing with time, (b) Constant contact angle (CCA) with contact radius decreasing with time. Fluid flow patterns in an evaporating droplets on hydrophilic rigid substrate- (c) Arrows represent a schematic of evaporation flux on droplet interface. The radial convective flow pattern towards the triple phase contact line carrying particles is represented by dashed lines. (d) Marangoni flows along a section of a droplet represented by arrows. The flow is induced by the surface tension gradient along the fluid-air interface.

As evaporation proceeds, multiple fluid flow paths are established within the droplet. The sessile shape of a droplet creates gradients in the evaporation flux at the interface, which drives radial convection flow inside the droplet. This convective flow transports suspended particles toward the triple-phase contact line

(TPCL) on a hydrophilic substrate, increasing particle concentration in that region. Consequently, solute concentration gradients develop, leading to surface tension gradients that drive Marangoni flow within the drying droplet. The ultimate fluid paths are dictated by the combined effects of convection, diffusion, Marangoni forces, and DLVO (DerjaguinLandauVerweyOverbeek) forces. Several reviews explain the causes and impacts of these flows. Evaporation induces clustering of the solute (beads) in the solvent (water), resulting in a disordered and dynamically changing process that moves toward a final equilibrium state. The dried deposition pattern can range from a “coffee ring” to uniform deposition or more complex patterns.

Droplets on soft hydrophobic substrates, like those in this study, display distinct flow lines during the early stages of evaporation. The evaporation flux is reported to be higher at the droplet apex than at the base. The greater height-to-base ratio of droplets on hydrophobic substrates contributes to higher thermal resistance between the apex and base, particularly in liquids with low thermal conductivity, such as water. Significant temperature gradients can develop, leading to thermo-capillary and buoyancy-driven flows within the droplet.

The substrate in our experimental study is a soft surface, which deforms slightly on the deposition of the drop. The PDMS surface produces a hydrophobic contact angle $\sim 120^\circ$. As discussed by [117, 118, 119, 120], this may lead to a large temperature difference of a few degrees across the height of the droplet i.e. the top of the droplet may be several degrees cooler than the bottom. This induces evaporative convection flow from top to bottom along the drop boundary. An opposing thermo-capillary flow is driven from the base of the droplet towards the top, due to vapor trapped at the wedge-shaped depression at the base of the soft PDMS substrate. The flow at the droplet interface is affected by the interplay of the flow lines of both, the liquid and the gas phase.

3.3 Image Processing:

Most real systems are chaotic, resulting in inherently noisy images. Filtering out noise effectively to extract useful information from images is a prominent area of research. Our system was highly dynamic, with particles constantly moving inside the droplet due to various forces, leading to significant noise in the images. Additionally, photographing a convex surface from above introduces uneven light intensity distribution in the captured image. Therefore, any topological analysis of these images

involves a three-step processing procedure before any useful study can be conducted: (i) thresholding for gray scaling and converting raw image data to binary data, (ii) appropriately tessellating the binary file on a suitable geometric grid, and (iii) validating the codes used in the process.

The top convex surface of the drying droplet was captured using a video camera placed vertically above to film the entire evaporation process. The video was then sliced to obtain top-view photos of the droplet interface at regular time intervals for analysis, Figure (3.2).

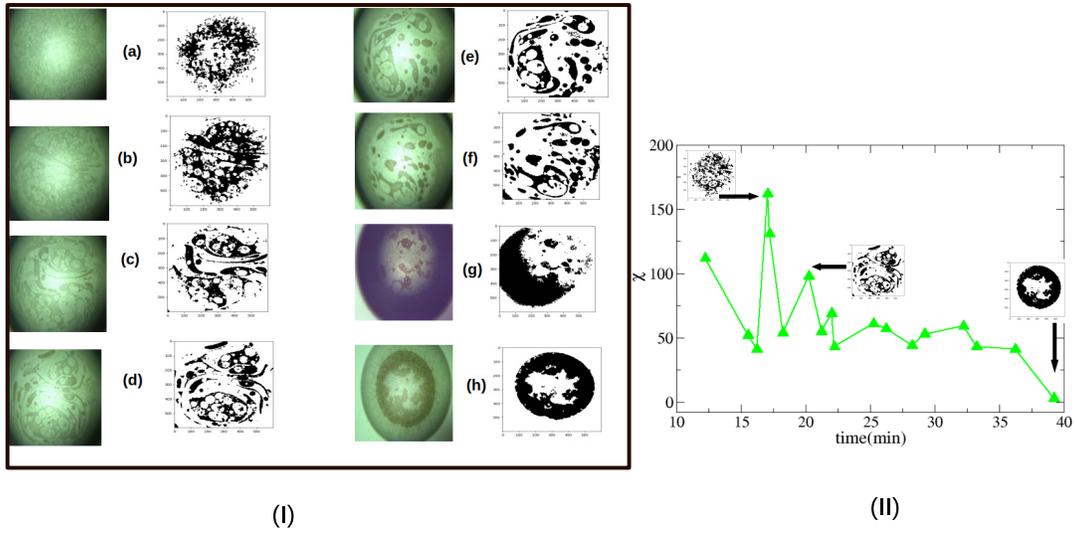


Figure 3.2: (I) Images of top surface of a droplet containing water and polystyrene beads during evaporation with their binary images at (a) 12.24 min, (b) 16.25 min, (c) 20.24 min, (d) 22.01 min, (e) 25.241 min, (f) 28.24 min, (g) 33.24 min, (h) 39.24 min. The binary images after applying a Gaussian filter are displayed alongside their original images; (II) Time variation of Euler Characteristic during droplet evaporation. The insets display the corresponding binary maps of the system.

3.3.1 Thresholding for grey scaling

The droplet of water containing micro-sized polystyrene beads evaporates in a constant contact angle mode (CCA), meaning the contact area gradually decreases as

evaporation proceeds. As the water evaporates, the droplet remains convex upward, with its curvature decreasing slowly over time. Each video slice corresponds to a specific moment in time. In each snapshot extracted from the video, the lighting from above appears uneven, with the central region being more brightly illuminated. When this image is gray-scaled with a constant threshold, the central portion appears washed out compared to the peripheral region. This issue can be corrected by applying an appropriate filter.

Using a constant threshold for gray-scaling such noisy images results in varying accuracy in binary mapping across different regions of the same image. Figure (3.3) presents the best binary images of the system at different times, produced with a constant threshold value. Comparing these with the corresponding actual photographs (Figure 3.2) reveals a clear lack of detail.

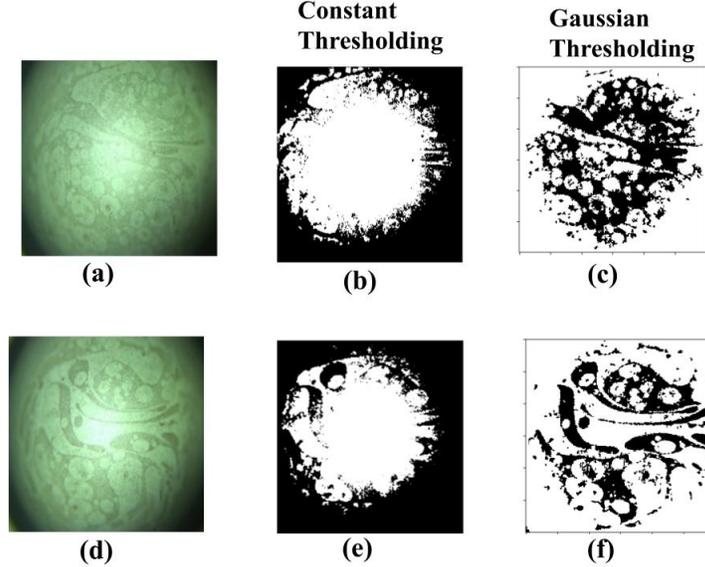


Figure 3.3: Best possible binarization of images of the droplet at (a)16.25 min and (d)20.24 min- without filter:(b and e); with Gaussian filter:(c and f), respectively.

A visual examination of the intensity profiles of the images (Figure 3.2) suggested that the intensity distribution over the pixels was almost Gaussian. Figure (3.4) supports this hypothesis, showing that the light intensity distribution in a real image follows a Gaussian distribution. This observation led us to use a Gaussian filter for thresholding the pixels to achieve better binarization. In Figure (3.4), the light intensity distribution is well-fitted by a Gaussian function. Therefore, a Gaussian

filter with suitable parameters has been applied to each image shown in the left column of Figure (3.2).

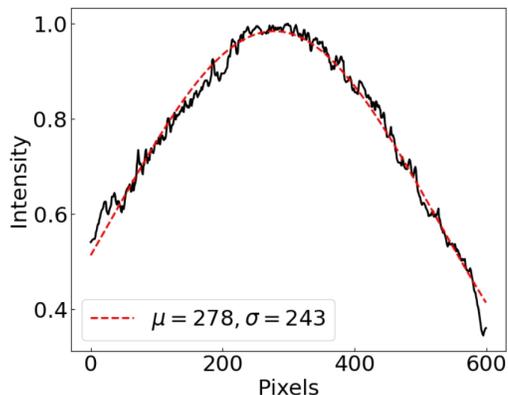


Figure 3.4: Intensity distribution of the images. The black graph displays the intensity distribution. The red graph is a fitted Gaussian curve.

3.3.2 Gaussian Filter

The Gaussian filter is a non-uniform low-pass filter. The kernel coefficients diminish with increasing distance from the kernels center. Central pixels have a higher weighting than those on the periphery.

Let $Y = [y_1, y_2, , y_n]$ be a band-limited digitized signal that is sampled at a frequency of f_s (sampling frequency) and let $N = [n_1, n_2, , n_N]$ be the noise sequence. Assume that $(n_i)_{i=1...N}$ is Gaussian distributed with mean μ_N and variance σ_N^2 . Let Y_N represent the signal Y contaminated by the additive white Gaussian noise N so that

$$Y_N = Y + N \quad (3.2)$$

An estimate of \hat{Y} needs to be made such that the absolute value of the mean square error in the estimate

$$|Y - \hat{Y}|^2 \quad (3.3)$$

is a minimum.

A Gaussian filter in the time domain is parameterized by its means μ_f and variance σ_f^2 and is represented by

$$G_f(\mu_f, \sigma_f^2, t) = \frac{1}{\sqrt{2\pi\sigma_f^2}} \exp^{-\frac{(t-\mu_f)^2}{2\sigma_f^2}} \quad (3.4)$$

The standard deviation of the Gaussian function plays an important role in its behavior. The values located between $\pm\sigma$ account for 68% of the set, while two standard deviations from the mean (blue and brown) account for 95%, and three standard deviations (blue, brown and green) account for 99.7%, Figure (3.5). This is very im-

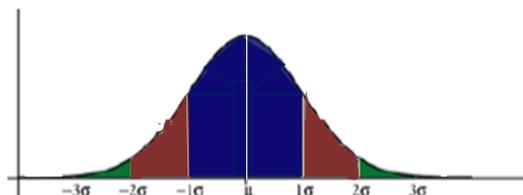


Figure 3.5: Distribution of Gaussian function values.

portant when designing a Gaussian kernel of fixed length. Larger values of σ produce a wider peak, i.e., a greater blurring. This requires that the kernel size must increase with increasing σ to maintain the Gaussian nature of the filter. The Gaussian kernel coefficients depend on the value of σ with the coefficients close to 0 at the edge of the filter. The kernel is rotationally symmetric with no directional bias. This allows separation along directional axes for fast computation. Thus in a 2-dimensional system, we need to use a 2-dimensional Gaussian function, which is simply the product of two 1-dimensional Gaussian functions, one for each direction. The Gaussian filter works by using the 2-dimensional distribution as a point-spread function that is achieved by convolving the distribution with the image. As the Gaussian distribution is non-zero everywhere, an infinitely large convolution kernel is theoretically required to produce a discrete approximation to the function. Fortunately, the distribution approaches very close to zero at about three standard deviations from the mean. 99% of the distribution falls within 3 standard deviations, which implies that we can normally limit the kernel size to contain only values within three standard deviations of the mean.

In the frequency domain Eq.(3.2) becomes

$$Y_N(k_\omega) = Y(k_\omega) + N(k_\omega) \quad (3.5)$$

and the Gaussian filter Eq.(3.6)

$$G(k_\omega) = \exp\left(\frac{-k_\omega^2 \sigma_f^2}{2}\right) \quad (3.6)$$

The estimate of the signal $\hat{Y}(k_\omega)$ due to filtering by Gaussian filter in Eq.(3.5) can be written as

$$\hat{Y}(k_\omega) = Y(k_\omega)G(k_\omega) + N(k_\omega)G(k_\omega) \quad (3.7)$$

The error in the filtered output is given by

$$\epsilon(k_\omega) = Y(k_\omega) - \hat{Y}(k_\omega) \quad (3.8)$$

i.e.,

$$\epsilon(k_\omega) = Y(k_\omega)[1 - G(k_\omega)] + N(k_\omega)G(k_\omega) \quad (3.9)$$

The first term on the *RHS* of Eq.(3.9) represents the contribution from signal distortion, and the second term represents the remainder noise after Gaussian smoothing.

Let the input and output signal-to-noise ratio(S), expressed in terms of power(P) be respectively

$$S_i = \frac{P_Y}{P_N} \quad (3.10)$$

$$S_o = \frac{P_Y}{P_Y - P_{\hat{Y}}} = \frac{P_Y}{P_\epsilon} \quad (3.11)$$

For a certain σ_f^2 , the Gaussian filter is able to filter the signal such that $S_o > S_i$, i.e., simultaneously remove the noise and not distort the signal. If σ_f^2 is increased, the cut-off frequency and the bandwidth of Gaussian filter decreases. This results in greater noise removal but increases signal distortion too. When $\sigma_f \rightarrow \infty$, $S_i = S_o$, i.e. an all pass filter is obtained. The search is for an optimal $\sigma_{f,opt}$ of the Gaussian filter such that S_o is maximized, and this goal is achieved through either curve or model fitting.

In the study, the function $T_{i,j}$, defined in Eq.(3.12), was used for thresholding every $(i, j)^{th}$ pixel of an image. In spite of best efforts, some significant details may be missing in the central region of c-e in Figure (3.2) after thresholding, due to stronger lighting there.

$$\begin{aligned}
T_{i,j} = & \alpha \frac{1}{0.45\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{X_{i,j}}{0.45}\right)^2\right] \\
& + 0.08 \frac{1}{0.15\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{X_{i,j} - 0.61}{0.15}\right)^2\right] \\
& - 0.2 \frac{1}{200\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{i - 600}{200}\right)^2\right] \\
& \frac{1}{200\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{j - 600}{200}\right)^2\right]
\end{aligned} \tag{3.12}$$

Here (i, j) defines the pixel location in the square array of pixels and $X_{i,j}$ is the distance of $(i, j)^{th}$ pixel from the center pixel of the image. The threshold value can be changed by varying the co-efficient α .

3.3.3 Image and tessellation:

Conventionally, images are digitized and stored as a rectangular array of values, where the image is sampled at each point on a two-dimensional grid to store intensity and implicit location information for each sample. Thus, a square grid is a natural choice for tessellating any binary image file. However, hexagonal grids are reported to be more effective for image processing due to their symmetric neighborhood and the absence of corner-to-corner connections, as seen in square grids (Figure 3.6). Regions composed of hexagonal cells (Figure 3.6a) do not suffer from the connectivity issues commonly found in regions composed of square or triangular cells [7, 121, 122]. For example, in the case of square cells, if 8-connectivity is chosen, the five square pixels shown in Figure (3.6b) will form a single region; however, if 4-connectivity is chosen, there will be two regions. Similarly, the six triangular cells in Figure (3.6c) illustrate comparable issues.

A triangular lattice functions as cell groups for a hexagonal lattice, with the centers of the hexagons corresponding to sites on the triangular lattice. In the following sub-section, the simulation of tessellating a triangular lattice to determine the Euler characteristic for different occupation probabilities of its sites will be described along with estimation of certain statistical characteristics of a triangular lattice. The outcome will be compared against established analytical results to validate the sim-

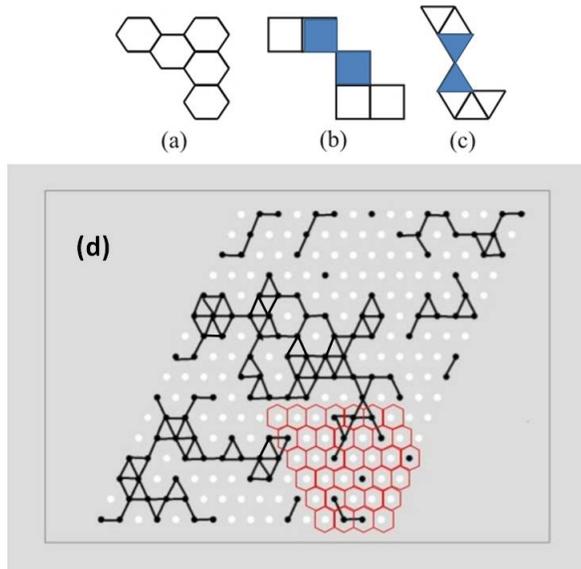


Figure 3.6: (a) Images composed of hexagonal cells do not present the connectivity problems, sometimes associated with images composed of squares or triangles. As can be seen from this figure (cases (b) and (c)), cells might appear connected by their corners. (d) A random distribution of a simulated triangular lattice of size 16×16 with $p = 0.48$ with black clusters being marked in the picture. Hexagonal lattice generated as a subset of triangular lattice, marked by red .

ulation. Next, the triangular tessellation code would be extended to a hexagonal tessellation of the image for determining the Euler characteristic.

3.3.4 Euler Characteristic on triangular lattice:

A triangular lattice was generated having constant lattice distance. The sites of the lattice were occupied randomly with probability p . The occupied sites were marked as black and the vacant sites were marked as white. Each site has six nearest neighbours which are all symmetric. Black(occupied) sites having black (occupied) nearest neighbors, form a cluster of black (occupied) sites; and white(vacant)sites having white(vacant) nearest neighbors, form a cluster of white(vacant) sites, Figure (3.6d). Both the clusters of black and white sites were counted and labelled using the Hoshen-Kopelman Algorithm[90]. For each probability p , the Euler Characteristic $\chi(p)$ was calculated using the definition in eq.(3.1). p was varied from 0 to 1 in increments of 0.1. For each value of p , configuration averages of the Euler Characteristic

was calculated over 1500 trials, till the results converged to within 0.01%. To check for finite size effects, the study was done for lattice sizes 16×16 , 32×32 , 64×64 , 128×128 , 256×256 , 512×512 , 1024×1024 .

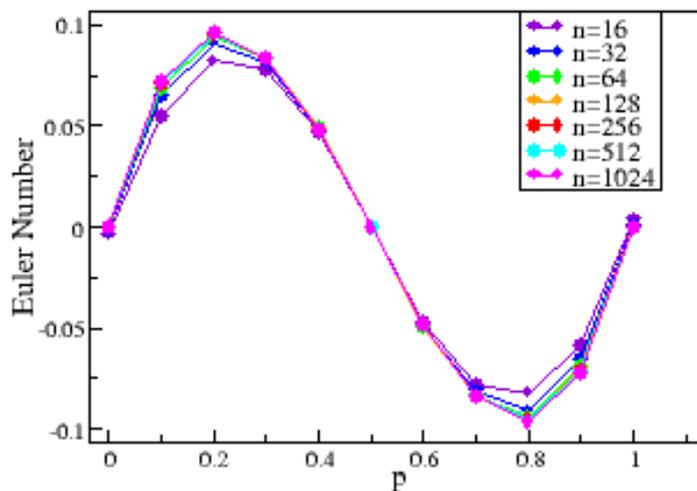


Figure 3.7: The variation of Euler Characteristic (χ) for triangular lattice of different sizes. In each cases, the value becomes zero near the percolation threshold, $p_c = 0.5$.

Since we used a triangular lattice to generate a hexagonal lattice for tessellating an image file and creating its binary version, it was crucial to validate the code for the triangular lattice. As part of our validation process, we first calculated the Euler Characteristic on a triangular lattice with variable occupation probabilities p , Figure (3.7). We also estimated several other features as functions of p , such as the cluster size distribution, the percolation threshold, and the fractal dimension of the percolating cluster. Finite-size scaling on different system sizes was performed to determine the percolation threshold, which was approximate $p_c = 0.5$. Our simulation results for the variation of the Euler Characteristic with p and the cluster size distribution were in good agreement with the analytical results on the triangular lattice reported by Sykes and Essam [9].

3.3.5 Extension to hexagonal tessellation:

After validating our algorithm to calculate the Euler Characteristic, we moved towards applying it on the image files that we had intended to. Images are typically stored in square pixels, but better resolution can be achieved using a binary map based on a hexagonal grid, as the latter does not have the diagonal connectivity problem. A triangular lattice serves as the cell group for hexagonal grids; connecting the centers of hexagonal grids forms a triangular lattice, as shown in Figure (3.6d). The hexagonal grid was created by suppressing alternate rows and columns of the square lattice of an image pixel file. This process can be done using any of the four possible combinations shown in Figure (3.8) by alternatively suppressing rows and columns in the square lattice [123]. We estimated the Euler Characteristic in every choice for an image and found out that for large matrices/high resolution pixel files, the value of the estimated Euler characteristic is almost the same for any of the choices opted. Thus, our algorithm to compute the Euler characteristic of the site percolation model for the triangular lattice was made suitable to compute the Euler Characteristic of an image in the hexagonal grid format, as both have the same 6-symmetric neighborhood structure. The site (i, j) on the lattice is nothing but the (i, j) th pixel that can be black /white. One can visualize the binary image as an entire lattice consisting of black objects on a white or void background or vice-versa. In our convention, if a grid point is black, it contains an object pixel and contributes to clusters of the object, and if it is white, it contains a background pixel contributing to the clusters of the background. The difference between the two complementary clusters in the image/lattice gives the Euler Characteristic of the image, as defined in eqn.(3.1).

To test the algorithm on images, the Euler Characteristic was estimated for static patterns with known values of Euler Characteristic. The alphabets A, B and C were chosen. The binary maps of the 3 alphabets were created by grey-scaling at a constant threshold value of 0.5. The binary images were sub-sampled, i.e. alternate rows and columns of the pixel array were suppressed, to get the hexagonal grid-like structure. Finally, the Euler Characteristic of the sub-sampled images were calculated using the same algorithm discussed in the earlier section for the case of triangular lattice. The value of the Euler Characteristic was calculated for all 4 possible ways of sub-samplings, and the results were found to be identical. This established that the final result was independent of the process of sub-sampling. However, this is probably true only for those images that have a high density of pixels. The values of the Euler Characteristic were estimated to be -1 , -2 and 0

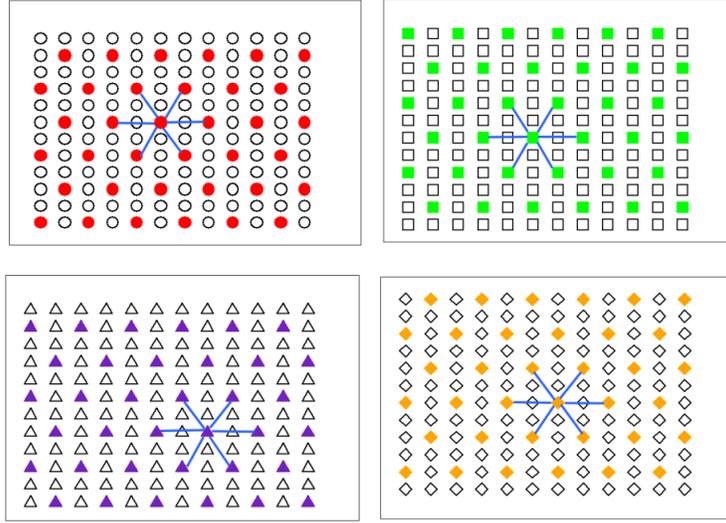


Figure 3.8: The four possible choices to suppress rows and columns of square lattice to turn into a hexagonal grid-like format.

for ‘A’, ‘B’, and ‘C’ respectively, as the background was also considered.

3.4 Evolving topological features of agglomeration patterns and the dynamical system

Our studied system here was a tiny water droplet of $\approx 0.5 \mu\text{L}$ volume, containing polystyrene nano-beads. The droplet evaporation time given the ambient conditions (discussed in Section 3.2) was ~ 40 minutes and within that time frame, the video captured from the top view showed intriguing patterns forming and changing before settling into an agglomerated cluster of coffee ring. These flow features and agglomeration patterns are very specific to the compositions of the droplets, substrate properties, and the ambient conditions. My interest was not to study the physical parameters, like contact angle, viscosity, morphology analysis in SEM, etc. that usually are the standard methods in the droplet community. I was more drawn to understand how the topological features were changing throughout this non-equilibrium dynamical process as a consequence of the evolution of the different physical forces present in the system. These flow patterns, observed experimentally, are visualized through aggregation patterns or the movement of discrete particles (in our system

- the polystyrene beads) that arrange themselves guided by the streamlines. Figures (3.2 a-h) show that the agglomeration patterns by polystyrene beads underwent various changes in their aggregation pattern as evaporation proceeded. Polystyrene beads have a very low negative surface charge in aqueous solution and the inter-particle interactions can thus be neglected. So, it may be said that the arrangement of the beads was mainly defined by the flow streams within the droplet.

To analyze and understand these topological features of the aggregation patterns in the drying droplet (from the sequence of binary images recorded from the flow) the topological invariant, ‘Euler Characteristic’ was used primarily in the study and then progressed into introducing topological ideas that explored the multi-scale time-evolution of the drying droplet through the image sequence. Many of the mathematical ideas in this section come from algebraic geometry and are motivated by the idea of topological persistence in data science. The focus was always to mitigate instrumental and computational errors while working with these noisy images coming from the disordered dynamical system.

