Testing topological conjugacy of time series *

Paweł Dłotko †, Michał Lipiński † ‡, and Justyna Signerska-Rynkowska † §

Abstract. This paper considers a problem of testing, from a finite sample, a topological conjugacy of two trajectories coming from dynamical systems (X, f) and (Y, g). More precisely, given $x_1, \ldots, x_n \subset X$ and $y_1, \ldots, y_n \subset Y$ such that $x_{i+1} = f(x_i)$ and $y_{i+1} = g(y_i)$ as well as $h : X \to Y$, we deliver a number of tests to check if f and g are topologically conjugated via h. The values of the tests are close to zero for systems conjugate by h and large for systems that are not. Convergence of the test values, in case when sample size goes to infinity, is established. We provide a number of numerical examples indicating scalability and robustness of the presented methods. In addition, we show how the presented method gives rise to a test of sufficient embedding dimension, mentioned in Takens' embedding theorem. Our methods also apply to the situation when we are given two observables of deterministic processes, of a form of one or higher dimensional time-series. In this case, their similarity can be accessed by comparing the dynamics of their Takens' reconstructions. Finally, we include a proof-of-concept study using the presented methods to search for an approximation of the homeomorphism conjugating given systems.

Key words. conjugacy, semiconjugacy, embedding, nonlinear time-series analysis, false nearest neighbors, similarity measures, k-nearest neighbors

MSC codes. Primary: 37M10, 37C15, Secondary: 65P99, 65Q306

1. Introduction.

Understanding sampled dynamics is of primal importance in multiple branches of science where there is a lack of solid theoretical models of the underlying phenomena [9, 10, 18, 20, 31]. It delivers a foundation for various equation–free models of observed dynamics and allows to draw conclusions about the unknown observed processes. In the considered case we start with two, potentially different, phase spaces X and Y and a map $h : X \to Y$. Given two sampled trajectories, referred to in this paper by *time series*, $x_1, \ldots, x_n \subset X$ and $y_1, \ldots, y_n \subset Y$ we assume that they are both generated by a continuous maps $f : X \to X$ and $g : Y \to Y$[1] in a way that $x_{i+1} = f(x_i)$ and $y_{i+1} = g(y_i)$. In what follows, we build a number of tests that allow to distinguish trajectories that are conjugated by the given map h from those that are not. It should be noted that the problem of finding

* **Funding:** The authors gratefully acknowledge the support of Dioscuri program initiated by the Max Planck Society, jointly managed with the National Science Centre (Poland), and mutually funded by the Polish Ministry of Science and Higher Education and the German Federal Ministry of Education and Research. J S-R was also supported by National Science Centre (Poland) grant 2019/35/D/ST1/02253.

† Dioscuri Centre in TDA, Institute of Mathematics, Polish Academy of Sciences, ul. Śniadeckich 8, 00-656 Warszawa, Poland, (pdlotko@impan.pl, michal.lipinski@impan.pl, j.signerska@impan.pl) (https://dioscuri-tda.org/members/pawel.html, https://sites.google.com/view/michal-lipinski, https://dioscuri-tda.org/members/justyna.html).

‡ Institute of Computer Science and Computational Mathematics, Jagiellonian University, ul. Łojasiewicza 6, 30-348 Kraków, Poland

§ Institute of Applied Mathematics, Gdańsk University of Technology, ul. Narutowicza 11/12, 80-233 Gdańsk, Poland

[1] For finite samples those maps always exist assuming that $x_i = x_j$ and $y_i = y_j$ if and only if $i = j$.

an appropriate h for two conjugated dynamical system is in general very difficult and goes beyond the scope of this paper. However, in Section 5 and in Appendix A we propose and validate a method of approximating map h for one dimensional systems f and g utilizing one of the proposed statistics.

The presented problem is practically important for the following reasons. Firstly, the proposed machinery allows to test for conjugacy, in case when the formulas that generate the underlying dynamics, as f and g above, are not known explicitly, and the input data are based on observations of the considered system.

Secondly, some of the presented methods apply in the case when the dynamics f and g on X and Y is explicitly known, but we want to test if a given map $h : X \to Y$ between the phase spaces has a potential to be a topological conjugacy. It is important as the theoretical results on conjugacy are given only for a handful of systems and our methods give a tool for numerical hypothesis testing.

