Persistence Diagrams to Visualise Damage to Biological Networks

Matthew H. J. Bailey[1] and Mark Wilson[1*]

[1]Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford, OX1 3QZ, United Kingdom.

*Corresponding author(s). E-mail(s): mark.wilson@chem.ox.ac.uk;

Abstract

One of the critical tools of persistent homology is the persistence diagram. We demonstrate the applicability of a persistence diagram showing the existence of topological features (here rings in a 2D network) generated over time instead of space as a tool to analyse trajectories of biological networks. We show how the time persistence diagram is useful in order to identify critical phenomena such as rupturing and to visualise important features in 2D biological networks; they are particularly useful to highlight patterns of damage and to identify if particular patterns are significant or ephemeral. Persistence diagrams are also used to analyse repair phenomena, and we explore how the measured properties of a dynamical phenomenon change according to the sampling frequency. This shows that the persistence diagrams are robust and still provide useful information even for data of low temporal resolution. Finally, we combine persistence diagrams across many trajectories to show how the technique highlights the existence of sharp transitions at critical points in the rupturing process.

Keywords: Network Theory, Persistence Diagram, Collagen IV, Ring Statistics
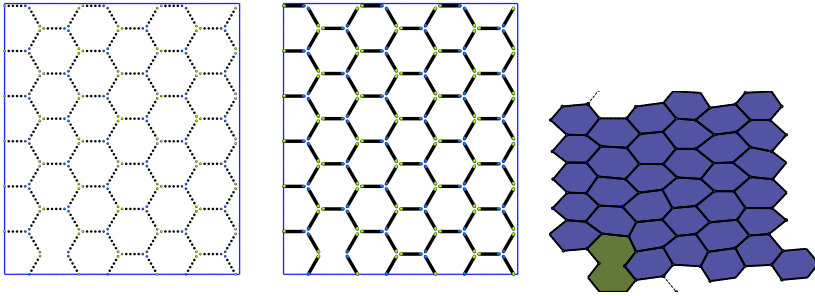
1 Introduction

Biological networks are often extremely complicated structures, often made of ordered arrangements of biopolymers. These structures can suffer damage or

degrade over time, with serious consequences for the functionality of the network and the organism containing that network. It is important to understand the precise nature of damage over the lifespan of a network; a better understanding of how defects or tears propagate over time could be used to inform methods to either halt or repair the damage. One such biopolymer network of scientific interest that loses function over time is that of collagen in the ocular lens capsule. The loss of flexibility owing to damage-induced changes in the lens capsule has consequences for the visual system losing its ability to accommodate different focal depths (Duncan et al (1997)), or the lens capsule can suffer immediate damage when ruptured during cataract surgery (Liu et al (2015)) or by exposure to focussed high energy sources.

Some biological systems consist of layers of two dimensional networks — depending on the network, these layers could themselves be highly ordered (made up of single repeating geometric units, said by Burd (2009) to be hexagons) or highly disordered as shown in microscope images from Barnard et al (1992) and Yurchenco and Ruben (1987). To avoid the computational complexity of considering the position of every atom, it can be a powerful approach is to simplify the biological network with a graph theory approach, treating molecules as edges and their interaction sites as nodes. The use of graph theory opens up an avenue to analyse the mid-range structure of 2D networks, an approach that has seen great success when used by Le Roux and Jund (2010) to analyse inorganic 2D networks such as glasses or by Ormrod Morley et al (2020) to study amorphous graphene.

The rings in a 2D network can be conceived of as being the next level up in a hierarchy of network features: a node at an interaction site is a zero-dimensional feature, an edge between two nodes is a one-dimensional feature, and the rings formed edges are two-dimensional features as shown in **Figure 1**. A ring representation of a 2D network allows for analysis of medium range order, including how those rings are correlated to one another (for example, are rings with many sides adjacent to rings with few sides or *vice versa* as measured by Sadjadi and Thorpe (2016)) or characterising networks by the distribution of the rings seen in them (referred to by authors such as Kumar et al (2014) or Le Roux and Jund (2010) as 'ring statistics'). Ring statistics are also used in studies of 3D networks, although the definition of a ring is less well defined and the ring structure is considerably harder to visualise. The most intuitive set of 3D features to treat with a persistence diagram are pores in zeolites as discussed by Lee et al (2017), although those are outside the scope of this work.

Framing biological networks like this is reminiscent of the persistent homology frameworks used to characterise proteins, 2D atomic glasses and zeolite structures (see refs. Xia et al (2015); Ormrod Morley et al (2021); Lee et al (2017)). Persistent homology approaches work by analysing objects called homology groups, generally constructed by placing a sphere of size r on each data point in a point cloud and expanding the size of r. When two spheres overlap, an edge is drawn between them and the simplices (polygons, polyhedra,

(a)   Atom positions, a
zero dimensional feature
with little spatial order-
ing information

(b) Molecules formed of
bonds between atoms, a
one dimensional feature
with short range spatial
ordering information

(c) Rings formed by the
gaps between molecules,
a two dimensional feature
with medium-range spa-
tial ordering information

Figure 1: Rings can be seen as medium-range order in a network, and built up as the highest dimensional feature. Here, the same network is represented in three different ways: atom positions, molecules and rings

or hyper-polyhedra) that are formed are analysed. The homology groups generated that way are often unwieldy (Steinberg (2019)), and are most often simplified into a few useful outputs (among others): Betti numbers β_n which measure the number of n-dimensional features; persistence diagrams which show how features appear and disappear as functions of r; and birth-death diagrams which show clusters of features appearing or disappearing. We can map the persistent homology framework onto the established network theory simply, as the first three Betti numbers correspond to the number of zero-, one- and two-dimensional features in the networks. In most persistent homology approaches, a persistence diagram charts at which value of a length scale (called a filtration parameter) various features first start to exist ('are born', in the language of persistent homology) or stop existing ('die') (Pereira and De Mello (2015)). Persistence diagrams have been criticised (see, for example, refs. Sørensen et al (2020) and Ormrod Morley et al (2021)) as they may be hard to interpret quantitatively, and they often reflect properties of a network that are difficult to visualise.

