



A topological perspective on weather regimes

Kristian Strommen¹ · Matthew Chantry¹ · Joshua Dorrington¹ · Nina Otter²

Received: 4 September 2021 / Accepted: 20 June 2022 / Published online: 6 July 2022
© The Author(s) 2022

Abstract

It has long been suggested that the mid-latitude atmospheric circulation possesses what has come to be known as ‘weather regimes’, loosely categorised as regions of phase space with above-average density and/or extended persistence. Their existence and behaviour has been extensively studied in meteorology and climate science, due to their potential for drastically simplifying the complex and chaotic mid-latitude dynamics. Several well-known, simple non-linear dynamical systems have been used as toy-models of the atmosphere in order to understand and exemplify such regime behaviour. Nevertheless, no agreed-upon and clear-cut definition of a ‘regime’ exists in the literature, and unambiguously detecting their existence in the atmospheric circulation is stymied by the high dimensionality of the system. We argue here for an approach which equates the existence of regimes in a dynamical system with the existence of non-trivial topological structure of the system’s attractor. We show using persistent homology, an algorithmic tool in topological data analysis, that this approach is computationally tractable, practically informative, and identifies the relevant regime structure across a range of examples.

1 Introduction

The discovery of the ‘butterfly effect’ (Lorenz 1963) effectively ended the idea that weather forecasting can be understood purely as the problem of integrating a deterministic system forward in time. Instead, the problem of accurate weather forecasting becomes one of determining, from a given initial state, the likely trajectories of the atmosphere on its underlying attractor (Slingo and Palmer 2011). Similarly, the problem of producing reliable climate projections can be understood as determining how, and to what extent, the likelihood of traversing different trajectories changes in the presence of an external forcing (Corti et al. 1999; Palmer 1999; Woollings et al. 2010b). As such, it becomes natural to

ask whether the climate attractor exhibits significant deviations from Gaussianity, since such deviations, even locally, may strongly constrain the available trajectories. In other words, understanding the ‘shape’ of the attractor becomes a problem of great practical importance.

The study of local non-Gaussianity in the atmosphere has classically been done under the guise of so-called weather (or circulation) regimes (Vautard 1990; Michelangeli et al. 1995; Corti et al. 1999; Lorenz 2006; Hannachi et al. 2017). The basic idea¹ is to determine a small number of dynamically relevant large-scale flow patterns (the regimes) that dominate the low-frequency variability and transition from one to another in an approximately Markovian manner (Baur 1951). We note that the use of the word ‘regime’ in this paper should not be confused with usages associated with transitions induced by changes in the system’s parameters, i.e., bifurcations. This concept of atmospheric weather regimes has been most famously applied to study the circulation of the wintertime Euro-Atlantic sector. Here, the dominant pattern of variability is a dipole pattern of pressure anomalies referred to as the North Atlantic Oscillation (NAO) (Hurrell et al. 2003), shown in Fig. 1. Knowing whether the NAO is in its positive or negative phase gives a good first-order approximation of winter weather in Europe

✉ Kristian Strommen
kristian.strommen@physics.ox.ac.uk

Matthew Chantry
matthew.chantry@physics.ox.ac.uk

Joshua Dorrington
joshua.dorrington@physics.ox.ac.uk

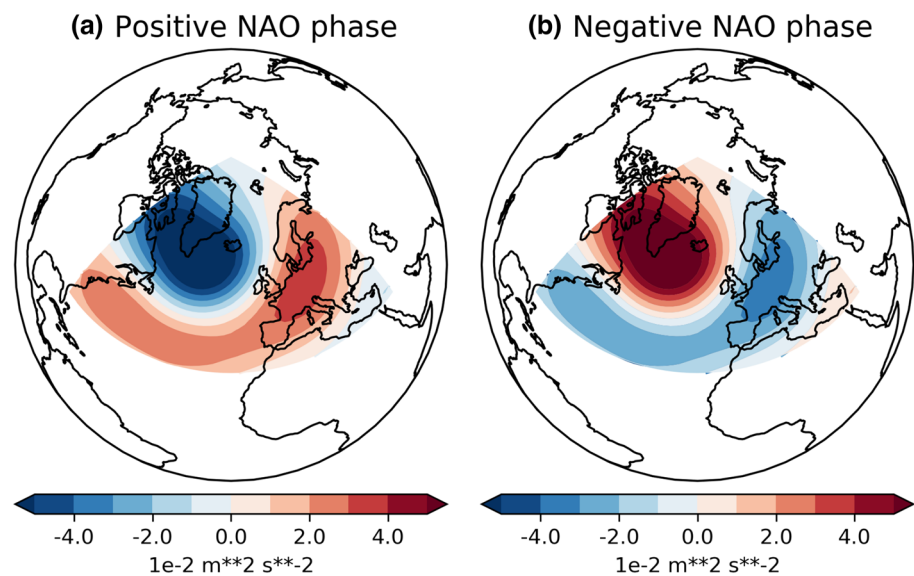
Nina Otter
n.otter@qmul.ac.uk

¹ Department of Physics, University of Oxford, Parks Rd, Oxford OX1 3PJ, UK

² Department of Mathematics, UCLA, Los Angeles, CA 90095, USA

¹ The overview given here has been somewhat simplified. For example, the Markovian assumption can be relaxed (Franzke et al. 2009; O’Kane et al. 2013), and regime systems can in principle give rise to Gaussian statistics (Majda et al. 2006).

Fig. 1 The positive (a) and negative (b) phases of the North Atlantic Oscillation (NAO), the dominant pattern of variability in the Euro-Atlantic wintertime circulation. The pattern is a dipole of atmospheric pressure anomalies, measured here as the first empirical orthogonal function of geopotential height at 500 hPa. The two phases can be viewed as one of several possible ways to decompose the Euro-Atlantic circulation into distinct regimes



and eastern North America, and is one of several possible regime views of the Euro-Atlantic circulation. Strong NAO events are frequently linked to extreme surface weather and increased predictability, such as the European winter of 2019/2020, the warmest on record to date (Hardiman et al. 2020). Simple dynamical systems such as the Lorenz '63 system (Lorenz 1963), the Charney–deVore (Charney and DeVore 1979) and Lorenz '96 systems (Lorenz 1996) have frequently been used to illustrate or study the use of such atmospheric weather regimes (Palmer 1999; Charney and DeVore 1979; Christensen et al. 2015).

However, despite being studied since the 1950's (Baur 1951), no clear-cut and generally accepted definition of a regime exists. Most definitions found in existing studies, often stated only implicitly, are based either on density considerations, where a regime corresponds to a region of above-average density in phase space ('clusters') (Stephenson et al. 2004; Vautard 1990; Straus 2010), or some temporal persistence criteria, whereby a regime is a dynamical phenomenon (e.g., a blocking anticyclone) with a clear lifecycle and a lifespan exceeding some prescribed threshold (Mo and Ghil 1987; Lorenz 2006; Franzke et al. 2008); some studies also combine the two (Grams et al. 2017; Falkena et al. 2020). A comprehensive overview of many of these regime methodologies can be found in Hannachi et al. (2017). When applied to the Euro-Atlantic circulation, these approaches typically produce anywhere between 2 and 7 regimes which are not, in general, easily comparable with each other. The resulting ambiguity has led some to question whether meaningful weather regimes really exist at all (Stephenson et al. 2004; Christiansen 2007; Fereday 2017).

Besides the lack of agreement in the definition of regimes, existing approaches suffer from two key technical problems. Firstly, the algorithms involved often require essentially

ad-hoc choices up front, such as the choice of cluster number in *K*-means clustering algorithms, or temporal persistence thresholds, which directly influence the output regimes. In such cases, an 'optimal' number of regimes is often determined post hoc as the number which maximises a chosen metric, such as the Bayesian information criterion (Falkena et al. 2020). This adds an additional layer of complexity to analysis, since different choices of the validating metric may give different maxima and it can be hard to motivate the choice of one metric over another (Christiansen 2007; Dorington and Strommen 2020). Secondly, and more seriously, the algorithms involved typically scale very poorly with the dimension of the data set, the so-called 'curse of dimensionality' (Radovanovic et al. 2010). This is particularly the case for more technical approaches based on computing exact solutions of the flow (e.g., fixed-points and periodic orbits), approaches which might be thought to offer potentially more robust candidate definitions of a regime. In a state-of-the-art application of these concepts to atmospheric modelling, Lucarini and Gritsun (2020) identified unstable periodic orbits (UPOs) corresponding to zonal and blocking events in a low resolution quasi-geostrophic model, considerably extending earlier work by Itoh and Kimoto (1996). However, the dimensionality of this model still sits well below that of weather and climate models, let alone the physical system itself. Since Majda et al. (2006) showed that the fixed points of a truncated system do not necessarily correspond to the regimes of the untruncated system, this represents a serious obstacle towards using such techniques to define regimes in atmospheric data. More problematically perhaps, UPOs and their stability are model features, limited by the accuracy of their associated models and such analysis cannot be directly applied to observational data.

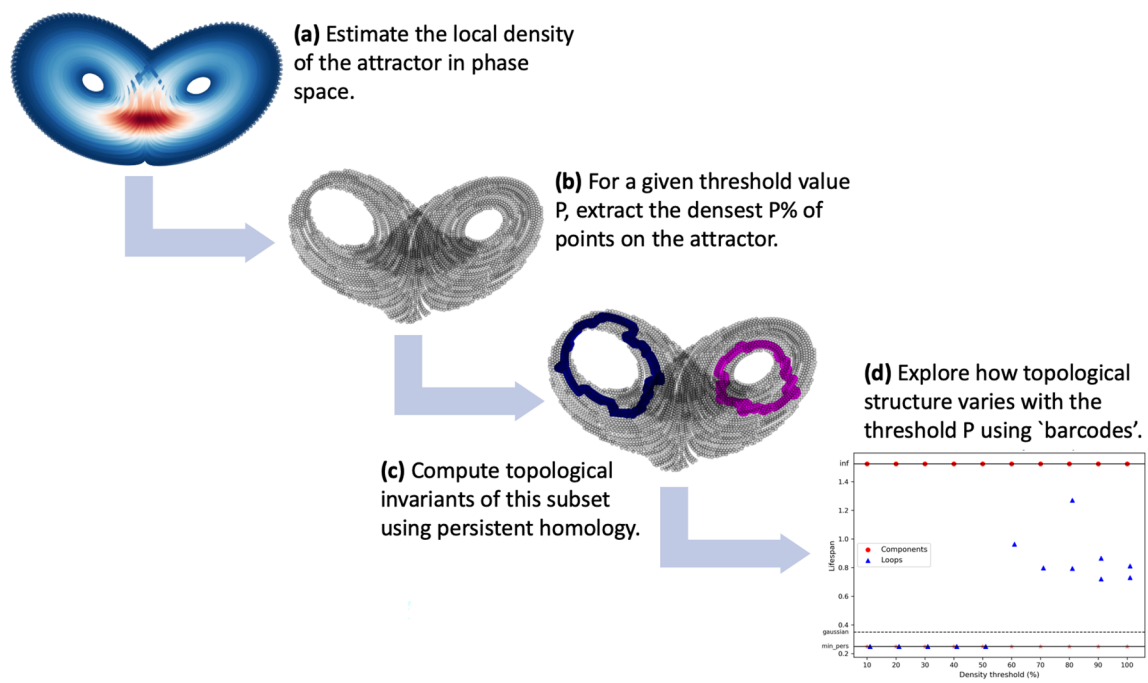


Fig. 2 A schematic of the pipeline we use to analyse topological features of dynamical systems, here exemplified using the Lorenz '63 system. In step (b), the picture shows the densest 80% of points from the attractor in (a). In step (c), the topological invariants are in this

case the two loops, identified and visualised here using persistent homology algorithms, and the one connected component made up of all the points

We posit that the 'curse of dimensionality' represents a serious obstacle to a more robust exploration of regimes in the atmospheric circulation, where the dimensionality of the phase space is orders of magnitudes larger than the number of measurements. Classical regime approaches circumvent the curse primarily by only considering a single atmospheric variable, such as pressure or zonal winds, and will often reduce the dimensionality even further using empirical orthogonal function (EOF) decompositions. Such low-dimensional projections of the circulation generally produce data sets where deviations from Gaussianity, or metastability in the case of Hidden Markov Model approaches, are extremely subtle and often not statistically significant (Stephenson et al. 2004; Franzke et al. 2008; Dorrington and Strommen 2020). On the other hand, several studies suggest that, for example, the multimodal behaviour observed in the North Atlantic jet involves not just changes to zonal winds, but the complex interplay between winds, pressure and temperatures (Novak et al. 2015, 2017). We therefore consider it plausible that truly robust and unambiguous regime detection may require consideration of multiple dimensions of data encoding several atmospheric variables at once. In other words, the regime structure, which appears highly subtle in low-dimensional projections, may be considerably more blatant in higher dimensions.

We propose that the emergence of the field of topological data analysis (Carlsson 2008) offers a new perspective on regimes, by bringing the attention away from specific dynamical properties, such as density and temporal persistence, and back to a more general consideration of the 'shape', i.e., the *topology*, of dynamical systems. Persistent homology (Otter et al. 2017) is a technique in topological data analysis that gives a principled way of studying the shape of datasets such as point clouds, digital images or networks. Given a point cloud, such as the one in Fig. 2b, one associates to it a nested sequence of spaces (called a 'filtration'), which are obtained by for instance thickening each point with a ball of radius ϵ and then allowing ϵ to vary. Persistent homology then gives a way to summarise how different types of topological invariants describing the shape, such as the number of components of the thickened point cloud, or the number of holes, evolve across the filtration (i.e., when varying ϵ). We provide an informal overview of persistent homology in Sect. 3.1, and a more technical, mathematical definition in Appendix A.

Many iconic topological features of dynamical systems, such as the two holes in the Lorenz '63 system (cf. Fig. 10), can effectively vanish in the limit of infinitely many points, thus in our work it turns out to be crucial to augment the standard filtration by taking into account not only the distance between points, but also density. To do this, we

associate to any dynamical system, a *bifiltration* of spaces, by considering nested sequences of spaces given by varying both the density of points as well as distances between them: see Fig. 8. We then proceed to study non-trivial topological features of the attractor, such as the number of connected components (i.e., well-separated regions of points), and holes (e.g., as in Lorenz '63). This bifiltration thereby gives a way of measuring the 'regime structure' in a robust and computationally tractable way.

We argue that this topological perspective, comes with two decisive advantages. Firstly, it provides a natural unifying framework for understanding several disparate regime systems, thereby offering a potential resolution to the ambiguity around lack of definitions. Secondly, its implementation in algorithms largely sidesteps the aforementioned key technical problems. Crucially, since homological computations do not essentially depend on the dimensionality of the data set being studied, our method avoids the 'curse of dimensionality'² and is therefore particularly well suited to the analysis of high-dimensional climate data. To justify these assertions, we will present a highly flexible methodology, depicted in Fig. 2, for detecting non-trivial topological structure in an arbitrary dynamical system of potentially very high dimensionality. By applying this methodology to a number of classic examples of regime systems (Lorenz '63, Lorenz '96, Charney–deVore and the North Atlantic eddy-driven jet), we show that their non-trivial topology is a clear unifying feature, and that the particular topological structure associated with these systems captures well their most familiar features. By contrast, any simplistic definition of a regime based on density and temporal persistence invariably fails to account for one or more of these systems. This suggests that regimes might be best understood as the results of varied attempts to capture the non-trivial topology of the underlying attractor. In particular, non-trivial topological invariants can be viewed as indicating the existence of regimes.

It is worth being clear up front that our method does not add further clarity to the analysis of atmospheric data sets that are approximately Gaussian. Rather, the anticipated benefit of our method is in its potential application to very high dimensional atmospheric data, where structure may be more clearly non-Gaussian. The fact that our method correctly identifies the relevant structure in four well-studied cases means such exploratory applications to high dimensional data can be carried out with more confidence. In particular, our methodology paves the way for a more comprehensive exploration of the 'shape' of the atmospheric circulation.

The potential of persistent homology as a tool for analysing dynamical systems was first suggested, among others, in Maletić et al. (2016), in which it was demonstrated that persistent homology can locate the holes of the Lorenz '63 system: see also Charó et al. (2021). Of particular relevance is the recent work of Yalınız and Budanur (2020), which uses persistent homology and UPOs in order to obtain a simplified representation of chaotic dynamical systems, an approach similar in spirit to our paper. Two recent applications of persistent homology to the real atmosphere are Muszynski et al. (2019), which studies 'atmospheric rivers' using a combination of homology and machine learning, and Tymochko et al. (2020), which used persistent homology to quantify a diurnal cycle in hurricanes. For a different application of topological ideas to ocean modeling, we can also recommend Stanley (2019). There are several additional lines of work, that have used methods from topology to study dynamical systems, such as Khasawneh and Munch (2016), and Kramár et al. (2016), to cite a few.

Finally, a cautionary note on language. The use of the word 'persistent' in 'persistent homology' comes from the way in which topological features that persist for a certain number of filtration values are considered to give meaningful information. In particular, there is no obvious relationship with the temporal persistence of regime states. To avoid ambiguity, in this manuscript the word 'persistence' will always refer to topological persistence, while when referring to temporal persistence of regime states, we will make use of the qualifier 'temporal'.

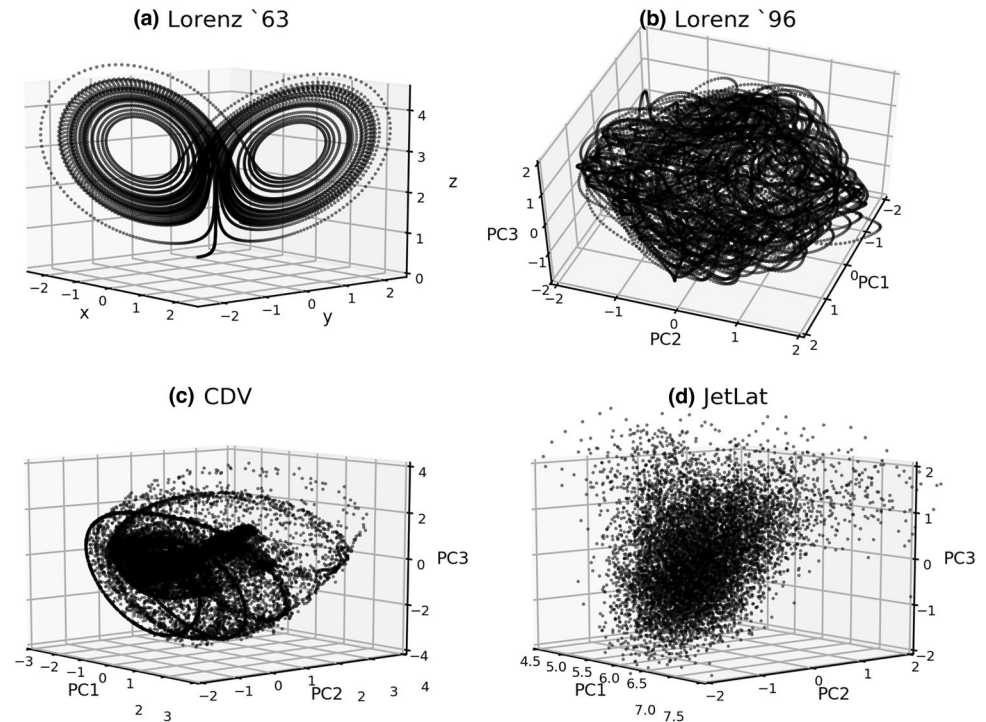
The paper is structured as follows. In Sect. 2 we provide details of the dynamical systems used, including the observational atmospheric data. In Sect. 3, we provide an informal introduction to persistent homology, and we motivate the need for a bifiltration. The formal, mathematical definitions of the concepts underlying persistent homology are included in Appendix A; readers willing to treat these formalities as a 'black box' should not find their understanding of the paper otherwise compromised. In Sect. 4 we detail the algorithmic procedure used to compute topological metrics: this section can be skipped without loss of continuity or basic understanding. The results of applying this methodology to our suite of data sets is shown and discussed in Sects. 5 and 6, with conclusions and future directions in Sect. 7.