Thirdly, those methods can be used to estimate the optimal parameters of the dynamics reconstruction. A basic way to achieve such a reconstruction is via time delay embedding, a technique that depends on parameters including the *embedding dimension* and the *time lag* (or *delay*). When the parameters of the method are appropriately set up and the assumptions of Takens' Embedding Theorem hold (see [29, 7]), then (generically) a reconstruction is obtained, meaning that for *generic* dynamical system and *generic* observable, the delay-coordinate map produces a *conjugacy* (dynamical equivalence) between reconstructed dynamics and the original (unknown) dynamics (cut to the limit set of a given trajectory)[2]. However, without the prior knowledge of the underlying dynamics (e.g. dimensions of the attractor), the values of those parameters have to be determined experimentally from the data. It is typically achieved by implicitly testing for a conjugacy of the time delay embeddings to spaces of constitutive dimensions. Specifically, it is assumed that the optimal dimension of reconstruction d is achieved when there is no conjugacy of the reconstruction in dimension d to the reconstruction in the dimension d', where $d' < d$, while there is a conjugacy between reconstruction in dimension d and reconstruction in dimension d'', where $d < d''$. Those conditions can be tested with methods presented in this paper.

The main contributions of this paper include:

- We propose a generalization of the FNN (*False Nearest Neighbor*) method [13] so that it can be applied to test for topological conjugacy of time series[3]. Moreover, we present its further modification called KNN method.
- We propose two entirely new methods: ConjTest and ConjTest$^+$. Instead of providing an almost binary answer to a question if two sampled dynamical systems are conjugate (which happens for the generalized FNN and the KNN method), their result is a continuous variable that can serve as a scale of similarity of two dynamics. This

[2] One should, though, be aware that the *generic set* in the classical Takens' Embedding Theorem might be a set of a small measure. However, recent advances in probabilistic versions of Takens' Theorem ([3]) assert that, under even milder assumptions, the delay-coordinate map provides injective (not necessary conjugacy) correspondence between the points of the original system in the subset of full measure and the points in the reconstructed space.

[3] Classical FNN method was used only to estimate the embedding dimension in a dynamics reconstruction using time delay embedding.

property makes the two new methods appropriate for noisy data.

- We present a number of benchmark experiments to test the presented methods. In particular we analyze how different methods are robust for the type of testing (e.g. noise, determinism, alignment of a time series).
- Additionally, in one dimensional setting, we propose a heuristic method for approximating the possible conjugating homeomorphism between two dynamical systems given by time series.

To the best of our knowledge there are no “explicit” methods available to test conjugacy of dynamical systems given by their finite sample in a form of time series as proposed in this paper. A number of methods exist to estimate the parameters of a time delay embedding. They include, among others, mutual information [12], autocorrelation and higher order correlations [1], a curvature-based approach [11] or wavering product [8] for selecting the time-lag, selecting of embedding dimension based on GP algorithm [2] or the above mentioned FNN algorithm, as well as some methods allowing to choose the embedding dimension and the time lag simultaneously as, for example, C-C method based on correlation integral [17], methods based on symbolic analysis and entropy [19] or some rigorous statistical tests [24]. However, the problem of topological conjugacy between the maps generating two given time series and finding the connecting homeomorphism which conjugates the two dynamical systems, due to its complexity, has been mainly approached using machine learning tools (see e.g. the recent work [6] which for the unknown map f and given time series generated by f, employed deep neural network for discovering the simple map g which could model the unknown dynamics f together with the map h conjugating f and g). Some theoretical ideas on finding conjugating homeomorphism (or, in general, a *commuter* between two maps) are discussed later in Section 5 together with related works.

Numerous methods providing some similarity measures between time series exist (see reviews [16]). However, we claim that those classical methods are not suitable for the problem we tackle in this paper. While those methods often look for an actual similarity of signals or correlation, we are more interested in the dynamical generators hiding behind the data. For instance, two time series sampled from the same chaotic system can be highly uncorrelated, yet we would like to recognize them as similar, because the dynamical system constituting them is the same. Moreover, methods introduced in this work are applicable for time series embedded in any metric space, while most of the methods are restricted to $\mathbb{R}$, some of them are still useful in $\mathbb{R}^d$.

The paper consists of four parts: Section 2 introduces the basic concepts behind the proposed methods. Section 3 presents four methods designed for data-driven evaluation of conjugacy of two dynamical systems. Section 4 explores the features of the proposed methods using a number of numerical experiments. Section 5 develops the method of estimating the possible conjugacy map $h : X \to Y$ for time series generated from dynamical systems (X, f) and (Y, g) in the case when the phase spaces X and Y are intervals in $\mathbb{R}$. Additional details of that procedure and proofs are contained in A. Lastly, in Section 6 we summarize most important observations and discuss their possible significance in real-world time series analysis.

Finally, it should be noted that in the continuous setting, topological conjugacy is very fragile; it may be destroyed by an infinitesimal change of parameters of the system once that causes bifurcation. However, two finite sample of the trajectories obtained from the system before and after bifurcation are very close and it would require much large change of parameters to detect problems with conjugacy. It is a consequence of the fact that the techniques proposed in this paper operates on finite data. Therefore, they can provide evidences that the proposed connecting homeomorphism is not a topological conjugacy of the two considered systems, but they will not allow to prove, in any rigorous sense, the conjugacy between them.