3.4.1 Time evolution of the Euler Characteristic through images

The principal objective was to study the time evolution of the agglomeration patterns through the topological invariant, the Euler Characteristic, and whether we can characterize the topological changes happening throughout, in terms of it. The computation process started with image processing and optimal thresholding to turn the image sequences into binary matrices (discussed in Section 3.3). In the system of the evaporating droplet of polystyrene beads in water, the polystyrene beads were considered as objects (black), while the water constitutes the background (white), Figure (3.2). The clustering of the beads during evaporation was disordered and dynamically changing while the system gradually advanced towards the final equilibrium state. These dynamically evolving flow features were mapped via the variation of the value of the Euler Characteristic with time. To estimate the Euler Characteristic for each image, the hexagonal grid was used with the sub-sampling described earlier, Section 3.3.5.

Figure (3.9) shows the variation of the Euler Characteristic for the images at different time intervals during the evaporation of the droplet. In the early times of the evaporation process, the system was very chaotic, with particles spread all over the droplet surface. This yields a high value for the Euler Characteristic(χ). As time

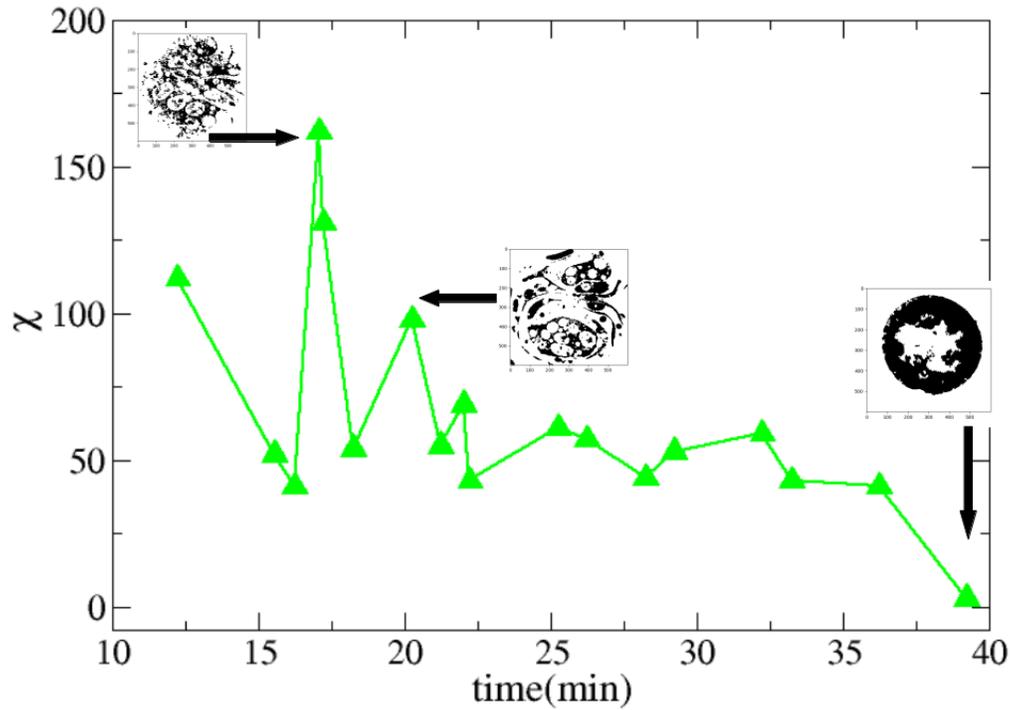


Figure 3.9: Time variation of the Euler Characteristic (χ) of the evaporating droplet system using a hexagonal grid. The insets display the corresponding binary maps at different time instants.

progressed, The Euler Characteristic dropped gradually. The plot, Figure (3.9) shows very little variation between 19 to 36 minutes indicating an almost stable aggregation pattern during this interval. Between 36-40 minutes the Euler Characteristic falls to ~ 0 when a single thick ring was observed as indicated in the Figure (3.9). The aggregation pattern of the particles remained in this configuration in the completely dried droplet.

Along with the variation of χ with time, the size distribution of the clusters of the objects in the images was also calculated at four different time instants. This gives more clear information about the clustering process. Clusters varying within 100 sites have been binned together. This distribution has been displayed in Figure (3.10). At the beginning of the evaporation, most of the beads were isolated, which resulted in the surge of the graph at cluster size with value 1, Figure (3.10). At an

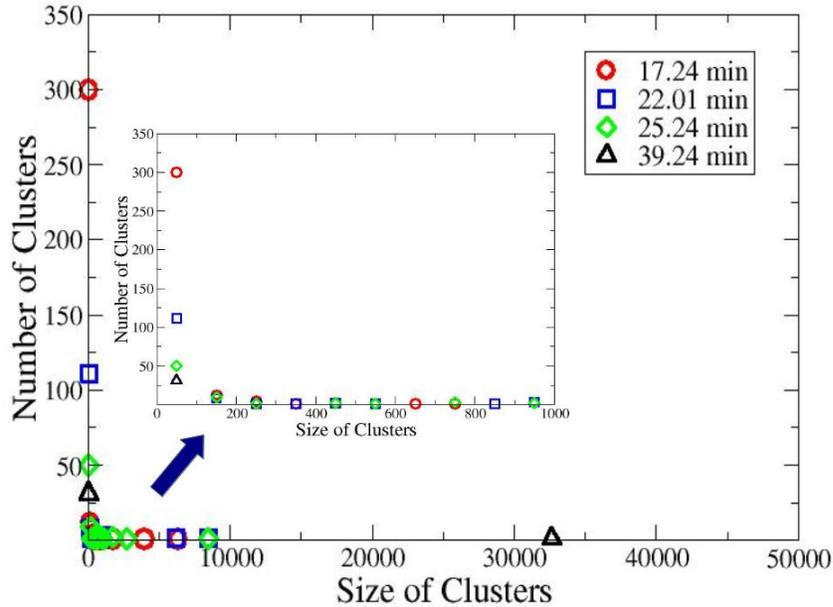


Figure 3.10: Size Distribution of clusters at time instants (a) 17.05 min, (b) 22.01 min, (c) 25.24 min, (d) 39.24 min. A zoomed view is shown in the inset.

intermediate time of the evaporation process, the clustering of beads started due to the interplay of the fluid flow processes in the drying droplet, this was reflected in a rough exponential distribution of cluster sizes, inset of Figure (3.10). Towards the end of evaporation, examination of Figure (3.2h) showed a complete clustering of all the beads to form an almost single big cluster. This is correctly represented by the presence of a single cluster of huge size in Figure (3.10).

It may be noted that the dynamical system studied was a drying droplet on a soft PDMS substrate and during the drying process the particle number, i.e., the number of polystyrene beads remained constant in the entire droplet but the curvature of the droplet decreased with time as the droplet flattened with progress in drying. As all the images were photographed from a vertical position atop the droplet, the number of particles that came into focus of the camera at a time was not constant but increased with time. This may be treated as analogous to increasing p , the probability of occupation by objects in a given system size, as discussed in Section(3.3.4). The total number of 'black' pixels expressed as a fraction of the total number of pixels contained in the system, may be defined as p in that case.

3.5 Introducing the approach for multiscale study of topological features:

Observation up to this was all right but it did not feel enough to satisfy the urge of extracting all the topological features that we could see with open eyes! There was something more that we were perhaps overlooking or failing to capture alone in the time variation of the topological invariant, Euler Characteristic(χ). At this time, the idea to look at a complex pattern at multiple scales of resolution came to us, which is quite analogous to the philosophy used in Topological data science based on the theories in algebraic geometry.

A binary image at a specific time was looked at different scales of resolution r , where $r > 0$. This can be thought as the r -neighbourhood of a set being combined along with the primary sets, for a given binary matrix. Given a set $S \subset \mathbb{R}^2$, we defined the neighbourhood of S at scale r (where $r > 0$) to be the set $S_r = \{x \in \mathbb{R}^2 : d(x, S) \leq r\}$, where $d(x, S)$ is the usual Euclidean distance of x from the set S . In our case, the set S can be considered to be a particular cluster of black pixels in the binary image, and the scale parameter r is a positive integer-valued. By this, at scale r each of the black(object) clusters present in the binary image matrix gets replaced by a cluster of black pixels including its r -neighbourhood with the already existing cluster. The same idea is also used in topological data analysis, where one replaces point sets by unions of r -balls centered at those points - and studies the topology of the resulting structure given by the union of those r -balls. It is implied that if the black clusters get replaced by their union with r -neighbourhoods ($r > 0$), the resulting cluster pattern will have a smaller number of components that will again be reflected in the Euler Characteristic χ_r of the binary image at scale r . Figure (3.11) shows a binary image during the droplet evaporation at different values of scale r . The ‘Scale 0’ image is basically the binary matrix yielded after optimal thresholding of the image. As the r values of the scale keep on increasing, more neighborhoods come into the union of clusters, decreasing the resolution. Looking at images where there exist multiple features (small scale and large scale) along with noises, it is very much required to consider this multiscale approach. The vortex-like feature which was not loud and prominent in scale 0 gets more clear as we consider the image at the gradually increasing scales. It is to be noted that here scale r has only integer values as it depicts the integer neighborhood and discrete nature of the pixel grid/lattice system. Therefore, we extended the analysis by studying the connectivity of each binary image at increasing scales r to gain insight into the flow - as recorded by the agglomeration patterns. Figure (3.12a) displays the variation of

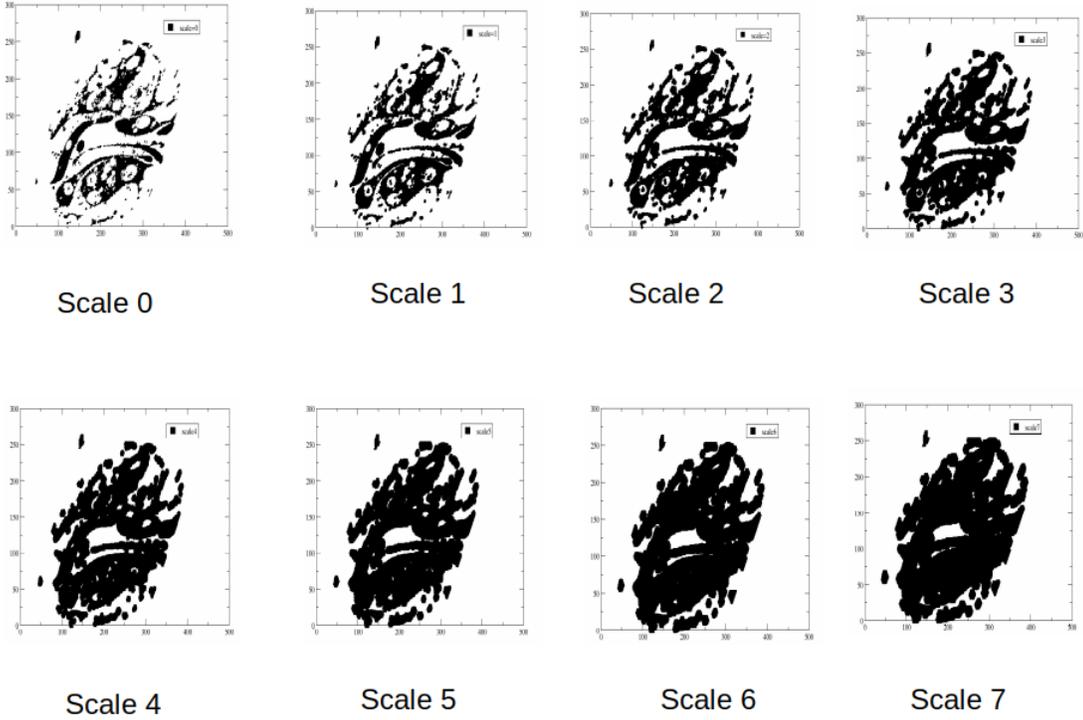


Figure 3.11: Top-view of drying droplet image as each black cluster is enlarged by a scale of r -neighborhood. Scale $r = 0$ corresponds to the original binary image at $t=20.24$ min. Scales in the figure vary between $r = 0$ to $r = 7$.

the Euler Characteristic (χ) with scale r for images at different time steps. There is a drop in the χ value around scale 2. This was due to the removal of very small clusters that got merged with comparatively larger ones. One may note that for a range of scale r , χ took negative values for certain window of time implying the number of loops present was more than the clusters of aggregation in the flow pattern. Again Figure (3.12b) shows the time variation of χ at different scales r . the variation in the curve as r increases and around scale 6, the curve shows minimum variation. Since at a large scale, the resolution becomes very poor, where almost all clusters get merged making a blob only! One may not analyze beyond that value of scale.

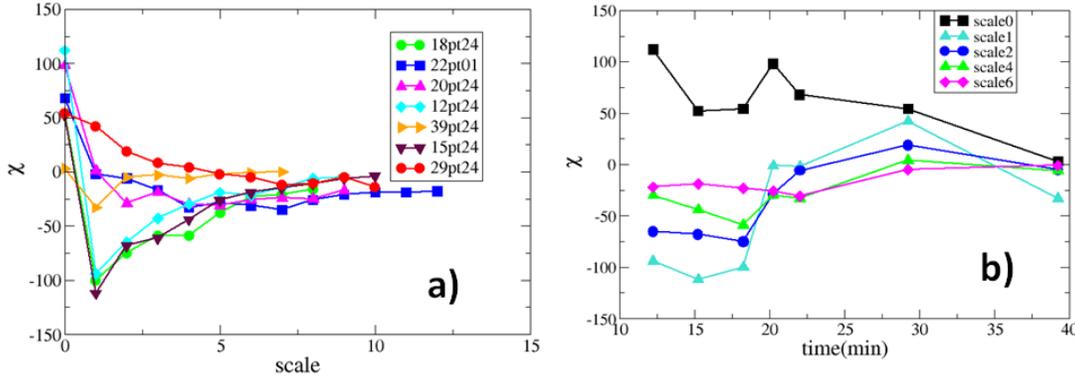


Figure 3.12: (a) Variation of χ with different scales for constant times. (b) Variation of χ with time at different r-neighbourhood scales.

3.6 Introduction to the Euler Characteristic Surface (ECS) :

Studying the time variation and scale variation of the Euler Characteristic did not give a compact overview of the dynamical system which is why we proposed to deliver the multiscale connectivity information of the images described above for different discrete time steps in the form of a surface which we call the ‘Euler Characteristic Surface’. This design summarizes all the spatio-temporal topological changes happening in the dynamical system into a condensed form through the Euler Characteristic surface $\chi_s(r, t)$. The values of the Euler Characteristic $\chi_s(r, t)$ at different scale r and time step t for the studied dynamical system which is the drying droplet, were plotted as a surface in \mathbb{R}^3 , Figure (3.13). Since the flow patterns or topological features that are emerging and changing with time in a droplet, or in any dynamical

system where we can see similar time evolution of connected patterns, are very individualistic and unique to the system itself, one may consider the Euler Characteristic Surface associated with the dynamical system as its *unique topological signature*.

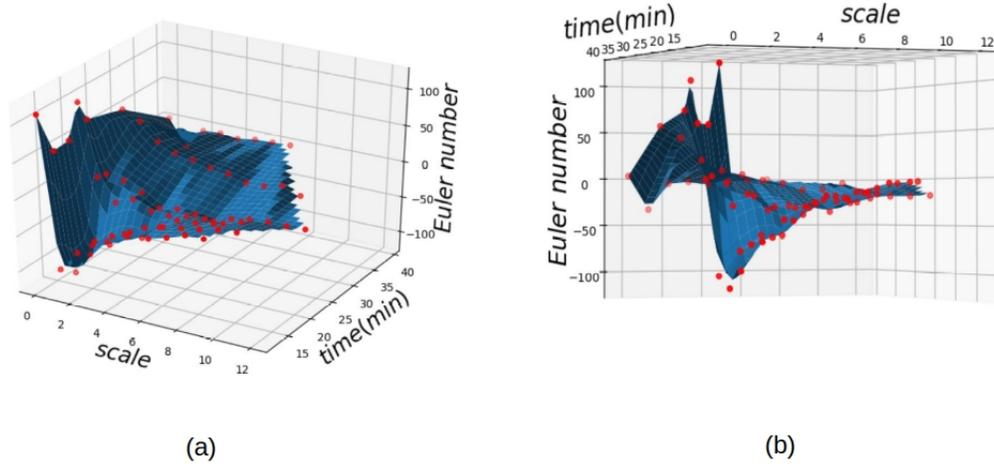


Figure 3.13: The Euler Characteristic surface in scale (r) and time (t) of the drying droplet. The experimental data points are marked by red circles, the blue surface is the best-fit surface constructed by the $r - t$ phase space. Two orientations of the same surface are displayed in figures(a) and (b).

3.6.1 Euler Characteristic Level Curves

Along with the Euler Characteristic Curve came its projected contour plot that we named as the ‘Euler Characteristic Level Curves’. Given an Euler Characteristic Surface as discussed above, we considered the contour curves on a 2d surface. These Euler Characteristic Level Curves are loci of constant χ_s on the surface $\chi_s(r, t)$, Figure (3.14a and b). Assuming that we can fit a smooth surface to the data generating the surface $\chi_s(r, t)$, the tangent lines to the level curve at a specific point are orthogonal to the gradient $\nabla\chi$ at that point. This gives the opportunity to develop algebraic expressions of the level curves in such a case. The level curves plotted may

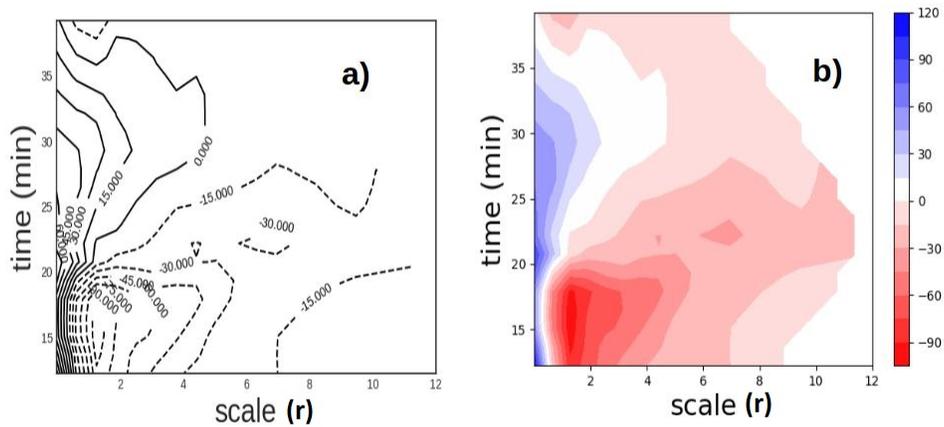


Figure 3.14: (a) The Euler Characteristic Surface Level Curves. Curves having the equal value of Euler Characteristic have χ displayed. Positive χ curves are bold lines and negative χ curves are dotted lines. (b) Color plot of the Euler Characteristic Level Curves. The color legend displays different χ values.

give a low-scale overview of the evolution of the Euler Characteristic(χ). The pinch points in the level curves give the position of scale and time where a sudden change in the value of χ , thus in the system's topology takes place.

3.6.2 Proposing the Euler Metric for dynamical systems

The Euler Characteristic Surface introduced in Section (3.6) is a function $\chi_s(r, t) : \mathbb{N} \times [0, \infty) \rightarrow \mathbb{Z}$. In this case, as we restricted to the scales $0 \leq r \leq R \in \mathbb{N}$ and time $0 \leq t \leq T \in \mathbb{R}$, $\chi_s(r, t)$ is a compactly supported function $\chi : [0, R] \times [0, T] \rightarrow \mathbb{Z}$. Given sufficiently many data points, one can fit a continuous (even polynomial) surface to the data points, so we can assume χ to be continuous. For a given $R \in \mathbb{N}$ and $T \in \mathbb{R}$, given a specific droplet drying situation one can calculate χ as a possible descriptor of the drying process with specific parameters and can calculate its L^2

norm

$$\|\chi\|_2 = \left(\int_{[0,R] \times [0,T]} \chi^2 \right)^{\frac{1}{2}} \quad (3.13)$$

Given two flows with parameter sets F_1 and F_2 for droplets of the same size, one can calculate the corresponding Euler Characteristic Surfaces χ_{F_1} and χ_{F_2} , and calculate the distance between these surfaces to be

$$d(\chi_{F_1}, \chi_{F_2}) = \|\chi_{F_1} - \chi_{F_2}\|_2 = \left(\int_{[0,R] \times [0,T]} (\chi_{F_1} - \chi_{F_2})^2 \right)^{\frac{1}{2}} \quad (3.14)$$

We call the above-defined distance the Euler distance or Euler Metric between the sequence of images describing these two flow situations. This is a well-defined metric and makes the space of Euler Characteristic Surfaces (or, equivalently, the corresponding space of sequences of images) into a metric space - even a Hilbert space. This makes the theoretical framework proposed by us amenable to the full range of available statistical techniques.

3.7 Discussion

Our aim in this work was to search for topological descriptors of a dynamically evolving complex flow system via the topological invariant Euler Characteristic. The drive was to investigate to what extent the time variation of the Euler Characteristic may help to map the time evolution of a dynamical process. Such studies were not done earlier, especially for time-varying dynamical systems. Our idea to analyze the connectivity information of agglomeration clusters, changing with time, can be used to understand the different physical flows that are running in the background. In developing the code to estimate the Euler Characteristic of images, the essentiality of finding an optimal filter for thresholding and choosing an appropriate grid structure for the binary image data was realized. The codes were verified on sample images and structures for accuracy.

The output of the scale-time-based variation of the Euler Characteristic of particle aggregation pattern on the droplet interface can be useful to understand

Table 3.1: Summary of topological tools studied

Sl.	Proposed Topological Tools	Flow Features Studied	Ref.
1.	Euler Characteristic	The time evolution of Euler Characteristic help study the clustering behavior of suspended beads.	Figure (3.9), Section 3.4.1.
2.	Euler Characteristic at scale $r > 0$	More stable to noisy images of the flow. A sequence $\{\chi_r\}_{r \in \mathbb{N}}$ help detect possible vortices in the flow.	Figure (3.11), Section 3.5.
3.	Euler Characteristic Surface	A multiscale topological summary of the agglomeration patterns in the flow.	Figure (3.13), Section 3.6.
4.	Euler Characteristic Level Curve	A low dimensional signature that organizes the multiscale connectivity information in a planar plot, which helps study flow stability.	Figure (3.14), Sub-section.3.6.1.
5.	Metric on the space of Euler Characteristic Surfaces	A proposed similarity-dissimilarity measure between distinct flow situations, that is amenable to statistical techniques.	Sub-section3.6.2.

the possible changes in the dynamics caused by the flow lines. Table.(3.1) gives a summary of the different topological characteristics that were introduced as tools to study topological features in a dynamical system.

3.7.1 Particle agglomeration features

The process of studying a flow pattern through sequence of binary images introduces possibilities of loss of information - both instrumental and in thresholding. As described in Section(3.5), we introduced several topological constructs to mitigate this - in the form of Euler Characteristic at scale $r > 0$, Section(3.5), the Euler Characteristic Surface $\chi_s(r, t)$, Section(3.6), and Euler Characteristic Level Curves, Subsection(3.6.1).

After all this, the natural response one may have is how these values of the Euler Characteristic may correlate to the physical forces causing the non-equilibrium system to gradually reach its equilibrium. Let us try to understand the hint that one may obtain from the topological description of the flow images obtained at different time instants of the evaporating droplet. The most observed topological features of aggregation patterns in this case, along with their effect on the Euler Characteristic are compared with actual images of aggregation of particles on droplet interface, in Figure(3.15). The Euler Characteristic Surface (ECS) of the drying

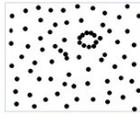
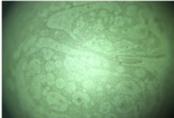
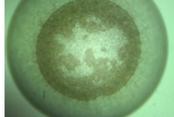
Pattern	Euler characteristic χ	Image
	Large +ve value; $N_b(82) - N_w(2) = 80$	
	Large -ve value; $N_b(1) - N_w(10) = -9$	
	Small -ve value; $N_b(1) - N_w(3) = -2$	
	Small -ve value; $N_b(1) - N_w(2) = -1$	

Figure 3.15: Comparison table illustrating a schematic of topological features (column 1) and their reflection in Euler Characteristic Numbers (column 2), compared with images of droplet interface features observed during different time instants of evaporation (column 3).

droplet starts from a crest when almost all the polystyrene beads are uniformly spread out in the droplet. Therefore, the number of black clusters N_b was very high where beads remained like disjoint islands in the ‘sea’ of practically a single white cluster. As drying proceeded, flow streams pushed the single beads into forming clusters, decreasing N_b at the cost of increasing N_w as the ‘sea’ starts breaking up into lakes. This is reflected in a sharp dip in the ECS as observed. A comparison of columns 1 and 2 of Figure (3.15), supports the plausibility of this argument. As

temperature gradient built up in the droplet after a while, the competition between Marangoni¹ and thermal buoyancy flows set in causing formation of vortices in the drying droplet [124]. Viewing under high magnification, one can see the vortices throw up a pair of ‘lobes’ - a comparison of droplet image with schematic in the third row of Figure (3.15)- which is reflected in the Euler Characteristic with its small negative value. Finally, towards the end of evaporation, the beads aggregated into an almost single ring-like structure bringing χ value to ~ 0 . The ECS also flattens out to coincide with the plane defined by the $r - t$ plane. It is obvious that the χ values obtained from the analysis of noisy images of the drying droplet will not be a perfect match to ideal geometrical features represented in column 1 of Figure (3.15). However, thus one may use ECS to interpret the flow changes that occurred in this microscopic fluid flow system.