2 Data

We will be making use of four data sets: three data sets generated using toy-models of the atmospheric circulation and one data set using actual atmospheric data. All the data sets used have been frequently studied as examples of regime systems, but the dynamics are strikingly different in each case. A key goal of this paper is to show that our

² As will be seen, our preprocessing includes a step which suffers from the curse, but the topological algorithms themselves do not. Potential improvements to our preprocessing will be discussed.

Fig. 3 Visualisations of the four regime systems considered in their ambient phase space. In **a** Lorenz '63, **b** Lorenz '96, **c** Charney–deVore (CDV), **d** the North Atlantic eddy-driven jet (JetLat). See main text for details



methodology gives the ‘right answer’ for these known cases. In this section we give a brief and informal overview of the data sets we use: more detailed descriptions of the three toy-models used, including defining equations, can be found in Appendix B.

2.1 Toy-models

We describe the toy-models in increasing order of dimensionality.

The first toy-model we consider is also the most well known, namely the Lorenz '63 system. First introduced and studied in Lorenz (1963), it is a chaotic dynamical system in three variables, essentially derived as a highly simplified model of convection. It has also been re-derived as a toy-model of the NAO (Molteni and Kucharski 2019). The attractor, visualised in Fig. 3a, famously resembles a butterfly, and is usually viewed as having two regimes corresponding to the two ‘wings’. Its regime behaviour has been extensively studied (Palmer 1994; Corti et al. 1999; Yadav et al. 2005). An interactive simulation showing the evolution of this system can be found at https://joshdorrington.github.io/L63_simulator/. Simply click on any point on the attractor to animate a trajectory initiated at that point.

The second toy-model considered is the Charney–deVore (CDV) model, and is defined using six variables. It was derived in Charney and DeVore (1979) based on a severe spectral truncation of the barotropic vorticity equation in a β -plane channel, and can be thought of as a crude model of

large-scale midlatitude blocking dynamics. It exhibits multimodality, typically interpreted in terms of two regimes corresponding to blocked and zonal flow. The attractor is visualised in Fig. 3c. An interactive simulation can be found at https://joshdorrington.github.io/cdv_simulator/. Simply click on any point on the attractor to animate a trajectory initiated at that point. The nature of the two regimes are discussed further in Sect. B.2.

Finally, we consider the Lorenz '96 model. It was introduced in Lorenz (1996) as an idealized, chaotic model of the atmosphere which is of greater complexity than the Lorenz '63 system (Karimi and Paul 2010). It is defined in our case by coupling eight variables $X_k, k = 1, \dots, 8$, interpreted as large-scale modes of variability, with 32 variables $Y_j, j = 1, \dots, 32$, representing small-scale modes of variability. Due to its interpretation of large-to-small scale coupling, Lorenz '96 has been utilised in several studies looking at different ways to parameterise unresolved sub-grid scale variability in forecast systems (Wilks 2005; Christensen et al. 2015; Vissio and Lucarini 2018; Gagne et al. 2020). Its regime structure has been considered in, e.g., Lorenz (2006) and Christensen et al. (2015), who both viewed it as having two distinct regimes: these are discussed further in Sect. 5.3. A three dimensional projection of the space to its first three EOFs is shown in Fig. 3b. A considerably more illuminating visualisation of the complex behaviour of this system, with the eight ‘large-scale modes’ suppressed, can be viewed at <http://youtu.be/rYnkHory39o>, which animates multiple projections of the 32 dimensional space onto 4 randomly

chosen dimensions. The behaviour is qualitatively similar when large-scale modes are included.

In order to visualise the results we obtain, it is necessary to pick a three-dimensional projection in the case of CDV and Lorenz '96. We use the first three EOFs in each case, which explain around 98% of the variance for CDV and around 80% for Lorenz '96. Furthermore, we found that our topological method produces qualitatively identical results when using the first three EOFs for CDV (as opposed to all six variables). For Lorenz '96, the work in Christensen et al. (2015) showed that the regime variability is concentrated in the first four EOFs, and here we found qualitatively identical results applying our method to this four-dimensional truncation. To maintain maximal consistency between our computations and our visualisations, we therefore present results obtained using these truncations. However, we stress that these truncations are carried out for visualisation purposes only: our methodology works equally well using the full, raw data sets in each case.

2.2 Observational data: the North Atlantic jet

To represent real atmospheric data, we make use of so-called reanalysis data. Actual observational data, whether from stations or satellite, are always unevenly distributed in time and space and therefore contain gaps. Reanalysis data fills in these gaps by blending observations with short-range weather forecasts using data assimilation methods. Here we make use of ERA20C (Poli et al. 2016), which covers the period 1900–2010, motivated by the desire to have as much data as possible. Because the period prior to 1979 suffers from a lack of satellite data, we validated our results using ERA-Interim (Dee et al. 2011), which covers the period 1979–2015. ERA-Interim data was found to produce qualitatively similar results to ERA20C, and so is not shown.

The general suitability of ERA20C for regime-based studies has been commented on in previous studies (Parker et al. 2019; Strommen 2020), and essentially relies on the fact that there is a long and consistent record of surface observations in the Euro-Atlantic sector, which will be our area of interest. The existence and properties of regimes in the winter-time Euro-Atlantic circulation has been extensively studied, either through the prism of pressure fields, typically geopotential height at 500 hPa, or winds, in the form of zonal winds at 850 hPa (hereafter ua850). Studies based on pressure data (Vautard 1990; Michelangeli et al. 1995; Dawson et al. 2012; Dorrington and Strommen 2020; Falkena et al. 2020) typically use clustering algorithms to classify distinct regimes. On the other hand, wind data is usually processed more directly in order to capture the variability of the North Atlantic eddy-driven jetstream, a relatively coherent stream of zonal winds. By measuring the location of the maximum wind-speed of the jet, one can define the latitude of the jet

on any given day: the histogram of this jet-latitude index is visibly and robustly trimodal, suggesting the existence of three distinct regimes (Woollings et al. 2010a). The differences between these two perspectives, which would a priori be expected to be equivalent, can be reconciled by taking into account the added variability coming from the speed of the jet, after which both pressure and wind data suggest three very robust regimes (Madonna et al. 2017; Strommen 2020). Applications to predictability have been studied in both contexts, see, e.g., Cassou (2008) and Strommen (2020).

In this paper we will be focused on seeing how our framework views these three jet regimes, and so define a data set we will refer to as 'JetLat'. This will be a 3-dimensional data set consisting of the daily jet latitude, and the daily values of the first and second principal components of ua850 anomalies. Data is always restricted to the North Atlantic region, defined by 15N–75N, 300E–360E, and the winter season December–January–February (DJF). The jet latitude was computed using the methodology of Parker et al. (2019), which also includes a discussion of the jet in ERA20C. The JetLat data set is visualised in Fig. 3d.

We note again that the deliberate choice to use a data set which explicitly contains the jet latitude, already known to be multimodal, is motivated by the desire to validate our methodology against known regime systems before applying it to less well-understood contexts. The question of locating these jet regimes using unprocessed data (i.e., data not containing any prior knowledge of the jet-latitude index) will be discussed in the conclusions, in Sect. 7. The results of applying our methodology to pressure data will be discussed in Sect. 5.

3 Persistent homology for dynamical systems

Over the last 20 years, methods from the mathematical area of topology have been increasingly used to study data analysis problems. In this section we discuss how some of these methods can be used to study dynamical systems.

In topology one is interested in studying properties of shapes that do not change when one continuously deforms the shape, for instance when one squeezes or bends it. If one considers an annulus as in Fig. 4, then no matter how the annulus is bent or stretched, it will still be composed of one piece, and have one loop. One says that the number of pieces and loops of a shape are topological invariants. On the other hand, deformations that are not allowed include cutting or gluing. If one was to cut the annulus in half, as illustrated in Fig. 4, one would break it into separate pieces with no loops. One can think of these invariants as giving a very coarse description of the shape of a space or of data.

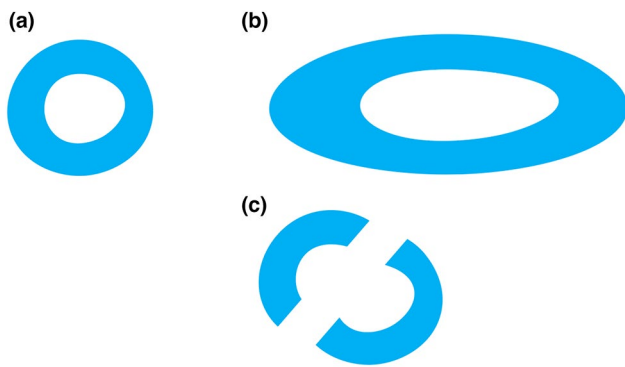


Fig. 4 **a** An annulus, and **b** a shape obtained by continuously deforming the annulus, which has the same number of pieces (components) and loops as the annulus. **c** A space obtained from the annulus by a deformation that is not continuous, and thus with a different number of pieces and loops

In particular, topological invariants often do not depend on the choice of of parametrisation, coordinates, or ambient dimension, and thus they are independent of many choices introduced during preprocessing steps. This is an aspect that is crucial in our work. We caution the reader that the interpretation of the topological invariants, thus what information they capture, and whether they are coarse or not, is context-specific, and depends on the specific application.

There are different ways to use topology to study data, see the survey Carlsson (2008) for an overview. Persistent homology, which is the method that we use in our work, is one of the standard techniques and has been very successful in many applications. In the remaining part of this section we first introduce persistent homology, and we then explain how it can be used to study the specific types of dynamical systems that we study in our work. Here we provide an informal description, while we provide rigorous definitions in Appendix A. Readers keen for even more extensive background may consult Otter et al. (2017) and references therein.

3.1 Persistent homology: informal overview

Given experimental data composed of points or vectors representing measurements, together with a measure of similarity (e.g., given by proximity, or correlation), in persistent homology one considers a thickening of the data set at increasing similarity scales, see Fig. 5 for an example. This process yields a nested sequence of increasingly thickened spaces, which are collectively called a ‘filtration’. One then analyses the evolution (so-called ‘persistence’) of the number of components, holes (or, equivalently, loops), voids, and higher-dimensional holes

(which we call ‘topological features’) across the filtration. The information captured by the filtration is therefore the information associated with continuously varying a free parameter (in this case the measure of similarity).

The barcode is an algebraic invariant that summarises how the topological features evolve across the filtration: the left endpoint of an interval in the barcode (the horizontal lines in Fig. 5c, d) represents the birth of a feature (the smallest distance value at which a component or hole appears in the filtration), while its right endpoint, roughly, represents the death of the same feature (the smallest distance value at which two components merge or a hole is filled in): the difference between the death and birth is referred to as the lifetime of the feature. When a feature is still ‘alive’ at the largest thickening scale that one considers, the lifetime interval is by convention set as an infinite interval. For instance, we can read off from Fig. 5c that there are two components that have significantly longer lifetimes than the others (corresponding to the cluster of points forming a figure-eight on the left of the figure, and the cluster of remaining points on the right), while from Fig. 5d we can infer that there are two holes that live much longer than the others, which correspond to the two holes in the figure-eight cluster. We provide a rigorous definition of holes and barcodes in Appendix A.

The interpretation of the intervals in the barcode that we have given here is only one of the possible applications of persistent homology to the study of data. In other types of applications, it might be the intervals of a certain length, and not necessarily the longest ones, that encode significant information, see for instance, Bendich et al. (2016); Bubenik et al. (2020). In particular, the interpretation of the barcode is application specific. We discuss how we interpret the barcode in our work in more detail in Sect. 4.4.

3.2 Computational complexity of PH

The theory behind (one-parameter) persistent homology is well-understood, and amounts to standard linear algebra. Conversely, the computation of the barcode is expensive, since the computational complexity can grow exponentially in the size of the input data, in the worst case. To sidestep such difficulties, in this work we use optimised algorithms, and sparsification techniques, see also Sect. 4.2. We refer the reader to the survey Otter et al. (2017) for a detailed discussion of the computational complexity of the main persistent homology algorithms.

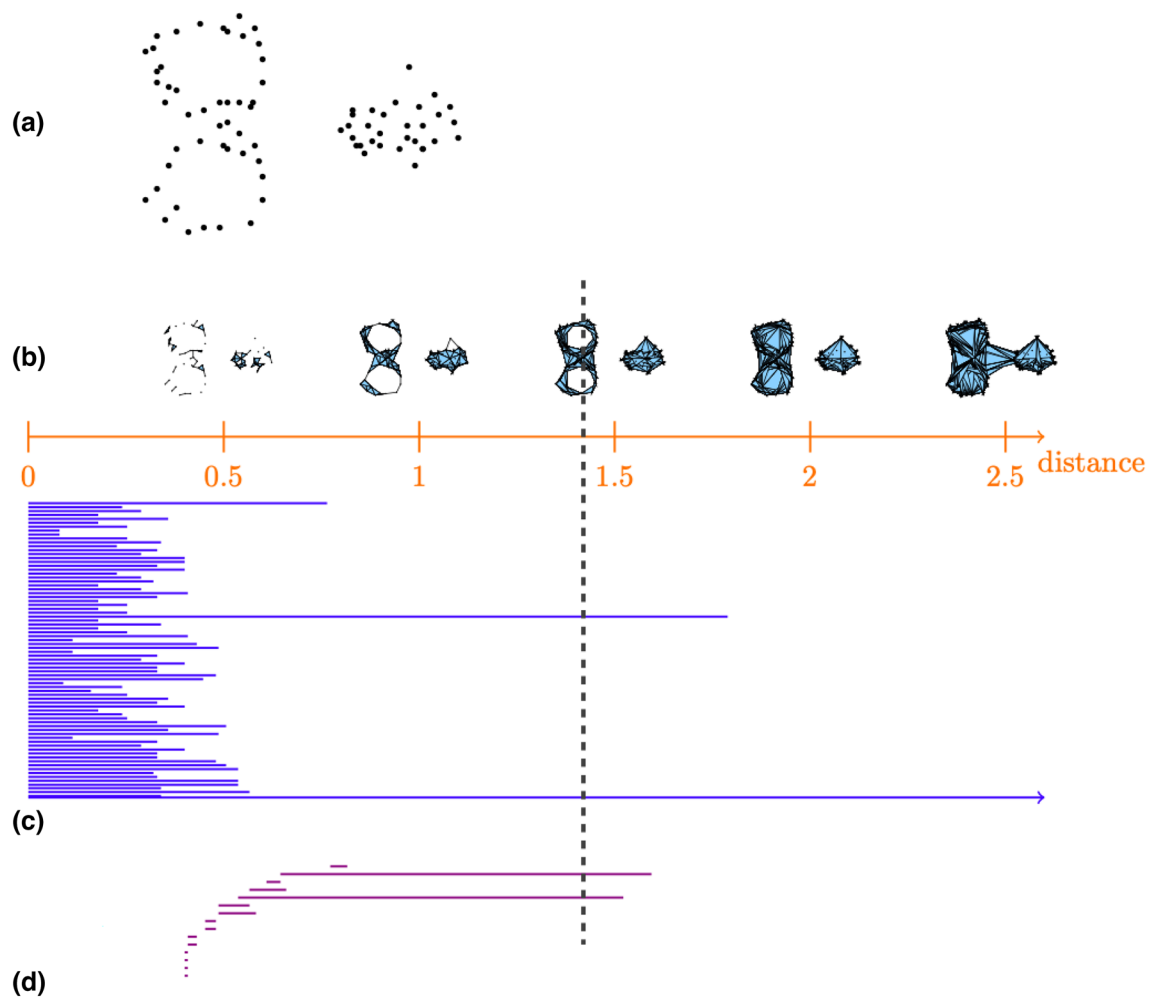


Fig. 5 **a** A set of points lying on a plane, with similarity given by proximity in the Euclidean distance. **b** A filtration of nested spaces, called a ‘Vietoris–Rips complex’, obtained by connecting points within a certain distance by an edge, and filling in resulting triangles. Barcodes describing how long the **c** components and **d** holes persist in the filtration. We illustrate using a gray vertical dashed line, how

we can read off information from the barcodes: at filtration value 1.4 there are two components, since the dashed gray line intersects two blue intervals in the barcode corresponding to the components, as well as two holes, since the gray line intersects two purple intervals in the barcode for the holes

The types of filtered spaces that we consider here rely on the computation of distances between points, and such distances have computational cost that is, in the worst case, linear in the dimension of the ambient space. Once the distances are computed, the computational complexity of persistent homology of these filtered spaces depends on the length of the input data, and, thus, on the number of samples in time, but not on the dimension of the ambient space.³ Thus, a consequence that this has for our work is that adding variables to our models increases

the computational cost only by a linear function of the dimension of the ambient space. This is one of the reasons that persistent homology is so effective for the study of dynamical systems, especially in atmospheric science where spaces are heavily undersampled and the dimensionality of the phase space is often orders of magnitudes larger than the number of measurements.

Experts in persistent homology will know that persistent homology computations scale badly when increasing the homological dimension. Concretely, the computational complexity⁴ is $\mathcal{O}(n^{k+1})$ in the worst case, where n is the number

³ That is, it doesn’t matter if the data set consists of, e.g., 100 monthly samples of atmospheric variables (x_1, x_2, x_3) , or 100 monthly samples of $(x_1, x_2, x_3, x_4, x_5, x_6)$.

⁴ For persistent homology this complexity is dominated by the building of the filtered Vietoris–Rips complex.

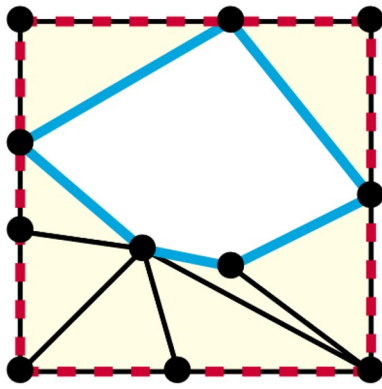


Fig. 6 Two 1-cycles that are representatives for the same hole: (cyan, solid) an optimal (i.e., minimal) representative, and (red, dashed) a not optimal one

of input points and k the homological dimension that we compute. In this work we focus on the study of components and loops ($k = 0, 1$), so we are unaffected by this exponential scaling. Higher-dimensional homology corresponds to higher dimensional analogues of loops, which for dimension 2 can be interpreted as a ‘void’ (e.g., the void enclosed by a sphere) and for dimensions ≥ 3 as higher-dimensional voids. We currently do not have obvious interpretations for what voids or higher-dimensional voids would mean for the types of systems that we study.

3.3 Optimal representatives of cycles

Given an interval in the barcode describing the lifetime of a component or hole, we are interested in studying the points in the data that correspond to such a component or hole. Such points are called ‘representatives’ for holes in dimension 0 (i.e., components) and in dimension 1 (i.e., holes). We refer the reader to Appendix A, and in particular Definition A.7 and the preceding paragraph for a definition of p -dimensional holes and p -cycles.

Ideally, we want to be able to choose representatives that are easily interpretable from a geometric point of view. For instance, we might want representatives for 1-dimensional holes to have minimal length in a suitable sense, see the illustration in Fig. 6.

Thus, we are interested in representatives that satisfy some minimality condition: for holes we compute optimal representatives (Dey et al. 2018) using the software Persloop (Jyamiti Research Group 2017), while for components, we use representatives to find all the points in a component. We note that finding optimal representatives for holes is a challenging problem; the software Persloop implements an algorithm that gives a heuristic

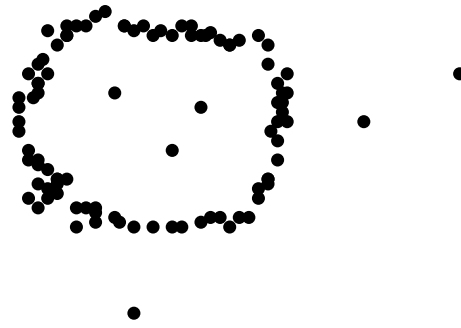


Fig. 7 A finite set of points in \mathbb{R}^2 for which a distance-based filtrations might fail to capture interesting topological information

approximation for 1-cycles in 3D, but which might fail to give meaningful 1-cycles on higher dimensional data sets.