2. Preliminaries.

2.1. Topological conjugacy. We start with a pair of metric spaces X and Y and a pair of dynamical systems: $\varphi : X \times \mathbb{T} \to X$ and $\psi : Y \times \mathbb{T} \to Y$, where $\mathbb{T} \in \{\mathbb{Z}, \mathbb{R}\}$. Fixing $t_X, t_Y \in \mathbb{T}$ define $f : X \ni x \to \varphi(x, t_X)$ and $g : Y \ni y \to \psi(y, t_Y)$. We say that f and g are *topologically conjugate* if there exists a homeomorphism $h : X \to Y$ such that the diagram

$$\begin{array}{ccc} X & \xrightarrow{f} & X \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle h} \\ Y & \xrightarrow{g} & Y \end{array} \tag{2.1}$$

commutes, i.e., $h \circ f = g \circ h$. If the map $h : X \to Y$ is not a homeomorphism but a continuous surjection then we say that g is *topologically semiconjugate* to f.

Let us consider as an example X being a unit circle, and f_α a rotation of X by an angle α. In this case, two maps, $f_\alpha, f_\beta : X \to X$ are conjugate if and only if $\alpha = \beta$ or $\alpha = -\beta$. This known fact is verified in the benchmark test in Section 4.1.

In our work we will consider finite time series $\mathcal{A} = \{x_i\}_{i=1}^n$ and $\mathcal{B} = \{y_i\}_{i=1}^n$ so that $x_{i+1} = f^i(x_1)$ and $y_{i+1} = g^i(y_1)$ for $i \in \{1, 2, \ldots, n-1\}$, $x_1 \in X$ and $y_1 \in Y$ and derive criteria to test (semi)topological conjugacy of f and g via h based on those samples and the given possible (semi)conjugacy h.

In what follows, a Hausdorff distance between $A, B \subset X$ will be used. It is defined as

$$\mathrm{d_H}(A, B) = \max\{\sup_{a \in A} d(a, B), \sup_{b \in B} d(b, A)\}$$

where d is metric in X and $d(x, A) := \inf_{a \in A} d(x, a)$.

2.2. Takens' Embedding Theorem. Our work is related to the problem of reconstruction of dynamics from one dimensional time series. For a fixed map $f : X \to X$ and $x_1 \in X$ take a time series $\mathcal{A} = \{x_i = f^{i-1}(x_1)\}_{i \geq 1}$ being a subset of an attractor $\Omega \subset X$ of the (box-counting) dimension m. Take $s : X \to \mathbb{R}$, a generic measurement function of observable states of the system, and one dimensional time series $\mathcal{S} = \{s(x_i)\}_{x_i \in \mathcal{A}}$, associated to $\mathcal{A}$. The celebrated Takens' Embedding Theorem [29] states that given $\mathcal{S}$ it is possible to reconstruct the original system with delay vectors, for instance $(s(x_i), s(x_{i+1}), \ldots, s(x_{i+d-1}))$, for sufficiently large *embedding dimension* $d \geq 2m + 1$ (the bound is often not optimal). The Takens' theorem implies that, under certain generic assumptions, an embedding of the attractor Ω into $\mathbb{R}^d$ given by

$$F_{s,f} : \Omega \ni x \mapsto \Big(s(x), s(f(x)), \ldots, s(f^{d-1}(x))\Big) \in \mathbb{R}^d \tag{2.2}$$

establishes a *topological conjugacy* between the original system (Ω, f) and $(F_{s,f}(\Omega), \sigma)$ with the dynamics on $F_{s,f}(\Omega) \subset \mathbb{R}^d$ given by the shift σ on the sequence space. Hence, Takens' Embedding Theorem allows to reconstruct both the topology of the original attractor and the dynamics.

The formula presented above is a special case of a reconstruction with a *lag* l given by

$$\Pi(\mathcal{A}, d, l) := \left\{ (s(x_i), s(x_{i+l}), \ldots, s(x_{i+(d-1)l})) \mid i \in \{1, 2, \ldots, n - d\,l\} \right\}.$$

From the theoretical point of view, the Takens' theorem holds for an arbitrary lag. However in practice a proper choice of l may strongly affect numerical reconstructions (see [14, Chapter 3]).

The precise statements, interpretations and conclusions of the mentioned theorems can be found in [7, 29, 25], and references therein.