This study described a general pattern of flow in the drying droplet on a hydrophobic surface. The images showed that it developed from a chaotic movement to a clearly identifiable stable pair of vortices and finally led to a single ring-like structure. This history captured in the images was quantified in the form of the time variation of the Euler characteristic and further, it got a clearer description by the Euler characteristic surface and Euler characteristic level curves. Later in the Chapter 4 the tools are further applied to compare drying results and flow patterns for droplets of different compositions under similar ambient conditions using the formalism developed here.

3.7.2 Flow stability features

Euler Characteristic Surface can also be used to infer the most stable and persistent topological feature present in the dynamical system. The plateau-like region in ECS where the variation in χ is very small compared to other regions may be considered for this. However, the Euler Characteristic Level Curves provide better information to understand the stability and frequency of flow lines. Examination of Figure (3.14a and b) reveals almost parallel contour lines over the time range up to 25 min approximately for the scale range between 0 – 4. This implies that without much variation the topological features persist over this length scale emphasizing their stability over other noisy features. The density of contour lines in the aforementioned time window indicates a very rapid change in χ values. The high density together

¹Marangoni flow in an evaporating droplet refers to the fluid motion within the droplet driven by surface tension gradients. These gradients arise due to variations in temperature or concentration across the droplet’s surface during evaporation.

with the parallel nature of lines indicate that though there is rapid dynamics leading to the organization and re-organization of particles, the essential topological feature is stable. The pinch points in the contour lines occurring at 15, 19, and 22.4 min approximately, Figure (3.14b), indicate change over from one topological feature to another.

The metric structure, “Euler Metric” proposed in Sub-section(3.6.2) can further be used to predict stability between families of dynamical systems where one may estimate the dynamical changes taking place due to the variation of a system parameter through estimation of this metric between them. This is studied later in Chapter (4).

3.8 Conclusion

This chapter thus presents the topological construct made to study complex disordered dynamical systems through the invariant Euler Characteristic. For our studied dynamical system – flow patterns in a drying droplet of polystyrene beads suspended in water, studying the evolution of the Euler Characteristic (χ) with time provides primary information about the change in clustering and connectivity along with time. In general, this approach can be used to extract information on topological changes of any complex system showing dynamical topological features via analyzing a time series of images of the system.

The highlight of the study was the birth of the idea of studying Euler Characteristic at multiple scale $r > 0$ of resolution, Section (3.5) for each image and establishing a novel topological construct- the Euler Characteristic Surface $\chi(r, t)$, Section(3.6). This proposed surface summarizes all topological features and acts as the system’s topological signature. The associate construct of the Euler Characteristic Surface, the Euler Characteristic Level Curves, Section(3.6.1) are also helpful to get low dimensional summary of the topological evolution. They also can be studied to identify persistent topological features and the pinch points of sudden topological changes in a dynamical system. The proposed metric- which is L^2 norm distance between two Euler Characteristic Surfaces, Section(3.6.2), has potential as a measure to distinguish between flow situations in dynamical systems.

Using the proposed topological tools, the objective to efficiently extract the topological information from dynamical systems through the time-sequence of binary

images was met, and also non-trivial conclusions were drawn about the flow dynamics solely from the observed aggregation patterns of polystyrene beads on the surface of the droplet. This indeed made us hopeful with our approach and the study was extended further for more dynamical systems with the same framework.

Chapter 4

Characterizing fluid dynamical systems using Euler Characteristic Surface and Euler Metric

4.1 Introduction

The potential of Euler Characteristic (EC) as a topological invariant has already been discussed in the last Chapter 3. In this chapter the effectiveness of the Euler Characteristic Surface(ECS) and Euler Metric (EM), to quantify differences between different dynamical systems is explored. The established topological prescription gets applied for the analysis for simulated fluid dynamical system and for real fluid dynamical systems. The aim was to investigate the efficiency of ECS and EM in the analysis of complex dynamical systems. The studied simulated system was a flow system whose evolution pattern was very sensitive to the variations in its single input parameter. The tools ECS and EM successfully captured the topological variation in the system along with characterizing different regimes of the flow system. Along with this an experimental configuration was done to study the comparison between real systems with the established construct. As earlier micron sized systems of drying droplets that were impregnated with large number of insoluble particles were studied, this time two different set of droplets having two different composition of medium and particles.

In Section 4.2 the description about the systems and the details of the computation process are given where the deterministic model of simulated dynamical flow pattern in Sub-section 4.2.1 and real stochastical systems of drying droplets in Sub-section 4.2.2 are discussed. The outcomes of the analysis with the tools of ECS and EM are discussed in Section 4.3. Studying topological measures of evolving patterns as functions of space and time in 4.3.1 helps to understand the critical points in the flow dynamics. In 4.3.2, application of the tools in distinguishing between complex real systems of drying droplets is explored. Finally, the pros and cons of the approach to study disordered dynamical systems with respect to their topological evolution are discussed in Section 4.4.

4.2 Material and Methods

In previous Chapter 3, it was hypothesised that if we have two different dynamical system, then the dissimilarity between the two can be quantified by constructing their corresponding Euler Characteristic Surfaces, followed by measuring the Euler Metric between the two ECSs. In order to demonstrate that, the topological tools of Euler Characteristic (χ), Euler Characteristic Surface $\chi_s(r, t)$ and the Euler Metric (EM) are used to study the patterns in - a simulation of a deterministic fluid flow evolution, and a stochastic fluid flow system. The objective of the application is to characterize and distinguish between dynamical fluid flow systems.

4.2.1 Simulated dynamical fluid flow system

As we were interested in dynamical systems that exhibit interesting complex features like vortices and loops, changing with time, we were searching for a model that could exhibit such flow patterns. At that time we stumbled upon an old literature about flow patterns in mixing [125]. When mixing, flow streamlines can form vortices around a constant axis of rotation, leaving islands of unmixed liquids along that axis. One way to address this issue is by varying the orientation of the axis of rotation. A common stirring device, the egg-beater, often shows variations in flow lines due to differences in design, typically referred to as ‘eggbeater flow’ [125]. In a basic eggbeater flow, fluid flow is induced by a hand-held eggbeater with two intermeshed rotating blades, one turning counterclockwise and the other clockwise. As an example of our deterministic dynamical fluid flow system, a modified egg-beater type of flow system was modeled.

The simulation followed an Eq.(5.8), where in each time step t , a point re-

orients itself simultaneously along both coordinate axes, on a two-dimensional phase plane. The point transformation was taken over a Poincare section of a $[0, 1] \times [0, 1]$ square window. The periodic nature of mixing was enforced by calculating the position **mod 1**. Taking the velocity profile on each coordinate direction to be $v_i = i(1 - i)$ for $i = x, y$, the motion of the particle was governed by the following set of recursive equations,

$$\begin{aligned} x_{t+1} &= x_t - ky_t(1 - y_t) \\ y_{t+1} &= y_t - kx_t(1 - x_t) \end{aligned} \tag{4.1}$$

Starting with a seed point, an iteration of the Eqs.(5.8), allowed upto a finite time step, resulted in the flow pattern made of with the displacements suffered by a single particle at discrete time-steps on a Poincare section. The flow pattern was found to be independent of the seed point in the simulation. The seed point in the simulation of the modified egg-beater flow system was chosen to be $(0.9, 0.2)$. The picture, yielded at any discrete time step thus depicted the flow pattern for a single particle generated up to that time step. Figure (4.1) displays the flow patterns at different time steps.

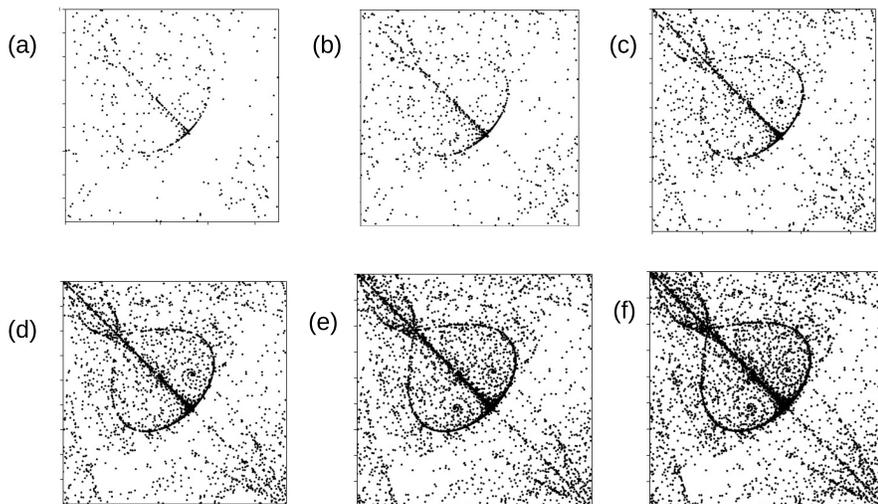


Figure 4.1: Patterns generated with $k = 4.1$ at time steps- (a) $t = 500$, (b) $t = 1000$, (c) $t = 2000$, (d) $t = 3000$, (e) $t = 5000$ and (f) $t = 7000$.

As one can see, in Eq.(5.8), the dynamical system depends solely on param-

eter k . Depending upon the value of k , final patterns generated can be different. We tried different values of k for trials. For higher k values, uniform particle flow over the whole phase plane was observed. For $1 < k < 4$ an almost linear flow movement along the diagonal of the Poincare section was obtained with no significant vortices or loops present. However for $4 < k < 5$ intriguing changes in flow pattern was observed, Figure (4.2). The patterns are plotted at the end of the final time-step of $t = 10000$. The physical significance of k can be described as the parameter affecting mixing frequency.

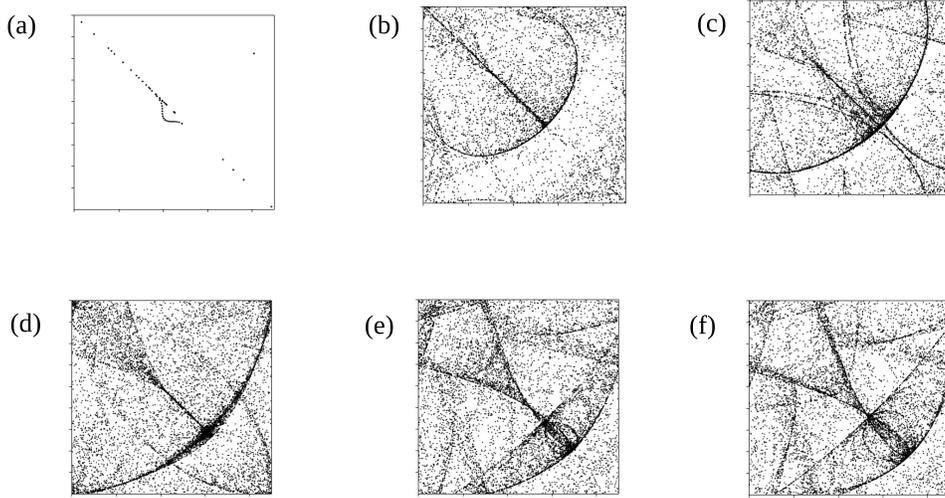


Figure 4.2: Final patterns generated after 10000 time steps for (a) $k = 4.00$, (b) $k = 4.20$, (c) $k = 4.40$, (d) $k = 4.60$, (e) $k = 4.80$ and (f) $k = 5.00$.

As discussed in Chapter(3), studying the flow patterns with multi-scale approach was followed. This gave an idea about the significant features of the patterns that existed over multiple scales r . Here, the flow patterns were analysed at increasing scale, from $r = 1$ to $r = 10$, where $r = 1$ indicates the original image simulated with radius of points measuring 1 pixel unit. As the scale increases, the radius of the black cluster increases including more neighbourhood around each black point, as shown in Figure (4.3). If one may notice, the importance of examining the pattern at different length scales is evident in the images in Figure (4.3). There exists persistent islands of voids at all scales studied and the beautifully symmetric swirls get prominent at some specific scales only.

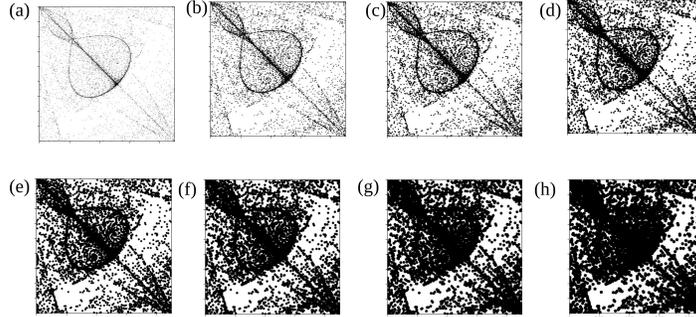


Figure 4.3: Final pattern after 10,000 time steps for $k = 4.1$ at different scales (a) $r = 1$, (b) $r = 2$, (c) $r = 3$, (d) $r = 4$, (e) $r = 5$, (f) $r = 6$, (g) $r = 7$, (h) $r = 8$.

We simulated families of dynamical systems with each system corresponding to a k value. For each of the dynamical systems the Euler Characteristic Surfaces were weaved that captured the spatio temporal topological evolution for each.

4.2.2 Stochastic flow systems: Flow patterns in drying droplets

Along with applying on simulated flow patterns, the earlier analysis on real droplet system was further progressed with configuring similar experimental droplet evaporations, with similar and different particles involved. The Topological tools EC, ECS, EM were applied to compare between the droplet systems. Before applying the tools the images went through appropriate image processing methods, as discussed in Chapter (3). The study compared flow patterns during droplet evaporation for - (i) two similar droplet systems and (ii) two dissimilar droplet systems.

One of the two droplet systems was the evaporating water-ethanol droplet containing suspended silica flakes (diameter ~ 70 nm) on PDMS (40:1) substrate. The other droplet system was the same droplet that we studied in Chapter (3), i.e. the evaporating water droplet containing insoluble suspended polystyrene beads (diameter ~ 500 nm) on PDMS (40:1) substrate. The key difference between two stochastic systems were their compositions. The objective was to study the evaporation features of these drying micro droplets and quantify the difference between them by the Euler metric. The evaporation was studied on soft hydrophobic substrate made of 40:1 polydimethylsiloxane (PDMS) as that facilitated more free motion

of the particles and comparably longer time of evaporation. For preparing the droplet solutions, in one case, polystyrene beads were mixed with de-ionized water and for the other case, nano-sized silica flakes were mixed with water-ethanol solution at 5 : 1 volume ratio. In both cases temperature was constant at ~ 25 °C and humidity fixed at ~ 50 %. Uniform suspensions were achieved by mixing the components on a magnetic stirrer. Droplets of volume $0.5 \mu\text{L}$ were deposited with micro-pipettes on the prepared PDMS substrates. The evaporation dynamics of each droplet was observed using an optical microscope with 20X magnification and stored as images using a computer interface. In a set of repeat experiments, videos of the evaporating droplets were captured from the top view by using a CCD camera. Snapshots from both processes are displayed in Figure (4.4). Image analysis for all the stud-

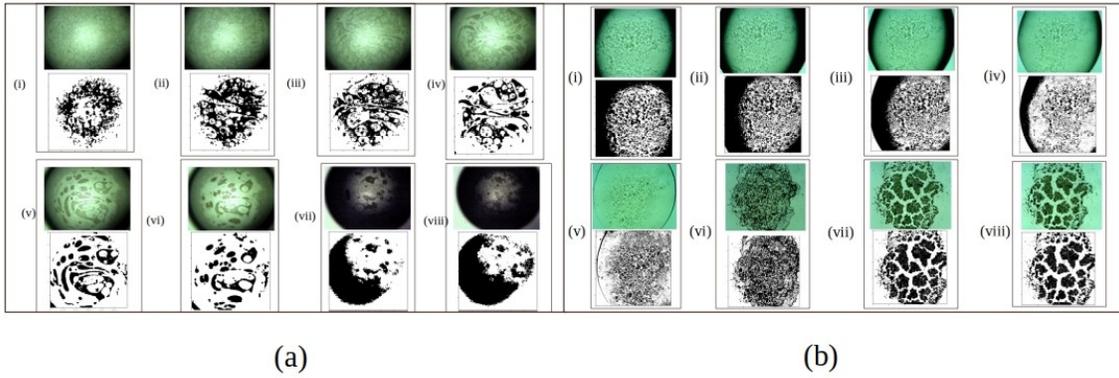


Figure 4.4: Top-view images of flow patterns at different time instants for 2 droplet drying systems. The binary image is displayed along with every real-time image. (a) Patterns of agglomeration while evaporation of water droplet containing polystyrene micro-beads on PDMS substrate (40:1) at (i) 12.24 min, (ii) 15.24 min, (iii) 18.24 min, (iv) 21.24 min, (v) 26.24 min, (vi) 29.24 min, (vii) 32.24 min, (viii) 36.24 min and (b) Agglomeration patterns for evaporation of water-ethanol droplet containing silica nano-particles on PDMS (40:1) substrate at (i) 3.12 min, (ii) 8.06 min, (iii) 10.23 min, (iv) 12.03 min, (v) 14.15 min, (vi) 15.42 min, (vii) 16.39 min, (viii) 17.35 min.

ied droplet evaporating systems were done, dealing individually with each image. Snapshots were sliced from the video of the top surface of the evaporating droplet, followed by processing for optimal thresholding to obtain a binary image as faithful as possible to the real snapshot, Figure (4.4). The binary image files were then processed to compute Euler characteristic(χ) at different times of the evaporation

process. Further, each snapshot was also used to compute Euler characteristic χ at increasing scale r for every chosen time point t . Thus we built Euler Characteristic Surfaces(ECS) for each droplet system.

4.3 Results and Discussion

In the following subsections we present the results and the corresponding analysis, done with the tools of EC, ECS and EM on the studied fluid dynamical systems. The outcome is discussed first for the deterministic simulated fluid flow system, followed by the same for the stochastic droplet drying systems.

4.3.1 Analysis of simulated fluid flow

First, we studied the simple variation of Euler Characteristic (χ) of the final flow pattern with changing the flow parameter k . The family of dynamical systems were simulated with changing parameter k between 4 to 5 and the final pattern generated after 10,000 time steps. k was varied with $dk = 0.1$ intervals but finer variation of k ($dk = 0.01$) was considered for regions with significant changes in the χ vs k plot, Figure (4.5). The change of χ (for the final pattern) is highly sensitive between $k = 4.00$ to $k = 4.4$, indicated as region A in Figure (4.5). In comparison, the change in χ from $k = 4.6$ to $k = 5.0$ is comparatively low, indicated as region B in same figure. This implies that the change in parameter k affects the final pattern's topology more in region A than that in region B. If one looks at the Figure(4.2), the flow patterns corresponding to these k values makes this clearly evident. Thus, the values of Euler Characteristic χ give us a qualitative idea about the topological changes happening in the final flow pattern with changing k . But this analysis does not capture the time evolution of the flow patterns as well as the multiscale topological evolution of each pattern. The importance of examining the connectivity of flow patterns at different length scales r , i.e at different resolutions has already been discussed in (3.5). It is evident from Figure (4.3), the intricate details of the patterns is prominent at certain scales and wash away in others while some persist amidst the variations. A point to notice in Figure (4.3) is that there exist two whorls in the bigger lobes, opposite to each other. The whorls persist through scale $r = 2$ to $r = 5$ but are not prominent beyond that resolution. Whereas, bottom islands below the lobes are more clear from scale $r = 5$ than at lower scale values. Thus different features may get highlighted at different scale of resolution and one can choose a suitable r scale to focus according to the point of interest without being distracted by other features. Therefore, capturing all possible topological information involves

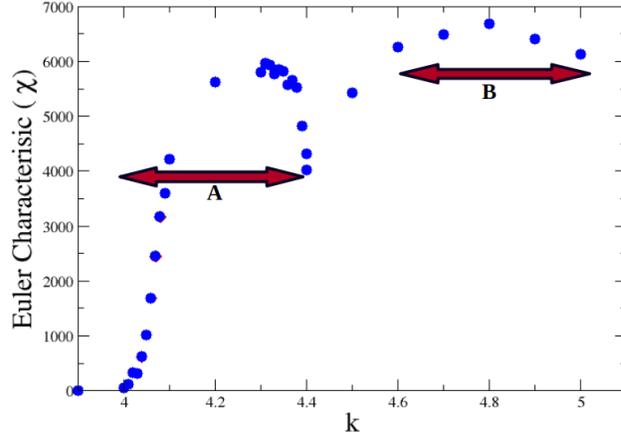


Figure 4.5: Plot of Euler Characteristic (χ) computed for final patterns i.e. after 10,000 time steps for different values of k in the simulated particle flow.

exploring the patterns at multi-scale resolutions.

The Euler Characteristic Surfaces(ECS), $\chi_s(r, t)$, were built for each of the dynamical system of evolving flow pattern with a constant k . To quantify the changes happening in the dynamical system described above for the variation of parameter k , the Euler Metric (EM), between ECSs were estimated. The expression of the Euler metric proposed in Chapter (3), may be recalled for this.

$$d(\chi_{x_1}, \chi_{x_2}) = \|\chi_{x_1} - \chi_{x_2}\|_2 = \left(\int_{[0,R] \times [0,T]} (\chi_{x_1} - \chi_{x_2})^2 \right)^{\frac{1}{2}} \quad (4.2)$$

where the Euler Characteristic Surfaces need to be constructed up to scale R and time T for the dynamical systems x_1 and x_2 . To compute the integral in Eq.(5.1), Monte Carlo integration discretization approach was followed where the absolute differences $|d\chi(r, t)|^2$ are computed between all the corresponding pair of points (r, t) in the two ECSs that are being compared. It is therefore essential that the pair of ECSs are of identical size and grid spacing.

Given two very differently evolving dynamical systems, their Euler Characteristic Surfaces are expected to be different in view. This can be verified with illustrating The Euler Characteristic Surfaces for the fluid flow pattern for the pa-

parameter with $k = 4.1$ to $k = 4.0$. As seen in Figures (4.6 a and b) two very different patterns were obtained after letting both the fluid dynamical systems evolve until the same number of time steps $t = 10,000$. As anticipated, the corresponding ECSs too, turned out to be remarkably different, Figures (4.6c and d).

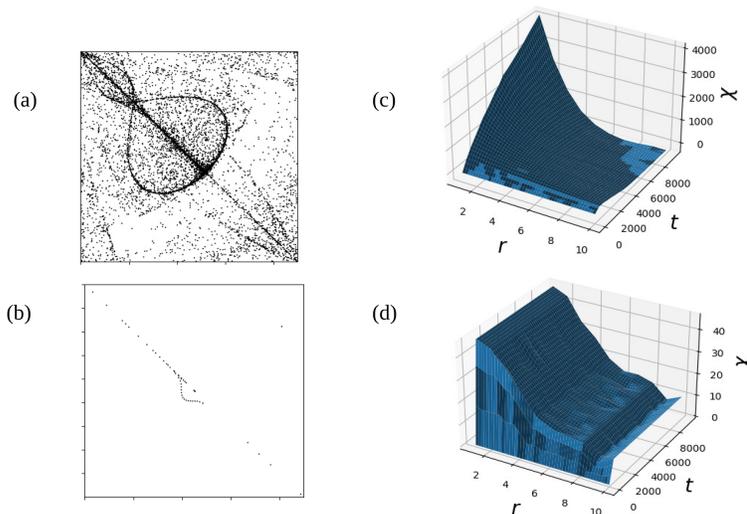


Figure 4.6: (a) Final pattern after 10,000 time steps for $k = 4.1$ at scale 2, (b) Final pattern after 10,000 time steps for $k = 4.0$ at scale 2, (c) Euler Characteristic Surface for $k = 4.1$ and (d) Euler Characteristic Surface for $k = 4.0$

The Euler Characteristic Surface (ECS) for a few other values of k s in between 4 to 5, were also constructed. Figures (4.7 a-d) display the final flow pattern with different k values along with their corresponding ECSs. Though some of the ECS may look roughly similar, a quantitative measure of the similarity or dissimilarity between them was needed to be estimated by computing the Euler Metric (EM) between two ECSs as defined in Eq.(5.1). Physically, the Euler Metric between any two ECSs is a measure of the L^2 norm distance between the surfaces as defined by Eq.(5.1). The EM value between two ECSs does not depend on which of the two surfaces is acting as the reference, besides it factors in the possible intersection of the surfaces themselves, as shown schematically in Figure (4.8). The Table (4.1) shows the estimated values of the Euler Metric(EM) between ECSs for flow dynamical systems with k and $k + dk$.

For changing the parameter k by an amount $dk = 0.1$, the variation in EM

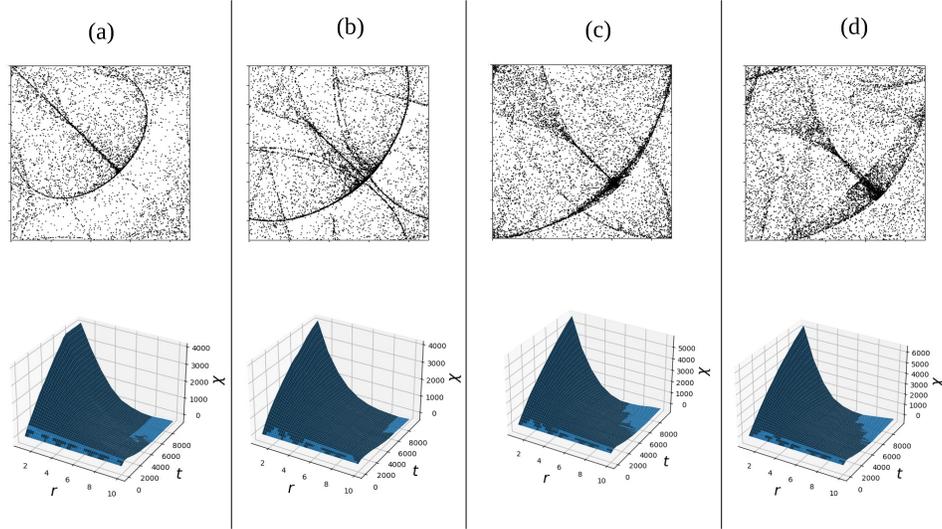


Figure 4.7: (a) Final pattern after 10,000 time steps for $k = 4.2$ and its corresponding ECS, (b) Final pattern after 10,000 time steps for $k = 4.4$ and its corresponding ECS, (c) Final pattern after 10,000 time steps for $k = 4.6$ and its corresponding ECS and (d) Final pattern after 10,000 time steps for $k = 4.8$ and its corresponding ECS.