Furthermore, analysing biological networks as a function of a length scale at a single point in time can potentially lose critical information. For a biological network, a meaningful length scale is already defined as approximately the length of one biomolecule, and the most meaningful information is in how those networks vary across time. To resolve this, we can look at persistence diagrams generated at different timescales and a fixed length scale. Related work by Dong et al (2021) discusses using persistence diagrams that are functions of both length scales and time scales to measure similarity between zeolites.
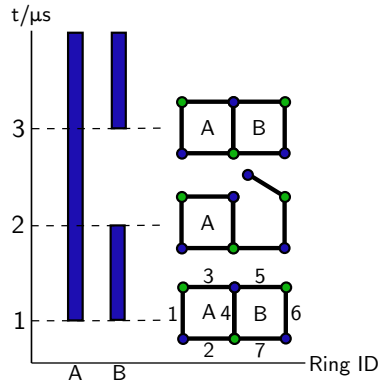
Figure 2: A schematic persistence diagram across time consisting of two rings. The left ring, labelled A is made up of molecules $\{1, 2, 3, 4\}$ and exists continually, leading to a solid persistence bar. The right ring, labelled B, is made up of molecules $\{4, 5, 6, 7\}$ and dies once during the simulation at $t = 2$, leading to a break in the persistence bar. When it reforms at $t = 3$ it is recognised as the same ring and there is a new segment to the bar

In this work, we take the idea of a persistence diagram (and the associated metrics for birth-death diagrams and ring lifetimes) and show their qualitative power for analysing ring statistics over time.

2 Methods

2.1 Persistence Diagrams

While criticism has been levelled at persistence diagrams for how much they rely on qualitative analysis (see refs. Sørensen et al (2020); Ormrod Morley (2020); Wasserman (2018)), if that is what's desired they can be extremely useful. The analysis of persistence diagrams is simplified if instead we track whether particular features in a network persistent across time instead of length scale. The medium-range order as represented by a ring structure in a network is especially interesting to us, as it is often hardest to track. In this analysis, we define a ring $\mathcal{R}$ as being constructed of a unique set of molecules — for example, a square might be made up of molecules $\{1, 2, 3, 4\}$ and a second triangle could be made up of molecules $\{4, 5, 6, 7\}$. We then can calculate an existence function $f(t; \mathcal{R})$ for each ring $\mathcal{R}$ over time t during a trajectory of interest. This function $f(t; \mathcal{R})$ has a major difference from the usual filtration functions used for persistent homology in that a single ring can be born, die, and live again. For example, the square $\{4, 5, 6, 7\}$ could exist at $t = 1$, die at $t = 2$, be reborn at $t = 3$ and die a final time at $t = 4$ as seen in the right bar of **Figure 2**. This leads to the persistence diagram in **Figure 2** striped in both directions, which will be useful for analysis of repair phenomena. The

lifespan of a given ring can be defined as $L(\mathcal{R}) = \int_0^\infty f(t;\ \mathcal{R})\mathrm{d}t$, and birth-death diagrams can be plotted using the first birth and final death of a given ring.

2.2 Polymer Model

To generate trajectories of interest, we use a simplified biopolymer model to mimic the self assembly process of collagen IV, first used by Bailey and Wilson (2021a). The biopolymers are made up of a linear chain of beads with bead-bead bonds modelled using a Morse potential to allow for bond breaking. The beads at either end of the polymer chain are marked as 'head beads' and experience Lennard-Jones attractions towards other head beads. The beads within the polymer are marked as 'body beads' and a truncated Lennard-Jones potential is used to mimic repulsion between molecule bodies and provide an excluded volume. This model has been successfully by Bailey and Wilson (2021a) used to mimic self-assembly phenomena of biological networks.

For the persistence diagram approach to analyse two different networks, both simulated using the polymer model: an ordered hexagonal network (inspired by a model proposed by Burd (2009)) and a disordered self-assembled network (inspired by studies performed by Yurchenco and Furthmayr (1984)).

For the ordered hexagonal networks we first started with an ordered hexagonal network with each edge being made up of one polymer. To break the symmetry of the network, a single polymer was removed to create a ten-sided ring. Over a time-scale of 100 μs this hexagonal network was stretched sinusoidally with a time period of 5.00 μs per cycle and the cycles increasing in amplitude from 10.0 % to 50.0 %. This method has been used by Bailey and Wilson (2021b) to mimic damage due to stretching in the ocular lens capsule.