2.3. Search for an optimal dimension for reconstruction. In practice, the bound in Takens' theorem is often not sharp and an embedding dimension less than $2m+1$ is already sufficient to reconstruct the original dynamics (see [3, 4]). Moreover, for time series encountered in practice, the attractor's dimension m is almost always unknown. To discover the sufficient dimension of reconstruction, the False Nearest Neighbor (FNN) method [13, 15], a heuristic technique for estimating the optimal dimension using a finite time series, is typically used. It is based on an idea to compare the embeddings of a time series into a couple of consecutive dimensions and to check if the introduction of an additional $d+1$ dimension separates some points that were close in d-dimensional embedding. Hence, it tests whether d-dimensional neighbors are (false) neighbors just because of the tightness of the d-dimensional space. The dimension where the value of the test stabilizes and no more false neighbors can be detected is proclaimed to be the optimal embedding dimension.

2.4. False Nearest Neighbor and beyond. The False Nearest Neighbor method implicitly tests semiconjugacy of d and $d+1$ dimensional Takens' embedding by checking if the neighborhood of d-embedded points are preserved in $d+1$ dimension. This technique was an inspiration for stating a more general question: given two time series, can we test if they were generated from conjugate dynamical systems? The positive answer could suggest that the two observed signals were actually generated by the same dynamics, but obtained by a different measurement function. In what follows, a number of tests inspired by these observations concerning False Nearest Neighbor method and Takens' Embedding Theorem, are presented.

3. Conjugacy testing methods.

In this section we introduce a number of new methods for quantifying the dynamical similarity of two time series. Before digging into them let us introduce some basic pieces of notation used throughout the section. From now on we assume that X is a metric space. Let $\mathcal{A} = \{x_i\}_{i=1}^n$ be a finite time series in space X. For $k \in \mathbb{N}$, by $\kappa(x, k, \mathcal{A})$ we denote the set of k-*nearest neighbors* of a point $x \in X$ among points in $\mathcal{A}$. Thus, the nearest neighbor of point x can be denoted by $\kappa(x, \mathcal{A}) := \kappa(x, 1, \mathcal{A})$. If $x \in \mathcal{A}$ then clearly $\kappa(x, \mathcal{A}) = \{x\}$. Hence, it is handful to consider also $\overline{\kappa}(x, k, \mathcal{A}) := \kappa(x, k, \mathcal{A} \setminus \{x\})$ and $\overline{\kappa}(x, \mathcal{A}) := \kappa(x, 1, \mathcal{A} \setminus \{x\})$[4].

[4] In case of non uniqueness, and arbitrary choice of a neighbor is made.

3.1. False Nearest Neighbor method. The first proposed method is an extension of the already mentioned FNN technique for estimating the optimal embedding dimension of time series. The idea of the classical FNN method relies on counting the number of so-called false nearest neighbors depending on the threshold parameter r. This is based on the observation that if the two reconstructed points

$$\mathbf{s}^1_d := (s(x_{k_1}), s(x_{k_1+l}), \ldots, s(x_{k_1+(d-1)l}))$$

and

$$\mathbf{s}^2_d := (s(x_{k_2}), s(x_{k_2+l}), \ldots, s(x_{k_2+(d-1)l}))$$

are nearest neighbors in the d-dimensional embedding but the distance between their (d+1)-dimensional counterparts

$$\mathbf{s}^1_{d+1} := (s(x_{k_1}), \ldots, s(x_{k_1+(d-1)l}), s(x_{k_1+dl}))$$

and

$$\mathbf{s}^2_{d+1} := (s(x_{k_2}), \ldots, s(x_{k_2+(d-1)l}), s(x_{k_2+dl}))$$

in $(d{+}1)$-dimensional embedding differs too much, then $\mathbf{s}^1_d$ and $\mathbf{s}^2_d$ were d-dimensional neighbors only due to folding of the space. In this case, we will refer to them as "false nearest neighbors". Precisely, the ordered pair $(\mathbf{s}^1_d, \mathbf{s}^2_d)$ of d-dimensional points is counted as false nearest neighbor, if the following conditions are satisfied: (I.) the point $\mathbf{s}^2_d$ is the closest point to $\mathbf{s}^1_d$ among all points in the d-dimensional embedding, (II.) the distance $|\mathbf{s}^1_d - \mathbf{s}^2_d|$ between the points $\mathbf{s}^1_d$ and $\mathbf{s}^2_d$ is less than σ/r, where σ is the standard deviation of d-dimensional points formed from delay-embedding of the time series and (III.) the ratio between the distance $|\mathbf{s}^1_{d+1} - \mathbf{s}^2_{d+1}|$ of $d+1$-dimensional counterparts of these points, $\mathbf{s}^1_{d+1}$ and $\mathbf{s}^2_{d+1}$, and the distance $|\mathbf{s}^1_d - \mathbf{s}^2_d|$ is greater than the threshold r. The condition (III.) is motivated by the fact that under continuous evolution, even if the original dynamics is chaotic, the position of two close points should not deviate too much in the nearest future (we assume that the system is deterministic, even if subjected to some noise, which is the main assumption of all the nonlinear analysis time series methods). On the other hand, the condition (II.) means that we consider only pairs of points which are originally not too far away since applying the condition (III.) to points which are already outliers in d dimensions does not make sense. Next, the statistic $\mathrm{FNN}(r)$ counts the relative number of such false nearest neighbors i.e. after normalizing with respect to the number of all the ordered pairs of points which satisfy (I.) and (II.). For discussion and some examples see e.g. [14].