$d(\chi_k, \chi_{k+dk})$ around any k is shown in Figure (4.9), the values are tabulated in Table (4.1). The highest $d(\chi_k, \chi_{k+dk})$ values occur between ECSs corresponding to $k = 4.0$ and 4.1, marked by **A** in Figure (4.9). The drastic changes in flow patterns at these k values, shown as insets in the figure, reflects effectively in the Euler Metric d . However if the Euler Metric ($d\chi_k, \chi_{k+dk}$) between the two ECSs is calculated for finer variation, $dk = 0.01$, around same $k = 4.0$, the EM value is considerably small, see region (B) of Figure (4.9). Thus for the studied simulated flow pattern, the EM between two ECSs corresponding to dynamical systems with k and $k + dk$, is a function of both k and dk , i.e. the EM $d(k, k + dk)$ is a function of the scale of resolution of parameter variation!

Another point that one can notice from Figure (4.9), there may be a region where for any given dk , the variation in EM remains almost constant, e.g. between points C and D. For variation with resolutions both $dk = 0.1$ and $dk = 0.01$, the topological fluctuations in the flow pattern were substantially stable. This is also evident in the snap shots of the final flow patterns in the insets of the Figure (4.9) when k ranges between 4.5 to 4.6. Thus it may inferred that the dispersion in the EM values for equal dk help in identifying stable regimes of the dynamical flow patterns.

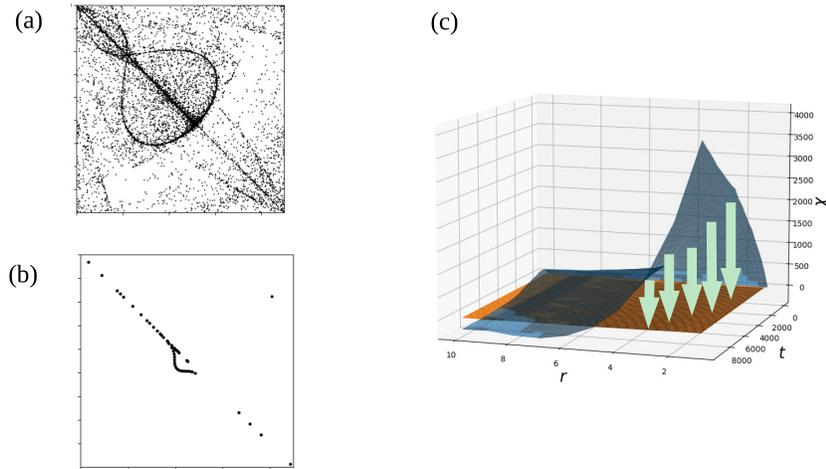


Figure 4.8: (a) Final pattern after 10,000 time steps for $k = 4.1$, (b) Final pattern after 10,000 time steps for $k = 4.0$, (c) Euler Metric : Arrows indicate the difference between corresponding points between the Euler Characteristic Surfaces for $k = 4.1$ and $k = 4.0$.

Further from Figure (4.9), the values of EM $d(\chi_k, \chi_k + dk)$ in the range between $k = 4.5$ and $k = 4.9$ have a comparatively smaller dispersion, indicating more stabilized flow dynamics around this region. On the other hand, the dispersion in EM around $k = 4.0$ is comparatively higher than the other values in the entire range, implying that a *drastic topological transition* occurred around that point, making it a critical region. *The ability to determine stability in flow dynamics with respect to perturbations or variations in a parameter k via EM computation, adds to the potential of the measure.*

4.3.2 Analysis on stochastic dynamical fluid system

For application on real dynamical systems, three droplet systems were mainly considered: two being evaporating water-ethanol droplet containing silica flakes (these are designated as Set 1 and Set 2); and the third one being an evaporating water droplet containing polystyrene beads. The aim behind choosing such systems were to examine as before – (a) How does ECS and EM quantify similarity between the same kind of dynamical systems and (b) How does the dissimilarity in droplet systems infused by the particles and medium reflect in our topological measure.

Table 4.1: Measures of Euler Metric($d\chi_k, \chi_{k+dk}$)

k values of ECSs	Euler Metric
4.00 - 4.10	2,68,450
4.10 - 4.20	61,469
4.20 - 4.30	41,485
4.30 - 4.40	1,01,774
4.40 - 4.50	82,017
4.50 - 4.60	38,078
4.60 - 4.70	35,261
4.70 - 4.80	19,505
4.80 - 4.90	21,582

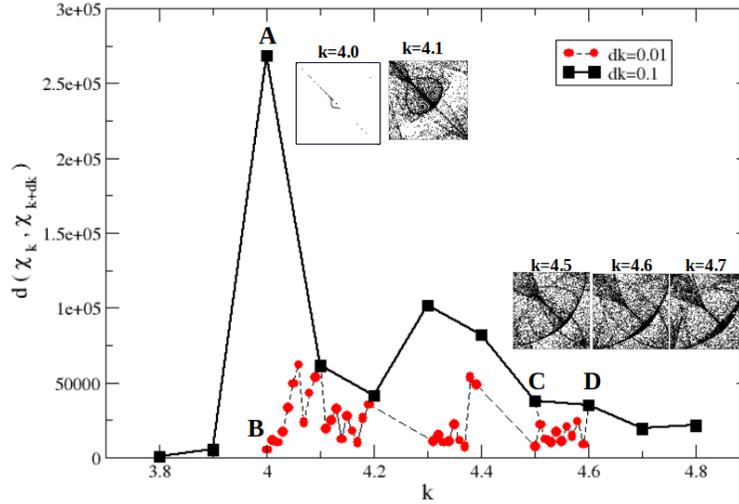


Figure 4.9: Euler Metric(d) computed for Euler Characteristic Surfaces corresponding to flow patterns with $k + dk$ and k . The insets depict flow patterns corresponding to significant k values at the end of 10,000 time steps. EM is computed at two values of variation dk as indicated in the legend. Point A marks EM between flow pattern with $k = 4.0$ and 4.1 , is a point of critical topological changes. The region between C and D are stable flow regions.

After optimally thresholding the images for each droplets, the prescription to build the Euler Characteristic Surfaces for each of the three sets of dynamical

Table 4.2: Measures of Euler Metric

Droplet Systems of ECSs	Euler Metric
silica flakes droplet and polystyrene beads droplet	1691
two different silica flakes droplets	1355

flow systems were followed. Since the estimation of Euler Metric between two Euler Characteristic Surfaces needs the necessity of both the system having similar ranges along time axis and scale (or resolution) axis and total evaporation time for each of the droplet systems was different, time (t) was normalised with respect to the total evaporation time (t_e) of the particular droplet. Evaporation time, t_e , was 19 minutes for the first droplet experiment with silica flakes (Set 1), 20 minutes for the second droplet experiment with silica flakes (Set 2); and 40 minutes for the droplet with polystyrene beads.

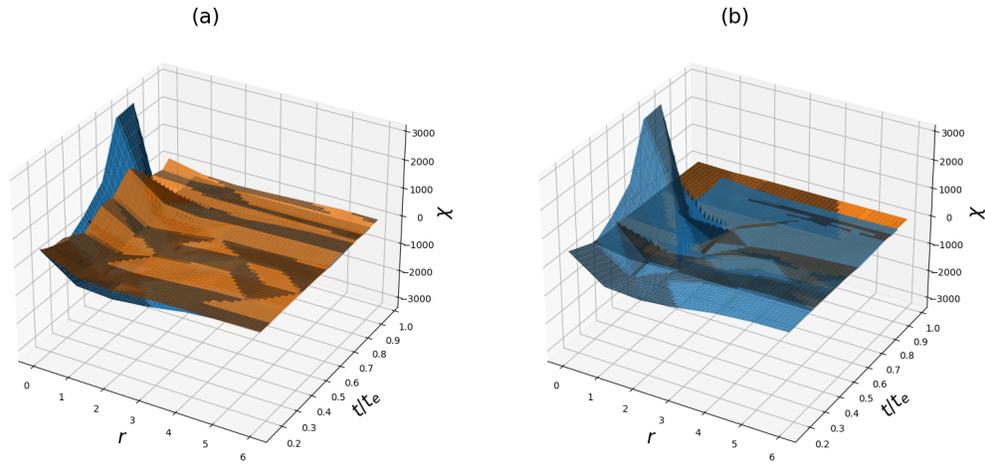


Figure 4.10: Euler Characteristic Surfaces – (a) for evaporating water-ethanol droplets containing silica flakes for Set 1(blue) and Set 2 (orange), (b) for evaporation of water droplet containing polystyrene beads (orange) and water-ethanol droplet containing silica flakes (blue).

As mentioned before, all analysis of the droplet flow systems were carried out on binary images of the real images, see Figure (4.4). Therefore, the analysis and result depends highly on the optimization of the entire image processing procedure. Suitable Gaussian filter was used to process the images, (discussed in Chapter (3), 3.3.2). After thresholding the binary matrices of the images were used to estimate the Euler Characteristic χ at different times and length scales r and the Euler Characteristic Surface(ECS), $\chi_s(r, t)$, was built for each of the droplet system. The ECSs of the corresponding 3 droplet systems are displayed in Figure (4.10). The two ECSs for the two water-ethanol-silica flake droplets, Figure (4.10a), almost coincide which is not surprising. Figure (4.10b) displays the ECS for the water-polystyrene droplet and a water-ethanol-silica flakes droplet and the Euler Characteristic Surfaces are clearly different to each other.

To quantify the distance between the Euler Characteristic Surfaces, the Euler Metric (EM) was computed for – (1) the two water-ethanol droplet systems containing silica flakes with same concentration, Figure (4.10a), and (2) the water-polystyrene droplet system with the water-ethanol-silica droplet system, Figure (4.10b). To note, in both cases normalised time (t/t_e) was considered. The estimated values of the Euler Metric (EM) are given in Table(4.2).

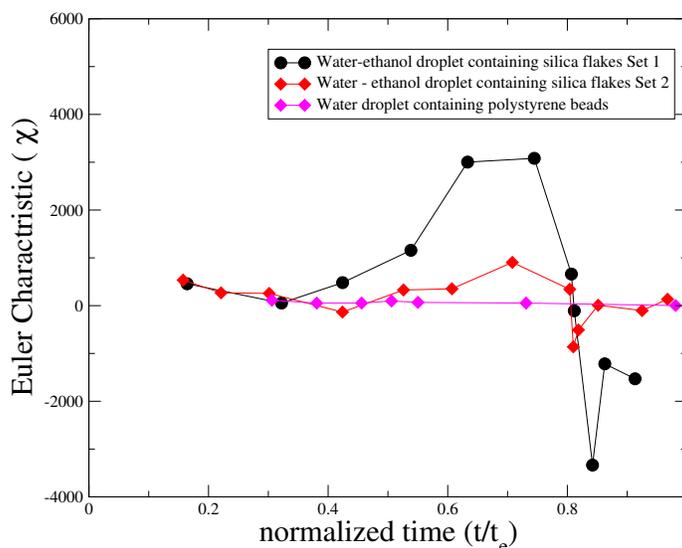


Figure 4.11: Euler Characteristic versus time at scale $r = 1$ for evaporating water-ethanol droplet containing silica flakes and for evaporation of water droplet containing polystyrene beads.

As one may expect that the value of the Euler Metric for ECSs between two similar droplet systems - the evaporating water-ethanol with silica flakes, to be close to zero, but it is not. For deterministic systems, the Euler metric for two similar dynamical systems is zero but in reality, replicating two similar dynamical systems for such non-equilibrium phenomena like evaporating droplets having dynamic particle flow within, is almost impossible! The flow patterns are very sensitive to even minor fluctuations in particle number, shape and size distribution, apart from ambient perturbations like air currents. Therefore, the value of estimated Euler Metric between the ECSs of two experimentally similar droplet system is justified. Despite this, the similarity in the variation of Euler characteristic(EC) against normalised time at scale $r = 1$, for the two droplets containing silica particles, is unmistakable, even though the magnitude of variation in χ is not the same, Figure (4.11). The shape change in the χ vs normalized time(t/t_e) implies that some critical topological changes occurred in the water-ethanol-silica systems around $(t/t_e) = 0.7$ and also around $(t/t_e) = 0.85$. Comparing with the snapshots taken from the video, the peak in χ , around $(t/t_e) = 0.7$, corresponds to the situation in Figure (4.4b,(iv)) where maximum number of crystals were counted. This followed by the crystals started joining together to form a large aggregates with small islands of fluid into them. The sudden dip in the χ versus t/t_e curve, for $t/t_e \approx 0.85$ corresponds to the situation in Figure (4.4b,(vi)), which is another critical point in the dynamics when the large crystalline aggregate with maximum holes within the network existed and after that point it broke into dense clusters, elevating the χ values. The system is more or less stable at other times. In contrast, the clear difference between the dynamics of water-polystyrene and the water-ethanol-silica flake system is evident from the distinct difference in the nature and magnitude of Euler characteristic variation of the droplets as shown in the Figure(4.11)(examine legends for different droplet systems).

Drawing a parallel between the simulated system and the real dynamical system of evaporating droplets, the variation of parameter k in the former may be taken analogous to changing the solution type of the two droplet systems. In both the situations, the ECS and EM have been able to quantify the topological change that follows the change in the dynamics of the systems. Along with helping to distinguish between dynamical system the measure further helped in identifying stability or critical regions in - the flow systems.

4.4 Conclusion

To conclude the study, the application of our proposed topological construct on dynamical systems to quantify the similarity or dissimilarity between them was found to be effective. The proposed hypothesis that the Euler Characteristic Surfaces(ECS) behave like topological signatures for dynamical systems, was verified. The ECSs of widely different dynamical systems were found to be widely different as well. Even for the same family of dynamical systems, Euler Characteristic Surfaces(ECS) along with the Euler metric(EM) captures the variations due to a slight perturbations of the single system parameter that often generate drastic changes in the dynamics. While analysing evolving patterns in dynamical systems with the tools, the Euler Characteristic(EC) and the Euler Characteristic Surfaces (ECS) give a qualitative idea about the topological changes happening in the systems, the the Euler Metric(EM) quantifies those changes. Though the effectiveness of the ECS and EM was more pronounced in the studied deterministic simulated system, it was also evident in studying very complex real systems of drying droplets.

The possibility of quantifying the similarity and dissimilarity between two dynamical systems through a simple topological measure, Euler Characteristic Surfaces and the Euler Metric between them, is indeed a novel approach in Topological Data Analysis and it shall be compared with the existing approaches in TDA in the next chapter. In a nutshell, the significance of the Euler metric studied till now can be summarized as :

- The EM is a powerful tool to sense the topological changes arising in a disordered dynamical flow.
- The EM can help distinguish between completely different flow systems. For any given fluid dynamical system:
 - The EM value is a function of both the system parameter and its variation.
 - The EM helps to distinguish between stable flow regimes as well critical points of drastic changes in flow topology.
- Drastic variation in the EM may act as pointers to underlying differences in physical principles responsible for changes in flow topology.

Chapter 5

The mathematical framework and stability of the Euler Characteristic Surface

5.1 Introduction

The last part of my dissertation, which was expanded after the successful endeavor with the Euler Characteristic Surface(ECS) and Euler Metric(EM) in Chapters(3,4), was to establish the detailed mathematical framework for the novel tools that we had introduced. For that, I delved into Homology from which the idea of Euler Characteristic Surface was inspired, studied the basics of the approaches used in Homology to describe the topology of complex data, and tried to find the stability of our tool with respect to the already established mathematical framework in Homology. Studying stability ensures that the introduced Euler Metric adheres to the rigorous framework of algebraic geometry providing consistency within the mathematical theory. Additionally, we aimed to explore and compare the usefulness of the ECS and EM with the already existing ones.

In recent times, studying data in terms of its topology has been a rapidly evolving field with approaches like Topological Data Analysis (TDA). Especially, for complex data understanding the shape of the data often turns out to be advantageous (We already saw its example in the last two chapters). The initial studies in TDA

started with explorations of static data clouds [126, 99, 127] but gradually it is advancing towards analyzing complex data from different kinds of dynamical systems, e.g. biological aggregation [128], social systems, gene networks [129, 130], economic/financial data [131, 132], power engineering [133] etc. Different approaches have been followed in building topological tools to study the shape of complex dynamical data sets, e.g., crocker plots [128, 134], vineyards [135] maps, crocker stacks [136], zigzag persistence [137, 138] etc. Most of these tools can be computationally complex and stability results need to be further investigated.

This chapter focuses on establishing the mathematical foundations of the topological tool, “Euler Characteristic Surface” (ECS) introduced in Chapter 3, that summarizes the multiscale topological information of a dynamical data set and thus helps in the characterization of and distinction between dynamical systems in Chapter 4. It should be mentioned that a slightly different form of Euler Characteristic Surfaces have been studied by different groups of authors [139, 140], but those works do not consider time varying data.

This chapter is outlined as follows. A preliminary idea about the concepts used from Homology is given in Section(5.2). Geometrical simplicial complexes such as the Alpha Complex are defined and the viability of using it in the analysis of large data sets is justified in 5.2.1. Concepts of persistent homology and its metrics - the p-Wasserstein metric and the bottleneck metric - that are often used to distinguish between data sets are discussed next in 5.2.2, followed by the pathway in establishing a relation between Euler Characteristic Surface and the persistence modules, and the stability of ECS under small perturbations of the data set proposed in 5.2.3. Besides the theoretical framework, the robustness of the tools of ECS and EM is demonstrated on two simulated dynamical data sets - (i) with a variable number of data points and (ii) a conserved number of data points in Section 5.3. Sub-section 5.3.1 (modified eggbeater flow model) compares the ECSs constructed via two different methods- coarse-graining and geometric simplicial complexes, along with testing the stability of the Euler Characteristic Surface through the Euler Metric measured with the dynamical data sets. The second model (Vicsek model), considered in 5.3.2, has been used for comparing the ECS and the Euler metric with the Persistent Homology tools. The focus of the work was to establish and validate the relations between Euler Characteristic Surface and Persistent Homology, building the theoretical foundation of our proposed topological construct.

5.2 Mathematical Framework

In this section I shall briefly discuss the mathematical concepts and terms that are related to our topological construction, ‘Euler Characteristic Surface’, and follow up establishing the theoretical framework of our topological construction.

5.2.1 Cell Complexes and Euler Characteristic Surfaces

To begin with let us first simplify what precisely a dynamical system means here and what is the basis of studying its topology. In this dissertation by dynamical system, we refer to *the time evolution of a finite set of points in some domain in Euclidean space*. The method that was used to explore these dynamical systems was to study the *time evolution of the topology of such a set of points* via the topological invariant Euler Characteristic χ . Along with the time evolution, analyzing the multi-scale topology of the set of points was achieved.

In order to extract multi-scale topological information, first, one needs to build cell complexes at a scale $r \geq 0$ with the given set of points.

Cell complex: Cell complexes [116] are topological spaces built inductively by attaching higher dimensional cells to the ‘lower dimensional skeleton’, where the ‘0-dimensional skeleton’ is the *vertex set* i.e. the point cloud to be studied. For example,

- In the case of studying pixellated digital images, the cell complex used was the complexes built from the set of pixels using the physically inspired ‘union of r -neighborhood method’, Chapter 3, which can be called as ‘coarse-graining’(following the algebraic geometry terminology).
- A cell complex can also be a *simplicial complex* which is commonly used for studying the topology of a finite set of points. A simplicial complex is a collection of simplices that glue together following certain rules, discussed later in this subsection.

Summarizing, Cell complexes break down complex spaces into simpler pieces for analysis.

Cell complex with coarse-graining:

In the last chapters, we dealt with both real image data and simulated data where we used discrete grids and followed the ‘union of r - neighborhood’ or ‘coarse-graining’ approach to get the multiscale pictures of the dynamical patterns. Here, the same approach will be described but in a more mathematical language, and then it will be compared with the popular method of building a cell complex which is a simplicial complex.

The construction rule for ‘coarse-graining’ is that given a finite set of points $F = \{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^N$, where one replaces a point by a ball of some non-trivial radius, the ‘scale’. More specifically, one replaces F by the set $F_r = \cup_{x \in F} B_r(x)$, which is the union of closed balls of radius $r > 0$ centered at each point of F . For pixelated images or discrete grid points, each black pixel/grid gets replaced by the union of that pixel with its nearest neighbors up to its r -th neighborhood, Figure (5.1). Unlike F , the set F_r encompasses all possible topological features, with the topology varying according to the value of r . For dynamical systems, the $x_i \in F$ is time-dependent, i.e. x_i is a function of t , and thus “coarse-grained” set can be represented as $F_{r,t}$. The cell complexes can therefore be thought of as a spatio-temporal data set. Once the cell complexes $F_{r,t}$ get built at different scale r and different time t , studying the dynamic topology of $F_{r,t}$ can be done by estimating the topological invariant - the Euler Characteristic for each complex.

Cell complex: Simplicial complex

Simplicial complexes can be thought of either as *geometric structures* in \mathbb{R}^d where ‘simplices’ are assembled following some specific rules or as a purely combinatorial structure called an *abstract simplicial complex*. By simplices, one means n -dimensional versions of triangles, e.g. a point being $0 - d$ simplex, a line being $1 - d$ simplex, a triangle being $2 - d$ simplex, etc. The definitions of the simplicial complex followed by the definition of its variations are given below.

Definition 1. *An abstract simplicial complex is a finite collection K of finite non-empty sets such that if σ is an element of K , so is every non-empty subset of σ . An element σ of K is called a simplex of dimension $|\sigma| - 1$. The dimension of the entire simplicial complex K is $\max_{\sigma \in K} \dim(\sigma)$. A subset $L \subset K$ is said to be a sub-complex of K if L is also an abstract simplicial complex.*

A geometric simplicial complex is clearly an abstract simplicial complex, and

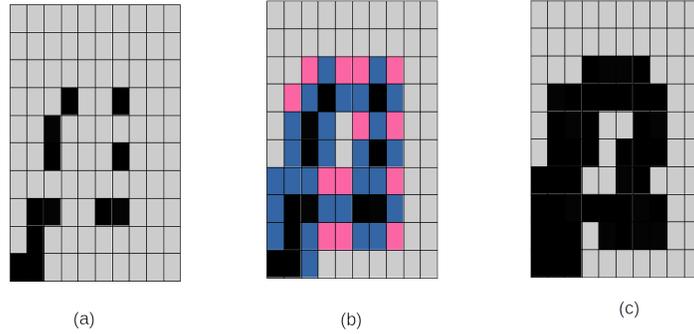


Figure 5.1: Schematic of Coarse-graining/union of r – neighborhood approach: (a) pixels at scale 0 (no-coarse graining applied) (b) The nearest neighborhood of the black pixels marked with blue color and the next nearest neighbors marked with pink color, (c) coarse-graining applied at scale 1 with diagonal connections attached with 0.5 probability.

any abstract simplicial complex can be viewed as a geometric simplicial complex (its geometric realization) ¹. Therefore, henceforth I shall only refer them as simplicial complexes. There are two commonly used methods of constructing simplicial complexes on a set of points, the Vietoris-Rips complex and the Čech complex. In my analysis I used Alpha complex which is a variation of the Čech complex construction, being an example of a Nerve Complex.

Definition 2. For a finite set of subsets $U = \{U_1, U_2, \dots, U_n\}$ in \mathbb{R}^n , the Nerve complex $\mathcal{N}(U)$ is defined by the rule, a subset $\sigma \subset U$ is a simplex iff $\bigcap_{i \in \sigma} U_i \neq \emptyset$.

In Figure (5.2a, b) the construction of the Nerve complex for given subsets is shown schematically considering the intersections of the subsets. Similarly, if one considers a set of points $P = \{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^d$, and $B(x_i, r) = \{x \in \mathbb{R}^d : \|x_i - x\| \leq r\}$ is the closed ball of radius r , then the Čech complex at scale r is the simplicial complex $\check{C}ech(P, r) = \{\sigma \subset P : \bigcap_{x_i \in \sigma} B(x_i, r) \neq \emptyset\}$. The Čech complex is the nerve of the closed r -balls around the set of points P and encodes the intersection pattern of the cells. Thus the Čech complex is homotopically equivalent ² to the

¹Every abstract simplicial complex of dimension d can be geometrically realized in \mathbb{R}^{2d+1}

²In algebraic topology, two topological spaces are said to be homotopically equivalent (or ho-

r -neighborhood of P , the cell complexes constructed through coarse-graining.

The problem with working with the Čech complex is that, for a set of points in \mathbb{R}^d , the Čech complex constructed on it may have arbitrarily high dimensions depending on the intersection pattern of the closed r -balls. This method can therefore lead to computational challenges, especially for large data sets. We shall resort to using the **Alpha complex** which bypasses this. For a finite set of points in \mathbb{R}^2 , the Alpha complex can be defined as below ³.