3.4 Multiparameter persistent homology

In many application problems, one might wish to study filtrations that depend on more than one parameter. For instance, consider the point cloud in \mathbb{R}^2 in Fig. 7. If one were to consider only the points belonging to higher-density regions, one could associate to these points a distance-based filtration, as illustrated in Fig. 5 and discussed in Sect. 3.1. Then, by computing the persistent homology of such a filtration, one could read off from the barcodes that the point cloud has a long-lived component, and a long-lived hole. For such a data set, it might be difficult in practice to choose the right density value, and therefore one would ideally wish to consider point clouds thresholded at all possible density values, thus obtaining a bifiltration, as illustrated in Fig. 8. Note that in the same way that a filtration can be thought of as a one-parameter family, a bifiltration is simply a two-parameter family, keeping track of the information associated with continuously varying the two free parameters (in this case the distance and density values).

The theory of persistent homology does not generalise to filtrations that depend on more than one parameter. In particular, there is no generalisation of the barcode, as described in Sect. 3.1 and illustrated in Fig. 5, for multifiltrations. Finding appropriate ways to quantify the ‘persistence’ of topological invariants, such as the number of components or holes, is currently one of the most active areas of research in TDA, and several researchers have proposed invariants that are computable, and capture in an appropriate sense what it means for topological features to be ‘persistent’, see, for instance, Harrington et al. (2019), Lesnick and Wright (2015), Vipond (2020). In Sect. 3.4.1 we discuss one such approach.

Fig. 8 A bifiltration obtained by decreasing density and increasing distance: given the finite set X of points in \mathbb{R}^2 illustrated in Fig. 7, and a density estimation, we consider subsets $X' \subset X$ of points having density above a certain threshold. For each subset X' of points we then construct a distance-based filtration by taking balls with increasing radii centered at the points. The yellow shading visualises the effect of placing a yellow disk of a given distance value around each point

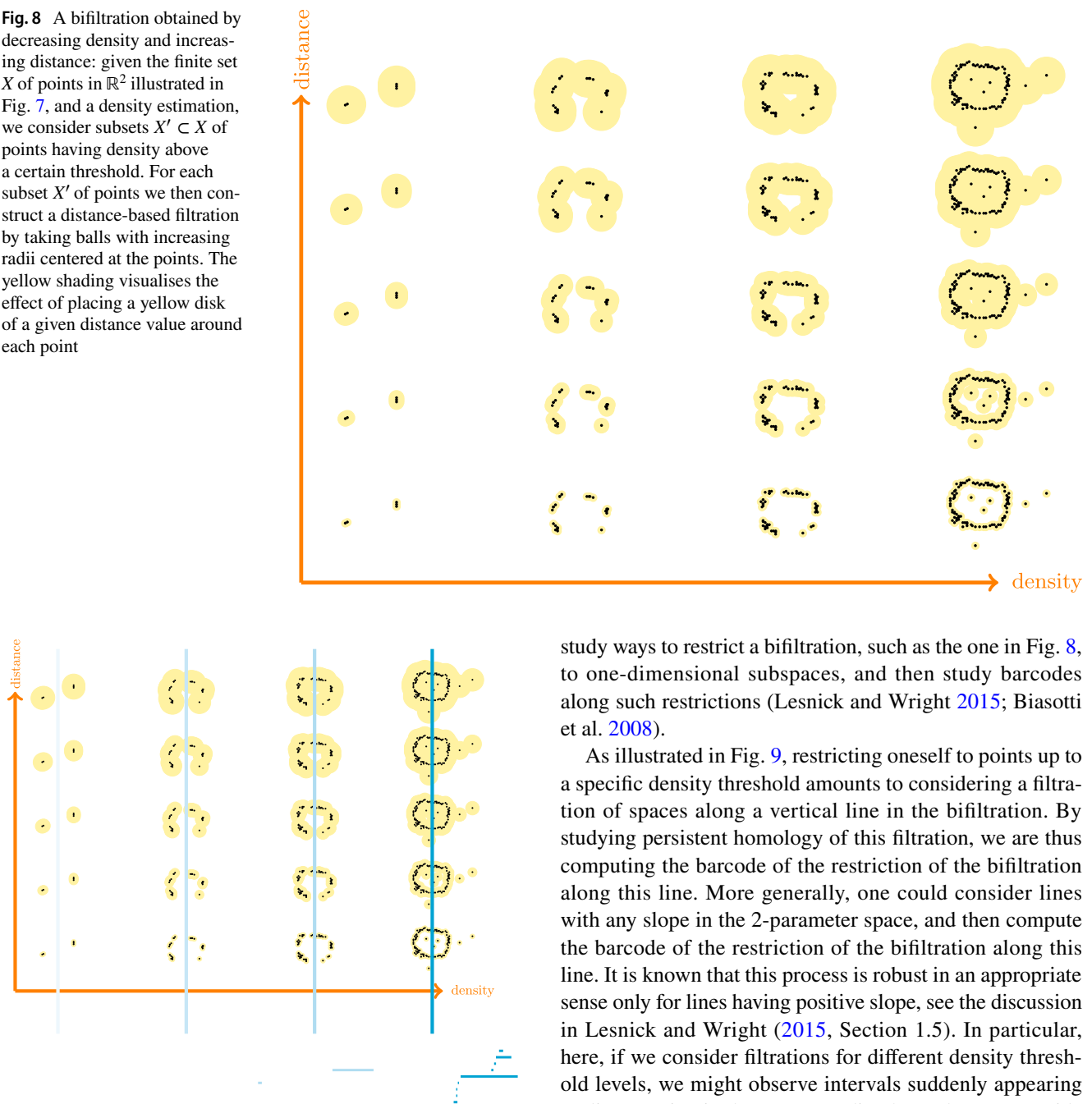


Fig. 9 Barcodes (the collection of blue intervals at the bottom of the figure) for holes along restrictions to vertical lines of the bifiltration from Fig. 8. The barcodes, for the holes, for the first two lines are empty due to the lack of any holes. We note that the barcodes for the components, which we are not depicting here, are not empty. The yellow shading visualises the effect of placing a yellow disk of a given distance value around each point

3.4.1 Barcodes along one-dimensional subspaces

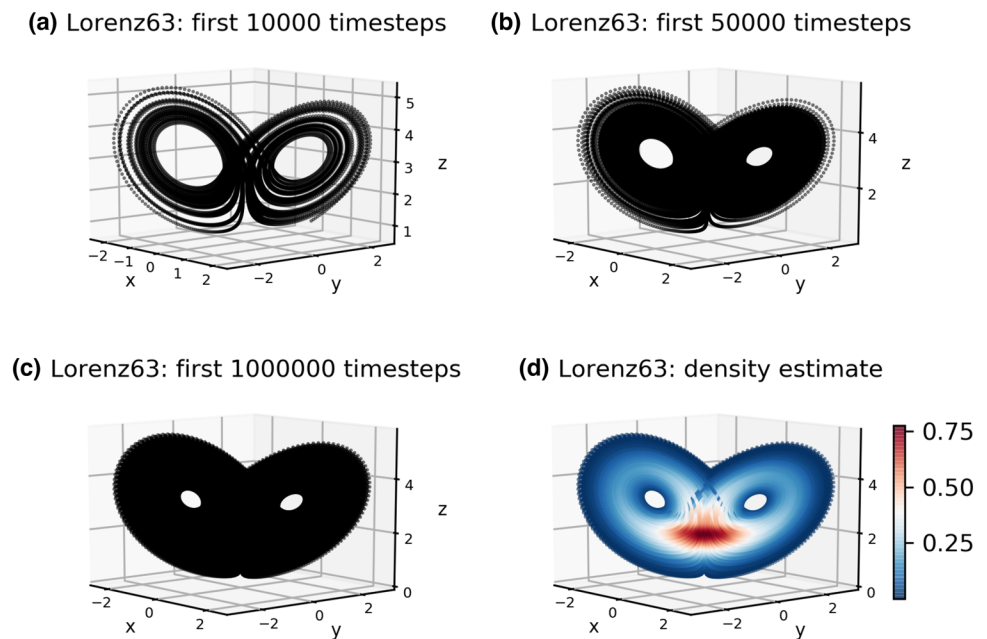
In one approach to defining invariants for multiparameter persistence that are suitable for applications, researchers

study ways to restrict a bifiltration, such as the one in Fig. 8, to one-dimensional subspaces, and then study barcodes along such restrictions (Lesnick and Wright 2015; Biasotti et al. 2008).

As illustrated in Fig. 9, restricting oneself to points up to a specific density threshold amounts to considering a filtration of spaces along a vertical line in the bifiltration. By studying persistent homology of this filtration, we are thus computing the barcode of the restriction of the bifiltration along this line. More generally, one could consider lines with any slope in the 2-parameter space, and then compute the barcode of the restriction of the bifiltration along this line. It is known that this process is robust in an appropriate sense only for lines having positive slope, see the discussion in Lesnick and Wright (2015, Section 1.5). In particular, here, if we consider filtrations for different density threshold levels, we might observe intervals suddenly appearing or disappearing in the corresponding barcodes. We provide a further example using the Lorenz '63 system in Fig. 11.

Lesnick and Wright implemented their methods (Lesnick and Wright 2015) in the software package RIVET (The RIVET Developers 2020), which is currently the only existing software package for the computation of multiparameter persistent homology. Unfortunately, the current implementation in RIVET is not memory-efficient enough for the types of data sets that we study in our work, since if one is interested in computing barcodes to study the lifetime of loops, the software can only handle data sets of a few hundred points. Thus, one main direction that we plan to pursue in future work, is to work on optimisations of the computations

Fig. 10 In **a–c**: the Lorenz '63 system visualised using 10,000, 50,000 and 1,000,000 timesteps respectively. In **d**, the Lorenz '63 system using 1,000,000 timesteps, with colours representing the density, as measured with the kernel density estimator



implemented in RIVET. In particular, in future work we plan to compute barcodes along restriction to lines with positive slopes, to obtain a method that is robust.

3.5 Bifiltrations for dynamical systems

The need to consider not just a filtration of distances, as in the standard method of one-parameter persistent homology, but a bifiltration of distance and density, can be motivated here in two ways. Firstly, and most fundamentally, the dynamical systems we are interested in are always *continuous*, and so no two regions on the attractor can be fully disconnected from each other. In fact, the connectedness of the attractor of a continuous dynamical system can be proved mathematically, given a suitable definition of ‘attractor’ (Gobbino and Sardella 1997), implying that persistent homology will never detect more than one long-lived connected component from a generic sample of the system. The second reason can be understood by considering the Lorenz '63 system. In Fig. 10 we demonstrate a particular feature of the system, namely that the size of the two iconic holes become smaller as one increases the sample size. This implies, somewhat paradoxically, that topological features may become harder to detect the more points you have. If the size of the features becomes comparable to the distance between consecutive points, then these features may, practically speaking, become impossible to detect computationally. These two observations suggest that a naive application of persistent homology to a continuous dynamical system may easily fail to detect both long-lived connected components and long-lived holes.

The basic underlying problem is that in one-parameter persistent homology one computes a filtration by increasing Euclidean distances between the points, ignoring any variations in density. However, the regimes classically identified with clustering methods typically correspond to regions of above-average density, suggesting that the connected components we are interested in should be relative to density. Furthermore, in the Lorenz '63 system, the reason any generic sample of the attractor yields visually clear holes is the fact that the regions of phase space close to the centre of the holes, i.e., near the fixed points, are very low density regions. Therefore, the holes in the system are only identified in data with respect to some chosen density threshold.

Persistent homology provides a solution to the problem of choosing a density threshold at which to study points: instead of trying to estimate the best value for the density parameter, we consider a *bifiltration* of distance and density. An example of such a bifiltration was given in Fig. 8. In Fig. 11 we illustrate this method for the Lorenz '63 system. We note that the bifiltration, and hence the corresponding topological properties that one observes, depend on a choice of density estimation function.

As a final remark, it may be possible to achieve good results by extending the filtration to other measures besides density. In particular, our tests suggest that using phase space velocities can, in some situations, be equally useful. Pre-filtering data based on phase space velocities has in fact been done in some earlier regime studies (Toth 1992; Straus et al. 2007b). Prior knowledge of the system of interest might inform more particular choices. We note that the use of a computationally costly measure of density may, to an extent, offset some of the computational gains of persistent

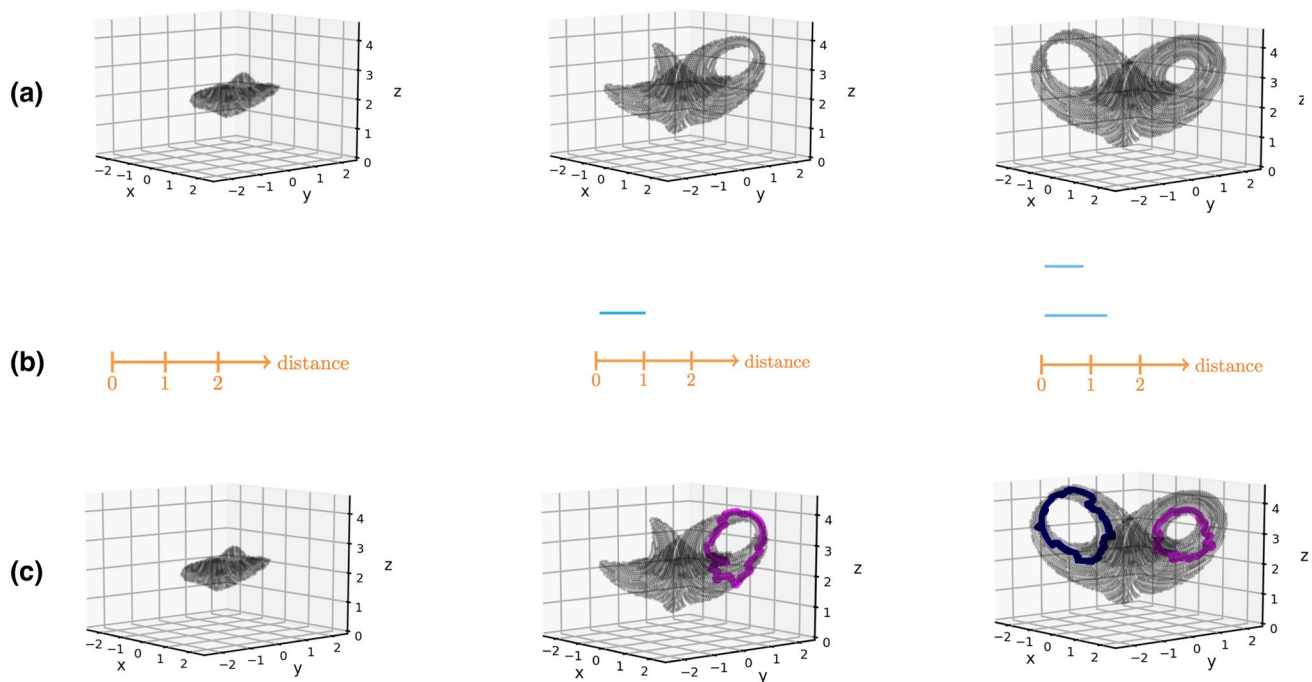


Fig. 11 We illustrate how we use the bifiltration to compute topological invariants on the Lorenz '63 system. **a** Subsample of 20% densest points (left), of 60% densest points (middle) and of 80% densest points (right). **b** Barcodes for holes for each of the points clouds in

a. We note that the barcode for the subsample of 20% densest points is empty. **c** Loops in the point cloud that are representatives for the holes detected by the barcodes

homology discussed in Sect. 3.2. The use of computationally cheaper measures, such as phase space velocities, may therefore be preferable for larger applications. For the context of this paper, however, we will only consider density.

4 Computational methodology

We now describe the full algorithm that we perform to analyse a given data-set sample, as outlined in the schematic Fig. 2. The basic method is the following:

- (1) Normalise each dimension in the data set to have unit variance. Denote this normalised data set by D .
- (2) Estimate the local density of D at every point in phase space: Fig. 2a.
- (3) Pick a percentage threshold $P\%$. Select the sub-sample D_P defined by the upper P th density percentile of D , i.e. the $P\%$ densest points of D : Fig. 2b.
- (4) Compute persistent homology for D_P and extract the topological features of interest: birth/death times for each cycle detected; the points belonging to each of the five longest-lived connected components; a topological representative of each of the five longest lived loops: Fig. 2c.

- (5) Repeat for values of P ranging from 10%, 20%, ..., 100% and examine the features that appear in the resulting bifiltration: Fig. 2d.

The essentially arbitrary choice to only show the 5 longest-lived cycles was made to ensure visual tidiness in plots. For the data sets considered here, no important information is lost by this restriction, though this will of course not be true in general. Note that depending on the choice of parameters (see Sect. 4.2), less than 5 cycles may be found. We also note that the normalisation in step 1 is important to ensure that interesting structure is not missed purely by virtue of existing along a direction in phase space with smaller magnitudes, such as loops that appear as 'squashed' ellipses in the raw data. Finally, in all our examples we use the percentage thresholds 10%, 20%, ..., 100%. This choice suffices to recover the main features of the systems we consider: applications to unknown dynamics may require finer thresholds. The methodology presented here is therefore best understood as a technique that can be flexibly adapted to new situations, much like classical one-parameter persistent homology. Further details on the other steps now follow.

4.1 Density estimation

The primary method used was a kernel density estimator (KDE) with a Gaussian kernel (Marron and Wand 2007), computed using inbuilt functions of the *scipy* python package (Virtanen et al. 2020), where we used the default option of Scott's Rule to determine the bandwidth. The bandwidth determines the minimal spatial scale of the features we compute. We are primarily interested in ignoring features occurring at scales comparable to the average distance between points at consecutive timesteps, which we will informally refer to as the 'grid-scale' of the data set. Using a KDE has two clear advantages for topological applications. Firstly, it produces smooth estimates, which avoids potential issues whereby outlier points remain even after a severe density threshold has been applied. Such outliers will often appear as spurious long-lived connected components, effectively just adding noise to the analysis. Secondly, KDE's are well-suited to represent multimodality, a key feature we want to capture.

A second, cruder method was also tested, which involved directly binning the space and counting the datapoints in each bin. To facilitate the computations, this density estimate was carried out in the space spanned by the first three EOFs, under the assumption that the resulting estimate would be accurate for the scales we were interested in studying. A fixed number of 160^d bins were used in each case, where d is the dimension of the data set.

In our results, the KDE produced good results in all cases except for the CDV system. As will be shown, the CDV system exhibits some very fine-scale structure in the form of 'thin', low-density loops that emerge within a larger, more chaotically-inhabited, low-density region. The Gaussian KDE we used was found to smear away a lot of this structure, while the direct binning method picked out these features easily. In the other data sets, both the KDE and direct binning methods produce qualitatively similar results, but the KDE exhibits a notably smoother estimate, as expected. For this reason, results obtained using the KDE are shown for all data sets except CDV, where the results obtained with direct binning are shown instead. It would clearly be of interest to address the question of whether a more appropriate density estimation method (e.g. choice of bandwidth) might yield good results in all cases, but this is left for future work.

4.2 Computation of persistent homology and representative cycles

In the present work, we compute persistent homology using the Vietoris–Rips complex (see Fig. 5) and the python package *Gudhi* (Developers). *Gudhi* takes as input both the data set and several user-specified input parameters, the choice of which we now outline.