To generate disordered networks we used an established self-assembly process that starts with a collection of the model polymers (Bailey et al (2020)). The polymers are initially placed on a random square grid, and allowed to form a proto-network, which was then equilibrated under a barostat at a constant stress ('2D pressure') p and temperature 100 K to fix the simulation cell size, and thus area per molecule. The use of different stresses allows networks with different areas, and hence topologies, to be constructed. This proto-network was heated to 1000 K using a Langevin thermostat to create a polymer liquid. This polymer liquid was cooled back down to 100 K over 100 μs to re-form a network, and ring statistics were tracked during the network forming process. Once the networks were cooled, we repeated the stretching simulation for used for the hexagonal networks to induce rupturing and once again tracked the ring statistics.

Ring statistics in a 2D graph can be tracked easily, as the excluded volume of the beads in the polymer model means that the polymers naturally form a planar graph with a defined embedding in a periodic plane under all but the most extreme conditions. The rings are then found as being nodes in the dual graph of the polymer graph; the dual graph can be found either by computing an anticlockwise ordering of edges around each node (de Berg (2000)), or

by computing the Delaunay triangulation of the points and re-joining the triangulation into rings (Bailey and Wilson (2021a)). In this work we use the latter method as we often wish to compute the Delaunay triangulation anyway for analysing networks (Dhume et al (2019)), and it can be simpler to treat algorithmically for a periodic simulation cell.

3   Results

With the simulation protocol established, we simulated both ordered and disordered networks to analyse their formation and damage. We will demonstrate the use of persistence diagrams to analyse ring structure first with individual snapshots to link the visualisation method with the simulated networks, and then move on to showing how persistence diagrams over time can be used to analyse the formation and rupturing behaviour of networks.

3.1   Ordered Networks

Snapshots of a single stretching simulation of an ordered hexagonal network are shown in **Figure 3** over the first 50.0 μs of simulation time. The snapshots were taken when the networks passed through an unstretched state during the simulation, and the maximum amount of stretching increased between each snapshot. For the first 20.0 μs the network topology remains fundamentally the same, with only small distortions to the positions of nodes in **Figure 3(a-c)**. After 30.0 μs however we see the ring structure start to break down, and rings with many sides begin to appear in **Figure 3(d)**. As the stretching simulation progresses, we see in **Figure 3(f)** that by 50.0 μs almost all hexagons have been destroyed and mostly extremely large rings remain.

The information accessible can be useful in identifying broad mechanisms by which key processes are taking place, but may be insufficient for obtaining sufficient statistical information. A systematic approach can be enabled by using a persistence diagram with the time filtration function $f(x, \mathcal{R})$ as shown in **Figure 4(a)**. This persistence diagram correlates well with the qualitative analysis performed earlier — the hexagon lifetimes are solid from 0 μs to 20.0 μs, and begin to break from 25.0 μs to 50.0 μs, with a few 'stragglers' lasting for a long time. Next, mid size rings form, as shown by redder lines, some of which live for a long time. For example, the 16-sided shape in the top left of the snapshots at 30.0 μs to 50.0 μs is represented as a long red line. A few decagons form as the stretching gets more intense, but they have short lifetimes. For stretching times from 40.0 μs onwards, we observe that very large rings with extremely short lifetimes are observed; this matches what we see in the snapshots **Figure 3(e)** and **Figure 3(f)**. Seeing the lifetime of these rings is critical as it undercuts a temptation to overinterpret the snapshots — looking at **Figure 3(e)**, it might seem that the 14-sided rings are important as there are 5 of them. However, analysing **Figure 4(a)** shows that those rings are only ephemeral. The short lifespan rings show up in the birth-death diagram of **Figure 4(b)** as red dots close to the diagonal line, whereas the long-lived
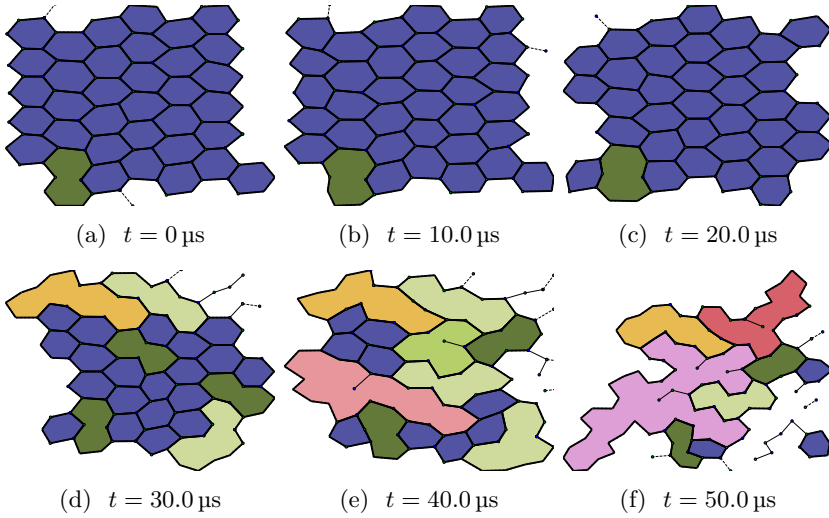
(a) $t = 0\,\mu s$ (b) $t = 10.0\,\mu s$ (c) $t = 20.0\,\mu s$