We generalize the FNN method to operate in the case of two time series (not necessarily created in a time-delay reconstruction) as follows. Let $\mathcal{A} = \{a_i\}_{i=1}^n \subset X$ and $\mathcal{B} = \{b_i\}_{i=1}^n \subset Y$ be two time series of the same length. Let $\xi : \mathcal{A} \to \mathcal{B}$ be a bijection relating points with the same index, i.e., $\xi(a_i) := b_i$. Then we define the directed FNN ratio between $\mathcal{A}$ and $\mathcal{B}$ as

$$\mathrm{FNN}(\mathcal{A}, \mathcal{B}; r) := \frac{\sum_{i=1}^n \Theta\left(\frac{\mathbf{d}_Y(b_i, \xi(\overline{\kappa}(a_i, \mathcal{A})))}{\mathbf{d}_X(a_i, \overline{\kappa}(a_i, \mathcal{A}))} - r\right) \Theta\left(\frac{\sigma}{r} - \mathbf{d}_X(a_i, \overline{\kappa}(a_i, \mathcal{A}))\right)}{\sum_{i=1}^n \Theta\left(\frac{\sigma}{r} - \mathbf{d}_X(a_i, \overline{\kappa}(a_i, \mathcal{A}))\right)} \tag{3.1}$$

where $\mathbf{d}_X$ and $\mathbf{d}_Y$ denote the distance function respectively in X and Y, σ is the standard deviation of the data (i.e. the standard deviation of the elements of the sequence $\mathcal{A}$), r is the parameter of the method and Θ is the usual Heaviside step function, i.e. $\Theta(x) = 1$ if $x > 0$ and 0 otherwise. Note that the distance $\mathbf{d}$ (i.e. $\mathbf{d}_X$ or $\mathbf{d}_Y$) might be defined in various ways however, as elements of time-series are usually elements of $\mathbb{R}^k$ (for some k), then $\mathbf{d}(x, y)$ is often simply the Euclidean norm $|x - y|$.

In the original FNN procedure we compare embeddings of a 1-dimensional time series $\mathcal{A}$ into d- versus $(d+1)$-dimensional space for a sequence of values of d and r. In particular, the following application of (3.1):

$$\text{FNN}(\mathcal{A}; r, d) := \text{FNN}(\Pi_d(\mathcal{A}), \Pi_{d+1}(\mathcal{A}); r), \tag{3.2}$$

coincides with the formula used in the standard form of FNN technique (compare with [14]). For a fixed value of d, if the values of FNN decline rapidly with the increase of r, then we interpret that dimension d is large enough not to introduce any artificial neighbors. The heuristic says that the lowest d with that property is the optimal embedding dimension for time series $\mathcal{A}$.

3.2. K-Nearest Neighbors. The key to the method presented in this section is an attempt to weaken and simplify the condition posed by FNN by considering a larger neighborhood of a point. As in the previous case, let $\mathcal{A} = \{a_i\}_{i=1}^n$ and $\mathcal{B} = \{b_i\}_{i=1}^n$ be two time series of the same length. Let $\xi : \mathcal{A} \to \mathcal{B}$ be a bijection defined $\xi(a_i) := b_i$. The proposed statistics, taking into account k nearest neighbors of each point, is given by the following formula:

$$\text{KNN}(\mathcal{A}, \mathcal{B}; k) := \frac{\sum_{i=1}^n \min\left\{e \in \mathbb{N} \;\middle|\; \xi\left(\overline{\kappa}(a_i, k, \mathcal{A})\right) \subseteq \overline{\kappa}(\xi(a_i), e + k, \mathcal{B})\right\}}{n^2}, \tag{3.3}$$

where n is the length of time series $\mathcal{A}$ and $\mathcal{B}$. We refer to the above method as KNN distance. The idea of the KNN method can be seen in the Figure 1.

Remark 3.1. In the above formula (3.3), for simplicity there is no counterpart of the parameters r that was present in FNN which controlled the dispersion of data and outliers. This means that one should assume that the data (perhaps after some preprocessing) does not contain unexpected outliers. Alternatively, the formula might be easily modified to include such a parameter.