- **max_edge**: This parameter determines the maximal distance threshold to consider in the filtration. Setting this as the maximal distance between any two points in the data set guarantees that the filtration terminates (i.e., ends with a single connected component at the end), so this parameter can always be chosen in a principled manner. Because all our data sets were normalised, we were able to set `max_edge = 5.0` for all data sets.
- **min_pers**: This parameter determines the minimal lifespan that a computed homological cycle needs to attain in order to be included in the final output from *Gudhi*. The choice of this parameter therefore determines the scales of the topological features one wants to consider, similar to the choice of bandwidth in the density filtration (cf. Sect. 4.1). While prior knowledge of the spatial scales of the system can be used to inform the choice of this parameter, the only downside in setting this parameter as very small is an associated increase in computational cost. Because all our data sets are normalised, we found that a parameter choice between 0.15 and 0.50 gave good answers at low cost for all data sets. The higher value was used for CDV, as the main features there exist at higher scales, while for systems like *JetLat*, with subtler behaviour, the smaller value was used.
- **sparse**: This parameter is internal to *Gudhi*'s algorithm, and determines the extent to which the computed Rips complex is sparsened before computing persistent homology. This was set to 0.7 for all data sets.
- **pre_sparse**: This parameter is fed in to the *Gudhi* `sparsify_point_set` routine, which is used to perform a preliminary sparsification of the data set prior to carrying out computations. The routine is built to sparsify data sets in a way which does not change the topology, e.g., by replacing densely connected regions with a sparser set of points covering the same region. Because our data sets are always filtered by density prior to computation, such a sparsification has no impact on our results, but allows the computations to be sped up significantly. In fact, for large data sets with a time dimension exceeding 30,000 timesteps, computations would typically run out of memory and crash. Setting an appropriate value of `pre_sparse`, which greatly reduces the number of points, was therefore crucial. In practice, we set this value as the smallest positive number which would allow the computations to finish at a reasonable rate. A value of 0.05 was found to be suitable for Lorenz '63, Lorenz '96, while 0.005 worked best for CDV. For the *JetLat* data set, where the total number of time-steps available is only around 10,000, this sparsification step was not necessary and hence not carried out.

After computing the filtration and homology at a given density threshold, the five longest-lived components and loops

were identified. Explicitly determining the points belonging to each connected component can be done easily using output from Gudhi, which gives the full filtration. By keeping track of which points are linked up as the filtration radius grows, basic python code suffices to determine all the components; the code used is freely available online (see the Data Availability Statement).

We note that obtaining a representative cycle of loops is significantly harder, as discussed in Sect. 3.3.

4.3 Sensitivity to parameter choices

Several tests were carried out to determine the sensitivity of our results to the parameter choices described in the previous section. A selection of density thresholds for the different data sets were chosen at random, and standard birth/death plots produced using Gudhi for the resulting filtered data sets. It was found that the qualitative features of these birth/death plots did not appreciably change in response to mild perturbations of the parameters, implying that the basic topological features, as summarized in our bifiltration plots, do not depend sensitively on our choices. The size and location of connected components was also found to be largely insensitive to such parameter changes.

On the other hand, the representatives of loops, as computed with PersLoop, were found to exhibit sensitive dependence, in particular on the `pre_sparse` parameter. A small perturbation of this parameter would often lead to the software not terminating properly, or producing a very different representative loop. A similar phenomenon was observed when keeping parameters fixed, but changing other aspects of pre-processing, such as the choice of density filtering or the use of EOF data versus raw data. The reader should therefore be cautioned that the representative cycles we show in our plots are not to be viewed as reliable output from a stable algorithm. Rather, they are included to demonstrate that the topological features seen in our bifiltration plots can, in principle, be visualised in the data itself, and really do correspond to the features one expects.

4.4 Significance testing and topological non-triviality

In this paper, we take the stance that Gaussian distributions should be viewed as having no interesting topological structure, in the sense of persistent homology, and no meaningful regimes. Note that while it is possible for genuine regime systems to produce Gaussian statistics in lower-dimensional projections (Majda et al. 2006), we take the view that in such cases it should still be possible to detect non-Gaussianity in the full, higher dimensional phase space. Therefore, in order to assess whether features identified in our bifiltration methodology are more than just sampling noise, we

implemented the following procedure. First we draw 10,000 random samples from a three-dimensional Gaussian distribution with unit variance. Secondly, we run this through our methodology described at the beginning of Sect. 4. The maximal lifespans of both the connected components and loops obtained at any of the density thresholds were kept: the whole procedure is then repeated ten times and the maximal lifespan obtained across all random draws is used as a measure of noise. Specifically, features with a lifespan close to this value, of around 0.4, are likely to be noise coming from grid-scale sampling variability, while features with a lifespan greatly exceeding this are likely to be indicative of significant non-trivial homology. In the context of this paper, we therefore define a dynamical system to have non-trivial topological structure if and only if its distance-density bifiltration produces cycles with lifespans exceeding that expected from Gaussian noise (i.e., lifespans exceeding 0.4, in the case where the dimensions of the systems have been normalised).

When carrying out this procedure, the connected components in Gaussian samples containing 3 or less points were not included, because one or two big outlier points can easily produce very long-lived ‘components’. For consistency, components with 3 or less points that are detected in any data set are always clearly marked in plots. Note that such outliers can also occur for filtered data unless `min_pers` is large, since there may randomly be some points fractionally closer to each other than any other points.

Finally, we note that because all our data sets are normalised prior to computing homology, the unit variance Gaussian offers an appropriate comparison for all the data sets that we consider.

5 Results

For each data set, we now produce a standard bifiltration plot summarising the lifespans of persistent cycles across a range of density thresholds. In addition, to visualise these topological features, particular density thresholds are hand-picked for each data set and plotted, together with a visualisation of either the connected components or the representatives of loops present at that threshold.

5.1 The Gaussian

As explained in Sect. 4.4, results from the unit variance Gaussian distribution are used to estimate the significance of features obtained for all other data sets, since any definition of regimes should exclude a Gaussian from having any. We therefore first present results for a randomly drawn sample of 10,000 points from such a distribution. These are shown in Fig. 12. As expected, no non-trivial

Fig. 12 A distance-density bifiltration of a unit variance Gaussian distribution. For each density threshold on the x -axis, the lifespan of the 5 longest-lived connected components (red dots if the component contains more than 3 points; red stars otherwise) and 5 longest-lived loops (blue triangles) are plotted. The stippled line shows the largest lifespan obtained across multiple Gaussian samples. The meaning of the \min_pers parameter is explained in Sect. 4.2

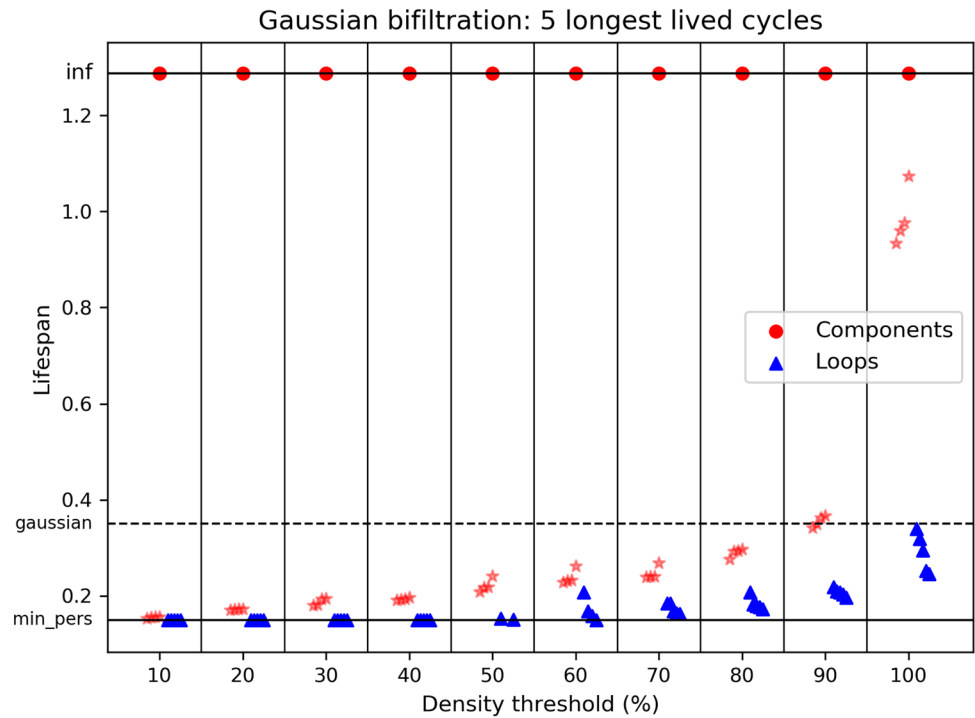
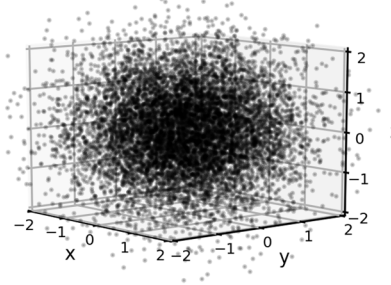
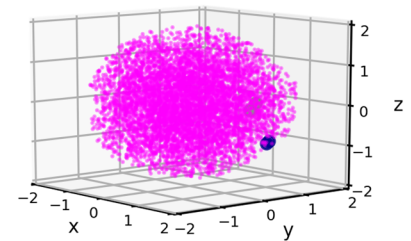


Fig. 13 In **a** a random sample of a unit variance Gaussian distribution. In **b** the 70% densest points of the sample; **c** the 40% densest points; **d** the 10% densest points. In **b–d**, points are coloured according to the connected component they live in: longest-lived (pink colour), 2nd longest lived (blue colour) and 3rd longest lived (green colour). Points belonging to components with 3 or less points have been made larger to aid visualisation

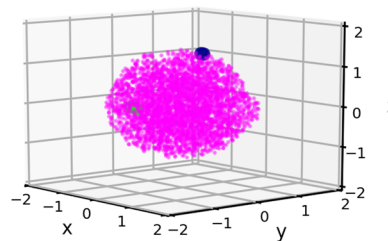
(a) Gaussian: 100% densest



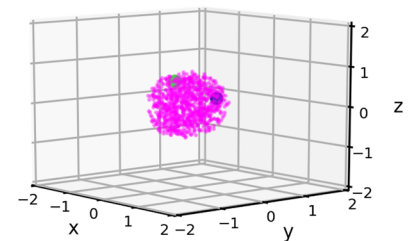
(b) Gaussian: 70% densest



(c) Gaussian: 40% densest



(d) Gaussian: 10% densest



topological features are detected in this data set, with each density threshold exhibiting only a single connected component (the red dot at infinity) and some spurious outliers (the red stars) at the ‘grid-scale’. The loops found (blue triangles) are all extremely close to the minimum persistence choice, implying that these were only barely registered by the algorithm and do not persist for notably longer than isolated outlier components.

The seeming change in behaviour at the 100% threshold, where no density filtering has been applied, is due to the existence of big outliers in the raw sample. This is clearly seen in Fig. 13, showing the Gaussian sample at various thresholds. Because even an extremely mild density threshold immediately removes the big outliers seen in Fig. 13a, the possible lifespan of small components with 3 or less points drops dramatically from the 100 to

Fig. 14 A distance-density bifiltration of the Lorenz '63 system. For each density threshold on the x -axis, the lifespan of the 5 longest-lived connected components (red dots if the component contains more than 3 points: red stars otherwise) and 5 longest-lived loops (blue triangles) are plotted. The stippled line shows the largest lifespan expected from Gaussian noise. The meaning of the `min_pers` parameter is explained in Sect. 4.2

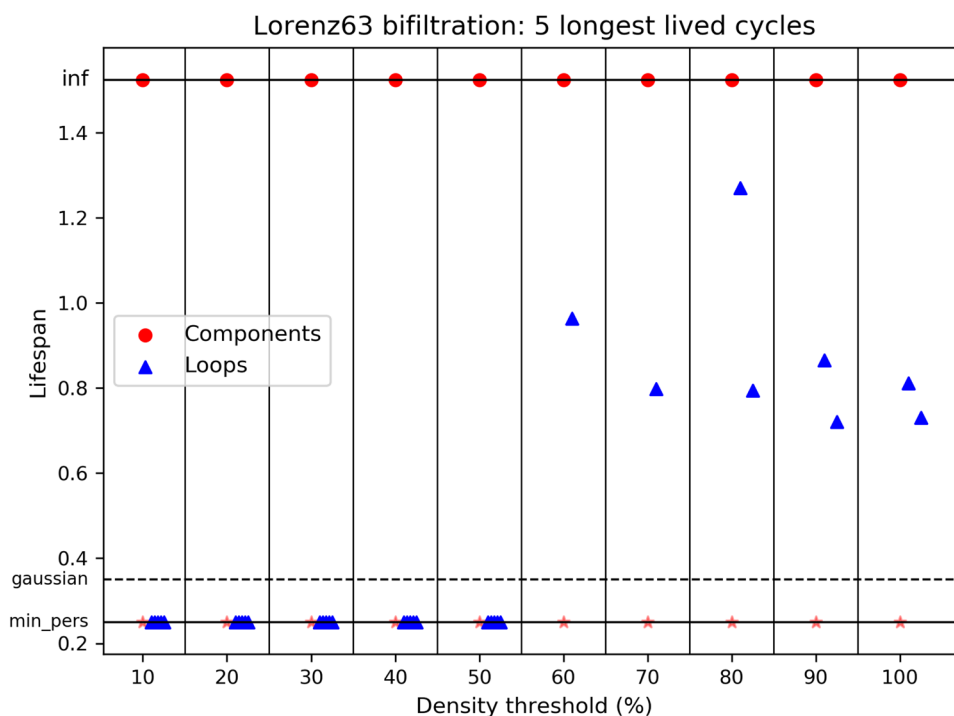
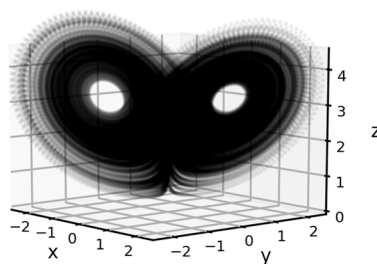
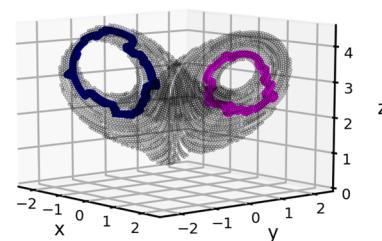


Fig. 15 In **a** a long integration of the Lorenz '63 system. In **b** the 80% densest points; **c** the 60% densest points; **d** the 20% densest points. In **b** and **c**, representatives of the 2 (respectively 1) longest-lived loops are overlain. The longest-lived loop is always in pink, the 2nd longest in blue. In **d**, the longest-lived connected component is marked in pink

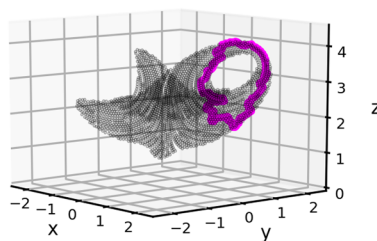
(a) Lorenz63: 100% densest



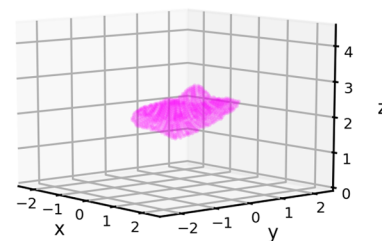
(b) Lorenz63: 80% densest



(c) Lorenz63: 60% densest



(d) Lorenz63: 20% densest

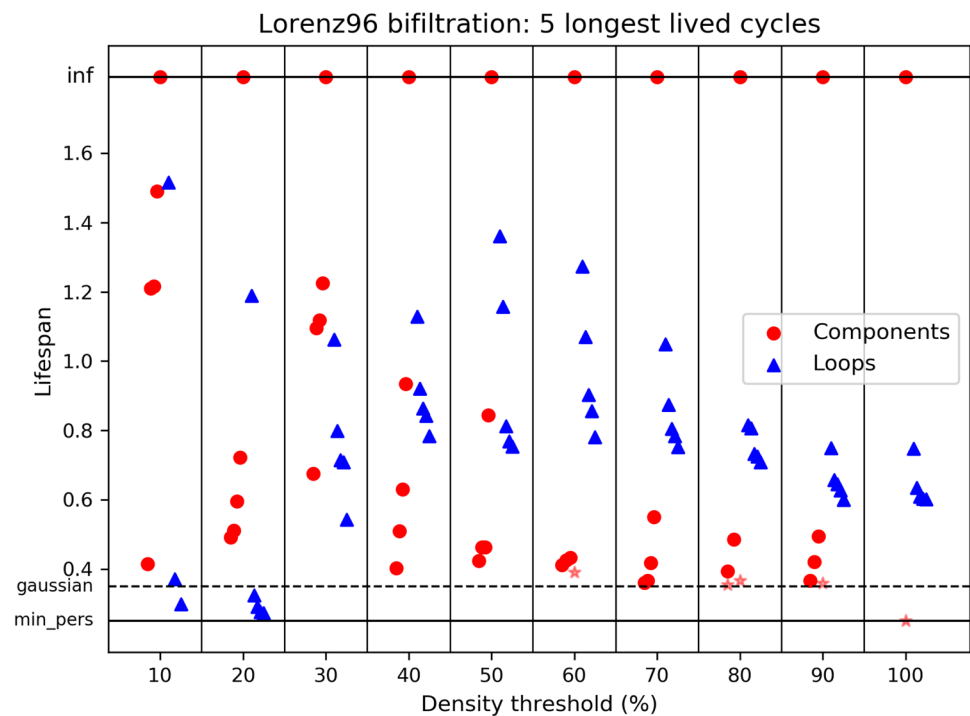


90% threshold. This is also why the longest-lived loops are found at the 100% threshold. As is clear from Fig. 13, these loops are just noise, and indeed any representatives of these produced by PersLoop (not shown) are visually confirmed as such. Figure 13b–d also highlight the 3 longest-lived connected components at each threshold. It can be seen that this yields one component containing almost all points, and two components consisting of one

or two points that simply happen to be fractionally further removed from the rest of the point mass.

These observations already confirm that our methodology correctly identifies the Gaussian as having no non-trivial topology at any density threshold. The comparison of Fig. 12 with the equivalent plots for other data sets, to which we now turn, will make this even clearer.

Fig. 16 A distance-density bifiltration of the Lorenz '96 system. For each density threshold on the x-axis, the lifespan of the 5 longest-lived connected components (red dots if the component contains more than 3 points; red stars otherwise) and 5 longest-lived loops (blue triangles) are plotted. The stippled line shows the largest lifespan expected from Gaussian noise. The meaning of the `min_pers` parameter is explained in Sect. 4.2



5.2 Lorenz '63

Figure 14 shows the bifiltration plot of the Lorenz '63 system. This plot can be understood by reference to Fig. 15, which visualises the system, and the longest-lived components/loops, at different thresholds. At low density thresholds, as shown in Fig. 15d, there is just one connected component, corresponding to the dense central region between the two wings. Because the density is concentrated in this area, as seen in Fig. 10d, there is no trace of the two wings until you move to higher thresholds. At the 60% threshold, Fig. 15c, enough points are included for one of the wings to emerge, at which point an extremely long-lived hole appears in the bifiltration plot: the representative produced by PersLoop confirms that this corresponds to the right wing. At the 70% threshold, the second wing also emerges, after which one retains two long-lived holes for all further density thresholds. Figure 15b confirms that the two holes found by Gudhi at this point correspond to the two holes in the wings. Note that the apparent asymmetry between the two loops is due to sampling variability.