Definition 3. *Consider a finite set of points in $F = \{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^2$, no four points lying on the same circle, and let $r > 0$. The alpha complex at scale r is defined as the nerve of $V_x \cap B(x, r)$, where V_x is the Voronoi cell corresponding to the point x .*

Figure (5.2c) shows the intersections between Voronoi cells and balls of radius r around a point cloud that generates the Alpha Complex shown in Figure (5.2d). The alpha complex has the same homotopy type of $F_r = \cup_{x \in F} B_r(x)$, and its construction ensures that it is in \mathbb{R}^2 [141] [142].

The Euler Characteristic of a cell complex can be defined as the alternating sum of the number of cells of each dimension. If \mathcal{K} is a finite dimensional cell complex, the Euler Characteristic is then defined by $\chi(\mathcal{K}) := \sum_{i=0}^{\dim(\mathcal{K})} (-1)^i \text{card}(\sigma_i)$, where σ_i is the collection of i -dimensional cells in \mathcal{K} . For simplicial complexes this simplifies into alternating sum of n -dimensional simplices where n runs from 0 to \dim of the simplicial complex. For the simplicial complexes that were studied, following Figure (5.2b) this comes down to the alternating sum of the number of points (different coloured balls), the number of lines and the number of triangles (filled with blue colours). The above definition can be reformulated as an alternating sum of Betti numbers (this is defined later in the section), which actually simplifies into the definition of the Euler Characteristic, $N_b - N_w$, that was used in our cell complexes built with “coarse-graining” or “union of r -neighborhood” method. To establish the

motopy equivalent) if they can be continuously deformed into each other. This concept is a central idea in homotopy theory, which studies spaces up to such deformations.

³the Alpha complex was defined here in \mathbb{R}^2 as our applications were in that setting, but the definition works in \mathbb{R}^n

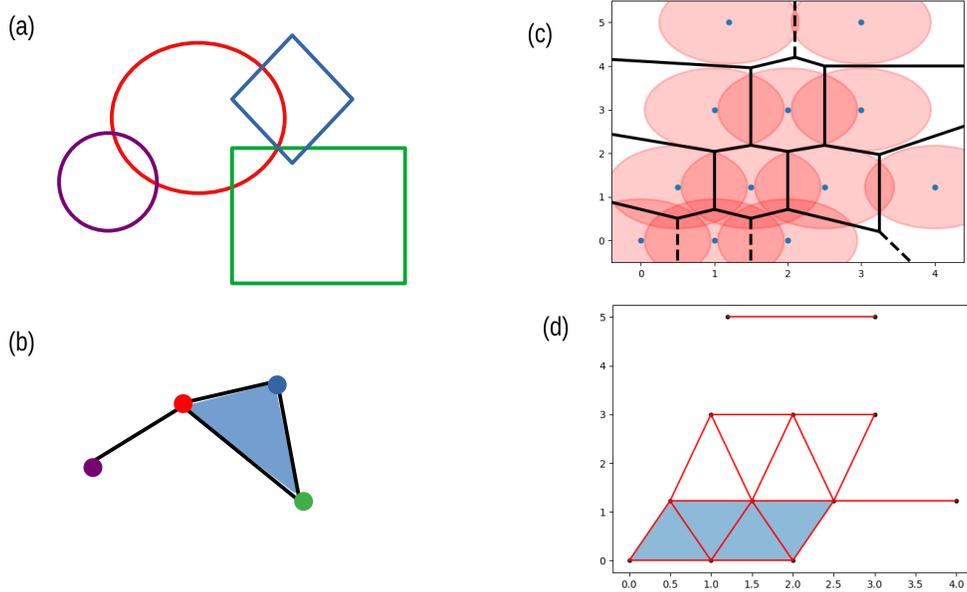


Figure 5.2: (a) The collection of subsets, (b) The corresponding Nerve complex, (c) Schematic to represent the intersection of Voronoi cells and balls of radius r around a point cloud and (d) The corresponding Alpha complex.

mathematical relationship of our tool, The Euler Characteristic Surface, with homology, I shall henceforth adopt the notion of simplicial complex and refer to it as $\chi(\mathcal{K}_{r,t})$, where $\mathcal{K}_{r,t}$ is the alpha complex at scale r on the dynamical system at the time instant t . Here both parameters r, t can vary continuously and one gets a two-dimensional surface as a map of the dynamical system. It is to be noted for Alpha complex the scale parameter r need not be an integer but any finite positive value. Figure (5.3) shows Alpha complexes of the same point cloud at different scales r .

The definition of Euler Metric (EM), which quantifies the difference between two ECSs describing two different dynamical systems, may also be generalized as the p -Euler Metric, by defining it using the L_p norm:

$$d(\chi_1, \chi_2) = \|\chi_1 - \chi_2\|_p = \left(\int_{[0,R] \times [0,T]} |\chi_1(r, t) - \chi_2(r, t)|^p \right)^{\frac{1}{p}} \quad (5.1)$$

with $p = 1, 2$. With $p = 2$, the Euler Metric (EM) is a Hilbert space metric. The

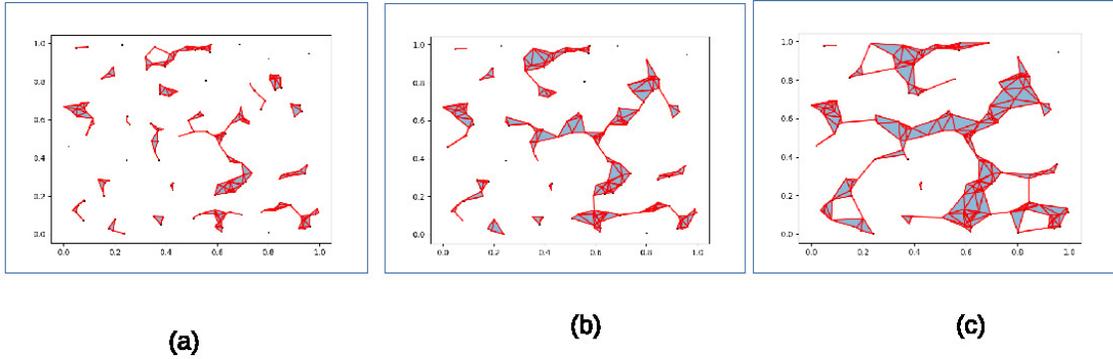


Figure 5.3: Alpha complexes of a point distribution generated using Vicsek model at scale – (a) $r = 0.03$, (b) $r = 0.045$, (c) $r = 0.055$.

intention behind this generalization can be understood later in Section(5.2.3) when the stability of the metric gets tested.

Our technique can be summarized by the following flowchart. Starting from a dynamical system $\{D\}$ represented by a dynamic point cloud F_t , we create the coarse-graining cloud $F_{r,t}$, from which create a family of complexes $\mathcal{K}_{r,t}$ which has the same homotopy as $F_{r,t}$. Then we create a topological signature of the dynamical system by the Euler Characteristic Surface $\chi_s r, t$.

$$\boxed{\mathcal{D} = F_t \mapsto F_{r,t} = \cup_{x_t \in F} B_r(x_t) \mapsto \mathcal{K}_{r,t} \mapsto \chi_s(r, t)}$$

5.2.2 Homology, Betti Surfaces and Persistence Diagrams

After defining the simplicial complexes, let us now be familiar with the concepts and terms of homology that will be used further to establish the mathematical stability of our topological framework. Homology provides a way to analyze and classify

topological spaces based on their structural features, such as connected components, holes, and voids. It assigns a sequence of Abelian groups⁴, called *homology groups* to a topological space. These groups capture information about the space’s k – *dimensional* features/holes.

Homology

Given a simplicial complex, one has k -chains, which are formal linear combinations of k -simplices for a given dimension k , with coefficients in a convenient field of choice (\mathbb{Z}_2 coefficients are chosen here, i.e. the coefficients are either 0 or 1). These k -chains are equipped with the natural addition operation and form an abelian group

C_k , [$C_k = \sum_{i=1}^N a_i \sigma_i$, where $\sigma_0, \sigma_1, \dots, \sigma_N$ are k -dimensional simplices of the simplicial

complex and a_i are the coefficients]. The boundary of a k -simplex is defined as the sum of all its $(k - 1)$ -dimensional faces, and this definition extends to all k -chains as well. This creates a boundary homomorphism ∂_k from chain group C_k to chain group C_{k-1} . A k -chain with empty boundary is called a k -cycle ($\partial_k(C) = 0$), forming a group Z_k of k -cycles, which is a subgroup of C_k . On other hand, a k -chain b is defined as a k -boundary if it is boundary of a $(k + 1)$ -chain, d . (i.e. $b = \partial_{k+1}(d)$). Thus a group B_k of k -boundaries is formed. Since, $\partial^2 = 0$, the B_k is a subgroup of Z_k . Finally, the k -th homology group H_k is the quotient of the k -th cycle group by the k -th boundary group [$H_k = Z_k/B_k$], and the k -th Betti number β_k is the rank of this homology group. This group measures the k -dimensional holes in the simplicial complex by identifying cycles that are not boundaries. For further details on the theory of homology, see [116].

Persistent Homology

The framework of homology can be expanded with studying multiscale topological features as often in dynamical systems some features are observed to persist over several length scales indicative of some inherent significance. While the detection of topological features, e.g. “holes” or lack of matrix in an appropriate dimension in a simplicial complex arising from a point cloud, can be done using the device of *homology*; for quantifying the “persistence” of such topological features/holes, one needs to further follow *persistent homology*. In Topological Data Analysis (TDA) one of the most established tools is Persistent Homology. A very brief idea of the

⁴An Abelian group is a set combined with an operation (like addition or multiplication) that satisfies four main rules: closure, associativity, identity, and inverse.

apparatus of Persistent Homology is given here.

The workflow of Persistent Homology begins with the process of filtration. For a static point cloud, the associated Alpha complexes for a given filtration (at different scales) make the set of sub-complexes, $\emptyset \subset K_0 \subset K_1 \cdots \subset K_n = K$ ($0, 1, 2, \dots, n$ being the different scales of filtration). One also gets a sequence of associated homology groups $H_k(K_0) \rightarrow H_k(K_1) \rightarrow \cdots \rightarrow H_k(K_n)$, and - for each value of scale $s < t$ - the induced homomorphism $f_k^{s,t} : H_k(K_s) \rightarrow H_k(K_t)$. The k^{th} persistent homology group $H_k^{s,t}$ is the image of $f_k^{s,t}$, and the corresponding k^{th} persistent Betti number $\beta_k^{s,t}$ is the rank of $H_k^{s,t}$. While the k^{th} persistent Betti number quantifies how many of the homology classes (cycles) of K_s still survive in K_t , the entire picture is graphically represented in a k -dimensional persistence barcode and k -dimensional persistence diagram. The k -dimensional persistence diagram is a multi-set of points (b, d) that records the k -dimensional homology classes that are ‘born’ at filtration B indicated by coordinate b and ‘die’ at filtration D indicated by coordinate d . Figure (5.5) shows a schematic of Persistence diagram. For further details on the theory of persistent homology, see [141].

These Persistence Diagrams are also used to compare between data sets. There exists a few measures that quantify the difference between two persistence diagrams corresponding to two sets of point clouds. Of the metrics that are usually considered in the space of persistence diagrams are the bottleneck metric and the p -Wasserstein metrics [141]. To estimate the difference, one creates bijections/-matchings between two persistence diagrams (including diagonal points) and takes the infimum of all the matching distances in the appropriate norm (ℓ_∞ or ℓ_p), Figure (5.4). In the case of Bottleneck distance, the matching distance for each match is the maximum of the d_∞ distances. For the Wasserstein metric the p -th root of the sum of all p -th powers of d_∞ distances in a match is considered as the matching distance. Finally, in both cases, the minimum of all the matching distances is considered as the metric value. The table in Figure (5.4) portrays the step-by-step estimation of the above metric distances. Formally, the p -Wasserstein distance (p can be $1, 2, \dots, \infty$) between two k -dimensional persistence diagrams \mathcal{A} and \mathcal{B} is given as: ⁵

$$W_p(\mathcal{A}, \mathcal{B}) = \inf_{\pi: \mathcal{A} \rightarrow \mathcal{B}} \left[\sum_{(b,d) \in \mathcal{A}} \|(b, d) - \pi(b, d)\|_\infty^p \right]^{\frac{1}{p}} \quad (5.2)$$

⁵Note that the ambient metric is the d_∞ metric on the plane.

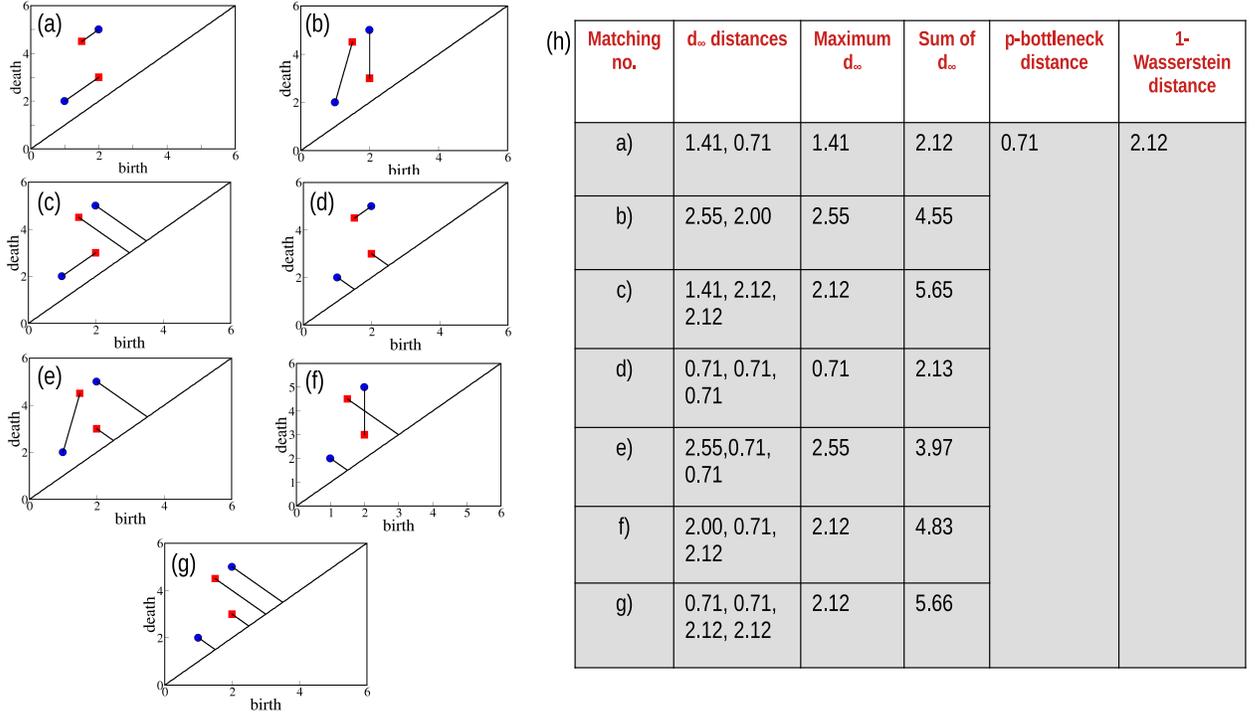


Figure 5.4: (a)-(g) represent different possible bijections or matching between two Persistence Diagrams (red and blue), (h) The table showcases the matching distances and the estimated metrics.

Betti Curve and Betti Surface

In the next Section, (5.2.3) the relationship between the Euler Characteristic Surface will be set up with Persistent Homology. For this purpose and for convenience of visualization, let us define - for an alpha complex at varying scale r and fixed time t_* - the k -dimensional Betti Curve $\beta_k(\mathcal{K}_{r,t_*})$. The k -dimensional Betti Curve at scale r and time t_* gives the value of k -th Betti number which measures the number of k -dimensional holes in the Alpha complex, \mathcal{K}_{r,t_*} . Similar to the construction of the Euler Characteristic Surface, one can define k -dimensional Betti Surfaces $\beta_k(\mathcal{K}_{r,t})$. Now, since Euler characteristic can be defined as the alternating sum of Betti numbers, one can yield the Euler Characteristic Surface from the Betti Surfaces with the

$$\text{relation: } \chi(\mathcal{K}_{r,t}) = \sum_{n=0}^{\dim(\mathcal{K})} (-1)^n \beta_n(\mathcal{K}_{r,t})$$

5.2.3 Comparison with Persistent Homology

Though Persistent Homology is a well-established topological tool in data science, it is computationally unwieldy and expensive in big data analysis, especially in the context of time series data that come from dynamical systems [143, 144], and the outputs of Persistent Homology are often not suitable for using directly into machine learning algorithms [145]. In this context, our apparatus of Euler Characteristic Surface(ECS) with the appropriate Euler Metric provides a computationally inexpensive tool. In this section, I try to set up the relation between our construction of Euler Characteristic Surface and Persistent Homology through Betti Surfaces. The relations further strengthen the mathematical stability of our tools.

Time-Slice wise stability of ECS against Persistent Homology

Firstly, I employed the methodologies detailed in sources [146] and [140] to analyze the stability of a time-slice of the Euler Characteristic Surface at a specific moment ($t = t^*$) with the information obtained from the corresponding time-sliced Betti Surfaces and hence from the corresponding persistence diagrams. This analysis is included to ensure the completeness of our exposition, particularly since our context involves the spatiotemporal Euler Characteristic Surface (which deals with time series data, unlike the context in [140]). Additionally this helped to determine which of the $L_p, (p = 1, 2)$ metric is the appropriate choice for ECS. The main idea of setting up the mathematical framework can be described by the following diagram:

$$\boxed{\left\{ \mathbf{PD}_n(\mathcal{K}) \right\}_{n=0}^{\infty} \longmapsto \left\{ \beta_n(\mathcal{K}_r) \right\}_{n=0}^{\infty} \longmapsto \chi(\mathcal{K}_r)}$$

The sequence of Betti curves/surfaces can be constructed from the corresponding persistence diagrams, and the Euler Characteristic curves/surfaces can be constructed from the corresponding Betti curves/surfaces.

Given a simplicial complex $\mathcal{K} = \{\mathcal{K}_r\}_{r>0}$ with the filtration function determined by the scale r , let $\beta_n(\mathcal{K}_r)$ be the n -th Betti number of \mathcal{K}_r . Let $\mathbf{PD}_n(\mathcal{K})$ be the n -dimensional persistence diagram of the filtered simplicial complex \mathcal{K} , and the complete persistent diagram is $\mathbf{PD}(\mathcal{K}) = \cup_{n=0}^{\dim \mathcal{K}} \mathbf{PD}_n(\mathcal{K})$. As mentioned earlier, the Euler Characteristic of the simplicial complex $\mathcal{K}_r = \mathcal{K}(r, t^*)$ with scale r at time-slice t^* can be expressed as alternating sums of Betti numbers $\chi(\mathcal{K}_r) = \sum_n (-1)^n \beta_n(\mathcal{K}_r)$.

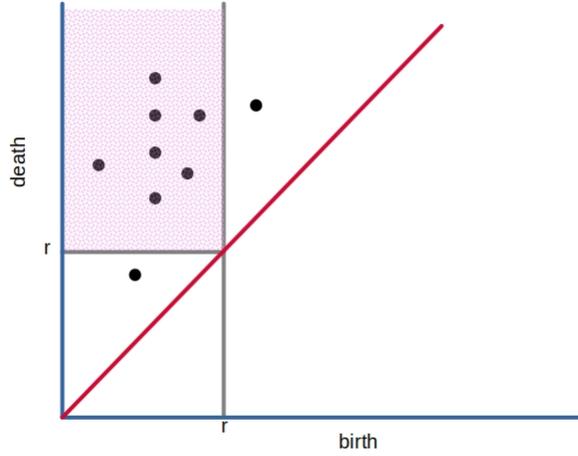


Figure 5.5: a schematic of persistence diagram. The points within the shaded box whose birth at $\leq r$ and death at $> r$ contribute to the simplicial complex at scale r and so to the Betti r curve.

Definition 4. The n -dimensional Persistent Betti r -curve of \mathcal{K} is defined as follows, where $\mathbb{1}_{[b,d]}$ is the indicator function of the interval $[b, d]$.⁶

$$\beta_n(\mathbf{PD}(\mathcal{K}), r) = \sum_{(b,d) \in \mathbf{PD}_n(\mathcal{K})} \mathbb{1}_{[b,d]}(r)$$

Lemma 1. $\beta_n(\mathcal{K}_r) = \beta_n(\mathbf{PD}(\mathcal{K}), r)$

Proof. From the Fundamental Lemma of Persistent Homology [141], $\beta_n(\mathcal{K}_r)$ equals the number of points (b, d) in the n -dimensional persistence diagram of \mathcal{K} such that $r \in [b, d]$ (see Figure (5.5)). This correlates the Betti curve with the Persistence diagram. \square

⁶The indicator function is defined as,
 $\mathbb{1}_{[b,d]}(x) = 0$ if $x \notin [b, d]$,
 $\mathbb{1}_{[b,d]}(x) = 1$ if $x \in [b, d]$

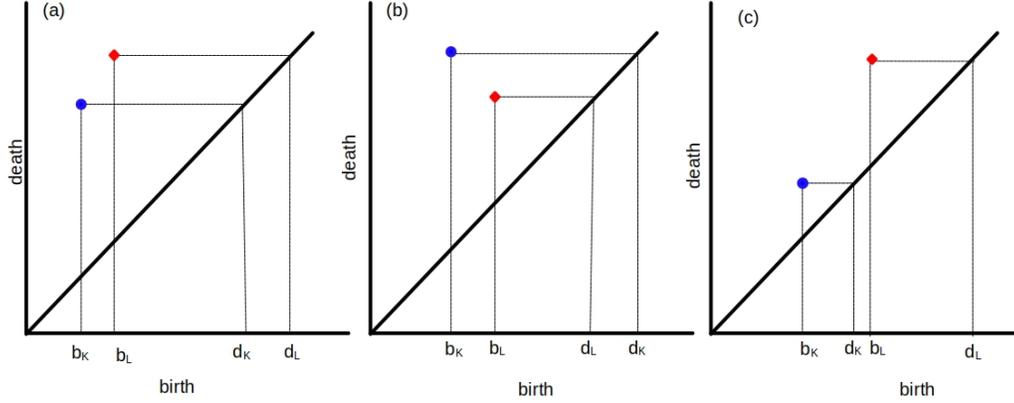


Figure 5.6: (a) and (b) are the possible orientations for (b_k, d_k) and (b_l, d_l) for a possible optimum matching between a pair of points from two different persistent diagrams. Matching (c) is not optimal.

Lemma 2. *For either of the configurations in Figure (5.6 (a) or (b)), the following inequalities hold:*

$$\|\mathbb{1}_{[b^{\mathcal{K}}, d^{\mathcal{K}}]}(r) - \mathbb{1}_{[b^{\mathcal{L}}, d^{\mathcal{L}}]}(r)\|_1 \leq 2d_\infty((b^{\mathcal{K}}, d^{\mathcal{K}}), (b^{\mathcal{L}}, d^{\mathcal{L}})) \quad (5.3)$$

$$\|\mathbb{1}_{[b^{\mathcal{K}}, d^{\mathcal{K}}]}(r) - \mathbb{1}_{[b^{\mathcal{L}}, d^{\mathcal{L}}]}(r)\|_2 \leq \sqrt{2d_\infty((b^{\mathcal{K}}, d^{\mathcal{K}}), (b^{\mathcal{L}}, d^{\mathcal{L}}))} \quad (5.4)$$

Proof. In each of cases 5.6 (a) or (b), the inequalities 5.3 and 5.4 can be verified. For example, for the configuration of Figure (5.6 (a)), $\|\mathbb{1}_{[b^{\mathcal{K}}, d^{\mathcal{K}}]}(r) - \mathbb{1}_{[b^{\mathcal{L}}, d^{\mathcal{L}}]}(r)\|_p^p = \int_{b^{\mathcal{K}}}^{b^{\mathcal{L}}} |1-0|^p + \int_{b^{\mathcal{L}}}^{d^{\mathcal{K}}} |1-1|^p + \int_{d^{\mathcal{K}}}^{d^{\mathcal{L}}} |0-1|^p = |b^{\mathcal{L}} - b^{\mathcal{K}}| + |d^{\mathcal{L}} - d^{\mathcal{K}}| \leq 2d_\infty((b^{\mathcal{K}}, d^{\mathcal{K}}), (b^{\mathcal{L}}, d^{\mathcal{L}}))$. A similar calculation works for the configuration of figure 5.6 (b).