Set $\overline{\kappa}(a_i, k, \mathcal{A})$ can be interpreted as a discrete approximation of the neighborhood of a_i. Thus, for a point a_i the formula measures how much larger neighborhood of the corresponding point $b_i = \xi(a_i)$ we need to take to contain the image of the chosen neighborhood of a_i. This discrepancy is expressed relatively to the size of the point cloud. Next we compute the average of this relative discrepancy among all points. Moreover, looking at the formula (3.3) immediately reveals that in the numerator we sum up n terms each of which takes values between 0 and n and it is not hard to give an example when all of these terms are n actually (like a standard n-simplex). Therefore, as we want KNN to be upper-bounded by 1, we put n^2 in the denominator of (3.3) as the normalization factor.

Note that neither f nor g appear in the definitions of FNN and KNN. Nevertheless, the dynamics is hidden in the indices. That is, $a_j \in \overline{\kappa}(a_i, k, \mathcal{A})$ means that a_i returns to its own vicinity in $|j - i|$ time steps.

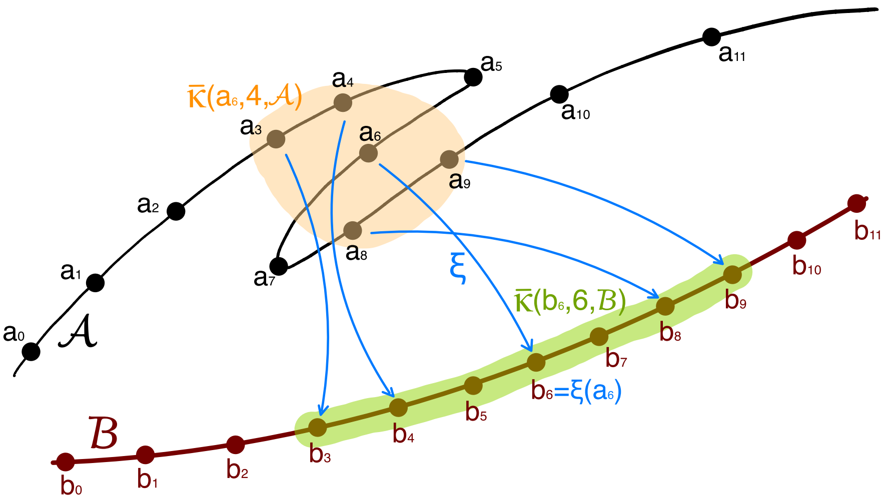

Figure 1. *Top (continuous) black line represents trajectory from which $\mathcal{A}$ is sampled (black dots). Bottom (continuous) trajectory is sampled to obtain $\mathcal{B}$ (burgundy dots). Set $U := \overline{\kappa}(a_6, 4, \mathcal{A})$ highlighted with orange color, represents 4-nearest neighbors of $a_6 \in \mathcal{A}$. The smallest k-neighborhood of b_6 that contains $\xi(U)$ is the one with $k = 6$. The corresponding $\overline{\kappa}(b_6, 4, \mathcal{B})$ is highlighted with green color. Hence, the contribution of point a_6 to the numerator of* $\mathrm{KNN}(\mathcal{A}, \mathcal{B}, 4)$ *is* $6 - 4 = 2$.

3.3. Conjugacy test. The third method tests the conjugacy of two time series by directly checking the commutativity of the diagram (2.1) which is tested in a more direct way compared to the methods presented so far. We no longer assume that both time series are of the same size, however, the method requires a *connecting map* $h : X \to Y$, a candidate for a (semi)conjugating map. Unlike the map ξ in FNN and KNN method map h may transform a point $a_i \in \mathcal{A}$ into a point in Y that doesn't belong to $\mathcal{B}$. Nevertheless, the points in $\mathcal{B}$ are crucial because they carry the information about the dynamics $g : Y \to Y$. Thus, in order to follow trajectories of points in Y we introduce $\tilde{h} : \mathcal{A} \to \mathcal{B}$, a discrete approximation of h:

$$\tilde{h}(a_i) := \kappa\left(h(a_i), \mathcal{B}\right). \tag{3.4}$$

The map $\tilde{h}$ simply assigns to a_i the closest element(s) of $h(a_i)$ from the time series $\mathcal{B}$. For a set $A \subset \mathcal{A}$ we compute the value pointwise, i.e. $\tilde{h}(A) = \{\tilde{h}(a) \mid a \in A\}$ (see Figure 2). Note that it may happen that $\tilde{h}(A)$ has less elements than A.