(d) $t = 30.0\,\mu s$ (e) $t = 40.0\,\mu s$ (f) $t = 50.0\,\mu s$

Figure 3: Ring snapshots of a hexagonal system being stretched until break-ing. The rings are coloured by their number of edges (for example, all hexagons are dark blue and all decagons are dark green), molecules that are not part of a ring are removed for clarity
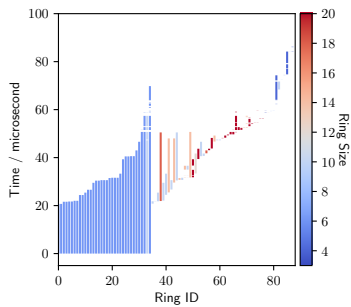
hexagons show up in the same diagram as blue dots all born at $t = 0\,\mu s$ on the left hand side. In **Figure 4(b)** we can see the evolution of the medium-range order as represented by the ring structure, as a series of hexagons die (blue dots) and are replaced with short-lived large rings (red dots).
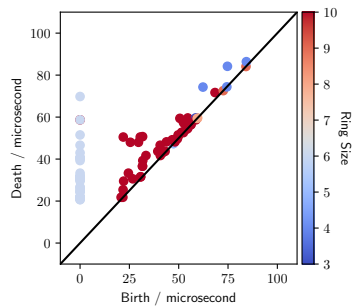
3.2   Disordered Networks

The analysis process can be repeated for more complex disordered networks, showing the power of barcode diagrams inspired from persistent homology approaches. The networks in this section were first self-assembled then rup-tured as discussed in **Section 2**. Pairs of birth-death diagrams and barcodes are shown in **Figure 5** for networks assembled under different initial condi-tions (here, the area available per molecule when the network formed, chosen systematically by applying a '2D pressure' to fix the simulation cell size).

The birth-death diagrams can then be used to identify a qualitative differ-ence in the way the networks form at different available areas per molecule. In **Figure 5(a)** and **Figure 5(d)**, only small rings with short lifespans are formed over the time-span $0\,\mu s$ to $40.0\,\mu s$. Snapshots of the networks at $40.0\,\mu s$ are shown in **Figure 5(g-i)**, just before the critical formation point. We can see that in the high area network there are many loose molecules which will soon assemble into rings.

Many rings with long lifespans then form over the time-span $40.0\,\mu s$ to $60.0\,\mu s$ with few rings formed in the final $60.0\,\mu s$ to $100\,\mu s$ as the system is

(a)   Persistence diagram (bar-code) of a hexagonal network stretched over 100 μs. Each bar is one ring, sorted left-to-right in order of the times they were first observed. The bars are coloured by the number of edges of each ring, and may be non-continuous if a ring was born, died, and was born again

(b)   A corresponding birth-death diagram to the lifetimes diagram for a hexagonal network being stretched. Each (birth, death) point corresponds to the first time a given ring was observed and the last time it was observed

Figure 4: Graphs traditionally used for persistent homology approaches used to analyse the trajectory shown in **Figure 3**

now stable. The gradient of **Figure 5(d)** represents the time between rings forming; the more horizontal line indicates that rings are formed rapidly and a more vertical indicates a long time-span between one ring being formed and the next.

The formation behaviour is no longer as sharply defined for the network formed at intermediate area per molecule in **Figure 5(b)**, and the network formed at low area per molecule in **Figure 5(c)** shows viable networks formed immediately. For these networks with less area per molecule, there is less room for molecules to 'dangle' and the network formation process has already begun.

A magnified view of the ring life lines is shown in **Figure 6** and emphasises a key difference this approach has from the traditional simplicial complex approach used in persistent homology. As discussed in **Section 1**, a dynamical system varying over time there is no guarantee that a given ring $\mathcal{R}$ that is born time t will still exist at a later time t'. Furthermore, dying at time t' does not preclude a ring from being born again even later at time t''. The loss of the mathematical continuity of simplices prevents us from using some further persistent homology analysis techniques (Stolz-Pretzer (2019)), such as using the barcodes to measure similarity of trajectories as performed by Chung et al (2021) or to identify spatially important information such as binding sites in the manner of Kovacev-Nikolic et al (2016). However, it is useful to identify
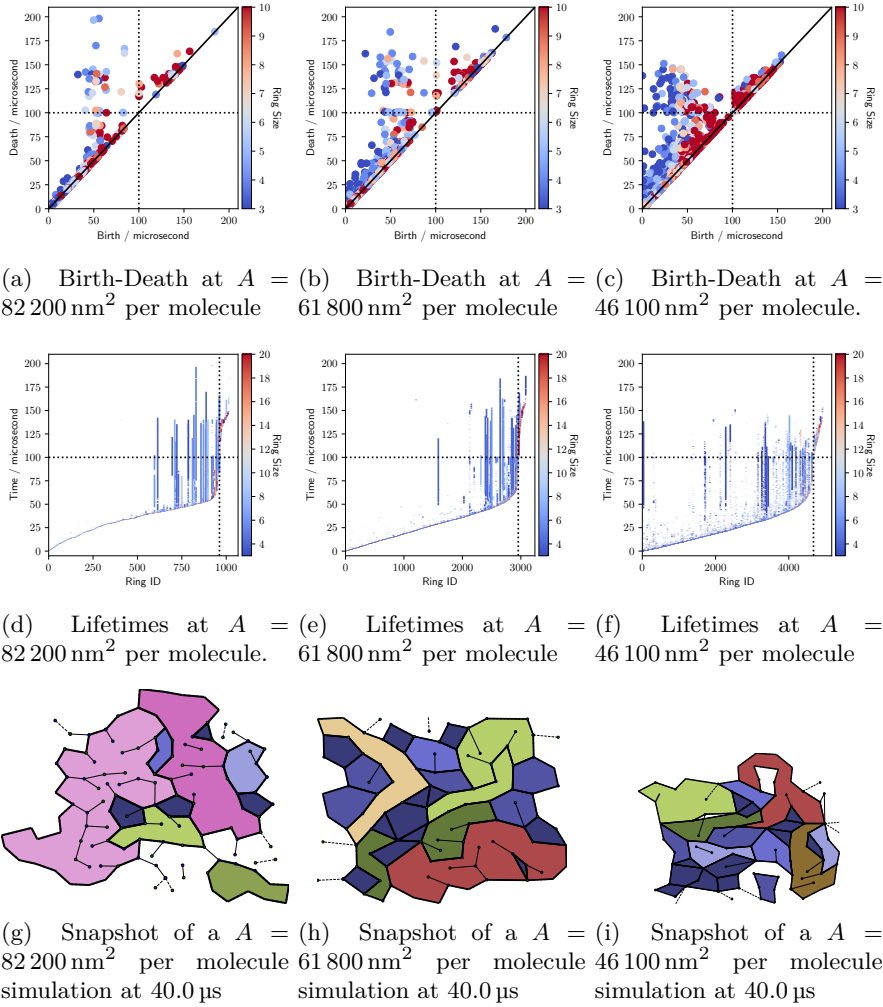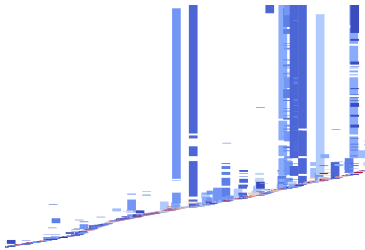
(a) Birth-Death at $A = 82\,200\,\mathrm{nm}^2$ per molecule