Two other points are worth observing in Fig. 14. Firstly, besides the key topological features coming from the wings, all other features have lifespans at the `min_pers` threshold, implying that these features exist only at or below the grid-scale of Lorenz '63. Secondly, these grid-scale features have lifespans below what is expected from a Gaussian bifiltration, demonstrating that our significance test has correctly classified these as noise. Furthermore, the lifespans of the two loops, and one connected component, greatly exceed

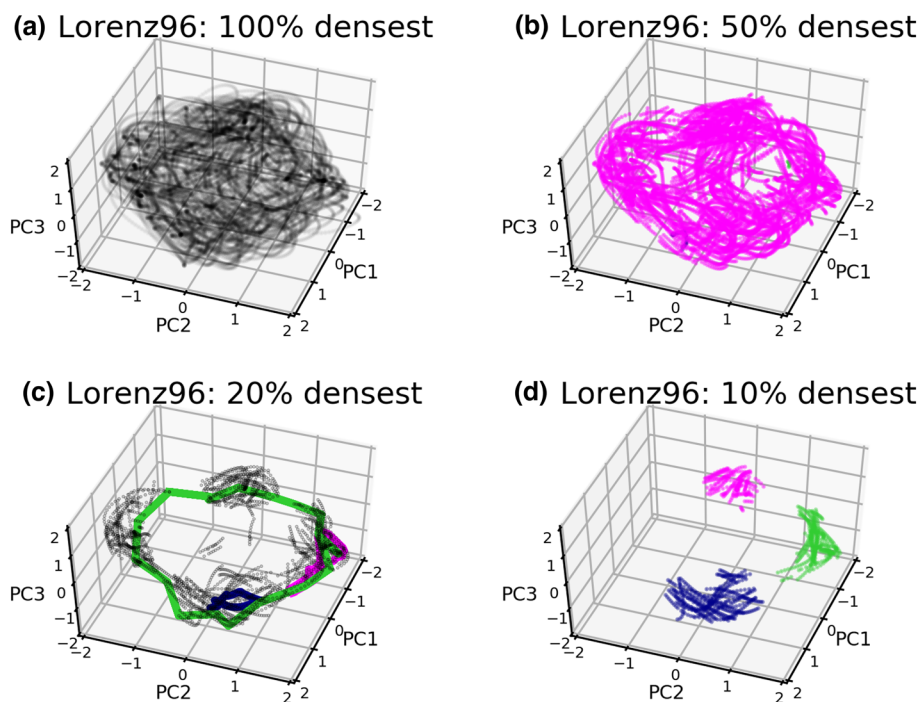
Gaussian noise. The conclusion from our methodology is therefore that the Lorenz '63 system has two significant holes, corresponding precisely to the two classical regimes defined by the wings (Palmer 1994), and is otherwise fully connected.

5.3 Lorenz '96

Figure 16 shows the bifiltration plot for the Lorenz '96 system, which suggests the existence of a considerable amount of significant topological structure. The apparent complexity of this structure is consistent with the impression obtained from animations of the dynamics (cf. Sect. 2.1), which show that the system is made up of a number of interweaving loops and components. Figure 17 shows some of this structure at different thresholds, though we remind the reader that because the homological computations in this case were done using a 4-dimensional EOF truncation, our 3-dimensional projections necessarily obscure some of the features. We also note that, due to the limitations of PersLoop, optimal loops were computed using the space spanned by the first three EOFs only, which also leads to some minor distortions.

The characteristic looping behaviour of the system is already visible in the unfiltered data set, Fig. 17a, reflecting the rotational symmetry in the defining equations. The looping trajectories result in regions which, after an appropriate density threshold is imposed, appear as holes in an otherwise connected space, as in Fig. 17b. The most prominent loop appearing in this manner is the one circling the full

Fig. 17 In **a** a long integration of the Lorenz '96 system. In **b** the 50% densest points; **c** the 20% densest points; **d** the 10% densest points. In **b** and **d**, representatives of the 3 longest-lived connected components are marked with colours, while in **c**, the 3 longest-lived loops are overlain in colour. In all cases, the longest-lived feature is in pink, the 2nd longest-lived in blue and the 3rd longest-lived in green



perimeter of the space, as seen in Fig. 17c. Note that PersLoop identifies this loop as the 3rd longest-lived at the 20% threshold. The representatives found for the longest and 2nd longest-lived loops are made to look particularly spurious due to the flattening of the 4th dimension, but are either way examples of the way in which PersLoop sometimes produces representatives that are far from optimal. For very severe density thresholds, such as the 10% threshold shown in Fig. 17d, the data set splits up into distinct components, implying significant local variations in density across the attractor.

To see how this topological structure relates to the more classical approach to regimes in Lorenz '96, recall the approach taken by Lorenz (Lorenz 2006), further expanded on in Christensen et al. (2015), which the reader should refer to for this discussion. In *ibid*, the dynamics are first projected onto the two-dimensional space spanned by the magnitudes of the concatenated principal component vectors $[PC1, PC2]$, $[PC3, PC4]$. Two local peaks in temporal persistence are identified in this space, clearly visible in Figure 7c of *ibid*, and these are used to define two regimes denoted *A* and *B*. Regime *A* corresponds to the bottom right-hand corner of the concatenated space, which is also where the density is concentrated (cf. subplot (a) of the same figure), while regime *B* corresponds to a very

low-density region in the top left-hand corner. In Fig. 18, points loosely corresponding to these two corners of phase space have been marked, with the top left-hand corner defined by $|[PC1, PC2]| < 3$, $|[PC3, PC4]| > 10$, and the bottom right-hand corner by $15 > |[PC1, PC2]| > 10$, $|[PC3, PC4]| < 5$, where vertical lines denote the vector magnitude.⁵ This clearly suggests that regime *A* corresponds to the densely populated loop around the outer perimeter, while regime *B* corresponds

Christensen et al. regimes in Lorenz96 EOF space

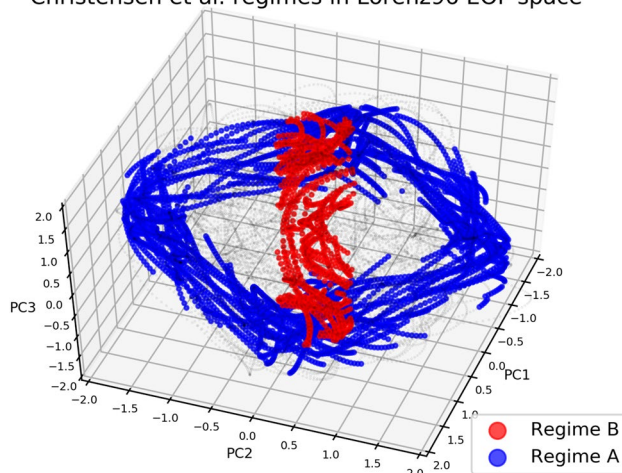


Fig. 18 The two regimes *A* (blue) and *B* (red), as defined in Christensen et al. (2015), marked in the space spanned by the first 3 principal components of Lorenz '96. The full data set is shown in transparent black in the background

⁵ We caution the reader that this subsetting was done using the raw PCs, as in Christensen et al. (2015), but that Fig. 18 uses *normalised* PCs.

Fig. 19 A distance-density bifiltration of the CDV system. For each density threshold on the x-axis, the lifespan of the 5 longest-lived connected components (red dots if the component contains more than 3 points; red stars otherwise) and 5 longest-lived loops (blue triangles) are plotted. The stippled line shows the largest lifespan expected from Gaussian noise. The meaning of the `min_pers` parameter is explained in Sect. 4.2

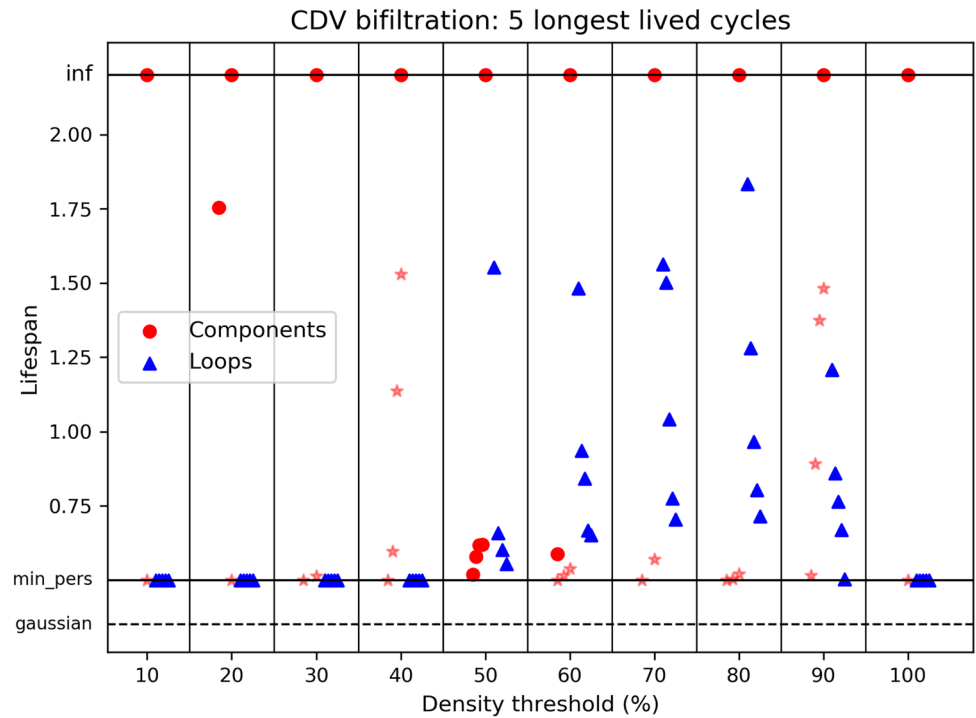
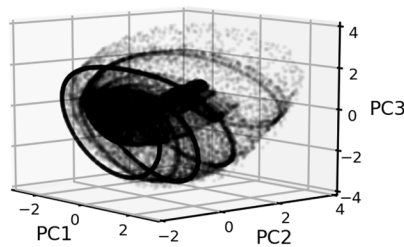
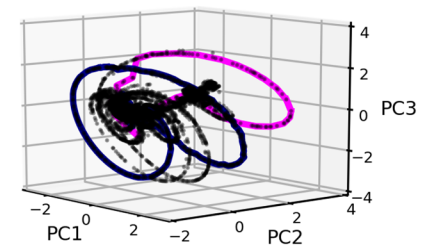


Fig. 20 In **a** a long integration of the CDV system. In **b**: the 70% densest points; **c** the 50% densest points; **d** the 20% densest points. In **b** and **c**, representatives of the 2 longest-lived loops are overlain in colour, while in **d**, the 2 longest-lived components are marked in colour. In all cases, the longest-lived feature is in pink and the 2nd longest-lived in blue

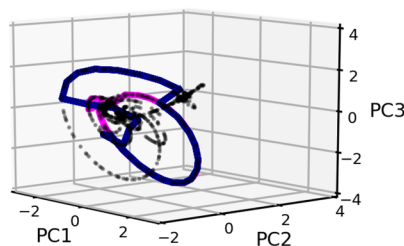
(a) CDV: 100% densest



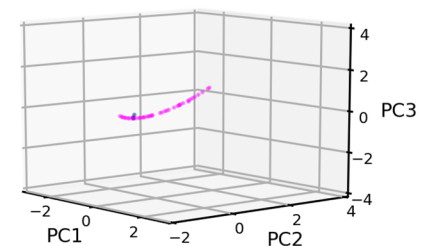
(b) CDV: 70% densest



(c) CDV: 50% densest



(d) CDV: 20% densest

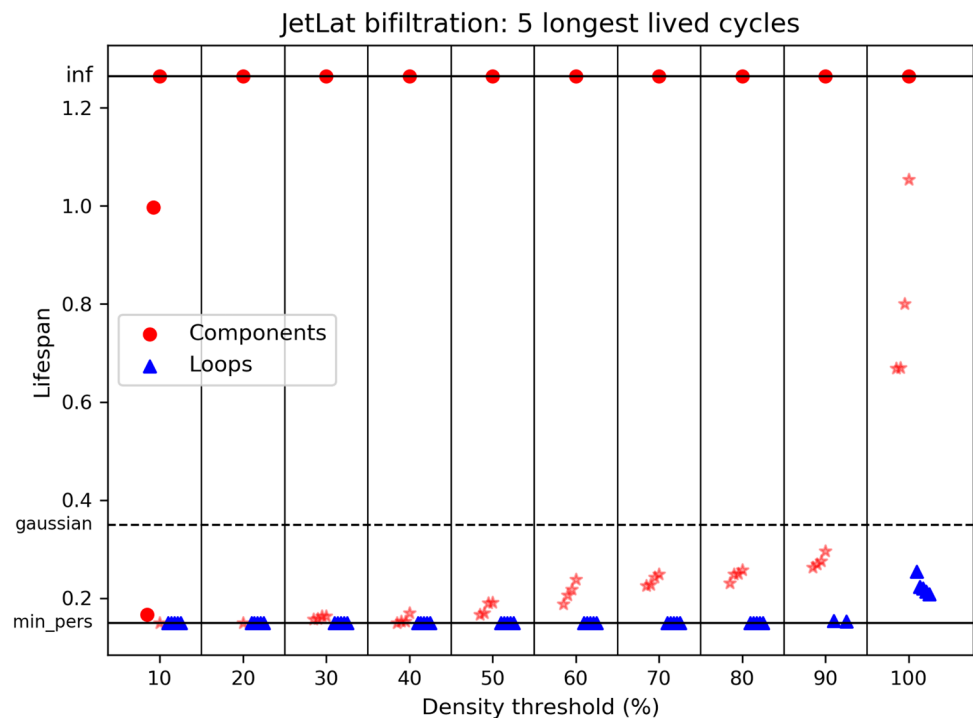


to the low-density hole in the centre; we remind the reader again that the squashing away of the fourth dimension gives the appearance of regime *B* spilling out into the perimeter. In other words, the regimes diagnosed in Christensen et al. (2015) correspond to topological features of the system that are detectable with persistent homology.

5.4 Charney–deVore

Figure 19 shows the bifiltration results for the CDV system: we remind the reader that the computations are done using the space spanned by the first three EOFs. The most notable features are a number of long-lived loops that emerge at

Fig. 21 A distance-density bifiltration of the JetLat data set. For each density threshold on the x-axis, the lifespan of the 5 longest-lived connected components (red dots if the component contains more than 3 points; red stars otherwise) and 5 longest-lived loops (blue triangles) are plotted. The stippled line shows the largest lifespan expected from Gaussian noise. The meaning of the `min_pers` parameter is explained in Sect. 4.2



density thresholds between 50% and 90%. The existence of such loops can already be seen by eye in the raw data set, shown in Fig. 20a. These loops correspond to low-dimensional, preferred trajectories shadowing unstable homoclinic orbits (Pomeau and Manneville 1980), separated by sparsely populated regions. The use of the direct binning method to estimate density effectively highlights these loops, and the representatives found by PersLoop, as in Fig. 20b, c, confirm that these are precisely the long-lived loops identified in Fig. 19.

In terms of connected components, the only threshold at which there appears to be more than one connected component with at least four points is the 20% threshold. However, manual inspection here reveals that this second component in fact contains *exactly* four points, and can therefore be considered as noise, as with the spurious components seen at the 40% and 90% thresholds. Therefore, from the perspective of the bifiltration, CDV can be thought of as a dense central region with low-density loops spiraling outward. This neatly matches the dynamics one observes in numerical simulation, and the theoretical understanding of the CDV system as chaotic transients bursting from a weakly unstable near-equilibria.

In the classical perspective (Charney and DeVore 1979), CDV has two persistent regimes associated with orbits slowing as they enter the neighbourhood of one of two fixed-points. One of these fixed points, associated with blocking, is located close to the dense central region, while the other more zonally symmetric fixed point lies close to the back left corner, when viewed as in Fig. 20, and the loops pass close

to this region. The regime dynamics in CDV are asymmetrical, in that the blocked regime is quasistationary and experiences almost deterministic evolution, while the zonal regime is characterised by turbulent chaotic behaviour (Pomeau and Manneville 1980). From this we can understand why the quasistationary blocking state is associated with a connected component, while the zonal state is not. Instead, the zonal regime can be understood as a consequence of the many looping trajectories visiting a common, disparate region of phase space.

5.5 The North Atlantic jet

We finally test our method using the JetLat data set, capturing variability of the North Atlantic eddy-driven jet. Figure 21 shows the result of the bifiltration computation, with Fig. 22 visualising select thresholds. The only evidence of non-trivial topology emerges when restricting to the 10% densest points, at which point the data set splits cleanly into two connected components, as shown in Fig. 22d. The lifespans of both components greatly exceed anything expected from Gaussian noise, and their sizes are also considerable, containing around 900 and 100 points each. Figure 23 shows composites of zonal wind anomalies of ERA20C across all days belonging to these two long-lived components, identifying the longest-lived one as the Central jet latitude mode and the 2nd longest-lived as the Northern jet latitude mode.

Since one dimension of the JetLat data set contains the jet latitude index, which is trimodal in and of itself, the *a priori* expectation might be that the data set should split into

Fig. 22 In **a** the raw JetLat data set. In **b** the 70% densest points; **c** the 50% densest points; **d** the 10% densest points. In **b–d**, representatives of the 3 longest-lived connected components are marked with colours. The longest-lived feature is in pink, the 2nd longest-lived in blue and the 3rd longest-lived in green

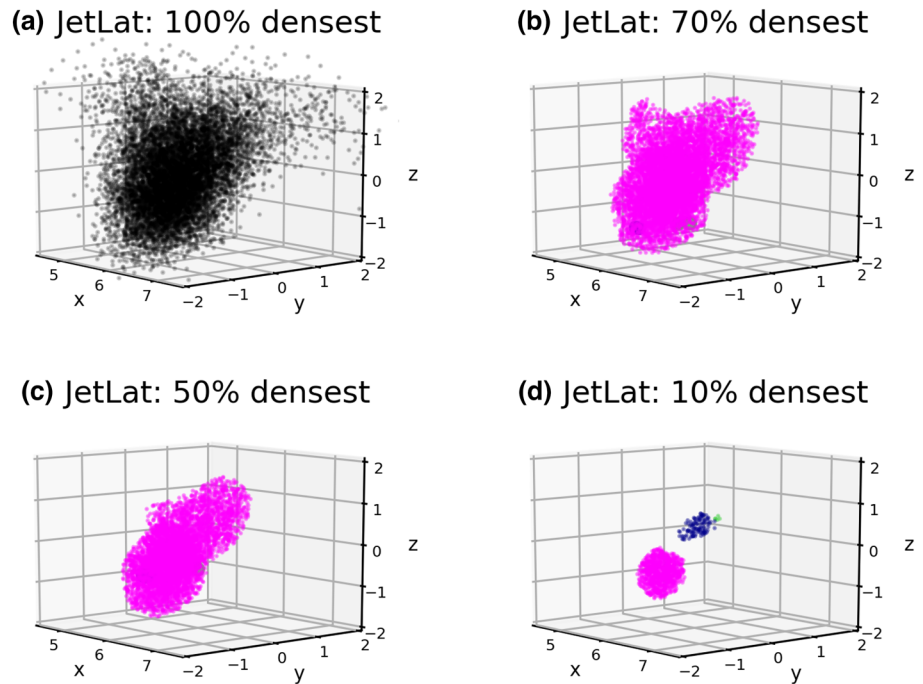
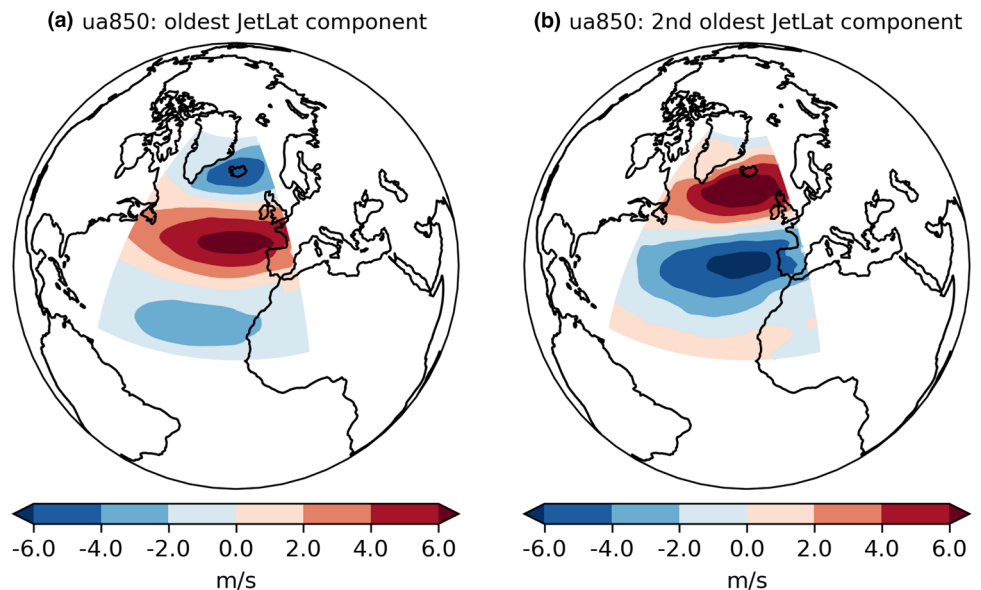