□

Theorem 1. *Let K and L be simplicial complexes with non-zero persistence diagrams in a finite number of dimensions and with each persistence diagram containing finitely many non-diagonal points. Then the Euler Characteristic Surface sliced at $t = t^*$ with the Euler 1-metric and the Euler 2-metric satisfies the following inequalities (the first one with respect to the 1-Wasserstein metric) at the finitely many non-zero persistence dimensions. In particular, where $N = \max\{\dim \mathcal{K}, \dim \mathcal{L}\}$,*

$${}^7\|\chi(\mathcal{K}(r, t^*)) - \chi(\mathcal{L}(r, t^*))\|_1 \leq 2 \left[\sum_{n=0}^N W_1(PD_n(\mathcal{K}), PD_n(\mathcal{L})) \right] \quad (5.5)$$

$$\|\chi(\mathcal{K}(r, t^*)) - \chi(\mathcal{L}(r, t^*))\|_2 \leq \sum_{n=0}^N \sum_i \sqrt{2d_\infty((b_i^K, d_i^K), (b_i^L, d_i^L))} \quad (5.6)$$

Proof. Following [140], an optimal matching between the two persistence diagrams with points $\{(b_i^K, d_i^K)\}, \{(b_i^L, d_i^L)\}$ can be considered, with the 1-Wasserstein or 2-Wasserstein metric. For an optimal matching between two persistence diagrams with either the 1-Wasserstein or 2-Wasserstein metrics, configurations of pairwise matching between points in Figure (5.6), (a) or (b) are possible. However, configuration 5.6 (c) is not possible under either of 1-Wasserstein or 2-Wasserstein metrics, as matching the points to the diagonal provides a better matching (as $|b_L - b_K|^p + |d_L - d_K|^p > |b_K - d_K|^p + |d_L - b_L|^p$ for $p = 1, 2$). Now, inequality (5.3) along with definition (4) and lemma (1) can be used to prove-

⁷See [140] in the context of Euler Characteristic Curves

$$\begin{aligned}
\|\chi(\mathcal{K}_r) - \chi(\mathcal{L}_r)\|_1 &= \left\| \sum_n (-1)^n (\beta_n(\mathbf{PD}(\mathcal{K}_r)) - \beta_n(\mathbf{PD}(\mathcal{L}_r))) \right\|_1 \\
&\leq \sum_n \|(\beta_n(\mathbf{PD}(\mathcal{K}_r)) - \beta_n(\mathbf{PD}(\mathcal{L}_r)))\|_1 \\
&\leq \sum_{n=0}^N \sum_i 2d_\infty((b_i^K, d_i^K), (b_i^L, d_i^L)) \\
&\leq 2 \sum_{n=0}^N W_1(PD_n(\mathcal{K}), PD_n(\mathcal{L}))
\end{aligned}$$

Similarly, using inequality (5.4)-

$$\begin{aligned}
\|\chi(\mathcal{K}_r) - \chi(\mathcal{L}_r)\|_2 &= \left\| \sum_n (-1)^n (\beta_n(\mathbf{PD}(\mathcal{K}_r)) - \beta_n(\mathbf{PD}(\mathcal{L}_r))) \right\|_2 \\
&\leq \sum_n \|(\beta_n(\mathbf{PD}(\mathcal{K}_r)) - \beta_n(\mathbf{PD}(\mathcal{L}_r)))\|_2 \\
&\leq \sum_{n=0}^N \sum_i \sqrt{2d_\infty((b_i^K, d_i^K), (b_i^L, d_i^L))}
\end{aligned}$$

□

Remark 1. *It is to be noted from Theorem 1 that the spatiotemporal Euler Characteristic Surface with the Euler 1-metric has better stability properties with respect to the corresponding persistence diagrams with the 1-Wasserstein metric, as inequality 5.5 is independent of the number of points of the persistence diagram. In comparison, for the Euler 2-metric, we have an inequality 5.6 that depends on the number of points of the persistence diagrams. This pattern follows the existing stability results for other topological summaries [147] [146], where the distance between the summaries is bound from above by the 1-Wasserstein distance between the corresponding persistence diagrams.*

Remark 2. *The spatiotemporal Euler Characteristic Surface has been defined using continuous filtration in terms of the scale parameter. In practice, the ECS is calculated at points of a discrete subset of scales S , so a proper choice of S is important to ensure that the discrete subset of scales preserves all the information about changes in the original continuous filtration. In particular, one needs to consider all the critical scales $r_1 < r_2 < \dots < r_m$, where the critical scales are those where at least one new simplex is created (as the underlying Delaunay complex is finite, there are only finitely many critical scales). This needs to be kept in mind while discretizing the left side of the inequalities in Theorem 1, to ensure that inequalities 5.5 and 5.6 are satisfied.*

Stability of ECS against perturbations of the data set

After establishing the stability of L_1 and L_2 Euler Metric with respect to the metrics in Persistent Homology, the next objective was to examine the temporal stability of the Euler Characteristic Surface construction. It was found that for finite-dimensional discrete dynamical systems where continuity properties can be assumed, one may have temporal stability of the Euler Characteristic by using a Wasserstein stability theorem as described below.

For this analysis one may view our setting to be finite simplicial complexes with sub-level filtrations based on (simplex-wise) monotone filtering functions⁸. For example, the Rips complex on a finite data set gives rise to such a filtration, where the filtration function is the diameter of the simplex. For our case, the Čech (Alpha) complex on a finite data set gives rise to such a filtration, where the filtration function is the radius of the smallest enclosing ball of a simplex (of the Delaunay triangulation).

To examine the temporal stability, a recent powerful and useful result, “Cellular Wasserstein Stability Theorem” was used which is stated below.

Theorem 2 (Skraba -Turner, [148]). *For monotone functions $f, g : \mathcal{K} \rightarrow \mathbb{R}$ on a finite simplicial complex \mathcal{K} , we have:*

$$W_p(\mathbf{PD}_n(f), \mathbf{PD}_n(g)) \leq \left[\sum_{\dim \sigma \in \{n, n+1\}} |f(\sigma) - g(\sigma)|^p \right]^{1/p} \quad (5.7)$$

⁸A filtering function $f : \mathcal{K} \rightarrow \mathbb{R}$ is said to be monotone when $\tau \subset \sigma$ implies that $f(\tau) \leq f(\sigma)$

The above, along with results from Section 5.2.3 gives the temporal stability result for Euler Characteristic Surfaces which is presented as a theorem in 3. Below we will say a discrete dynamical system X is (temporally) uniformly continuous if for each $\epsilon > 0$ there is $\delta > 0$ such that for each point $x_t \in X$ we have $d(x_{t+\delta}, x_t) < \epsilon$.

Theorem 3. *Let $X_t = \{x_1(t), \dots, x_M(t)\}$ be a (temporally) uniformly continuous-discrete dynamical system with finitely many points in \mathbb{R}^2 . Then for a sufficiently small $0 < \epsilon < \frac{1}{2} \min_{i \neq j} \|x_i(t_0) - x_j(t_0)\|_2$, there is $\delta > 0$ such that if $\mathcal{K}(r, t)$ be the Čech filtration⁹ on X_t , we have $\|\chi(\mathcal{K}(r, t_0)) - \chi(\mathcal{K}(r, t_0 + \delta))\|_1 \leq \frac{1}{3}M(M+1)(M+2)\epsilon$.*

Proof. One can note that for any acute-angled triangle T in \mathbb{R}^2 and for any $\epsilon > 0$, there is an $\delta_T > 0$ such that after perturbing the vertices by δ_T , the circumradius of T is changed by at most ϵ (this follows from the continuity of the circumradius function of a non-degenerate triangle).

Let $\epsilon > 0$ be less than $\frac{1}{2} \min_{1 \leq i \neq j \leq M} \|x_i(t_0) - x_j(t_0)\|_2$. Choose $\delta > 0$ such that for $x_i(t) \in X_t$ we have $d(x_i(t_0 + \delta), x_i(t_0)) < \min \delta_T$, where the minimum is over all acute-angled triangles formed by points of X_{t_0} . Take the filtering function on a simplex in theorem 2 to be the radius of the smallest enclosing circle of the points defining the simplex, with f corresponding to points at time t_0 and g corresponding to points at time $t_0 + \delta$.

Then, by theorem 2, we have for $n = 0, 1$ the following inequalities (for the inequalities below, note that the smallest enclosing circle of a finite set of points in \mathbb{R}^2 is realized by either two of those points (diametrically opposite) on the circle, or by three of those points on the circle):

$$W_1(\mathbf{PD}_0(f), \mathbf{PD}_0(g)) \leq \binom{M}{1}\epsilon + \binom{M}{2}\epsilon = \binom{M+1}{2}\epsilon = \frac{1}{2}M(M+1)\epsilon$$

$$W_1(\mathbf{PD}_1(f), \mathbf{PD}_1(g)) \leq \binom{M}{2}\epsilon + \binom{M}{3}\epsilon = \binom{M+1}{3}\epsilon = \frac{1}{6}M(M^2-1)\epsilon$$

⁹Note that the Čech and Alpha filtrations give the same persistent homology, by the nerve theorem [141].

Finally, from theorem 1 we have:

$$\|\chi(\mathcal{K}(r, t_0)) - \chi(\mathcal{K}(r, t_0 + \delta))\|_1 \leq 2 \sum_{n=0}^1 W_1(PD_n(\mathcal{K}), PD_n(\mathcal{L})) = \frac{1}{3}M(M+1)(M+2)\epsilon$$

In the inequality above the sum was done over $n = 0, 1$ as the Čech filtration has the same persistent homology as the Alpha filtration, therefore dimension $n=0,1$ is sufficient for our datasets.

□

5.3 Application on simulated data

To examine the theoretical background explained in the previous section, I then used two different kinds of time-varying point sets generated from two kinds of simulated systems. The first objective with simulated data was to explore how the construction of ECS gets affected by the choices of methods to construct the cell complexes and with perturbations in Sub-section 5.3.1, where the number of points in the vertex set grows with time. Following that I shall verify the relations established earlier in 5.2.3 and 5.2.3 with simulated data sets in 5.3.2 where the number of points in the vertex set is constant in every time step.

5.3.1 Eggbeater Flow: Dependence on parameters, Stability of ECS

Recalling the previous study in the Chapter(4), the analysis was done for simulated flow patterns of a fluid mixing model by constructing Euler Characteristic Surfaces using “coarse-graining” or the “union of r -neighbourhood” approach. Here, the analysis gets further extended with the study on the Euler Characteristic Surface using simplicial complex and exploring its mathematical characteristics on the same fluid mixing model of modified egg beater flow [125] that was used in the Chapter(4). The simulation model generates a point cloud at each time step following the equation Eq.(5.8) that describes the path of the flow at that time instant within a Poincare

section of length 1×1 . So, the point cloud that we are studying is the same as in Chapter(4).

$$\begin{aligned}x_{t+1} &= x_t - ky_t(1 - y_t) \\y_{t+1} &= y_t - kx_t(1 - x_t)\end{aligned}\tag{5.8}$$

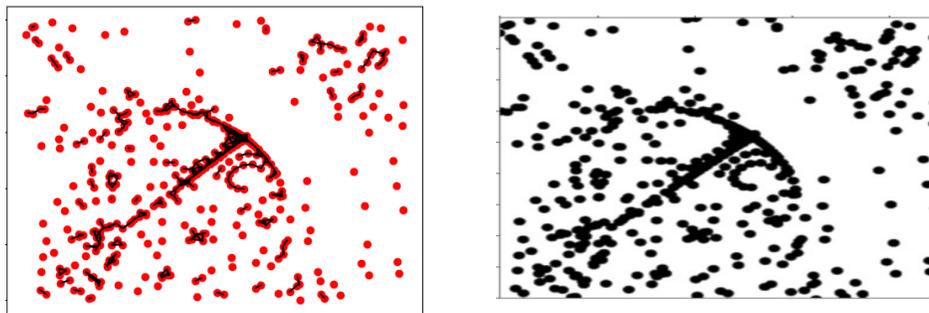


Figure 5.7: (a) Alpha Complex of flow pattern of modified egg-beater flow at time-step $t=500$ with $k=4.1$, at scale $r=0.01447$ estimated $\chi=161$, (b) Same point set at same scale ($r = 10$) with “coarse-graining” or “union of r -neighbourhood” approach, estimated $\chi = 148$.

The seed point was chosen similar to our previous simulation which is $(0.9, 0.2)$ and the process was allowed until 10,000 time steps for a constant value of k . The same families of dynamical systems for different values of k ranging between $k = 4.0$ to $k = 5.0$ were studied but this time the point cloud at each time step of the simulated flow system for a constant k was used to make the Alpha complex. The scales of filtration used to construct the Alpha complexes were almost comparable to the scales used for discrete grid systems where we had built the cell complex via the “union of r -neighborhood” or “coarse-graining” approach. The values of the scales in Alpha Complex were converted from the integer values of the scales in the “ r -neighborhood” into Euclidean distances by considering the approximate pixel/grid size. The Euler Characteristic(χ) was estimated in Alpha complexes using the alternative sum of numbers of zero-dimensional (point), one-dimensional (line),

and two-dimensional (triangle) simplices. In the case studied, the complexes have two homology groups, H_0 and H_1 (As the point cloud is embedded in 2-dimensional Euclidean space). If the “coarse-graining” or “union of r-neighborhood” method the value of the Euler Characteristic was estimated by the difference between the number of black clusters (N_b) and the number of white clusters (N_w) in the binary pixelated cell complexes, $\chi = N_b - N_w$. Figure (5.7) shows the estimated values of Euler Characteristic (χ) at different scales using both approaches. The values show a small deviation at higher scales (poorer resolution) as the conversion from pixels (with r-neighborhood included) to Euclidean radius value in the Alpha complex can not be very precise at higher scales (due to the discrete nature of the pixels). For Alpha complexes on a point cloud embedded in a continuous Euclidean space, the estimation of Euler characteristic becomes exact at all scales since there is no discretization required.

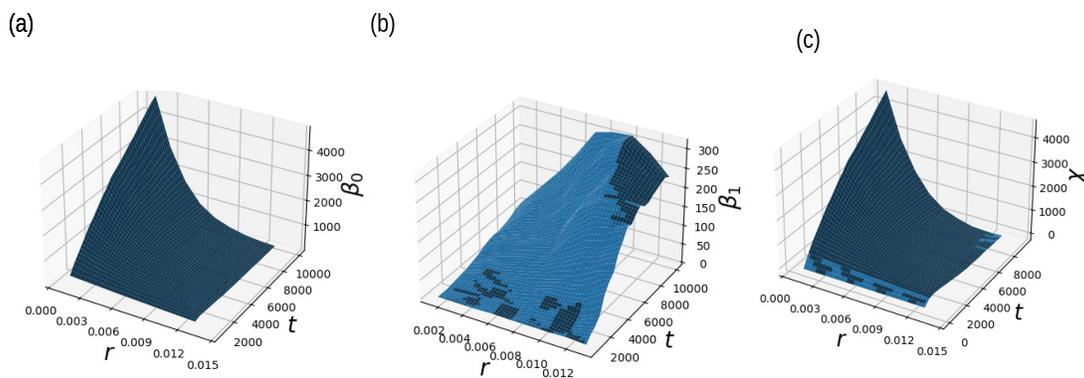


Figure 5.8: For Flow pattern with $k = 4.1$ - (a) 0-dimensional Betti surface, (b) 1-dimensional Betti surface, (c) The Euler Characteristic surface.

The Euler Characteristic Surfaces carry a summary of topological features present in all dimensions from the simplicial complexes generated from a dynamical system. It must be noted that if one is interested in studying a single-dimensional topological feature in detail, for example the the 0-dimensional simplices represent-

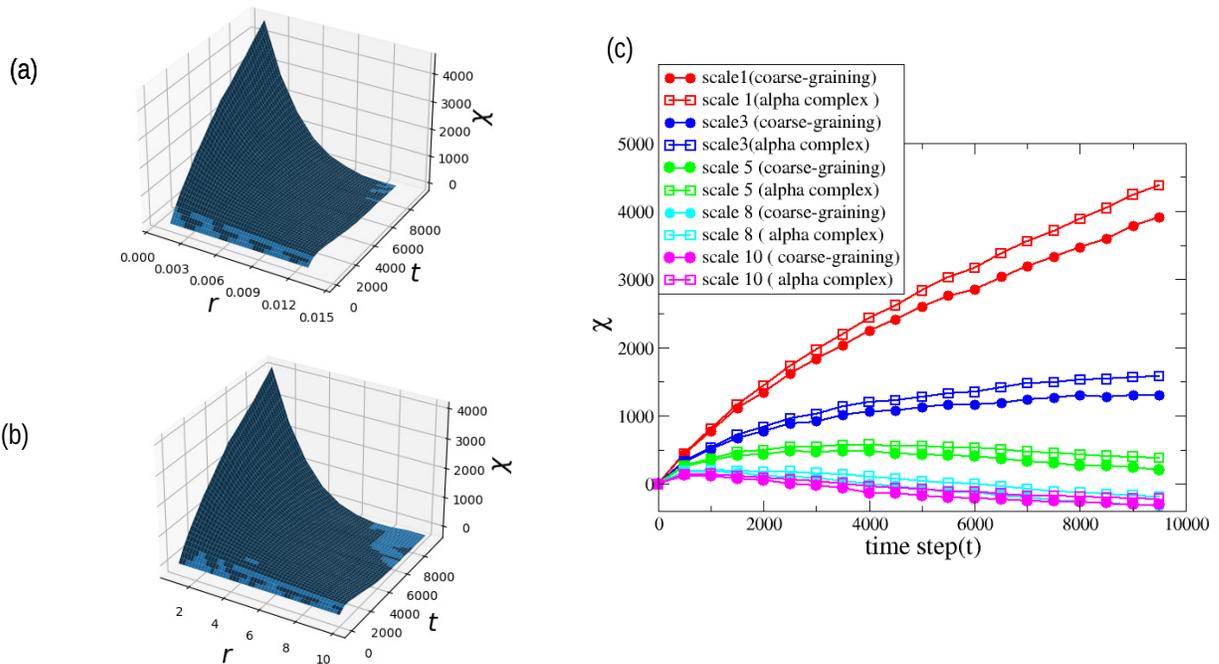


Figure 5.9: (a) Euler characteristic surface constructed through Alpha complexes, (b) Euler characteristic surface constructed through 'coarse-graining' or union of r -neighborhood approach on the pixelated grid, (c) The values of estimated Euler Characteristic (χ) at different scales using the two methods.

ing the connected components/ clusters, or the 1-dimensional simplices representing the loops, it can be done by constructing individual Betti surfaces similar to ECS construction. As seen in Figure (5.8), the spatiotemporal evolution of n -dimensional Betti numbers can be summarised into surfaces that we name the n -Betti surface $\beta_n(r, t)$. For the simulated flow pattern discussed above, the Betti surfaces corresponding to critical point $k = 4.1$ is displayed in Figure (5.8), where we can see the β_0 surface is qualitatively quite similar to the Euler Characteristic Surface but the β_1 surface representing the loops is clearly different and smaller in magnitudes. The metric distance between two Betti surfaces of two different dynamical systems can also be estimated in a similar fashion that was done for the Euler metric, i.e. estimating the L_2 norm and the L_1 norm distance between two Betti surfaces of similar dimension.

Despite the small discrepancy between the estimated χ values with the two

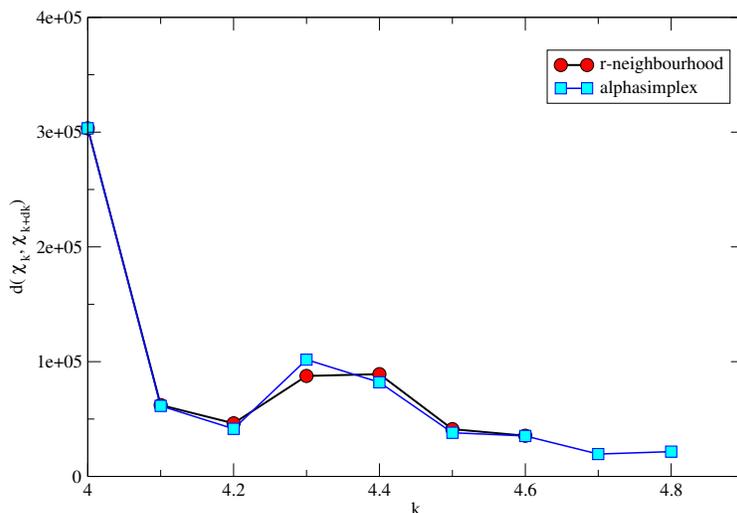


Figure 5.10: Euler metric estimated between surfaces with k and $k+dk$ where $dk=0.1$, using both approaches.

different kinds of cell complexes, in the over all the Euler Characteristic Surfaces constructed from the two methods appeared to be similar, Figure (5.9). Further, using Euler L_p ($p = 1, 2$) Metrics one can quantify their similarity or dissimilarity. Euler Characteristic Surfaces, for the dynamical systems of modified egg beater flow models using equations 5.8, were constructed this time with different k values, using the traditional Alpha Complex approach. Followed by the Euler Metric between surfaces of dynamical systems corresponding to parameter k and $k + dk$, were estimated in a similar manner that was done using “coarse-graining” or “union of r-neighborhood” in Chapter(4), Section(4.2.1). Values of the Euler Metric from both methods are shown in the plot in Figure (5.10). The estimation of the Euler Metric in both methods turns out to match each other regardless of the discrepancy in the estimation of Euler Characteristic. *Thus it can be concluded that the proposed Euler Metric is independent of the methods or approaches used to compute Euler Characteristic(χ).*

Next, the question that came to our mind is how high the scale of filtration should be taken to construct the ECSs, and whether these surfaces are stable with respect to small perturbations. For that, the stability of the Euler Characteristic Surfaces with respect to a small perturbation in the scale of resolution was tested. For a stable tool, a small perturbation in the scale should result in a small change in outcome or measure too. Therefore, the Euler Metric estimated between two Euler Characteristic Surfaces, one expanded up to scale R with the other ECS expanded up

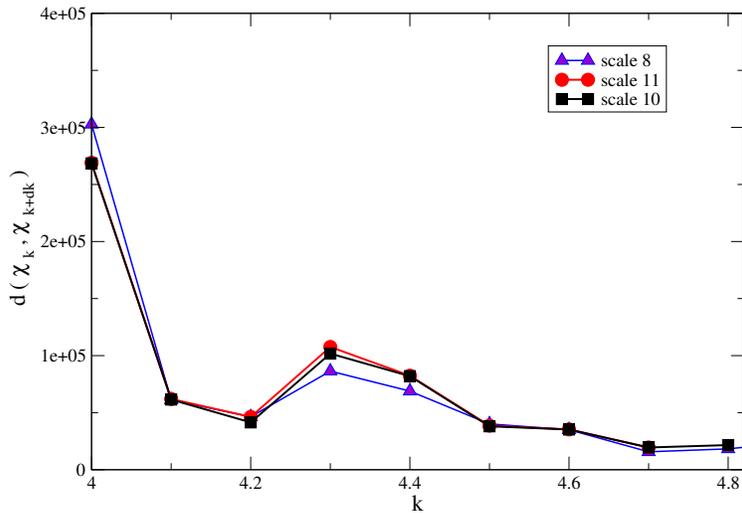


Figure 5.11: Euler metric estimated between surfaces keeping R upto scale 10(black)and upto scale 11 (red).

to $R + dR$, were compared keeping the other parameters constant. Given the change in scale, dR is small, the change in the measure of the Euler Metric turns out to be small too if the Euler Characteristic Surfaces are constructed with finer variations in scale and up to scales high enough to cover almost all possible topological features (in particular, critical scales should not be missed. See remark 2). The results of perturbing the scale on Euler Characteristic Surfaces and thus on Euler Metric are displayed in Figure (5.11). As one can see that the values of the Euler Metric between surfaces constructed up to scale 10 and scale 11, do differ very small compared with scale 8, where the change is louder.

5.3.2 Vicsek Model: A case study

The simulation model for egg-beater flow was alright to study the properties of ECS where the number of points was not needed to be constant at each time step but to compare the ECS with modules in Persistent homology and verifying the mathematical relations set up in Section 5.2.3, a dynamical system with constant number of point at every time step of evolution was the best choice. At that point, we decided to model the dynamical system using the well-known *Vicsek model* [149]. This model, renowned for illustrating collective behavior in living systems, offers a simple yet fascinating approach to observing a wide range of topological patterns.

In Persistent Homology, different metrics and approaches have been introduced to study the similarity and dissimilarity of shapes between two point clouds, the common ones being the “Bottleneck distance” and the “Wasserstein distance”. Substantial works applying these tools on static point sets [150, 151, 152] have been produced. Recently, they have been studied for time-varying metric spaces also, with approaches like those proposed in [136]. However, the frameworks might not be very much successful in interpreting time-series data. In the analysis below the stability of our proposed metric EM against the 1-Wasserstein distance, proposed in theorem(1) will be tested with the simulated data generated using the Vicsek model. For this, we first constructed the alpha complexes of the time-varying point clouds. The number of points (living matter) in each time step was kept constant by employing the periodicity within a box of length L . The classical Vicsek equation describing collective behavior is described in Eq.5.9 where each point/living matter travels with a constant magnitude of the velocity(v) and its orientation of velocity(θ) is influenced by its neighboring particles along with a noise parameter η .

$$x_i(t + 1) = x_i(t) + v_i(t + \nabla t) \nabla t \quad (5.9a)$$

$$\theta_i(t + \nabla t) = \frac{1}{N} \sum_{|x_i - x_j| \leq R} \theta_j(t) + U\left(\frac{-\eta}{2}, \frac{\eta}{2}\right) \quad (5.9b)$$

Different configurations of parameters η , R , v , L and N can be used to model different sets of dynamics. In this study the distribution of the points in this model was varied by two scenarios - (i) by varying noise parameter η for a constant sphere of interaction with radius R and (ii) by varying the radius of the sphere of interaction R for different noises ranging from $\eta = 0$ to 5. The box length L was kept at 1.0 unit, with the number of particles $N = 500$ moving with a velocity of magnitude $|v| = 0.03$. For lower values of noise or zero noise, the randomly distributed point set gradually transformed into aggregated clusters along with the order in motion taking place in the system. When the noise became high, it dominated the neighborhood interaction, and chaotic motion was prominent in the point sets with no order observed. This is discussed in detail in the following paragraphs.