(b) Birth-Death at $A = 61\,800\,\mathrm{nm}^2$ per molecule

(c) Birth-Death at $A = 46\,100\,\mathrm{nm}^2$ per molecule.

(d) Lifetimes at $A = 82\,200\,\mathrm{nm}^2$ per molecule.

(e) Lifetimes at $A = 61\,800\,\mathrm{nm}^2$ per molecule

(f) Lifetimes at $A = 46\,100\,\mathrm{nm}^2$ per molecule

(g) Snapshot of a $A = 82\,200\,\mathrm{nm}^2$ per molecule simulation at $40.0\,\mu\mathrm{s}$

(h) Snapshot of a $A = 61\,800\,\mathrm{nm}^2$ per molecule simulation at $40.0\,\mu\mathrm{s}$

(i) Snapshot of a $A = 46\,100\,\mathrm{nm}^2$ per molecule simulation at $40.0\,\mu\mathrm{s}$

Figure 5: Birth-death diagrams and persistence barcodes (lifetimes) for networks with simulation cells equilibrated at different areas per molecule. The area per molecule was set by first forming a preliminary network and applying a '2D pressure' via a barostat to fix the simulation cell volume. The preliminary network was then melted at that fixed volume and reformed. Each diagram shows results from a single trajectory, while the top and bottom diagrams of each pair represent different analyses of the same trajectory
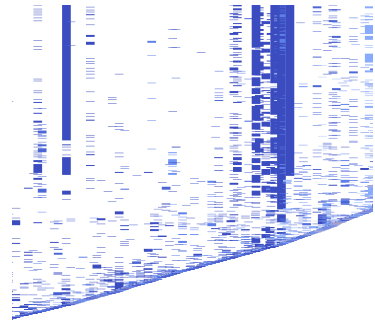
repair phenomena which are often of key interest in biological networks. We can see that rings in **Figure 6(a)** sometimes are born and die a few times before they finally have a stable lifespan, but broadly their life lines are solid and uninterrupted. However, one or two rings (for example, the left-most line)

Area per polymer /nm²	Unique Hexagons	Rebirths
46 100	62.8	4.81
61 800	56.7	3.66
82 200	26.4	3.39

Table 1: The number of non-ephemeral unique hexagons observed on average for simulations with a given area per polymer, and the average number of times each unique hexagon is reborn

(a) A highlight of **Figure 5(d)** showing rings 400 to 750 sputtering into existence

(b) A highlight of **Figure 5(f)** showing rings 1600 to 2100 sputtering into existence even more

Figure 6: Magnified sections of the persistence barcodes in **Figure 5** showing detail on the life-lines of some rings

are dead for extended periods of time but are reborn and continue to exist afterwards. For the higher density network shown in detail in **Figure 6(b)** we see a key difference from the low density network. For a high density network, the rings are close to one another and the polymer edges are packed in tightly. This means that few rings with long continuous lifespans are observed, but many microscopic damage / repair events happen continuously. This can be quantified with the data in **Table 1**, showing there are more unique hexagons observed for as the area per molecule decreases and the hexagons that are observed are reborn more often.