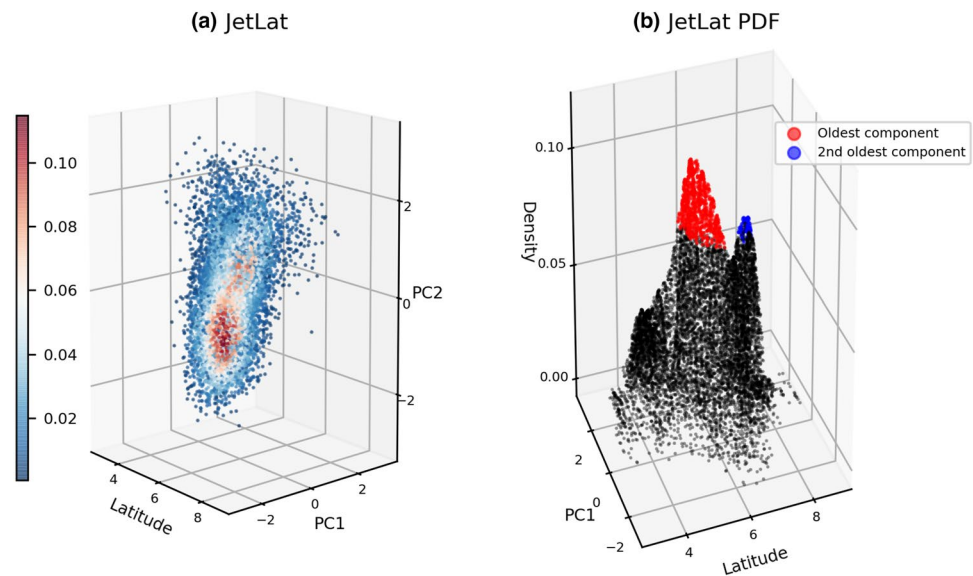
Fig. 23 Composites of zonal wind anomalies at 850 hPa for the ERA20C data set across all winter days between 1900 and 2010 that **a** belong to the longest-lived JetLat component; **b** belong to the 2nd longest-lived JetLat component



three connected components, not two. However, making the density filtration finer did not change the result, suggesting this is a robust outcome of our methodology. To understand why this happens, Fig. 24 shows the JetLat probability distribution function (pdf), as computed using the kernel density estimator. In panel (a), the raw data set is plotted with colours indicating density, while in (b), density is plotted as a function of jet latitude and $PC1$ (the first two dimensions of JetLat). In this latter panel, the points corresponding to the two long-lived components at the 10% threshold have been coloured in, with red being the longest-lived and blue

the 2nd longest-lived. While panel (a) already suggests that there are two, rather than three, clearly marked peaks in density, panel (b) most clearly explains what is happening. In and of itself, the jet latitude index is clearly trimodal, but the situation changes when it is extended out across multiple dimensions. While the Northern peak remains clearly separated from the Central peak, the Southern peak becomes smeared out across the space spanned by the two principal components, leaving it resembling a ‘shoulder’, rather than a clear peak. Because our density thresholds amount to taking horizontal slices across this space, the bifiltration is able to

Fig. 24 Visualising the pdf of the JetLat data set. In **a**, the 3-dimensional data with colours indicating the density, as estimated with a kernel density estimator. In **b**, density as a function of the two first dimensions of the JetLat data set, i.e. jet latitude and the first principal component of ua850. In **b**, points belonging to the longest-lived connected component are coloured red, while points belonging to the 2nd longest-lived component are coloured blue



find the Central and Northern peaks, but not the Southern. The implications of this are discussed in the next section.

Note that when computing a bifiltration using the first 3, 4 or 10 principal components of geopotential height anomalies at 500 hPa (Z500), the features detected are all at the level of Gaussian noise. This was found to be the case both when using data defined over the ‘jet domain’ 15N–75N, 300E–360E, and when using the larger domain 30N–90N, 80W–40E more commonly used in Z500-based studies (Straus et al. 2007a; Dawson et al. 2012). This is consistent with the findings of Stephenson et al. (2004), namely that Z500 space is close to Gaussian.

6 Discussion

6.1 Strengths and weaknesses of our methodology

The results in the previous section suggest that the bifiltration methodology succeeds in identifying whether a data set has non-trivial topological structure or not. In particular, it rejects a Gaussian distribution as having any, and correctly detects the relevant structure for four examples of data sets generally considered to have regimes. We further showed that the topological structure encodes, in different ways, the regime behaviour. For Lorenz ’63, the regimes correspond to two holes; for Lorenz ’96 to a loop and a hole; for CDV to a dense, connected region and several loops emanating from this; and for JetLat, to two dense, connected components. The four systems considered are thereby clearly distinguished through their differing homology.

Besides scaling extremely well with the dimension of the data set, our method has the desirable feature that it does

not require ad hoc parameter choices that directly influence the regime structure. This is essentially because in our perspective, the regime structure corresponds to the non-trivial topological features of the attractor. Because the attractor itself is fixed, so are its topological features, and our methodology simply provides a way of probing these features. The various parameter choices we made in the implementation of our methodology, outlined in Sect. 4, can be thought of as determining how precisely one probes the topology. For example, using a different set of density thresholds, or a different density estimator, amounts to probing the topology more or less finely. In particular, no assertion is ever being made about a choice of parameters which is ‘optimal’, unlike in the case of, e.g., K-means clustering, where in many cases a choice of K has to be made and justified at some point.⁶

The main apparent shortcoming of the methodology, besides the instabilities associated with trying to compute optimal representatives of loops, was the inability to identify three distinct regimes in the JetLat data set. As explained in Sect. 5.5, this is due to the fact that, when viewed across multiple dimensions, the Southern jet latitude mode appears less as a distinct peak and more as an extended shoulder, which the horizontal density slices of our filtration cannot easily capture. The obvious way to attempt to remedy this is to consider slices with positive slope. As explained in Sect. 3.5, this is also required to make the evolution of topological features continuous across the bifiltration, implying that this is a natural way to improve our methodology for

⁶ In some cases, the physical existence of uniquely defined regimes is incidental to the application at hand, and the need to fix K can then be sidestepped by assessing the robustness of the analysis across a range of K-values: see, e.g., Amini and Straus (2019).

stability reasons alone. We hope to examine this, using the RIVET software (cf. Sect. 3.4.1), in future work. It has also been noted (Hazelton 2003) that Gaussian kernels can sometimes flatten peaks too much: a more thorough examination of optimal density estimators for our data sets is for this reason another avenue of future work.

While the failure to detect the Southern jet mode should probably be viewed as a shortcoming, we would also suggest that this failure may shed some light on a few curious features in the literature. Firstly, many studies have tried to diagnose regimes in the Euro-Atlantic sector, and, depending on the choice of input data, pre-processing steps and diagnostics, these studies have suggested there may be anywhere between 2 and 7 regimes (see Hannachi and Iqbal 2019; Dorrington and Strommen 2020; Dawson et al. 2012; Madonna et al. 2017; Falkena et al. 2020 and Grams et al. 2017 respectively for examples of each number). While the ambiguity between the choices 3, 4 and 5 is at least in part due to the confounding influence of the jet speed (Dorrington and Strommen 2020), and the choice of 2 regimes usually corresponds to the North Atlantic Oscillation dipole (Woollings et al. 2008; Hannachi and Iqbal 2019), the striking divergence in the number of regimes across studies using similar techniques is still somewhat puzzling. Our results suggest that one possible reason for this is that, depending on what angle one views the Euro-Atlantic circulation from, different regimes may appear either as clearly distinct peaks or more ambiguous and hard to detect shoulders. The use of different spatial domains across different studies likely adds to this issue.

Secondly, in Strommen (2020), the ability of a numerical weather forecast model to make skillful predictions of the Euro-Atlantic circulation was studied from the perspective of the three jet latitude regimes. It was found that the model was able to skillfully detect changes in the Northern mode compared to the Southern and Central modes, but was not able to robustly separate between the Southern and Central modes. In other words, from the perspective of the forecast model, the jet appeared to behave as if it had 2, not 3, regimes. By considering Fig. 24b, it is perhaps not surprising that an imperfect model may struggle to reproduce the more subtle behaviour of the Southern shoulder, and produce a cruder approximation of the pdf as having just two peaks. A comparison between this figure and an equivalent one for model data (not shown) does suggest the model has a notably flatter Southern peak.

6.2 Why a simpler definition of regime fails

We have shown that non-trivial topological structure, as measured with a bifiltration of homology, provides a unifying way of understanding the main examples of non-linear dynamical systems generally considered as exemplifying

regime behaviour. Because this comes at the cost of introducing an extra level of abstraction, it is reasonable to ask if a similar unification could be achieved using the more common ways of understanding regimes, namely density peaks (i.e., clustering) or temporal persistence. We will now show that, on the face of it, no such simpler unification appears possible.

To see this, first notice that while for JetLat, the two regimes correspond clearly to local maxima in density, both Lorenz '96 and CDV are examples where the density of the two regimes are wildly different. For Lorenz '63, while a bimodal pdf can be obtained by time-averaging (Corti et al. 1999), Fig. 10d makes it clear that the regions defined by the two regimes (i.e., the two wings) are, in the raw data set, not local density maxima. Hence a definition of regimes as local density maxima/minima or clustering will invariably fail to account for one of these systems.

Next, one might consider a criteria based on any of the closely related concepts of temporal persistence, average residence times or phase space velocities.⁷ However, also here one finds that the behaviour of the different systems differs dramatically. In the Lorenz '63 system, temporal persistence peaks (and velocities are smallest) at the dense region connecting the two wings, while temporal persistence is in general minimal in the wings themselves, where velocities peak; the exception being the extremely rare trajectories that pass sufficiently close to either fixed point. On the other hand, for Lorenz '96, both regimes correspond to peaks in temporal persistence/residence time, as mentioned already, while in CDV the two regimes are broadly asymmetric in terms of their temporal persistence and velocities, with the blocking regime featuring high temporal persistence/low velocities and the zonal regime favouring low temporal persistence/high velocities. Even in the real atmosphere, the behaviour does not appear to be uniform. Already in Woollings et al. (2010a), where the jet latitude regimes were first presented, it was noted that the forcing on the jet by transient eddies, thought to be a key driver in generating temporal persistence, appears to be operating similarly at all latitudes, not just at the peaks of the trimodal distribution. In other words, the extent to which the three jet latitude modes can be characterised as having higher-than-average temporal persistence is ambiguous. This ambiguity is further supported by the results of Faranda et al. (2017), which examined the closely related 4-regime picture of the Euro-Atlantic. By computing a measure of both local temporal persistence and local density, they locate the four regimes in distinct quadrants of

⁷ Phase space velocity here means the distance traversed in phase space between consecutive time-steps. In particular, this is not the same as atmospheric velocities of, e.g., winds.

temporal persistence-density space, implying the regimes all have strikingly different characteristics.

A simplistic definition of regimes based on temporal persistence, residence times or velocities will, therefore, inevitably fail to capture the behaviour in one or more of these systems. It is also clear from this discussion that the situation cannot be salvaged by using a definition combining both notions. Hence it seems, to these authors, not to be possible to find a definition of regimes, using density or temporal persistence alone, that unifies all the systems we considered. While an alternative definition of regimes based on fixed points, UPOs or other ‘exact solution’ techniques might seem plausible, computing such solutions is extremely computationally demanding, and state-of-the-art techniques are only able to handle systems of significantly lower dimensionality than existing climate models (Lucarini and Gritsun 2020). More crucially, these techniques are inherently model features, in that they rely on being able to integrate the model dynamics. Given that models are known to exhibit systematic biases in their regime structure (Fabiano et al. 2020), inferring conclusions about the real atmosphere based on results obtained from models would require considerable care. It is therefore not currently clear how such ‘exact solution’ techniques can be applied to observational data sets.

Some readers may reasonably question whether it is in fact important to have a unified framework for understanding regimes, and that the word ‘regime’ is perhaps best understood as a context-dependent phrase that captures a wide variety of ways to simplify complex, non-linear dynamics. Indeed, it is possible that there exist dynamical systems that appear to exhibit regime behaviour that cannot be accounted for by topological means. Nevertheless, the fact that four very different systems do allow for such a topological characterisation lends confidence to this being possible in a wide variety of cases. Furthermore, we believe that the more ad hoc regime approach common in atmospheric science - and the lack of any clear unifying framework - has in general undermined confidence both in their practical usage and even their existence (Stephenson et al. 2004; Christiansen 2007; Ferrel 2017). The existence of non-trivial topological structure underpinning four quintessential examples found in the literature may help bolster confidence that the various attempts to diagnose regimes in the atmosphere are really characterising genuine features of the climate attractor.

7 Conclusions and further directions

In this paper we have argued that the unifying feature across the most well-known examples of regime systems is their non-trivial topological structure. We showed that, using persistent homology, one can compute topological invariants which encode such non-trivial structure. By carrying out

this computation for four classical regime systems (Lorenz '63, Lorenz '96, Charney–deVore and the North Atlantic jet), we showed that the information encoded in these topological invariants captures the key features of each system associated with their regimes. It was pointed out that these systems also exhibit widely differing behaviour in terms of the density and temporal persistence of their regimes, suggesting that no simple definition of regime structure based solely on these notions is likely to be general enough to capture all of them.

These results justify our suggestion that the notion of a regime in a dynamical system can be understood as the results of varied attempts to capture the non-trivial topology of the underlying attractor. This approach can be obviously adjusted to relate to local regions of phase space only, to account for, e.g., the Euro-Atlantic sector as a particular region in the larger climate attractor. Our methodology shows that besides being an approach which captures a sufficiently wide variety of behaviour, it has the important quality of being computationally tractable for the size of data sets typically used in meteorology and climate science. Furthermore, far from being simply a mathematically neat abstraction, we argue that this topological perspective on regimes offers concrete practical benefits, for two main reasons.

To understand the first reason, it is helpful to recall, as discussed in the introduction, that the *raison d'être* of regimes is to understand questions of predictability across multiple timescales. An overemphasis on properties related to density (as in clustering methods) or temporal persistence may end up obfuscating analysis, not only because regime systems can have a wide variety of behaviour with respect to these notions, but, crucially, because the most salient information may be located in entirely different aspects of the system. The CDV system is an instructive example in this regard. While its classical regimes are associated with fixed points, the most striking impact of these is the tight, looping behaviour it generates (cf. Fig. 20). Knowing that the system is on such a narrowly defined trajectory provides significantly more information than simply knowing that the system is in the vicinity of a fixed-point. From a topological perspective, where no knowledge of fixed-points is implicit, these loops are what stand out as the major feature of CDV, implying that focusing attention on such features can highlight information which is otherwise being overlooked. This potential of topological methods to obtain efficient, simplified representations of chaotic dynamics was also noted in Yalnız and Budanur (2020) using different ideas.

The second reason is the various technical benefits of persistent homology algorithms. Unlike many existing algorithms for regime analysis, such as *K*-means clustering, persistent homology is effectively non-prescriptive. That is, the only parameters required for the algorithm are generic to the system, such as a measure of the spatial scales of the

system, as opposed to parameters that explicitly influence the diagnosed regimes, such as the choice of K in K -means. Homological techniques are therefore particularly well-suited to studying systems where prior knowledge of regime structure is less clear. The ability of our technique to capture the regime behaviour associated to several classical systems lends confidence in its ability to locate relevant structure in such contexts. In addition, the excellent scaling properties of homological algorithms with the dimension of the data means that these algorithms are especially beneficial when analysing very high dimensional data, such as climate data.

There are some important shortcomings to the methodology we have presented, which point to future work. The use of a kernel density estimator, which scales poorly with dimension, to generate our bifiltration, is clearly undesirable and to some extent compromises the excellent scaling obtained from the use of persistent homology. There are several possible avenues of investigation here, including the use of more optimal density estimators; estimating densities using low-dimensional projections (as we did with the CDV data); and even direct Monte Carlo sampling techniques. It is also possible to use computationally cheaper metrics in place of density when subsetting data. For example, preliminary testing suggests that good results can be obtained by only retaining points where the local phase space velocity is ‘small’, and there is precedent for such an approach in the literature (Toth 1992; Straus et al. 2007b). The other key limitation is in the specific software used. As explained in the main text, it would be ideal to replace the crude horizontal density slices we used with more flexible slices of positive slope (cf. Sect. 3.4.1), both to improve the stability of the bifiltration and to enable phenomena like the southern jet latitude ‘shoulder’ to be clearly separated from the central and northern peaks (cf. Sect. 6.1). This would require optimisation of the algorithms used in software like RIVET. Improvements to software capable of producing stable optimal representatives of homology classes (such as PersLoop) will also be necessary in order to allow for confident visualisations of any topological structure detected in atmospheric data.

The apparent subtlety of regime structure gleaned from low-dimensional projections of the atmospheric circulation has been a longstanding source of uncertainty and ambiguity. The topological perspective we present here does not add further insight into such near-Gaussian data sets, as these would be classified as being just a single connected component with no further structure. Instead, if taken at face value, our perspective suggests that the varied approaches to characterising Euro-Atlantic weather regimes are indicative of non-trivial topological structure in the associated region of the climate attractor. In fact, there are tantalising clues in the literature that genuinely non-trivial loops in the attractor might be detectable when taking into account sufficiently

many variables, and that such loops relate to the regime behaviour of the Euro-Atlantic sector (cf. Novak et al. 2017, Figures 4 and 5). It is the hope of these authors that persistent homology may be a tool capable of detecting such topological features in the atmosphere using unprocessed, but very high-dimensional, observational data.

Appendix A: Persistent homology

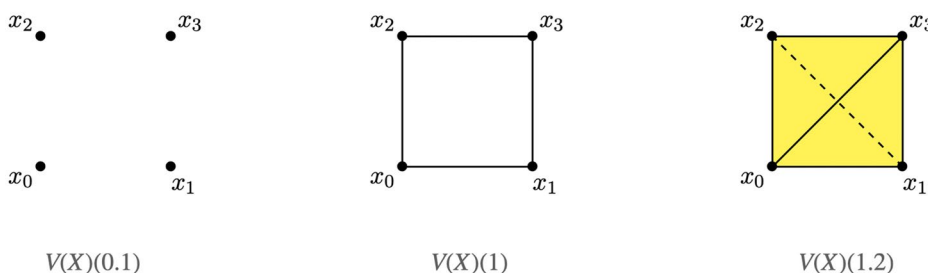
In this section we provide the definition of barcodes for the persistent homology of a finite metric space with respect to the Vietoris–Rips complex. Most of the material in this section only requires a linear algebra background (with the exception of Proposition A.10 for which more advanced algebraic notions are needed). We refer the reader to Otter et al. (2017, Section 4.1) and references therein for alternative definitions of barcodes, as well as further details and intuition on the concepts presented here.

Definition A.1 A **simplicial complex** $K = (V, \Sigma)$ is given by a set V together with a collection Σ of subsets of V satisfying that (i) $\{v\} \in \Sigma$ for all $v \in V$ and (ii) if $\sigma \in \Sigma$ and $\tau \subset \sigma$, then $\tau \in \Sigma$. We call the elements of V the **vertices** of the simplicial complex, while the elements of Σ are called **simplices**. A p -**simplex** is a simplex with cardinality $p + 1$, and we say that p is the **dimension** of such a simplex.