Analysing point clouds with and without noise(η) for a constant R

As the noise parameter (η) is an important one to decide the dynamics of the simulated point clouds, two different situations of the dynamical point sets were considered for the analysis-(i)the time-varying point clouds with noise, $\eta = 0$ and (ii) the

time-varying point clouds with noise, $\eta = 0.9$. In both cases, the other parameters were kept similar, $R = 0.05, N = 500$, and $L = 1$. Alpha complexes with scale r , varying from 0.005 to 0.065 unit, were constructed for each in the time series of point clouds. The scale of filtration to build these complexes is non-trivial as it highly influences the connectivity of clusters and loops. In Sub-section(5.2.1), Figure (5.3) shows the point cloud of the simalon with noise, $\eta = 0$ at time step $t = 11$, where the Alpha complex of the same point cloud at different scale r carried different topological features. Hence, while constructing the Alpha complexes, finer variation of scale was considered which makes the Euler Characteristic Surfaces stable, capturing sufficient topological information.

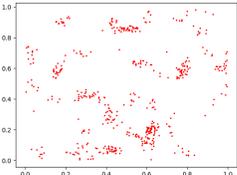
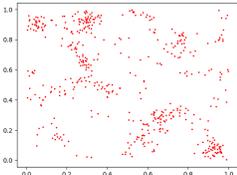
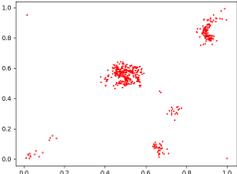
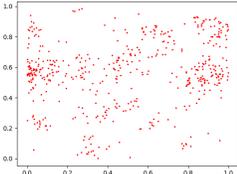
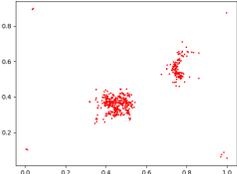
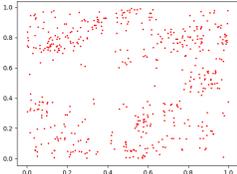
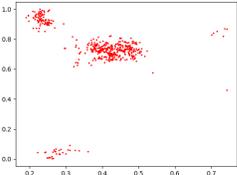
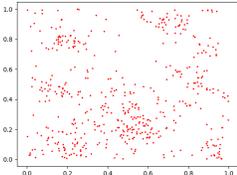
Verification of proofs A numerical verification of the stability results corresponding to L_1 Euler Metric, Eq. 5.5, was performed with these real simulated data-sets via the construction of the Euler Characteristic Surfaces and their time slices. Both the Euler L_1 metric and 1-Wasserstein distance were estimated at suitably spaced intervals of time-slices where the point sets had different geometry.

Table.5.1 shows supporting evidence for the theory discussed in Section 5.2.3. Here, for a particular time step t^* and between two dynamical systems K and L , let $A = \|\chi(\mathcal{K}(r, t^*)) - \chi(\mathcal{L}(r, t^*))\|_1$ and $B = 2 \left[\sum_{0 \leq n \leq 1} W_1(PD_n(\mathcal{K}), PD_n(\mathcal{L})) \right]$, where W_1 is 1-Wasserstein distance between n the persistence diagrams of \mathcal{K} and \mathcal{L} i.e. the dynamic point cloud with $\eta = 0$ and the same with $\eta = 0.9$. According to Eq.(5.5), $A \leq B$ which agrees with the values in Table.5.1, at all time-steps considered.

The temporal stability of our Euler Characteristic Surface proposed in theorem (3) for the uniformly continuous-discrete dynamical system was also verified with the point clouds at close time intervals. As the number of points is constant in every time step of the simulation and the point cloud moves with a finite and small velocity ($|v| = 0.03$), the change in Euler Characteristic Curves was hypothesized to be finite and small. Figure (5.12) shows the Euler Characteristic Curves, which are the time-slices of the Euler Characteristic Surface at different instants of time. One can see that for closely separated time intervals, the distance between the Euler Characteristic Curves is small (especially in the lower values of filtration of scales).

Further to quantify the temporal stability, the L_1 Euler metric between two Euler Characteristic Curves with different time-interval (dt) apart was calculated

Table 5.1: Measures between point clouds for $\eta = 0.0$ and $\eta = 0.90$

time(t^*)	point cloud at $\eta = 0.0$	point cloud at $\eta = 0.9$	A	B
11			1.3869	2.3825
151			4.2975	6.3003
351			5.3040	7.5766
451			5.1494	7.2694

for the dynamic point clouds in both of the cases i.e. $\eta = 0.9$ and $\eta = 0$. The smaller the time interval between two point clouds, the closer the Euler Characteristic curves are in Figure (5.12). This hinted that the metric distance between two Euler Characteristic Curves, corresponding to two point clouds that are small time-gap apart, should also be small. Figure (5.13) shows the values of the L_1 Euler metrics between two Euler Characteristic Curves, dt time-interval apart, plotted against time-interval dt . As one can see in the snapshots given in Table (5.1) when the simulation had zero noise, $\eta = 0$, the system quickly ordered forming dense and few aggregations of clusters/flocks. For comparatively larger noise, $\eta = 0.90$, the system did not follow such ordered alignment, making the point cloud more scattered and

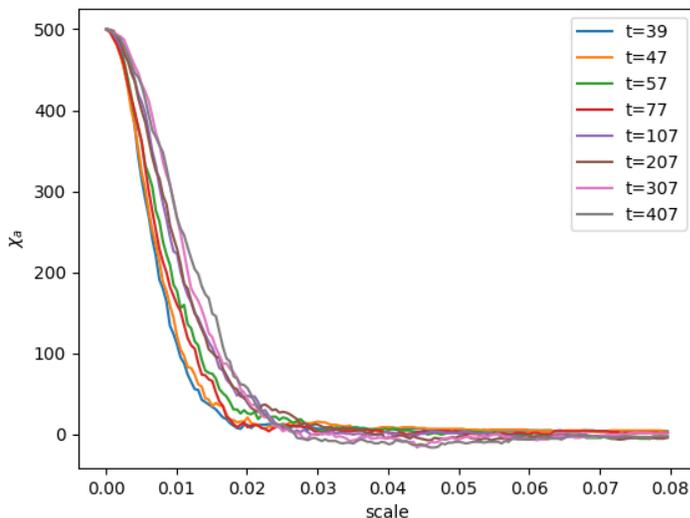


Figure 5.12: Slices of Euler Characteristic Surfaces at different time steps for the ECS corresponding to $\eta = 0.9$.

sparse. This distinction of behavior is reflected in the Euler Metric, Figure (5.13). For zero noise the the slope of the Euler L_1 metric vs time-interval curve is sharp and quickly reaches the equilibrium state (after sufficient dt later when the system settled into a state with ordered flocks, making no further change in the topology and hence in EM values). For noise $\eta = 0.9$ the metric distance fluctuates initially and then takes a longer time to reach a chaotic stable state. The chaos sustains without much fluctuation in the distribution of the point cloud. Also, a slightly higher value of the Euler Metric is observed for this case. The plot further shows that when dt is small, the L_1 Euler Metric value is small (≈ 0.25) which justifies the theorem about temporal stability of Euler Charactric Surface, theorem 3 and matches the quantitative bound, taking $m = 500$ and ϵ to be small.

Analysing Point Clouds with different noises(η), for different R values

The behavior of the Euler Metric in Figure (5.14) with different values of noise investigated to continue more analysis to understand whether there exists any reflection of the order-disorder transition in the Vicsek Model on our topological marker Euler Metric. By ‘order’ it is meant the traditional order parameter v_a which is the average

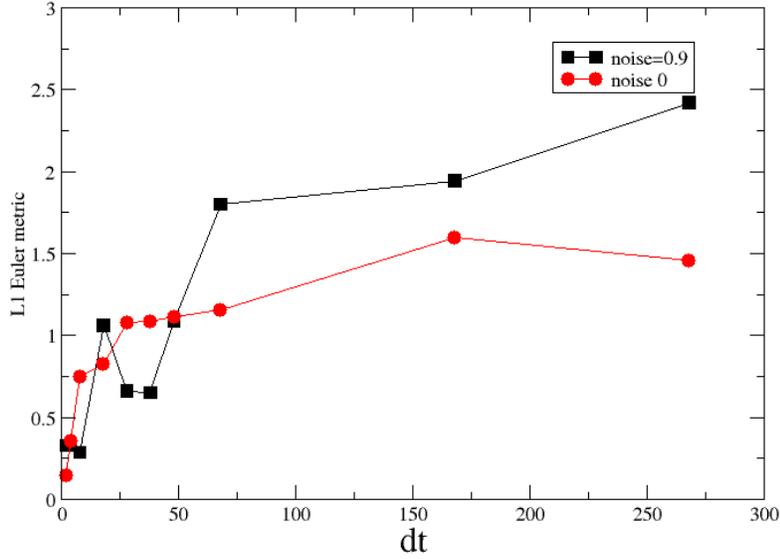


Figure 5.13: The L_1 Euler metric between ECS time-slices with dt distance apart from time slice $t = 39$.

normalized velocity of the particles/points. Therefore $v_a = \frac{1}{N |v|} \left| \sum_{i=1}^N v_i \right|$. When the system is perfectly ordered or all particles are aligned in the same direction, the order parameter $v_a = 1$, and as the disorder increases, it moves more towards zero. Figure (5.14a) shows that with the increase in noise η how the dynamical system moves from order to disorder. We tried different variations in interaction radius R , measuring how far one particle/ point will be influenced by the orientation/movement of the neighbouring particles. For very large R the system quickly reached its equilibrium state (that may be ordered or disordered depending upon the value of noise) as the interactions were very large. Now, to analyze from the topological aspect, as usual, the Alpha complex of the dynamical point clouds corresponding to similar values of η and R was generated followed by the construction of Euler Characteristic Surfaces. Next, the ECSs corresponding to a specific η were compared with the ECS corresponding to $\eta = 0$, keeping the interaction radius R constant. The L_2 Euler metric (Eq.5.1) was estimated for the comparison and the results are displayed

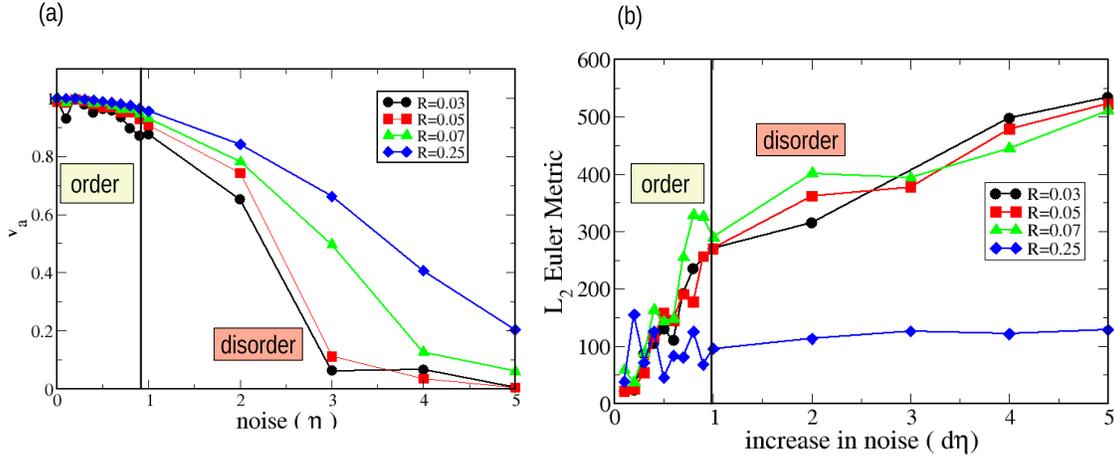


Figure 5.14: (a) L_2 Euler metric estimated between ECSs of noise η and zero noise. (b) The order parameter vs noise (η) is plotted for different interaction radius R .

in Figure (5.14b). The plot shows that the slope of the Euler Metric remarkably changes around noise $\eta \approx 0.9$ where the order-disorder transition happens in correspondence to Figure (5.14a). This result suggests that once the system enters into disorder with the increase in noise, the metric distance varies less, making a plateau-like region in the plot of the Euler metric. Thus, the variation Euler metric with noise carries a signal to the physical phase transition from order to disorder. This indeed needs to be studied more in detail in the future to explore whether quantitative changes in Euler Characteristic Surfaces measured by the Euler Metric have the potential to act as an alternative marker to characterize phase transition in collective motion phenomena.

5.4 Conclusion

This final work in my dissertation thus successfully established the robustness of the proposed construction of the spatiotemporal topological map of a dynamical system - the Euler Characteristic Surface (ECS) along with the measure that can differentiate between two ECSs - the Euler Metric. The correlation between the

tools of ECS construction with the standard measure of Persistent Homology (PH), commonly used by the TDA community, was set up and it helped to strengthen the mathematical framework of the constructs.

Classically, PH examines the dynamical system at different length scales at any fixed time, whereas most of the dynamical systems often exhibit intriguing topological changes across both scale and time. Recently, the time evolution of such systems has been analyzed using methods like Crocker Plots and Persistent Vineyards, which integrate Persistent Homology into dynamic contexts. Despite their utility, these methods are computationally intensive, and their stability requires further examination. Our proposed topological tool, the ECS, aims to encapsulate information about both the scale and temporal dynamics of a system within a single map. By demonstrating the stability and robustness of ECS against perturbations, it is established that ECS and EM can work as a powerful topological tool for analyzing dynamical systems. The ECS construction offers a streamlined approach to capturing complex topological changes, potentially transforming our understanding and analysis of dynamical behavior.

The highlights of this chapter may be summarized thus:

- Construction of the ECS using appropriate simplicial complexes based on the data points (filtered by a scale parameter) gives more precise results and is preferred over using discretized grids to map data cloud and calculate Euler Characteristic. However, in the case of extracting data from digital images, one can not bypass the usage of discrete grids/pixels. Since it was found that ECS constructed from both “coarse-graining” in grids and “simplicial complex” in Euclidean space were closely similar, and the Euler Metric remained unaffected by these choices, researchers can select the cell complex or ECS construction method that best suits their model.
- The Euler Characteristic Surfaces are robust to small perturbations in scale and time provided they are built up with fine resolutions in scale and time and up to sufficiently large scales that cover almost all topological features (all critical scales) and changes present in the dynamical system. This is indeed an important characteristic of the tool that establishes its stability.
- Recreating the Euler Characteristic Surfaces via Betti Curves from Persistence Diagrams and establishing the correlation between a time-slice of the ECS and

Persistence Diagrams make the ECS an alternative and more powerful tool of analysis in Topological Data Analysis. Additionally, it also contains the temporal development of the system, unlike Persistence Diagrams.

- The relations between the metric distances, the p-Wasserstein distance of Persistence Homology (PH), and the L_1 and L_2 Euler Metric of ECS lends to the stability of the Euler Characteristic Surface(ECS). The Euler L_2 Metric has a Hilbert space structure making it suitable for direct applications in machine learning algorithms, while better stability results were obtained using the Euler L_1 Metric. However, for systems with a finite number of points, this is not a major disadvantage as L_1 is embeddable in Hilbert space with bounded distortion.
- Similar to the previous study in Chapter(4), The ECS with the Euler Metric(EM) acted successfully to deliver an efficient summary of dynamical systems and thus to quantify the similarities and dissimilarities between two different dynamical systems/time-varying data sets.
- The ECS construction with the L_1 and L_2 metric further gave cues to order-disorder transitions in the collective motion of particles as observed in the analysis using the Vicsek Model dynamics.

Chapter 6

Conclusion and future directions

In this dissertation, I tried exploring dynamical systems by studying their geometry and topology, intrigued by the quest of how the patterns in a complex dynamical system encapsulate the principles of physics driving these systems. Can studying the topological patterns in turn help one find a pattern in the systems? How can one quantify these patterns and whether quantifying them will be fruitful at all?

I started the study with simple static real crack mosaics that were created by the dynamic process of desiccation and tried analyzing the combinatorial topology along with the geometric properties of the mosaics(Chapter2). Through the study, it was learned that classifying planar mosaics with respect to their individual material characteristics is not possible by only considering the combinatorial topology measures (n, v) and thus included the geometrical measures, “Angular defect” D , and “Iso-perimetric ratio” or “shape parameter” λ . A 4-parameter tuple (n, v, D, λ) was proposed to effectively describe a crack mosaic. This made similar materials form clusters in 3D spaces of (n, v, D) and (n, v, λ) , aiding in material identification from crack patterns. Simulated crack patterns from models like Voronoi, Gilbert, and Iterative cell division were compared to real crack patterns, providing insights into crack formation mechanisms. Additionally, in-house codes were developed to attempt measuring a few non-convex polygon features present in real crack patterns. The study further inspired research into the time-development trajectory of crack mosaics in topology-geometry space of (n, v, D, λ) in simulated columnar joints where the metamorphosis from the less symmetric Gilber tessellation to more symmetric Voronoi

tessellation was summarized in the (n, v, D) and (n, v, λ) trajectories.

From static polygonal patterns, I further moved into understanding complex flow patterns in droplet evaporation. In studying the flow features, the need to consider the multiscale (length-wise/resolution-wise) evolution of the patterns along with the temporal evolution was realized. A novel topological approach was emanated – “Euler Characteristic Surface”(ECS) along with “Euler Metric”(EM). The ECS carries the information of the spatio-temporal evolution of the topology in a dynamical system and the EM further quantifies the distance between two ECSs. The dynamic flow patterns in a drying droplet of polystyrene beads in water were analyzed through image processing followed by studying the evolution of the Euler Characteristic (χ) over time and scale. The study quantified the changes in clustering and connectivity in the aggregation features of the particles (that were the consequences of the flow within the droplet) and marked different phases/ modes of evaporation with different dominant forces (Chapter3). It was hypothesized that the Euler Characteristic Surface $\chi(r, t)$, which summarizes topological features, may serve as the system’s topological signature. The associated contour plot of the ECS, “Euler Characteristic Level Curves”(ECLC) can provide a low-dimensional summary of topological evolution and identify persistent features and sudden changes in the process.

The concept, “Euler Characteristic Surface” and “Euler Metric”, was further employed in studying more fluid dynamical systems, with both simulated and real data sets (Chapter4). The ECS and EM were found to be effectively quantifying the similarity and dissimilarity between dynamical systems. Additionally, the steady and the critical domains in the parameter variation for the simulated egg beater flow were identified by simply constructing ECSs and measuring the Euler Metric(EM) between the dynamical flows. Real droplet systems with different particle-fluid compositions and the same ambiance were compared with respect to ECSs and EM and the outcomes of the analysis were justified, especially the ECSs corresponding to two different droplet systems and similar droplet systems effectively portraying the difference or similarity respectively. The study established that the Euler Characteristic Surface and Euler Metric are strong descriptors for characterizing similarity and dissimilarity in dynamical systems.

All done, The mathematical picture of our ECS and EM was set up in Chapter (5) where the stability of the topological construct with respect to the already established modules in Homology was studied. The construction of ECS using sim-

plicial complex, Alpha complex was done and compared with the construction of ECS through coarse-graining or union of r - neighborhood method(used in Chapters (3,4)). Theorems were proposed and proved that describe the spatial and temporal stability of the ECS. This mathematical stability vastly strengthens the construction of ECS and EM. Simulated data using the Vicsek model was used to further verify the proposed theorems. In the process of analyzing the data, additionally, a cue for phase transition from order to disorder through ECSs and EM was uncovered, which further supported how sensitive and strong this topological construction can be!

Below, I summarise the key findings of this whole journey.

- Examining dynamical systems through the lens of topology and geometry can reveal valuable insights into the underlying physical processes. For instance, it can elucidate different phases and dominant flows in droplet evaporation, the geometric transformations in columnar joints driven by thermal flow, the critical flow regimes in the modified egg-beater model, and the patterns of aggregation in the Vicsek model, all of which align cohesively with the core philosophy of this study.
- Our novel tool - Euler Characteristic Surface and Euler metric worked fairly well as a topological descriptor and marker in the dynamical systems studied. It created a multiscale temporal topological summary for a dynamical system, accompanied by the Euler Characteristic Level Curves that carry that summary in low dimensional projection. The Euler metric(EM) quantified the similarity and dissimilarity between two systems in terms of their topology. These measures were not only able to characterize and distinguish between complex dynamical systems but also have the ability to carry cues for physical phase transition and critical points in the studied dynamical systems.
- Computing ECS and EM is moderately less expensive for large complex datasets. Solving differential equations numerically in dynamical systems or using other tools of topological data analysis (TDA), like Persistence Homology (PH) is sometimes difficult for large complex systems. The simplistic formulation of Euler Characteristic Surface along with its computational advantage make it suitable for possible applications in various practical situations involving big data.
- The verified quantifiable stability of the Euler Metric(EM) establishes the ro-

bustness of the construction- the Euler Characteristic Surface(ECS). Establishing the relation between the ECS and Persistent Homology(PH) completes mathematical framework.

- There still exists some limitations - Comparison of ECSs through Euler Metric has been limited to systems having a comparable scale of resolution and temporal stretch, in this dissertation. How the framework can be used for dynamical systems with widely different filtrations/ scales of resolution and temporal span, needs further consideration. Also, the optimal removal of noise and data extraction is the core of the construction, for working with real data, on which the outcome of the application of our tools vastly depends.
- Lastly, The ECS is a very new tool and it indeed needs diverse application on various dynamical systems for more interpretation!

6.0.1 Future directions:

There are indeed certain avenues for further exploration and research. The cues of order-disorder phase transition that we observed in the Vicsek model, that may be examined further for more precise results with trialing over a larger number of configurations. It will be certainly computationally very expensive and needs a clear framework before proceeding.

Also, can this approach of studying topological information further be used for “parameter recovery” in dynamical systems? Given a sufficient amount of information on different scenarios in a dynamical system through constructing corresponding Euler Characteristic Surfaces, whether parameter recovery can be done for an unknown scenario in real data sets. That could be beneficial for machine learning algorithms.

This stability study of the ECS in Chapter(5) was limited to the stability with respect to scale and time, separately, one at a time. In the future, the mathematical framework can further be extended to study the total stability of the Euler Characteristic Surface with respect to perturbation in time and scale, both simultaneously.

Mathematically, one may also think of constructing a higher dimensional version of the Euler Characteristic Surface where one may adapt filtering functions

$f : \mathcal{K} \rightarrow \mathbb{R}^{N+1}$, with N coordinates representing N spatial filters and one coordinate representing time and can proceed with similar manner. But, before applying it to a physical system one must reflect on where it can be applicable and whether that extra information provides a better understanding of the dynamical system.

Lastly, this overall construct was studied for the few dynamical systems that we discussed in this dissertation. There exist different dynamical systems where studying the evolution of topology through the proposed construct may be explored, particularly I am interested studying collective motion in active matter and brain connectivity with the same philosophy that I proposed here.

Bibliography

- [1] Wikimedia Commons Contributors, Topology joke, https://commons.wikimedia.org/wiki/File:Topology_joke.jpg, accessed: 2024-08-26 (2023).
- [2] R. Pujar, A. Kumar, K. Rao, S. Sadhukhan, T. Dutta, S. Tarafdar, G. U. Kulkarni, Narrowing desiccating crack patterns by an azeotropic solvent for the fabrication of nanomesh electrodes, *Langmuir* 35 (49) (2019) 16130–16135.
- [3] Flickr User: UD Research Institute, Columnar joints image, <https://www.flickr.com/photos/udri/49887068846>, accessed: 2024-08-26 (2023).
- [4] Flickr User: Bioversity International, River bed image, <https://www.flickr.com/photos/bioversity/16380021422>, accessed: 2024-08-26 (2023).
- [5] P. Budkewitsch, P.-Y. Robin, Modelling the evolution of columnar joints, *Journal of Volcanology and Geothermal Research* 59 (3) (1994) 219–239.
- [6] T. Khatun, T. Dutta, S. Tarafdar, islands in sea and lakes in mainland phases and related transitions simulated on a square lattice, *The European Physical Journal B* 90 (11) (2017) 213.
- [7] S. Tarafdar, T. DUTTA, T. Khatun, S. Dutta, S. Sen, Euler number and percolation threshold on a square lattice with diagonal connection probability and revisiting the island-mainland transition, *Frontiers in Physics* 7 (2019) 61.

- [8] T. Khatun, M. D. Choudhury, T. Dutta, S. Tarafdar, Electric-field-induced crack patterns: Experiments and simulation, *Physical Review E* 86 (1) (2012) 016114.
- [9] M. F. Sykes, J. W. Essam, Exact critical percolation probabilities for site and bond problems in two dimensions, *Journal of Mathematical Physics* 5 (8) (1964) 1117–1127.
- [10] D. Stauffer, A. Aharony, *Introduction to percolation theory*, CRC press, 2018.
- [11] S. H. Strogatz, *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*, CRC press, 2018.
- [12] K. T. Alligood, T. D. Sauer, J. A. Yorke, D. Chillingworth, *Chaos: an introduction to dynamical systems*, *SIAM Review* 40 (3) (1998) 732–732.
- [13] H. Koorehdavoudi, P. Bogdan, A statistical physics characterization of the complex systems dynamics: Quantifying complexity from spatio-temporal interactions, *Scientific reports* 6 (1) (2016) 27602.
- [14] E. O. Wilson, *Consilience: The unity of knowledge*, Vol. 31, Vintage, 1999.
- [15] S. H. Strogatz, Exploring complex networks, *nature* 410 (6825) (2001) 268–276.
- [16] R. Albert, A.-L. Barabási, Statistical mechanics of complex networks, *Reviews of modern physics* 74 (1) (2002) 47.
- [17] Y. Holovatch, R. Kenna, S. Thurner, Complex systems: physics beyond physics, *European Journal of Physics* 38 (2) (2017) 023002.
- [18] A. F. Siegenfeld, Y. Bar-Yam, An introduction to complex systems science and its applications, *Complexity* 2020 (1) (2020) 6105872.
- [19] H.-J. Vogel, Topological characterization of porous media, in: *Morphology of condensed matter: Physics and geometry of spatially complex systems*, Springer, 2002, pp. 75–92.