Denote the discrete k-approximation of the neighborhood of a_i in $\mathcal{A}$, namely the k nearest neighbors of a_i, by $U_i^k := \kappa(a_i, k, \mathcal{A}) \subset \mathcal{A}$. Then we define

$$\mathrm{ConjTest}(\mathcal{A}, \mathcal{B}; k, t, h) := \frac{\sum_{i=1}^{n} \mathrm{d_H}\left((h \circ f^t)(U_i^k),\ (g^t \circ \tilde{h})(U_i^k)\right)}{n\,\mathrm{diam}(\mathcal{B})}, \tag{3.5}$$

where $\mathrm{d_H}$ is the Hausdorff distance between two discrete sets and $\mathrm{diam}(\mathcal{B})$ is the diameter of the set $\mathcal{B}$. The idea of the formula (3.5) is to test at every point $a_i \in \mathcal{A}$ how two time series together with map h are close to satisfy diagram (2.1) defining topological conjugacy. First, we approximate the neighborhood of $a_i \in \mathcal{A}$ with U_i^k and then we try to traverse the

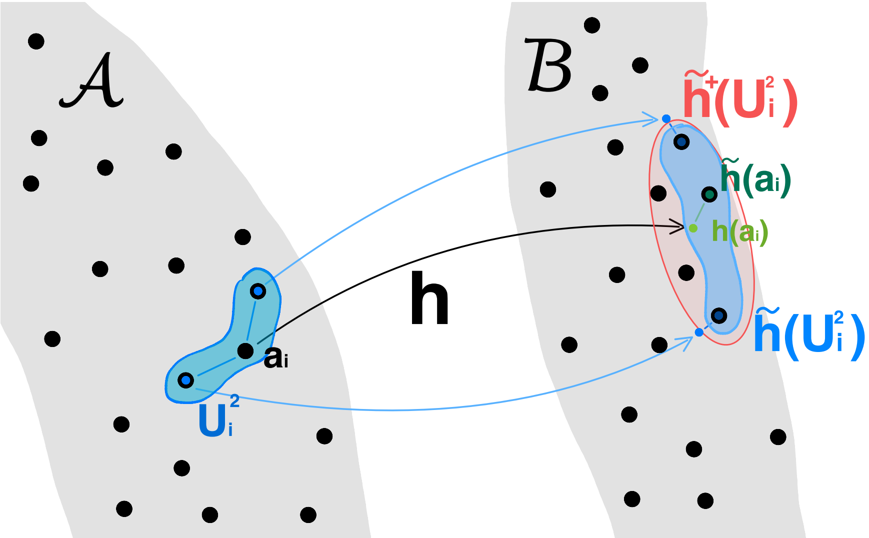

Figure 2. *A pictorial visualization of a difference between h, $\tilde{h}$ and $\tilde{h}^+$. Map h transforms a point $a \in \mathcal{A}$ into a point $h(a) \in Y$. Map $\tilde{h}$ approximates the value of the map h by finding the closest point in $\mathcal{B}$ for $h(a)$. The discrete neighborhood $U_i^2 \subset \mathcal{A}$ of a_i consists of three points and its image under $\tilde{h}$ has three points as well. However, $\tilde{h}^+(U_i^2)$ counts five elements, as there are points in $\mathcal{B}$ closer to $\tilde{h}(a_i)$ then points in $\tilde{h}(U_i^2)$.*

diagram in two possible ways. Thus, we end up with two sets in Y, that is $(h \circ f^t)(U_i^k)$ and $(g^t \circ \tilde{h})(U_i^k)$. We measure how those two sets diverge using the Hausdorff distance.

The extended version of the test presented above considers a larger approximation of $\tilde{h}(U_i^k)$. To this end, find the smallest k_i such that $\tilde{h}(U_i^k) \subset \kappa(h(a_i), k_i, \mathcal{B})$. The corresponding superset defines the enriched approximation (see Figure 2):

$$\tilde{h}^+(U_i^k) := \kappa(h(a_i), k_i, \mathcal{B}). \tag{3.6}$$

We use it to define a modified version of (3.5).

$$\operatorname{ConjTest}^+(\mathcal{A}, \mathcal{B}; k, t, h) := \frac{\sum_{i=1}^{n} \mathrm{d_H}\left((h \circ f^t)(U_i^k),\ g^t\left(\tilde{h}^+(U_i^k)\right)\right)}{n \operatorname{diam}(\mathcal{B})}. \tag{3.7}$$

The extension of ConjTest to ConjTest$^+$ was motivated by results of Experiment 4A described in Subsection 4.4. The experiment should clarify the purpose of making the method more complex.

We refer collectively to ConjTest and ConjTest$^+$ as ConjTest methods.

The forthcoming results provide mathematical justification of our method, i.e. “large” and non-decreasing values of the above tests suggest that there is no conjugacy between two time-series.