We note that what we call “simplicial complex” is usually called “abstract simplicial complex” in the literature. Every simplicial complex with a finite set of vertices can be realised as a subset of Euclidean space in a canonical way, by identifying the vertices with the standard basic unit vectors, and one can intuitively think of a simplicial complex as a subspace of Euclidean space obtained by gluing together vertices, edges and higher dimensional simplices along their common faces. In what follows we encourage the reader to keep this intuition in mind, and to think of 0-simplices as points in Euclidean space, 1-simplices as closed straight line segments, 2-simplices as triangle-shaped closed convex subspaces, and so on.

Example A.2 Consider the simplicial complex with set of vertices $V = \{a, b, c\}$ and set of simplices $\Sigma = \{\{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}\}$. We can realise (V, Σ) as a subset of \mathbb{R}^3 by identifying the vertex a with the vector $(1, 0, 0)$, the vertex b with the vector $(0, 1, 0)$ and the vertex c with the vector $(0, 0, 1)$. The 1-simplex $\{a, b\}$ is then identified with the straight line segment connecting the points $(1, 0, 0)$ and $(0, 1, 0)$, while the 1-simplex $\{b, c\}$ is identified with the straight line segment connecting the points $(0, 1, 0)$ and $(0, 0, 1)$.

Fig. 25 The Vietoris–Rips simplicial complexes from Example A.4



Definition A.3 Let (X, d) be a finite metric space, and let ϵ be a non-negative real number. The **Vietoris–Rips complex at scale ϵ** is the simplicial complex $V(X)(\epsilon)$ whose set of vertices is given by X , and such that $\sigma \subset X$ is in $V(X)(\epsilon)$ if and only if $d(x_i, x_j) \leq \epsilon$ for all $x_i, x_j \in \sigma$.

Example A.4 Consider $X = \{x_0, x_1, x_2, x_3\}$ with the following distances: $d(x_0, x_1) = 1$, $d(x_0, x_2) = 1$, $d(x_0, x_3) = 1.2$, $d(x_1, x_2) = 1.1$, $d(x_1, x_3) = 0.5$ and $d(x_2, x_3) = 0.6$. We then have that $V(X)(0.1)$ is a simplicial complex with the four 0-simplices $\{x_0\}$, $\{x_1\}$, $\{x_2\}$ and $\{x_3\}$, and with no higher dimensional simplices. On the other hand, $V(X)(1)$ is a simplicial complex with the four 0-simplices $\{x_0\}$, $\{x_1\}$, $\{x_2\}$ and $\{x_3\}$ and the four 1-simplices given by $\{x_0, x_1\}$, $\{x_1, x_3\}$, $\{x_0, x_2\}$ and $\{x_2, x_3\}$. Further, $V(X)(1.2)$ is the simplicial complex with the four 0-simplices $\{x_0\}$, $\{x_1\}$, $\{x_2\}$ and $\{x_3\}$, six 1-simplices $\{x_0, x_1\}$, $\{x_0, x_2\}$, $\{x_0, x_3\}$, $\{x_1, x_2\}$, $\{x_1, x_3\}$ and $\{x_2, x_3\}$, four 2-simplices $\{x_0, x_1, x_2\}$, $\{x_1, x_2, x_3\}$, $\{x_0, x_1, x_3\}$, $\{x_0, x_2, x_3\}$ and one 3-simplex $\{x_0, x_1, x_2, x_3\}$. We provide an illustration for these simplicial complexes in Fig. 25.

As we see in Example A.4, as we increase the scale value we are adding more and more simplices to the simplicial complex. In general, we have that if $\epsilon \leq \epsilon'$ then the set of simplices of $V(X)(\epsilon)$ is contained in that of $V(X)(\epsilon')$.

Definition A.5 Let \mathbb{F}_2 be the field with two elements, and $K = (V, \Sigma)$ a simplicial complex. For each $p = 0, 1, 2, \dots$ we define C_p to be the \mathbb{F}_2 -vector space with basis given by the p -simplices of K . Furthermore, we define linear maps $d_{p+1} : C_{p+1} \rightarrow C_p$ for all $p = 0, 1, 2, \dots$ as follows:

$$d_{p+1}(\sigma) = \sum_{\substack{\tau \in \Sigma_p \\ \text{s.t. } \tau \subset \sigma}} \tau,$$

where σ is any $p + 1$ -simplex in K . We set $d_0 : C_0 \rightarrow 0$ to be the map sending every element of C_0 to 0, where 0 denotes the zero vector space. The collection of vector spaces C_p and linear maps d_p for $p = 0, 1, 2, \dots$ is called the **simplicial chain complex** of K .

Example A.6 We consider the Vietoris–Rips simplicial complex $V(X)(1)$ from Example A.4. We have that C_0 is a vector space of dimension 4 with basis given by the four 0-simplices $\{x_0\}$, $\{x_1\}$, $\{x_2\}$ and $\{x_3\}$, while C_1 is a vector space of dimension 4 with basis given by the four 1-simplices $\{x_0, x_1\}$, $\{x_1, x_3\}$, $\{x_0, x_2\}$ and $\{x_2, x_3\}$. Since there are no simplices of dimension 2 or higher, C_p is the trivial vector space for $p \geq 2$. The map $d_1 : C_1 \rightarrow C_0$ is defined as follows: we have $d_1(\{x_0, x_1\}) = \{x_0\} + \{x_1\}$, $d_1(\{x_1, x_3\}) = \{x_1\} + \{x_3\}$, and similarly $d_1(\{x_0, x_2\}) = \{x_0\} + \{x_2\}$, and $d_1(\{x_2, x_3\}) = \{x_2\} + \{x_3\}$. Keeping the geometric intuition in mind, we can think of d_1 as sending each closed straight-line segment to the sum of the two points in its boundary. Furthermore, we have that $d_p = 0$ for any $p \neq 1$.

As one can easily compute, we have that $d_p \circ d_{p+1} = 0$ for all p , and therefore the image of the map d_{p+1} is contained in the kernel of the map d_p . The elements of the image of d_{p+1} are called p -boundaries, while the elements of the kernel of d_p are called p -cycles. We note that one can more generally define a simplicial chain complex for any field, but one would need to take care of defining the maps d_p differently, to ensure that the composition of any two consecutive maps yields the zero map. In the computations that we perform in this manuscript we use the field with two elements.

Definition A.7 Let p be a natural number. The p **th simplicial homology** of a simplicial complex K is the \mathbb{F}_2 -vector space $H_p(K) = \ker(d_p) / \text{im}(d_{p+1})$. One calls the rank of $H_p(K)$ the p **th Betti number** of K , or alternatively, the number of p -dimensional holes of K . We use the notation $\beta_p(K)$ to denote the p th Betti number of a simplicial complex K , or simply β_p when the simplicial complex is clear from the context.

The p th Betti number thus measures the number of p -cycles that are not p -boundaries. Intuitively, for $p = 1$, a cycle which is not a boundary corresponds to a loop which doesn't contain anything in its interior, in other words a hole. For instance, for the Vietoris–Rips complex $V(X)(1)$ from Example A.6, we have that $\beta_1 = 1$: we

have that $\{x_0, x_1\} + \{x_1, x_3\} + \{x_0, x_2\} + \{x_2, x_3\}$ is in the kernel of d_1 and thus is a 1-cycle, however, it is not in the image of d_2 , since there are no simplices of dimension 2 or higher. On the other hand, if we consider the Vietoris–Rips complex $V(X)(1.1)$, then we have that $\beta_1 = 0$. Here we again have that $\{x_0, x_1\} + \{x_1, x_3\} + \{x_0, x_2\} + \{x_2, x_3\}$ is in the kernel of d_1 , but now it is also the 1-boundary of $\{x_0, x_1, x_2\} + \{x_1, x_2, x_3\}$. We discuss some of these computations in detail in the following example.

Example A.8 We compute p th simplicial homology for the simplicial complex in Example A.6. For $p = 0$, we have that $H_0(V(X)(1)) = C_0/\text{im}(d_1)$. The map d_1 has a 3-dimensional image spanned by $\{x_0\}$, $\{x_1\} + \{x_3\}$ and $\{x_0\} + \{x_1\} + \{x_2\}$. Thus, the quotient $C_0/\text{im}(d_1)$ is 1-dimensional.

For $p = 1$, we have that $H_1(V(X)(1)) = \ker(d_1)/\text{im}(d_2) \cong \ker(d_1)$, since d_2 is the zero map. The kernel of d_1 is 1-dimensional, and is spanned by $\{x_0, x_1\} + \{x_1, x_3\} + \{x_0, x_2\} + \{x_2, x_3\}$. Furthermore, we have that the p th simplicial homology vector space is zero-dimensional for $p \geq 2$. Thus, we have that the simplicial complex $V(X)(1)$ has one 0-hole, one 1-hole, and no higher dimensional holes. A choice of basis vector for $H_0(V(X)(1))$ is given by $[\{x_0\} + \{x_1\} + \{x_2\} + \{x_3\}]$, while a generator for $H_1(V(X)(1))$ is $[\{x_0, x_1\} + \{x_1, x_3\} + \{x_0, x_2\} + \{x_2, x_3\}]$, where the square brackets denote equivalence classes in the respective quotient vector spaces. Intuitively, we can think of the generator of $H_1(V(X)(1))$ as consisting of the “loop” formed by gluing together the four straight line segments corresponding to the 1-simplices along their common boundary points.

Any map of simplicial complexes $f : K \rightarrow K'$, namely a map between the sets of vertices sending simplices to simplices, induces a linear map $H_p(f) : H_p(K) \rightarrow H_p(K')$ between the respective homology vector spaces. Here we are interested in inclusion maps between Vietoris–Rips complexes associated to a metric space X , at different scale values.

By computing simplicial homology of $V(X)(\epsilon)$ for any real number $\epsilon \geq 0$, we obtain what is called a persistence module:

Definition A.9 A **persistence module** is a collection of \mathbb{F}_2 -vector spaces $\{M_\epsilon\}_{\epsilon \in \mathbb{R}_{\geq 0}}$ together with a collection of \mathbb{F}_2 -linear maps $\phi(\epsilon, \epsilon') : M_\epsilon \rightarrow M_{\epsilon'}$ for all $\epsilon \leq \epsilon'$ in $\mathbb{R}_{\geq 0}$ such that $\phi(\epsilon', \epsilon'') \circ \phi(\epsilon, \epsilon') = \phi(\epsilon, \epsilon'')$

whenever $\epsilon \leq \epsilon' \leq \epsilon''$. The linear maps $\phi(\epsilon, \epsilon')$ are called **structure morphisms**.

Computing p th simplicial homology of the Vietoris–Rips complexes at all scales, we obtain a persistence module given by $\{H_p(V(X)(\epsilon))\}_{\epsilon \geq 0}$, together with the structure morphisms $H_p(\iota_{\epsilon, \epsilon'}) : H_p(V(X)(\epsilon)) \rightarrow H_p(V(X)(\epsilon'))$ for all $\epsilon \leq \epsilon'$ which are given by the linear maps induced by the inclusions $\iota_{\epsilon, \epsilon'} : V(X)(\epsilon) \rightarrow V(X)(\epsilon')$. While the p -th Betti numbers capture information about the number of p -dimensional holes at each scale, and for instance the number of connected components, or loops, for $p = 0$ or 1, respectively, p th persistent homology also captures information about how the connected components or loops change as we increase the scale parameter. More precisely, one has that such a persistence module satisfies a finiteness condition, which ensures that the following holds:

Proposition A.10 Let (X, d) be a finite metric space, and let $V(X)(\epsilon)$ be the Vietoris–Rips complex of V at scale ϵ . The persistence module $(\{H_p(V(X)(\epsilon))\}_{\epsilon \geq 0}, \{H_p(\iota_{\epsilon, \epsilon'})\}_{\epsilon \leq \epsilon'})$ is finitely generated, and we have that there exists an $m \in \mathbb{N}$ and unique (up to reindexing) intervals $[b_i, d_i]$ for $i = 1, \dots, m$ such that

$$\bigoplus_{\epsilon \geq 0} H_p(V(X)(\epsilon)) \cong \bigoplus_{i=1}^m \mathbb{I}([b_i, d_i]). \quad (\text{A1})$$

Here, for non-negative reals $a < b$, we denote by $\mathbb{I}([a, b])$ an interval module, namely a persistence module with

$$\mathbb{I}([a, b])(\epsilon) = \begin{cases} \mathbb{F}_2, & \epsilon \in [a, b) \\ 0, & \text{otherwise} \end{cases}$$

and the structure morphisms $\phi_{\epsilon, \epsilon'}$ are given by the identity linear map whenever $a \leq \epsilon \leq \epsilon' < b$, and the zero map otherwise.

Remark A.11 We note that in Eq. (A1) the direct sums of vector spaces are endowed with additional structure that comes from identifications encoded by the structure morphisms: these direct sums are graded modules over certain monoid rings, and the finiteness condition is understood as being a condition on these modules. We refer the reader to Corbet and Kerber (2018) and references therein for more details.

We can understand the decomposition in Eq. A1 as follows: there exists a choice of basis vectors at each scale value ϵ , such that we can represent the information given by the homology vector spaces and linear maps between them in a diagram consisting of disjoint intervals, called a *barcode*. For such a choice of basis vectors, we say that $x \in H_p(V(X)(\epsilon))$ is *born at* ϵ if it is not in the image of $H_p(\iota_{\epsilon', \epsilon})$ for any $\epsilon' < \epsilon$. Similarly, we say

that $0 \neq x \in H_p(V(X)(\epsilon))$ dies at ϵ'' , for $\epsilon'' > \epsilon$ if ϵ'' is the smallest scale value so that $H_p(t_{\epsilon, \epsilon''})(x) = 0$.

In Fig. 5 in the main text we provide an example of barcode: in purple we depict the barcode plot the persistence module $(\{H_1(V(X)(\epsilon))\}_{\epsilon \geq 0}, \{H_1(t_{\epsilon, \epsilon'})\}_{\epsilon \leq \epsilon'})$, and in blue the barcode for the persistence module $(\{H_0(V(X)(\epsilon))\}_{\epsilon \geq 0}, \{H_0(t_{\epsilon, \epsilon'})\}_{\epsilon \leq \epsilon'})$, where X is the finite set of points in Fig. 5a.

Example A.12 We consider again our running example, namely the Vietoris–Rips complexes from Example A.4. We have that the decomposition of $(\{H_0(V(X)(\epsilon))\}_{\epsilon \geq 0}, \{H_0(t_{\epsilon, \epsilon'})\}_{\epsilon \leq \epsilon'})$ consists of the intervals $[0, 0.5)$, $[0, 0.6)$, $[0, 1)$ and $[0, \infty)$. We can interpret each interval in the decomposition as describing the lifetime of one connected component, with the left endpoint representing the scale value at which it first appears, and the right endpoint represents the time at which it merges with another component: at scale 0 the Vietoris–Rips simplicial complex consists only of four 0-simplices, and thus for $\epsilon = 0$ we have four connected components. When we reach $\epsilon = 0.5$ the components corresponding to $\{x_1\}$ and $\{x_3\}$ merge, and thus as a result we have that at this scale level there are only three connected components present. Similarly, we have that at scale 0.6 two further components merge, and similarly at scale value 1, leaving only one connected component.

For $p = 1$, the decomposition of $(\{H_1(V(X)(\epsilon))\}_{\epsilon \geq 0}, \{H_1(t_{\epsilon, \epsilon'})\}_{\epsilon \leq \epsilon'})$ consists of a single interval $[1, 1.1)$. Intuitively, this interval can be thought of as describing the lifetime of the square formed by gluing together the four 1-simplices $\{x_0, x_1\}$, $\{x_1, x_3\}$, $\{x_0, x_2\}$, $\{x_2, x_3\}$ as discussed in Example A.8. The left endpoint of the interval is the lowest scale value at which the square appears in the collection of Vietoris–Rips complexes, while 1.1 is the scale value at which it is filled in by the two 2-simplices $\{x_0, x_1, x_2\}$ and $\{x_1, x_2, x_3\}$.

An alternative way of depicting the intervals in the decomposition in Eq. A1 is what is called a persistence diagram plot: this is a two-dimensional scatter plot in which each interval is represented by the point with coordinates given by its left and right endpoint. We note that in the barcode plot one needs to choose an order to stack the intervals, while in the persistence diagram plot one needs to choose a way to depict the multiplicity of the points.

Appendix B: Description of toy-models

We give further details on the three toy-models used.

Lorenz '63

The Lorenz '63 system, first introduced and studied in Lorenz (1963), is a chaotic dynamical system in three variables x, y, z defined by the equations

$$\begin{aligned}\dot{x} &= \sigma(y - x), \\ \dot{y} &= x(\rho - z) - y, \\ \dot{z} &= xy - \beta z,\end{aligned}$$

and represents a highly simplified model of Rayleigh–Bérnard convection. It was also re-derived as a toy model of the NAO (Molteni and Kucharski 2019). The attractor famously resembles a butterfly, usually viewed as having two regimes corresponding to the two ‘wings’; its regime behaviour has been extensively studied (Palmer 1994; Yadav et al. 2005). Here we use the standard choice of constants σ, β and ρ , namely $\sigma = 10, \beta = 8/3, \rho = 28$. We generate a timeseries of 20,000 points by integrating the equations with a forward Euler scheme at a timestep $dt = 5 \times 10^{-5}$.

Charney–deVore

The Charney–deVore (CDV) model, first derived in Charney and DeVore (1979), provided one of the first examples of multiple invariant measures in an atmospheric model, and can be thought of as a crude model of large-scale midlatitude blocking dynamics. It is based on a severe spectral truncation of the barotropic vorticity equation in a β -plane channel, as in Eq. (B2), where Ψ is a streamfunction, γh is an orographic profile and Ψ^* is an external forcing.

$$\frac{\partial}{\partial t} \nabla^2 \Psi = -J(\Psi, \nabla^2 \Psi + \gamma h) - \beta \frac{\partial \Psi}{\partial x} - C(\Psi - \Psi^*) \quad (\text{B2})$$

While in Charney and DeVore (1979) the main focus is on a three-mode truncation of the system, where a marginally less severe truncation keeping three zonal and two meridional modes is applied, the p.d.e. reduces to the six-equation o.d.e. system shown in Eq. B3, containing quadratic non-linearities and linear coriolis, orographic, and relaxation terms.

$$\begin{aligned}\dot{x}_1 &= \tilde{\gamma}_1 x_3 - C(x_1 - x_1^*) \\ \dot{x}_2 &= \beta_1 x_3 - \alpha_1 x_1 x_3 - \delta_1 x_4 x_6 - C(x_2 - x_2^*) \\ \dot{x}_3 &= -\beta_1 x_2 - \gamma_1 x_1 + \alpha_1 x_1 x_2 + \delta_1 x_4 x_5 - C(x_3 - x_3^*) \\ \dot{x}_4 &= \tilde{\gamma}_2 x_6 + \epsilon \cdot (x_2 x_6 - x_3 x_5) - C(x_4 - x_4^*) \\ \dot{x}_5 &= \beta_2 x_6 - \alpha_2 x_1 x_6 - \delta_2 x_3 x_4 - C(x_5 - x_5^*) \\ \dot{x}_6 &= -\beta_2 x_5 - \gamma_2 x_4 + \alpha_2 x_1 x_5 + \delta_2 x_2 x_4 - C(x_6 - x_6^*)\end{aligned} \quad (\text{B3})$$

A parameter set where this model produces chaotic dynamics was found in Crommelin et al. (2004), and we use those same parameters here (see *ibid* for a full discussion of the

constant values and meaning of each term in Eq. B3). An interactive simulation showing the evolution of this system can be found at https://joshdorrington.github.io/cdv_simulator/.

This system has been introduced as it exhibits multimodality (i.e., regimes) in a model which is significantly more complex and more physically interpretable than the Lorenz '63 system, and which also has particularly challenging phase-space structure. The regime dynamics are of Pomeau–Manneville type (Pomeau and Manneville 1980) in that they consist of long-lived quasi-stationary periods in the vicinity of a weakly unstable fixed point, punctuated by a 'bursting' behaviour and a transition to chaotic flow. These chaotic transients shadow unstable homoclinic orbits radiating from the fixed point, and so lend considerable structure to the model attractor, with a series of strongly preferred looping trajectories. A timeseries of 20,000 points was generated by integrating the equations with a forward Euler scheme at a timestep $dt = 2 \times 10^{-4}$. In order to visualise the data in three dimensions, a truncation of the six dimensional space is required. Because around 98% of the variance is explained by the first three empirical orthogonal functions (EOFs), we use these to define a truncated space. Homological computations were found to be essentially unchanged when using the truncated space or all six dimensions, so the truncated space is used in all computations.