- [20] C. Scholz, F. Wirner, J. Götz, U. Rüde, G. E. Schröder-Turk, K. Mecke, C. Bechinger, Permeability of porous materials determined from the euler characteristic, *Physical review letters* 109 (26) (2012) 264504.
- [21] J. Schmalzing, K. M. Górski, Minkowski functionals used in the morphological analysis of cosmic microwave background anisotropy maps, *Monthly Notices of the Royal Astronomical Society* 297 (2) (1998) 355–365.
- [22] F. A. Santos, E. P. Raposo, M. D. Coutinho-Filho, M. Copelli, C. J. Stam, L. Douw, Topological phase transitions in functional brain networks, *Physical Review E* 100 (3) (2019) 032414.
- [23] E. C. de Amorim Filho, R. A. Moreira, F. A. Santos, The euler characteristic and topological phase transitions in complex systems, *Journal of Physics: Complexity* 3 (2) (2022) 025003.
- [24] L. Snidaro, G. L. Foresti, Real-time thresholding with euler numbers, *Pattern Recognition Letters* 24 (9-10) (2003) 1533–1544.
- [25] R. Gilmore, Topological analysis of chaotic dynamical systems, *Reviews of Modern Physics* 70 (4) (1998) 1455.
- [26] M. Ghil, D. Sciamarella, Dynamical systems, algebraic topology and the climate sciences, *Nonlinear Processes in Geophysics* 30 (4) (2023) 399–434.
- [27] A. Zomorodian, G. Carlsson, Computing persistent homology, in: *Proceedings of the twentieth annual symposium on Computational geometry*, 2004, pp. 347–356.
- [28] G. Carlsson, Topology and data, *Bulletin of the American Mathematical Society* 46 (2) (2009) 255–308.
- [29] G. Carlsson, Topological methods for data modelling, *Nature Reviews Physics* 2 (12) (2020) 697–708.
- [30] L. Goehring, A. Nakahara, T. Dutta, S. Kitsunozaki, S. Tarafdar, *Desiccation cracks and their patterns: Formation and Modelling in Science and Nature*, John Wiley & Sons, 2015.

- [31] S. Bohn, L. Pauchard, Y. Couder, Hierarchical crack pattern as formed by successive domain divisions., *Physical Review E* 71 (4) (2005) 046214.
- [32] S. Bohn, J. Platkiewicz, B. Andreotti, M. Adda-Bedia, Y. Couder, Hierarchical crack pattern as formed by successive domain divisions. ii. from disordered to deterministic behavior, *Physical Review E* 71 (4) (2005) 046215.
- [33] S. Bucklow, The description of craquelure patterns, *Studies in conservation* 42 (3) (1997) 129–140.
- [34] L. Krzemień, M. Łukomski, Ł. Bratasz, R. Kozłowski, M. F. Mecklenburg, Mechanism of craquelure pattern formation on panel paintings, *Studies in Conservation* 61 (6) (2016) 324–330.
- [35] Y. Couder, L. Pauchard, C. Allain, M. Adda-Bedia, S. Douady, The leaf venation as formed in a tensorial field, *The European Physical Journal B-Condensed Matter and Complex Systems* 28 (2) (2002) 135–138.
- [36] M. F. Laguna, S. Bohn, E. A. Jagla, The role of elastic stresses on leaf venation morphogenesis, *PLoS Comput Biol* 4 (4) (2008) e1000055.
- [37] Y. Bar-Sinai, J.-D. Julien, E. Sharon, S. Armon, N. Nakayama, M. Adda-Bedia, A. Boudaoud, Mechanical stress induces remodeling of vascular networks in growing leaves, *PLoS computational biology* 12 (4) (2016) e1004819.
- [38] E. Aker, K. J. Måløy, A. Hansen, Dynamics of stable viscous displacement in porous media, *Physical Review E* 61 (3) (2000) 2936.
- [39] R. Morgado, F. A. Oliveira, G. G. Batrouni, A. Hansen, Relation between anomalous and normal diffusion in systems with memory, *Physical review letters* 89 (10) (2002) 100601.
- [40] H. A. Knudsen, E. Aker, A. Hansen, Bulk flow regimes and fractional flow in 2d porous media by numerical simulations, *Transport in Porous Media* 47 (1) (2002) 99–121.
- [41] B. Zhao, C. W. MacMinn, B. K. Primkulov, Y. Chen, A. J. Valocchi, J. Zhao, Q. Kang, K. Bruning, J. E. McClure, C. T. Miller, et al., *Comprehensive com-*

- parison of pore-scale models for multiphase flow in porous media, *Proceedings of the National Academy of Sciences* 116 (28) (2019) 13799–13806.
- [42] L. De Arcangelis, A. Hansen, H. Herrmann, S. Roux, Scaling laws in fracture, *Physical Review B* 40 (1) (1989) 877.
- [43] A. Hansen, J. Schmittbuhl, Origin of the universal roughness exponent of brittle fracture surfaces: stress-weighted percolation in the damage zone, *Physical review letters* 90 (4) (2003) 045504.
- [44] S. Santucci, K. J. Måløy, A. Delaplace, J. Mathiesen, A. Hansen, J. Ø. H. Bakke, J. Schmittbuhl, L. Vanel, P. Ray, Statistics of fracture surfaces, *Physical review E* 75 (1) (2007) 016104.
- [45] J. Schmittbuhl, A. Hansen, G. G. Batrouni, Roughness of interfacial crack fronts: stress-weighted percolation in the damage zone, *Physical review letters* 90 (4) (2003) 045505.
- [46] G. G. Batrouni, A. Hansen, J. Schmittbuhl, Heterogeneous interfacial failure between two elastic blocks, *Physical Review E* 65 (3) (2002) 036126.
- [47] S. Pradhan, A. Hansen, Failure properties of loaded fiber bundles having a lower cutoff in fiber threshold distribution, *Physical Review E* 72 (2) (2005) 026111.
- [48] A. Hansen, P. C. Hemmer, S. Pradhan, *The fiber bundle model: modeling failure in materials*, John Wiley & Sons, 2015.
- [49] A. Moreira, C. Oliveira, A. Hansen, N. Araújo, H. Herrmann, J. Andrade Jr, Fracturing highly disordered materials, *Physical review letters* 109 (25) (2012) 255701.
- [50] K. Mecke, *Integralgeometrie in der statistischen Physik: Perkolation, komplexe Flüssigkeiten und die Struktur des Universums*, Deutsch, 1994.
- [51] C. H. Arns, J. Mecke, K. Mecke, D. Stoyan, Second-order analysis by variograms for curvature measures of two-phase structures, *The European Physical Journal B-Condensed Matter and Complex Systems* 47 (3) (2005) 397–409.

- [52] K. R. Mecke, Integral geometry in statistical physics, *International Journal of Modern Physics B* 12 (09) (1998) 861–899.
- [53] K. R. Mecke, Additivity, convexity, and beyond: applications of minkowski functionals in statistical physics, in: *Statistical Physics and Spatial Statistics*, Springer, 2000, pp. 111–184.
- [54] K. R. Mecke, T. Buchert, H. Wagner, Robust morphological measures for large-scale structure in the universe, arXiv preprint astro-ph/9312028 (1993).
- [55] K. Jacobs, S. Herminghaus, K. R. Mecke, Thin liquid polymer films rupture via defects, *Langmuir* 14 (4) (1998) 965–969.
- [56] E. Bouchaud, Scaling properties of cracks, *Journal of Physics: Condensed Matter* 9 (21) (1997) 4319.
- [57] J. Baer, T. Kent, S. Anderson, Image analysis and fractal geometry to characterize soil desiccation cracks, *Geoderma* 154 (1-2) (2009) 153–163.
- [58] T. Ohnishi, O. Okada, H. Shirakata, Morphological similarity of road networks and cracks, *Physica A: Statistical Mechanics and its Applications* 392 (18) (2013) 4127–4133.
- [59] J. Bouchaud, E. Bouchaud, G. Lapasset, J. Planes, Models of fractal cracks, *Physical review letters* 71 (14) (1993) 2240.
- [60] R. P.-y. Wei, *Fracture Mechanics: Perspectives and Directions (twentieth Symposium)*, Vol. 1020, Astm International, 1989.
- [61] P. Meakin, The growth of rough surfaces and interfaces, *Physics Reports* 235 (4-5) (1993) 189–289.
- [62] B. D. Ripley, *Statistical inference for spatial processes*, Cambridge university press, 1988.
- [63] P. J. Diggle, *Statistical analysis of spatial and spatio-temporal point patterns*, CRC press, 2013.

- [64] J. Moller, R. P. Waagepetersen, Statistical inference and simulation for spatial point processes, CRC Press, 2003.
- [65] D. Stoyan, H. Stoyan, Estimating pair correlation functions of planar cluster processes, *Biometrical Journal* 38 (3) (1996) 259–271.
- [66] D. Stoyan, H. Stoyan, Improving ratio estimators of second order point process characteristics, *Scandinavian Journal of Statistics* 27 (4) (2000) 641–656.
- [67] D. Stoyan, H. Stoyan, A. Tscheschel, T. Mattfeldt, On the estimation of distance distribution functions for point processes and random sets, *Image Analysis & Stereology* 20 (1) (2001) 65–69.
- [68] S. N. Chiu, D. Stoyan, W. S. Kendall, J. Mecke, *Stochastic geometry and its applications*, John Wiley & Sons, 2013.
- [69] K. R. Mecke, D. Stoyan, Morphological characterization of point patterns, *Biometrical Journal: Journal of Mathematical Methods in Biosciences* 47 (4) (2005) 473–488.
- [70] R. J. Adler, On excursion sets, tube formulas and maxima of random fields, *Annals of Applied Probability* (2000) 1–74.
- [71] K. J. Worsley, Local maxima and the expected euler characteristic of excursion sets of χ^2 , f and t fields, *Advances in Applied Probability* 26 (1) (1994) 13–42.
- [72] K. J. Worsley, Estimating the number of peaks in a random field using the hadwiger characteristic of excursion sets, with applications to medical images, *The Annals of Statistics* (1995) 640–669.
- [73] K. J. Worsley, Boundary corrections for the expected euler characteristic of excursion sets of random fields, with an application to astrophysics, *Advances in Applied Probability* 27 (4) (1995) 943–959.
- [74] C. A. Andresen, A. Hansen, R. Le Goc, P. Davy, S. M. Hope, Topology of fracture networks, *Frontiers in Physics* 1 (2013) 7.

- [75] S. M. Hope, P. Davy, J. Maillot, R. Le Goc, A. Hansen, Topological impact of constrained fracture growth, *Frontiers in Physics* 3 (2015) 75.
- [76] L. Y. Yi, K. J. Dong, R. P. Zou, A. B. Yu, [Coordination number of the packing of ternary mixtures of spheres: Dem simulations versus measurements](#), *Industrial & Engineering Chemistry Research* 50 (14) (2011) 8773–8785. doi:[10.1021/ie200765h](https://doi.org/10.1021/ie200765h). URL <https://doi.org/10.1021/ie200765h>
- [77] G. R. Irwin, Analysis of stresses and strains near the end of a crack transversing a plate, *Trans. ASME, Ser. E, J. Appl. Mech.* 24 (1957) 361–364.
- [78] B. Grünbaum, G. Shephard, *Tilings & Patterns*, Dover, 2016.
- [79] G. Domokos, D. J. Jerolmack, F. Kun, J. Török, Plato’s cube and the natural geometry of fragmentation, *Proceedings of the National Academy of Sciences* 117 (31) (2020) 18178–18185.
- [80] R. Osserman, The isoperimetric inequality, *Bulletin of the American Mathematical Society* 84 (6) (1978) 11821238.
- [81] T. Khatun, T. Dutta, S. Tarafdar, Topology of desiccation crack patterns in clay and invariance of crack interface area with thickness, *The European Physical Journal E* 38 (8) (2015) 1–11.
- [82] Y. Akiba, J. Magome, H. Kobayashi, H. Shima, Morphometric analysis of polygonal cracking patterns in desiccated starch slurries, *Physical Review E* 96 (2) (2017) 023003.
- [83] R. Seghir, S. Arscott, Controlled mud-crack patterning and self-organized cracking of polydimethylsiloxane elastomer surfaces, *Scientific reports* 5 (1) (2015) 1–16.
- [84] S. Sadhukhan, A. Kumar, G. U. Kulkarni, S. Tarafdar, T. Dutta, A spring network simulation in three dimensions for designing optimal crack pattern template to fabricate transparent conducting electrodes, *Bulletin of Materials Science* 42 (5) (2019) 1–14.

- [85] M. Kraus, N. Pourmoghaddam, J. Schneider, G. Siebert, Break-calibrating stochastic tessellations for the prediction of fracture patterns of thermally prestressed glass, Glass Performance Days, Tampere (2019).
- [86] sidefx.com, Working with glass, <https://www.sidefx.com/docs/houdini/destruction/glass.html>, accessed : 2021-02-24.
- [87] pxfuel.com, brown soil, ground, cracked, dry, desert, nature, drought, land, dirt, arid climate, <https://images.app.goo.gl/aN7ZvP5gpEQUDMk56>, accessed :2021-02-21.
- [88] G. Bradski, The OpenCV Library, Dr. Dobb's Journal of Software Tools (2000).
- [89] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu, the scikit-image contributors, [scikit-image: image processing in Python](https://doi.org/10.7717/peerj.453), PeerJ 2 (2014) e453. doi:10.7717/peerj.453.
URL <https://doi.org/10.7717/peerj.453>
- [90] J. Hoshen, R. Kopelman, Percolation and cluster distribution. i. cluster multiple labeling technique and critical concentration algorithm, Physical Review B 14 (8) (1976) 3438.
- [91] C. B. Barber, D. P. Dobkin, H. Huhdanpaa, [The quickhull algorithm for convex hulls](https://doi.org/10.1145/235815.235821), ACM Trans. Math. Softw. 22 (4) (1996) 469–483. doi:10.1145/235815.235821.
URL <https://doi.org/10.1145/235815.235821>
- [92] A. Roy, R. Haque, A. Mitra, S. Tarafdar, T. Dutta, Combinatorial topology and geometry of fracture networks, Physical Review E 105 (3) (2022) 034801.
- [93] L. Goehring, S. W. Morris, Z. Lin, Experimental investigation of the scaling of columnar joints, Physical Review E 74 (3) (2006) 036115.
- [94] L. Goehring, Evolving fracture patterns: columnar joints, mud cracks and polygonal terrain, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 371 (2004) (2013) 20120353.

- [95] R. A. Haque, A. Roy, A. J. Mitra, S. Tarafdar, T. Dutta, Evolution of tiling-like crack patterns in maturing columnar joints, *Langmuir* 38 (22) (2022) 7035–7045.
- [96] D. Brutin, *Droplet wetting and evaporation: from pure to complex fluids*, Academic Press, 2015.
- [97] S. Tarafdar, Y. Y. Tarasevich, M. Dutta Choudhury, T. Dutta, D. Zang, Droplet drying patterns on solid substrates: from hydrophilic to superhydrophobic contact to levitating drops, *Advances in Condensed Matter Physics* 2018 (2018).
- [98] L. Euler, *Elementa doctrinae solidorum*, *Novi commentarii academiae scientiarum Petropolitanae* (1758) 109–140.
- [99] R. J. Adler, O. Bobrowski, M. S. Borman, E. Subag, S. Weinberger, Persistent homology for random fields and complexes, in: *Borrowing strength: theory powering applications—a Festschrift for Lawrence D. Brown*, Vol. 6, Institute of Mathematical Statistics, 2010, pp. 124–144.
- [100] R. J. Adler, K. Bartz, S. C. Kou, A. Monod, Estimating thresholding levels for random fields via euler characteristics, *arXiv preprint arXiv:1704.08562* (2017).
- [101] R. J. Adler, J. E. Taylor, et al., *Random fields and geometry*, Vol. 80, Springer, 2007.
- [102] J. Schmalzing, K. M. Górski, Minkowski functionals used in the morphological analysis of cosmic microwave background anisotropy maps, *Monthly Notices of the Royal Astronomical Society* 297 (2) (1998) 355–365.
- [103] P. Pranav, R. Van de Weygaert, G. Vegter, B. J. Jones, R. J. Adler, J. Feldbrugge, C. Park, T. Buchert, M. Kerber, Topology and geometry of gaussian random fields i: on betti numbers, euler characteristic, and minkowski functionals, *Monthly Notices of the Royal Astronomical Society* 485 (3) (2019) 4167–4208.

- [104] M. Kerscher, K. Mecke, J. Schmalzing, C. Beisbart, T. Buchert, H. Wagner, Morphological fluctuations of large-scale structure: The pscz survey, *Astronomy & Astrophysics* 373 (1) (2001) 1–11.
- [105] J. M. Kilner, S. J. Kiebel, K. J. Friston, Applications of random field theory to electrophysiology, *Neuroscience letters* 374 (3) (2005) 174–178.
- [106] H. Lee, M. K. Chung, H. Kang, B.-N. Kim, D. S. Lee, Discriminative persistent homology of brain networks, in: 2011 IEEE international symposium on biomedical imaging: from nano to macro, IEEE, 2011, pp. 841–844.
- [107] K. R. Mecke, D. Stoyan, *Statistical physics and spatial statistics: the art of analyzing and modeling spatial structures and pattern formation*, Vol. 554, Springer Science & Business Media, 2000.
- [108] C. H. Arns, M. A. Knackstedt, K. Mecke, Reconstructing complex materials via effective grain shapes, *Physical Review Letters* 91 (21) (2003) 215506.
- [109] K. Mecke, Morphological thermodynamics of composite media, *Fluid Phase Equilibria* 150 (1998) 591–598.
- [110] H. H. Khanamiri, C. F. Berg, P. A. Slotte, S. Schlüter, O. Torsæter, Description of free energy for immiscible two-fluid flow in porous media by integral geometry and thermodynamics, *Water Resources Research* 54 (11) (2018) 9045–9059.
- [111] H. Hansen-Goos, R. Roth, K. Mecke, S. Dietrich, Solvation of proteins: linking thermodynamics to geometry, *Physical review letters* 99 (12) (2007) 128101.
- [112] O. Bobrowski, P. Skraba, Homological percolation and the euler characteristic, *Physical Review E* 101 (3) (2020) 032304.
- [113] J. E. McClure, T. Ramstad, Z. Li, R. T. Armstrong, S. Berg, Modeling geometric state for fluids in porous media: Evolution of the euler characteristic, *Transport in Porous Media* 133 (2020) 229–250.
- [114] E. C. de Amorim Filho, R. A. Moreira, F. A. Santos, The euler characteristic and topological phase transitions in complex systems, *Journal of Physics: Complexity* 3 (2) (2022) 025003.

- [115] A. Smith, V. M. Zavala, The euler characteristic: A general topological descriptor for complex data, *Computers & Chemical Engineering* 154 (2021) 107463.
- [116] A. Hatcher, *Algebraic Topology*, Cambridge University Press, 2002.
- [117] J. M. Stauber, S. K. Wilson, B. R. Duffy, K. Sefiane, Evaporation of droplets on strongly hydrophobic substrates, *Langmuir* 31 (12) (2015) 3653–3660.
- [118] H. Gelderblom, A. G. Marin, H. Nair, A. Van Houselt, L. Lefferts, J. H. Snoeijer, D. Lohse, How water droplets evaporate on a superhydrophobic substrate, *Physical Review E* 83 (2) (2011) 026306.
- [119] Z. Pan, S. Dash, J. A. Weibel, S. V. Garimella, Assessment of water droplet evaporation mechanisms on hydrophobic and superhydrophobic substrates, *Langmuir* 29 (51) (2013) 15831–15841.
- [120] K. Gleason, S. A. Putnam, Microdroplet evaporation with a forced pinned contact line, *Langmuir* 30 (34) (2014) 10548–10555.
- [121] R. C. Staunton, N. Storey, A comparison between square and hexagonal sampling methods for pipeline image processing, in: *Optics, Illumination, and Image Sensing for Machine Vision IV*, Vol. 1194, International Society for Optics and Photonics, 1990, pp. 142–151.
- [122] L. Middleton, J. Sivaswamy, *Hexagonal image processing: A practical approach*, Springer Science & Business Media, 2006.
- [123] K. P. Sankar, T. Sanjay, E. Rajan, Hexagonal pixel grid modelling and processing of digital images using clap algorithms, in: *International Conference on Systemic*, 2004.
- [124] R. Bennacer, K. Sefiane, Vortices, dissipation and flow transition in volatile binary drops, *Journal of fluid mechanics* 749 (2014) 649–665.
- [125] J. Franjione, J. Ottino, F. Smith, Symmetry concepts for the geometric analysis of mixing flows, *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences* 338 (1650) (1992) 301–323.

- [126] G. Carlsson, Topology and data, *Bulletin of the American Mathematical Society* 46 (2) (2009) 255–308.
- [127] G. Carlsson, A. Zomorodian, The theory of multidimensional persistence, in: *Proceedings of the twenty-third annual symposium on Computational geometry*, 2007, pp. 184–193.
- [128] C. M. Topaz, L. Ziegelmeier, T. Halverson, Topological data analysis of biological aggregation models, *PloS one* 10 (5) (2015) e0126383.
- [129] S. Mandal, A. Guzmán-Sáenz, N. Haiminen, S. Basu, L. Parida, A topological data analysis approach on predicting phenotypes from gene expression data, in: *International Conference on Algorithms for Computational Biology*, Springer, 2020, pp. 178–187.
- [130] R. Rabadán, Y. Mohamedi, U. Rubin, T. Chu, A. N. Alghalith, O. Elliott, L. Arnés, S. Cal, Á. J. Obaya, A. J. Levine, et al., Identification of relevant genetic alterations in cancer using topological data analysis, *Nature communications* 11 (1) (2020) 3808.
- [131] M. Gidea, Y. Katz, Topological data analysis of financial time series: Landscapes of crashes, *Physica A: Statistical Mechanics and its Applications* 491 (2018) 820–834.
- [132] P. T.-W. Yen, S. A. Cheong, Using topological data analysis (tda) and persistent homology to analyze the stock markets in singapore and taiwan, *Frontiers in Physics* 9 (2021) 572216.
- [133] U. Islambekov, M. Yuvaraj, Y. R. Gel, Harnessing the power of topological data analysis to detect change points, *Environmetrics* 31 (1) (2020) e2612.
- [134] M. Ulmer, L. Ziegelmeier, C. M. Topaz, A topological approach to selecting models of biological experiments, *PloS one* 14 (3) (2019) e0213679.
- [135] D. Cohen-Steiner, H. Edelsbrunner, D. Morozov, Vines and vineyards by updating persistence in linear time, in: *Proceedings of the twenty-second annual symposium on Computational geometry*, 2006, pp. 119–126.

- [136] L. Xian, H. Adams, C. M. Topaz, L. Ziegelmeier, Capturing dynamics of time-varying data via topology, arXiv preprint arXiv:2010.05780 (2020).
- [137] G. Carlsson, V. De Silva, D. Morozov, Zigzag persistent homology and real-valued functions, in: Proceedings of the twenty-fifth annual symposium on Computational geometry, 2009, pp. 247–256.
- [138] A. Myers, D. Muñoz, F. A. Khasawneh, E. Munch, Temporal network analysis using zigzag persistence, EPJ Data Science 12 (1) (2023) 6.
- [139] G. Beltramo, P. Skraba, R. Andreeva, R. Sarkar, Y. Giarratano, M. Bernabeu, Euler characteristic surfaces, Foundations of Data Science 4 (2022) 505–536.
- [140] P. Dlotko, D. Gurnari, Euler characteristic curves and profiles: a stable shape invariant for big data problems, Giga Science 12 (2023) 1–16.
- [141] H. Edelsbrunner, J. Harer, Computational Topology: An Introduction, American Mathematical Society, 2010.
- [142] Ž. Virk, Introduction to Persistent Homology, Založba UL FRI, 2022.
- [143] N. Otter, M. A. Porter, U. Tillmann, P. Grindrod, H. A. Harrington, A roadmap for the computation of persistent homology, EPJ Data Science 6 (2017) 1–38.
- [144] M. Kerber, Persistent homology: state of the art and challenges, International Mathematische Nachrichten 231 (15-33) (2016) 1.
- [145] C. S. Pun, S. X. Lee, K. Xia, Persistent-homology-based machine learning: a survey and a comparative study, Artificial Intelligence Review 55 (7) (2022) 5169–5213.
- [146] Y.-M. Chung, A. Lawson, Persistence curves: a canonical framework for summarizing persistence diagrams, Adv Comut. Math. 48:6 (2023) 1–42.
- [147] M. Carriere, M. Cuturi, S. Oudot, Sliced wasserstein kernel for persistence diagrams, Proceedings of the 34th International Conference on Machine Learning 70 (2017) 664–673.

- [148] P. Skraba, K. Turner, Wasserstein stability for persistence diagrams, arXiv arXiv:2006.16824v5 (2022) 1–41.
- [149] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, O. Shochet, Novel type of phase transition in a system of self-driven particles, *Physical review letters* 75 (6) (1995) 1226.
- [150] S. Agami, Comparison of persistence diagrams, *Communications in Statistics-Simulation and Computation* 52 (5) (2023) 1948–1961.
- [151] M. Tsizh, V. Tymchyshyn, F. Vazza, Wasserstein distance as a new tool for discriminating cosmologies through the topology of large-scale structure, *Monthly Notices of the Royal Astronomical Society* 522 (2) (2023) 2697–2706.
- [152] M. K. Chung, C. G. Ramos, F. B. De Paiva, J. Mathis, V. Prabhakaran, V. A. Nair, M. E. Meyerand, B. P. Hermann, J. R. Binder, A. F. Struck, Unified topological inference for brain networks in temporal lobe epilepsy using the wasserstein distance, *NeuroImage* 284 (2023) 120436.