Turning our attention to the top right quartile of **Figure 5(a-c)** and the top half of **Figure 5(d-f)** we can see how this framework is useful to analyse the difference in rupturing phenomena. We can see in **Figure 5(a)** and **Figure 5(d)** that many rings are destroyed at approximately $t = 150\,\mu s$, although a few rings live for an unusually long time. At the end of the simulation a few small rings are briefly re-formed, leading to a point placed high on the birth-death diagram but no obvious corresponding line on the lifetime diagram. This is because certain tiny rings exist for one sampling point right

at the end of the simulation, below the resolution of the barcode diagram. The rupturing at approximately 150 μs is seen for networks at all densities, although differences can be seen in how many new rings are formed during the rupturing process. For the low density network, only very few entirely new rings are formed during the rupturing process but for the high density network many rings are formed (seen as more points in the top right corner of the birth-death diagram) owing to the increased density of polymers allowing for easier repair and reforming of rings.

3.3 Timestep

Persistence diagrams are useful in that they allow us to identify patterns in data that exist in small regions of interest (usually spatial, here temporal) or large regions of interest; either of these regimes may be most useful for further analysis depending on the topic at hand (Stolz-Pretzer (2019)). As samples of trajectories are taken with finite gaps of time between them and disk space is not infinite, the temporal sampling rate of a trajectory is a controllable variable. Similarly, the sampling rate for our persistence diagram analysis can be some integer fraction of the sampling rate from the trajectory. Analysing networks at different sampling rates can reveal important patterns in the data, for example low-frequency sampling rates can smooth out ephemeral patterns in the data, and high-frequency sampling rates can draw attention to short-lived phenomena. Alternatively, low-frequency sampling rates can save the computational expense of tracking many unique and short-lived rings. Importantly, the similarity between **Figure 7(a-c)** reinforces the idea that high resolution temporal data is not required for persistence diagrams to be a useful analysis tool. Even at the coarsest sampling frequency, key behaviours can still be seen in the persistence diagram and the bars can show which regions of time are interesting enough to study in more detail.

In **Figure 7** we can see the effect of different sampling frequencies. While the overall pattern remains the same, they differ in a few ways that are useful for overall analysis of a trajectory. For example, the same critical points of formation beginning at about 40.0 μs and rupturing at 150 μs exist in all diagrams. While the lines of critical behaviour become clearer with more samples, the fact that they exist at all timesteps reinforces their importance. At a very small timestep, short-lived rings (or absences of rings) can be missed out. A short timestep will count a superset of the rings identified using a large timestep, but include many more rings with lifespans less than the window size Δt. This also means that repair phenomena can be missed if the timestep is too short compared to the timescale of repair. This is especially obvious if we highlight the number of times a given ring is repaired against the timescale, as is shown in **Figure 8**. These show that if the timescale is increased, repair phenomena can be missed entirely; this manifests as the graphs trailing off to show on average one birth per ring and that they are alive continuously over their lifespan for long sampling timescales. However, at lower sampling
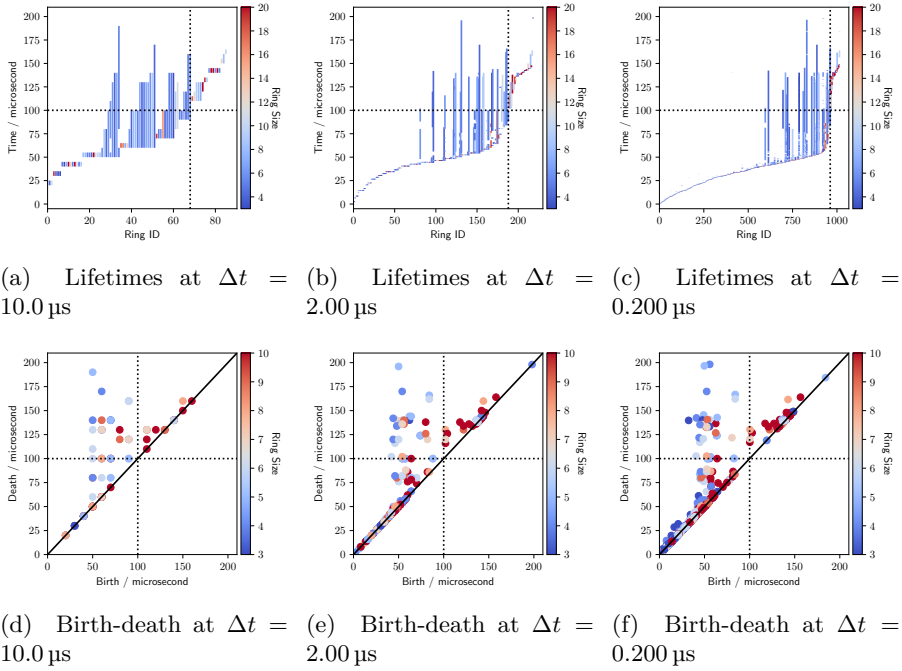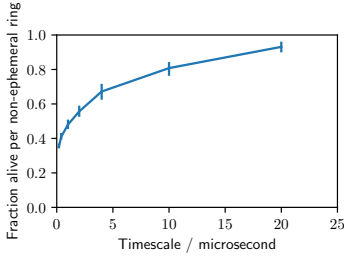
(a)   Lifetimes at Δt = 10.0 μs