Lorenz '96

The Lorenz '96 model was introduced in Lorenz (1996) as an idealized, chaotic model of the atmosphere which is of greater complexity than the Lorenz '63 system (Karimi and Paul 2010). It is defined in our case by coupling eight variables $X_k, k = 1, \dots, 8$, representing large-scale variability with 32 variables $Y_j, j = 1, \dots, 32$, representing small-scale variability, using the following equations:

$$\begin{aligned} \dot{X}_k &= -X_{k-1}(X_{k-2} - X_{k+1}) - X_k + F \\ &\quad - \frac{hc}{b} \sum_{j=(k-1)+1}^{kJ} Y_j; \quad k = 1, \dots, K, \end{aligned} \quad (\text{B4})$$

$$\begin{aligned} \dot{Y}_j &= -cbY_{j+1}(Y_{j+2} - Y_{j-1}) - cY_j \\ &\quad + \frac{hc}{b} X_{\text{int}[(j-1)/J]+1}; \quad j = 1, \dots, JK. \end{aligned} \quad (\text{B5})$$

Cyclic boundary conditions are then imposed: $X_{k+K} = X_k, Y_{j+JK} = Y_j$. The parameters are chosen as in Christensen et al. (2015), which also discusses the meaning of the different constants.

Due to the interpretation of the equations in terms of large-scale modes coupled to small-scale modes, Lorenz '96 has been utilised in several studies looking at different ways

to parameterise unresolved sub-grid scale variability in forecast systems (Wilks 2005; Christensen et al. 2015; Vissio and Lucarini 2018; Gagne et al. 2020). Its regime structure has been considered in, e.g., Lorenz (2006) and Christensen et al. (2015). The analysis of *ibid* also makes it clear that the key regime variability is concentrated in the first four EOFs. Computations are therefore always done using the subspace spanned by these.

Acknowledgements We thank Sayan Mandal and Tamal Dey for their assistance with the software package PersLoop, as well as Hannah Christensen for sharing Lorenz '96 data and helpful conversations. We thank Mason A. Porter for helpful feedback. We also thank six anonymous reviewers for extensive feedback which helped improve the paper dramatically.

Funding KS was funded by a Thomas Philips and Jocelyn Keene Junior Research Fellowship in Climate Science at Jesus College, Oxford. JD is funded by NERC Grant NE/L002612/1. MC was supported by a grant from the Office of Naval Research Global.

Data Availability Python code supporting the computations of this paper, along with the data used in this study, can be found at <https://github.com/KristianJS/BifiltPH>. An attempt has been made to make the code user friendly and readable, to allow interested readers to apply our methodology to new data sets.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Amini S, Straus DM (2019) Control of storminess over the Pacific and North America by circulation regimes. *Clim Dyn* 52:4749–4770. <https://doi.org/10.1007/s00382-018-4409-7>
- Baur F (1951) Extended-range weather forecasting. In: *Compendium of meteorology*. Springer, pp 814–833
- Bendich P, Marron JS, Miller E, Pieloch A, Skwerer S (2016) Persistent homology analysis of brain artery trees. *Ann Appl Stat* 10:198–218. <https://doi.org/10.1214/15-AOAS886>
- Biasotti S, Cerri A, Frosini P, Giorgi D, Landi C (2008) Multidimensional size functions for shape comparison. *J Math Imaging Vis*
- Bubenik P, Hull M, Patel D, Whittle B (2020) Persistent homology detects curvature. *Inverse Probl* 36:025008
- Carlsson G (2008) Topology and data. *Bull AMS* 46:255–308
- Cassou C (2008) Intraseasonal interaction between the Madden-Julian Oscillation and the North Atlantic Oscillation. *Nature* 455:523–527. <https://doi.org/10.1038/nature07286>

- Charney JG, DeVore JG (1979) Multiple flow equilibria in the atmosphere and blocking. *J Atmos Sci* 36:1205–1216. [https://doi.org/10.1175/1520-0469\(1979\)036<1205:MFEITA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1979)036<1205:MFEITA>2.0.CO;2)
- Charó GD, Chekroun MD, Sciamarella D, Ghil M (2021) Topological effects of noise on nonlinear dynamics
- Christensen HM, Moroz IM, Palmer TN (2015) Simulating weather regimes: impact of stochastic and perturbed parameter schemes in a simple atmospheric model. *Clim Dyn* 44:2195–2214. <https://doi.org/10.1007/s00382-014-2239-9>
- Christiansen B (2007) Atmospheric circulation regimes: can cluster analysis provide the number? *J Clim*. <https://doi.org/10.1175/JCLI4107.1>
- Corbet R, Kerber M (2018) The representation theorem of persistence revisited and generalized. *J Appl Comput Topol* 2:1–31. <https://doi.org/10.1007/s41468-018-0015-3>
- Corti S, Molteni F, Palmer TN (1999) Signature of recent climate change in frequencies of natural atmospheric circulation regimes. *Nature* 398:799–802. <https://doi.org/10.1038/19745>
- Crommelin DT, Opsteegh JD, Verhulst F (2004) A mechanism for atmospheric regime behavior. *J Atmos Sci* 61:1406–1419. [https://doi.org/10.1175/1520-0469\(2004\)061<1406:AMFARB>2.0.CO;2](https://doi.org/10.1175/1520-0469(2004)061<1406:AMFARB>2.0.CO;2)
- Dawson A, Palmer TN, Corti S (2012) Simulating regime structures in weather and climate prediction models. *Geophys Res Lett*. <https://doi.org/10.1029/2012GL053284>
- Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmasda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen I, Kållberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette J-J, Park B-K, Peubey C, de Rosnay P, Tavolato C, Thépaut J-N, Vitart F (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q J R Meteorol Soc* 137:553–597. <https://doi.org/10.1002/qj.828>
- Developers TG. GUDHI, Geometry understanding in higher dimensions. <https://gudhi.inria.fr/>
- Dey TK, Hao T, Mandal S (2018) Persistent 1-cycles: definition, computation, and its application. *Comput Topol Image Context* 2019:123–136
- Dorrington J, Strommen KJ (2020) Jet speed variability obscures Euro-Atlantic regime structure. *Geophys Res Lett* 47:e2020GL087907. <https://doi.org/10.1029/2020GL087907>
- Fabiano F, Christensen HM, Strommen K, Athanasiadis P, Baker A, Schiemann R, Corti S (2020) Euro-Atlantic weather regimes in the PRIMAVERA coupled climate simulations: impact of resolution and mean state biases on model performance. *Clim Dyn*. <https://doi.org/10.1007/s00382-020-05271-w>
- Falkena SK, de Wiljes J, Weisheimer A, Shepherd TG (2020) Revisiting the identification of wintertime atmospheric circulation regimes in the Euro-Atlantic sector. *Q J R Meteorol Soc* 146:2801–2814. <https://doi.org/10.1002/qj.3818>
- Faranda D, Messori G, Yiou P (2017) Dynamical proxies of North Atlantic predictability and extremes. *Sci Rep*. <https://doi.org/10.1038/srep41278>
- Fereday D (2017) How persistent are North Atlantic-European sector weather regimes? *J Clim*. <https://doi.org/10.1175/JCLI-D-16-0328.1>
- Franzke C, Crommelin D, Fischer A, Majda AJ (2008) A hidden Markov model perspective on regimes and metastability in atmospheric flows. *J Clim* 21:1740–1757. <https://doi.org/10.1175/2007JCLI1751.1>
- Franzke C, Horenko I, Majda AJ, Klein R (2009) Systematic metastable atmospheric regime identification in an AGCM. *J Atmos Sci* 66:1997–2012
- Gagne DJ II, Christensen HM, Subramanian AC, Monahan AH (2020) Machine learning for stochastic parameterization: generative adversarial networks in the Lorenz '96 model. *J Adv Model Earth Syst* 12:e2019MS001896
- Gobbino M, Sardella M (1997) On the connectedness of attractors for dynamical systems. *J Differ Equ* 133:1–14. <https://doi.org/10.1006/jdeq.1996.3166>
- Grams CM, Beerli R, Pfenninger S, Staffell I, Wernli H (2017) Balancing Europe's wind-power output through spatial deployment informed by weather regimes. *Nat Clim Change* 7:557–562. <https://doi.org/10.1038/NCLIMATE3338>
- Hannachi A, Iqbal W (2019) Bimodality of hemispheric winter atmospheric variability via average flow tendencies and kernel EOFs. *Tellus Ser A Dyn Meteorol Oceanogr*. <https://doi.org/10.1080/16000870.2019.1633847>
- Hannachi A, Straus DM, Franzke CLE, Corti S, Woollings T (2017) Low-frequency nonlinearity and regime behavior in the Northern Hemisphere extratropical atmosphere. *Rev Geophys* 55:199–234. <https://doi.org/10.1002/2015RG000509>
- Hardiman SC, Dunstone NJ, Scaife AA, Smith DM, Knight JR, Davies P, Claus M, Greatbatch RJ (2020) Predictability of European winter 2019/20: Indian Ocean dipole impacts on the NAO. *Atmos Sci Lett* 21:e1005. <https://doi.org/10.1002/asl.1005>
- Harrington HA, Otter N, Schenck H, Tillmann U (2019) Stratifying multiparameter persistent homology. *SIAM J Appl Algebra Geom* 3:439–471
- Hazell ML (2003) Variable kernel density estimation. *Aust N Z J Stat*. <https://doi.org/10.1111/1467-842X.00283>
- Hurrell JW, Kushnir Y, Otterson G, Visbeck M (2003) An overview of the North Atlantic oscillation. *The North Atlantic Oscillation: Climatic Significance and Environmental Impact* 134:263. <https://doi.org/10.1029/GM134>
- Itoh H, Kimoto M (1996) Multiple attractors and chaotic itinerancy in a quasigeostrophic model with realistic topography: implications for weather regimes and low-frequency variability. *J Atmos Sci* 53:2217–2231
- Jyamiti Research Group: Persloop (2017). <https://github.com/Sayan-m90/Persloop-viewer>
- Karimi A, Paul MR (2010) Extensive chaos in the Lorenz-96 model. *Chaos*. <https://doi.org/10.1063/1.3496397>
- Khasawneh FA, Munch E (2016) Chatter detection in turning using persistent homology. *Mech Syst Signal Process* 70–71:527–541. <https://doi.org/10.1016/j.ymssp.2015.09.046>
- Kramár M, Levanger R, Tithof J, Suri B, Xu M, Paul M, Schatz MF, Mischaikow K (2016) Analysis of Kolmogorov flow and Rayleigh-Bénard convection using persistent homology. *Physica D Nonlinear Phenom*. 334:82–98. <https://doi.org/10.1016/j.physd.2016.02.003>
- Lesnick M, Wright M (2015) Interactive visualization of 2-D persistence modules, arXiv e-prints. [arXiv:1512.00180](https://arxiv.org/abs/1512.00180)
- Lorenz EN (1963) Deterministic nonperiodic flow. *J Atmos Sci* 20:130–141
- Lorenz EN (1996) Predictability: a problem partly solved
- Lorenz EN (2006) Regimes in simple systems. *J Atmos Sci* 1:2056–2073. <https://doi.org/10.1175/JAS3727.1>
- Lucarini V, Gritsun A (2020) A new mathematical framework for atmospheric blocking events. *Clim Dyn* 54:575–598
- Madonna E, Li C, Grams CM, Woollings T (2017) The link between eddy-driven jet variability and weather regimes in the North Atlantic-European sector. *Q J R Meteorol Soc* 143:2960–2972. <https://doi.org/10.1002/qj.3155>
- Majda AJ, Franzke CL, Fischer A, Crommelin DT (2006) Distinct metastable atmospheric regimes despite nearly Gaussian statistics: a paradigm model. *Proc Natl Acad Sci* 103:8309–8314
- Maletić S, Zhao Y, Rajković M (2016) Persistent topological features of dynamical systems. *Chaos*. <https://doi.org/10.1063/1.4949472>
- Marron JS, Wand MP (2007) Exact mean integrated squared error. *Ann Stat*. <https://doi.org/10.1214/aos/1176348653>

- Michelangeli P-A, Vautard R, Legras B (1995) Weather regimes: recurrence and quasi stationarity. *J Atmos Sci* 52:1237–1256. [https://doi.org/10.1175/1520-0469\(1995\)052<1237:WRRASQ>2.0.CO;2](https://doi.org/10.1175/1520-0469(1995)052<1237:WRRASQ>2.0.CO;2)
- Mo KC, Ghil M (1987) Statistics and dynamics of persistent anomalies. *J Atmos Sci* 44:877–902. [https://doi.org/10.1175/1520-0469\(1987\)044<0877:sadopa>2.0.co;2](https://doi.org/10.1175/1520-0469(1987)044<0877:sadopa>2.0.co;2)
- Molteni F, Kucharski F (2019) A heuristic dynamical model of the North Atlantic Oscillation with a Lorenz-type chaotic attractor. *Clim Dyn* 52:6173–6193. <https://doi.org/10.1007/s00382-018-4509-4>
- Muszynski G, Kashinath K, Kurlin V, Wehner M (2019) Topological data analysis and machine learning for recognizing atmospheric river patterns in large climate datasets. *Geosci Model Dev*. <https://doi.org/10.5194/gmd-12-613-2019>
- Novak L, Ambaum MH, Tailleux R (2015) The life cycle of the North Atlantic storm track. *J Atmos Sci* 72:821–833. <https://doi.org/10.1175/JAS-D-14-0082.1>
- Novak L, Ambaum MHP, Tailleux R (2017) Marginal stability and predator-prey behaviour within storm tracks. *Q J R Meteorol Soc* 143:1421–1433. <https://doi.org/10.1002/qj.3014>
- O’Kane TJ, Risbey JS, Franzke C, Horenko I, Monselesan DP (2013) Changes in the metastability of the midlatitude southern hemisphere circulation and the utility of nonstationary cluster analysis and split-flow blocking indices as diagnostic tools. *J Atmos Sci* 70:824–842. <https://doi.org/10.1175/JAS-D-12-028.1>
- Otter N, Porter MA, Tillmann U, Grindrod P, Harrington HA (2017) A roadmap for the computation of persistent homology. *EPJ Data Sci* 6:17
- Palmer TN (1994) Chaos and predictability in forecasting the monsoon. *Proc Indian Natl Sci Acad* 60:57–66
- Palmer TN (1999) A nonlinear dynamical perspective on climate prediction. *J Clim* 12:575–591. [https://doi.org/10.1175/1520-0442\(1999\)012<0575:ANDPOC>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<0575:ANDPOC>2.0.CO;2)
- Parker T, Woollings T, Weisheimer A, O’Reilly C, Baker L, Shaffrey L (2019) Seasonal predictability of the Winter North Atlantic oscillation from a jet stream perspective. *Geophys Res Lett*. <https://doi.org/10.1029/2019GL084402>
- Poli P, Hersbach H, Dee DP, Berrisford P, Simmons AJ, Vitart F, Laloyaux P, Tan DGH, Peubey C, Thépaut J-N, Trémolet Y, Hólm EV, Bonavita M, Isaksen I, Fisher M (2016) ERA-20C: an atmospheric reanalysis of the twentieth century. *J Clim* 29:4083–4097. <https://doi.org/10.1175/JCLI-D-15-0556.1>
- Pomeau Y, Manneville P (1980) Intermittent transition to turbulence in dissipative dynamical systems. *Universality in Chaos*, Second Edition 197:327–335. <https://doi.org/10.1201/9780203734636>
- Radovanovic M, Nanopoulos A, Ivanovic M (2010) Hubs in space: popular nearest neighbors in high-dimensional data. *J Mach Learn Res* 11:2487–2531
- Slingo J, Palmer T (2011) Uncertainty in weather and climate prediction. *Philos Trans R Soc A Math Phys Eng Sci* 369:4751–4767. <https://doi.org/10.1098/rsta.2011.0161>
- Stanley GJ (2019) Neutral surface topology. *Ocean Model* 138:88–106. <https://doi.org/10.1016/j.ocemod.2019.01.008>
- Stephenson DB, Hannachi A, O’Neill A (2004) On the existence of multiple climate regimes. *Q J R Meteorol Soc* 130:583–605
- Straus DM (2010) Synoptic-eddy feedbacks and circulation regime analysis. *Mon Weather Rev* 138:4026–4034. <https://doi.org/10.1175/2010MWR3333.1>
- Straus DM, Corti S, Molteni F (2007a) Circulation regimes: chaotic variability versus SST-forced predictability. *J Clim* 20:2251–2272. <https://doi.org/10.1175/JCLI4070.1>
- Straus DM, Corti S, Molteni F (2007b) Circulation regimes: chaotic variability versus SST-forced predictability. *J Clim* 20:2251–2272
- Strommen K (2020) Jet latitude regimes and the predictability of the North Atlantic Oscillation. *Q J R Meteorol Soc* 146:2368–2391. <https://doi.org/10.1002/qj.3796>
- The RIVET Developers: RIVET (2020). <https://github.com/rivetTDA/rivet/>
- Toth Z (1992) Quasi-stationary and transient periods in the Northern Hemisphere circulation series. *J Clim* 5:1235–1247
- Tymochko S, Munch E, Dunion J, Corbosiero K, Torn R (2020) Using persistent homology to quantify a diurnal cycle in hurricanes. *Pattern Recognit Lett* 133:137–143. <https://doi.org/10.1016/j.patrec.2020.02.022>
- Vautard R (1990) Multiple weather regimes over the North Atlantic: analysis of precursors and successors. *Mon Weather Rev* 118:2056–2081. [https://doi.org/10.1175/1520-0493\(1990\)118<2056:MWROTN>2.0.CO;2](https://doi.org/10.1175/1520-0493(1990)118<2056:MWROTN>2.0.CO;2)
- Vipond O (2020) Multiparameter persistence landscapes. *J Mach Learn Res* 21:1–38
- Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt SJ, Brett M, Wilson J, Millman KJ, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey CJ, Polat İ, Feng Y, Moore EW, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P (2020) SciPy 1.0 contributors: SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 17:261–272. <https://doi.org/10.1038/s41592-019-0686-2>
- Vissio G, Lucarini V (2018) A proof of concept for scale-adaptive parametrizations: the case of the Lorenz ’96 model. *Q J R Meteorol Soc* 144:63–75. <https://doi.org/10.1002/qj.3184>
- Wilks DS (2005) Effects of stochastic parametrizations in the Lorenz ’96 system. *Q J R Meteorol Soc* 131:389–407. <https://doi.org/10.1256/qj.04.03>
- Woollings T, Hoskins B, Blackburn M, Berrisford P (2008) A new Rossby wave-breaking interpretation of the North Atlantic oscillation. *J Atmos Sci* 65:609–626. <https://doi.org/10.1175/2007JAS2347.1>
- Woollings T, Hannachi A, Hoskins B (2010a) Variability of the North Atlantic eddy-driven jet stream. *Q J R Meteorol Soc* 136:856–868. <https://doi.org/10.1002/qj.625>
- Woollings T, Hannachi A, Hoskins B, Turner A (2010b) A regime view of the North Atlantic oscillation and its response to anthropogenic forcing. *J Clim* 23:1291–1307. <https://doi.org/10.1175/2009JCLI3087.1>
- Yadav RS, Dwivedi S, Mittal AK (2005) Prediction rules for regime changes and length in a new regime for the Lorenz model. *J Atmos Sci* 62:2316–2321. <https://doi.org/10.1175/JAS3469.1>
- Yalnız G, Budanur NB (2020) Inferring symbolic dynamics of chaotic flows from persistence. *Chaos* 30:033109. <https://doi.org/10.1063/1.5122969>

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.