(b)   Lifetimes at Δt = 2.00 μs

(c)   Lifetimes at Δt = 0.200 μs

(d)   Birth-death at Δt = 10.0 μs

(e)   Birth-death at Δt = 2.00 μs

(f)   Birth-death at Δt = 0.200 μs

Figure 7: Lifetimes and birth death diagrams for the trajectory previously seen in **Figure 5(a)** and **Figure 5(d)** (which was sampled at $\Delta t = 2.00$ μs)

timescales the repair phenomena become clear as each ring is repeatedly reborn and is only alive for a fraction of its total lifespan. This does not plateau at short timespans, indicating that there are possibly infinitely many microscopic damage and repair phenomena before a ring is destroyed for good as a consequence of the thermal motion of the connecting sites.
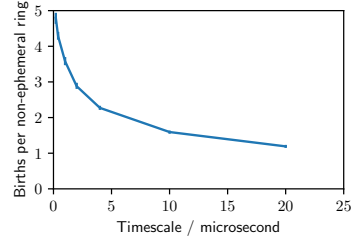
3.4   Combined Trajectories

Finally, we can use the birth-death diagrams to analyse families of simulations and see how critical behaviour emerges from them. The self assembly then rupture simulations were repeated 20 times with different random seeds and the birth death diagrams (as first seen in **Figure 5(a-c)**) for each trajectory were combined. The resulting combined birth-death diagrams are shown in **Figure 9**, and the critical behaviours are now more obvious. We can see in all three subfigures lines just after 100 μs in x and y axes, representing minor damage to rings when the rupturing process begins.

The change in ring formation is also clearer, with many rings being born only after 40.0 μs in **Figure 9(a)**, but starting from 0 μs in **Figure 9(c)**. Next, the different in rupturing behaviour is clearer with a clear line through 150 μs in **Figure 9(c)** for all rings, which is not present in the first two figures. Finally,
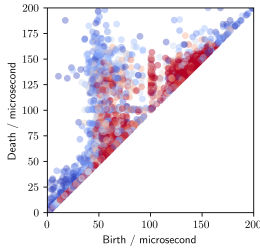
(a) The fraction of their total lifespan that hexagonal rings are alive for in the $A = 46\,100\,\text{nm}^2$ per molecule simulations
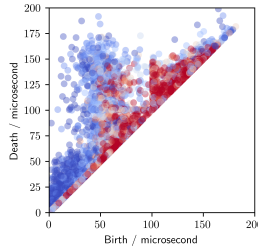
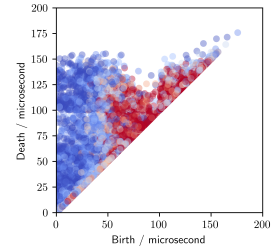(b) The number of times each hexagonal ring is reborn during the $A = 46\,100\,\text{nm}^2$ per molecule simulations

Figure 8: Two metrics of repair, the fraction of their lifespans spent alive and the number of rebirths, for hexagonal rings. These figures only include non-ephemeral rings — *i.e.* rings that are alive in more than one sample. The error bars represent a bootstrapped 95.0 % confidence interval averaged across 20 simulations

(a) $A = 82\,200\,\text{nm}^2$ per molecule

(b) $A = 61\,800\,\text{nm}^2$ per molecule

(c) $A = 46\,100\,\text{nm}^2$ per molecule

Figure 9: Birth-death diagrams at specific formation areas repeated over many trajectories

these diagrams allow us to see with great clarity which rings form at different times of a simulation; the first region of births from 0 µs to 55.0 µs shows that mainly small rings are formed, while larger rings are formed from 55.0 µs to 150 µs and any rings that form after that are small rings in the detritus of the ruptured network. This can be seen in the coloured regions going blue-red-blue from bottom-left to top-right of each figure, with relatively sharp boundaries between each region for simulations with low area per molecule and softer boundaries for simulations with larger area per molecule.

4 Conclusion

In conclusion, we have demonstrated how the tools of persistent homology, especially the persistence diagram and the concept of tracking births and deaths of individual features, can be useful to analyse temporally varying data. The use of temporal persistence diagrams analysing unique rings being born and dying (possibly repeatedly) over the course of a simulation gives us the ability to clearly identify which rings are long-lived and which are ephemeral. It also allows us to observe how medium-range structure (as represented by rings) changes over time, and which topological features are favoured when. The key difference from traditional persistence diagrams, in which time is not a factor, is the fact that given rings can be born, die, and be born again. This is critical for analysis of repair phenomena in biological networks which have, until now, been difficult to visualise and effectively quantify. Critically, the density of the temporal data does not appear to mask significant network changes from being identified. The method employed aids as a lens to analyse the networks, and this provides further insights into how a network changes over time, allowing us to assess the timescale of network behaviours. We anticipate that this analysis can be extended from 2D biological networks to other network formation and damage analysis tasks, such as glass ageing or zeolite formation.

Declarations

4.1 Funding

4.2 Availability of Data and Materials

This paper conforms to the RCUK data management requirements.

4.3 Code Availability

All code used to generate data in this paper is available on GitHub git-hub.com/WilsonGroupOxford under the Network-Monte-Carlo, rings and WormLikeCurve repositories.

4.4 Authors' Contributions

Matthew H. J. Bailey: *Data Curation, Investigation, Software, Visualization, Writing — Original Draft*

Mark Wilson: *Conceptualisation, Project Administration, Supervision, Writing — Review & Editing*

References

Bailey MHJ, Wilson M (2021a) Self assembly of model polymers into biological random networks. Comput Struct Biotechnol J 19:1253–1262. https://doi.org/10.1016/j.csbj.2021.02.001

Bailey MHJ, Wilson M (2021b) Simulation of defects, flexibility and rupture in biopolymer networks. RSC Advances

Bailey MHJ, Ormrod Morley D, Wilson M (2020) Simplified computational model for generating biological networks. RSC Adv 10(63):38,275–38,280. https://doi.org/10.1039/D0RA06205G

Barnard K, Burgess SA, Carter DA, et al (1992) Three-dimensional structure of type IV collagen in the mammalian lens capsule. J Struct Biol 108(1):6–13. https://doi.org/10.1016/1047-8477(92)90002-R

de Berg M (2000) Computational geometry : algorithms and applications, 2nd edn. Springer, Berlin ; London

Burd HJ (2009) A structural constitutive model for the human lens capsule. Biomech Model Mechanobiol 8(3):217–231. https://doi.org/10.1007/s10237-008-0130-5

Chung YM, Hu CS, Lo YL, et al (2021) A Persistent Homology Approach to Heart Rate Variability Analysis With an Application to Sleep-Wake Classification. Front Physiol 12(March):1–18. https://doi.org/10.3389/fphys.2021.637684, https://arxiv.org/abs/arXiv:1908.06856

Dhume RY, Shih ED, Barocas VH (2019) Multiscale model of fatigue of collagen gels. Biomech Model Mechanobiol 18(1):175–187. https://doi.org/10.1007/s10237-018-1075-y

Dong Z, Pu J, Lin H (2021) Multiscale persistent topological descriptor for porous structure retrieval. Comput Aided Geom Des 88:102,004. https://doi.org/10.1016/j.cagd.2021.102004

Duncan G, Michael Wormstone I, Davies PD (1997) The aging human lens: Structure, growth, and physiological behaviour. Br J Ophthalmol 81(10):818–823. https://doi.org/10.1136/bjo.81.10.818

Kovacev-Nikolic V, Bubenik P, Nikolić D, et al (2016) Using persistent homology and dynamical distances to analyze protein binding. Stat Appl Genet Mol Biol 15(1):19–38. https://doi.org/10.1515/sagmb-2015-0057,

https://arxiv.org/abs/arXiv:1412.1394

Kumar A, Sherrington D, Wilson M, et al (2014) Ring statistics of silica bilayers. J Phys Condens Matter 26(39):1–6. https://doi.org/10.1088/0953-8984/26/39/395401

Le Roux S, Jund P (2010) Ring statistics analysis of topological networks: New approach and application to amorphous GeS2 and SiO2 systems. Comput Mater Sci 49(1):70–83. https://doi.org/10.1016/j.commatsci.2010.04.023

Lee Y, Barthel SD, Dłotko P, et al (2017) Quantifying similarity of pore-geometry in nanoporous materials. Nat Commun 8(1):15,396. https://doi.org/10.1038/ncomms15396

Liu X, Wang L, Du C, et al (2015) Mechanism of lens capsular rupture following blunt trauma: a finite element study. Comput Methods Biomech Biomed Engin 18(8):914–921. https://doi.org/10.1080/10255842.2014.975798

Ormrod Morley D (2020) Structural Disorder in Two-Dimensional Network-Forming Materials. Dphil, University of Oxford

Ormrod Morley D, Thorneywork AL, Dullens RPA, et al (2020) Generalized network theory of physical two-dimensional systems. Phys Rev E 101(4):042,309. https://doi.org/10.1103/PhysRevE.101.042309

Ormrod Morley D, Salmon PS, Wilson M (2021) Persistent homology in two-dimensional atomic networks. J Chem Phys 154(12). https://doi.org/10.1063/5.0040393

Pereira CM, De Mello RF (2015) Persistent homology for time series and spatial data clustering. Expert Syst Appl 42(15-16):6026–6038. https://doi.org/10.1016/j.eswa.2015.04.010

Sadjadi M, Thorpe MF (2016) Ring correlations in random networks. Phys Rev E 94(6). https://doi.org/10.1103/PhysRevE.94.062304, https://arxiv.org/abs/arXiv:1607.06568

Sørensen SS, Biscio CA, Bauchy M, et al (2020) Revealing Hidden Medium-range Order in Amorphous Materials using Topological Data Analysis. Sci Adv 6(37). https://doi.org/10.1126/sciadv.abc2320

Steinberg L (2019) Topological Data Analysis and its Application to Chemical Systems. Phd, University of Southampton

Stolz-Pretzer BJ (2019) Global and Local Persistent Homology for the Shape and Classification of Biological Data. Dphil thesis, University of Oxford

Wasserman L (2018) Topological Data Analysis. Annu Rev Stat Its Appl 5(1):501–532. https://doi.org/10.1146/annurev-statistics-031017-100045

Xia K, Zhao Z, Wei GW (2015) Multiresolution persistent homology for excessively large biomolecular datasets. J Chem Phys 143(13). https://doi.org/10.1063/1.4931733

Yurchenco PD, Furthmayr H (1984) Self-assembly of basement membrane collagen. Biochemistry 23(8):1839–1850. https://doi.org/10.1021/bi00303a040

Yurchenco PD, Ruben GC (1987) Basement membrane structure in situ: evidence for lateral associations in the type IV collagen network. J Cell Biol 105(6):2559–2568. https://doi.org/10.1083/jcb.105.6.2559