Theorem 3.2. *Let $f : X \to X$ and $g : Y \to Y$, where $X \subset \mathbb{R}^{d_X}$ and $Y \subset \mathbb{R}^{d_Y}$, be continuous maps ($d_X$ and d_Y denote dimensions of the spaces). For $y_1 \in Y$ define $\mathcal{B}_m := \left\{b_i := g^{i-1}(y_1) \mid i \in \{1, \ldots, m\}\right\}$.*

Suppose that Y is compact and that the trajectory of y_1 is dense in Y, i.e. the set $\mathcal{B}_m$ becomes dense in Y as $m \to \infty$. If g is semiconjugate to f with h as a semiconjugacy map, then for every fixed n, t and k

$$\lim_{m\to\infty} \mathrm{ConjTest}(\mathcal{A}_n, \mathcal{B}_m; k, t, h) = 0, \tag{3.8}$$

where $\mathcal{A}_n := \{a_i := f^{i-1}(x_1) \mid i \in \{1, \ldots, n\}\}$, $x_1 \in X$, is any time-series in X of a length n.

Moreover, the convergence is uniform with respect to n and with respect to the choice of the starting point x_1 (i.e. the "rate" of convergence does not depend on the time-series $\mathcal{A}_n$).

Proof. Since g is semiconjugate to f via h, $h : X \to Y$ is a continuous surjection such that for every $t \in \mathbb{N}$ we have $h \circ f^t = g^t \circ h$. Fix $t \in \mathbb{N}$ and $k \in \mathbb{N}$ and let $\varepsilon > 0$. We will show that there exists M such that for all $m > M$, all $n \in \mathbb{N}$ and every finite time-series $\mathcal{A}_n := \{a_i := f^{i-1}(x_1) \mid i \in \{1, \ldots, n\}\} \subset X$ of length n (where $x_1 \in X$ is some point in X) it holds that

$$\mathrm{ConjTest}(\mathcal{A}_n, \mathcal{B}_m; k, t, h) < \varepsilon. \tag{3.9}$$

Note that $|b_2 - b_1| \leq |\mathcal{B}_m|$ for any $m \geq 2$ (with $|\mathcal{B}_m|$ denoting cardinality of the set $\mathcal{B}_m$), which we will use at the end of the proof. As g is continuous and Y is compact, there exists δ such that $|g^t(y) - g^t(\tilde{y})| < \varepsilon\,|b_2 - b_1|$ for every y, $\tilde{y} \in Y$ with $|y - \tilde{y}| < \delta$. As $\mathcal{B} = \{y_1, g(y_1), \ldots, g^m(y_1), \ldots\} = \{b_1, \ldots, b_m, \ldots\}$ is dense in Y, there exists M such that if $m > M$ then for every $n \in \mathbb{N}$, every $x_1 \in X$ and every $i \in \{1, 2, \ldots, \}$ there exists $j_m(i) \in \{1, 2, \ldots, m\}$ such that

$$|b_{j_m(i)} - h(a_i)| < \delta,$$

where $a_i = f^{i-1}(x_1) \in \mathcal{A}_n$.

Thus for $m > M$, we always (independently of the point $a_i \in X$) have

$$|h(f^t(a_i)) - g^t(\tilde{h}(a_i))| = |g^t(h(a_i)) - g^t(\tilde{h}(a_i))| < \varepsilon\,|b_2 - b_1|$$

as $g^t(h(a_i)) = h(f^t(a_i))$ and $|\tilde{h}(a_i) - h(a_i)| < \delta$. Consequently,

$$\mathrm{d_H}\left((h \circ f^t)(U_i^k),\ (g^t \circ \tilde{h})(U_i^k)\right) < \varepsilon\,|b_2 - b_1|,$$

where $U_i^k = \kappa(a_i, k, \mathcal{A}_n)$ and $\tilde{h}(U_i^k) = \{\kappa(h(a_j), \mathcal{B}_m) \mid a_j \in U_k^i\}$. Therefore

$$\frac{\sum_{i=1}^n \mathrm{d_H}\left((h \circ f^t)(U_i^k),\ (g^t \circ \tilde{h})(U_i^k)\right)}{n\,\mathrm{diam}(\mathcal{B}_m)} < \frac{n\,\varepsilon\,|b_2 - b_1|}{n\,\mathrm{diam}(\mathcal{B}_m)} \leq \varepsilon$$

since $|b_2 - b_1| \leq \mathrm{diam}(\mathcal{B}_m)$ for every $m \geq 1$. This proves (3.9). ■

The compactness of Y and the density of the set $\mathcal{B} = \{y_1, g(y_1), \ldots, g^m(y_1), \ldots\}$ in Y is needed to obtain the uniform convergence in (3.8) but, as follows from the proof above, these assumptions can be relaxed at the cost of possible losing the uniformity of the convergence:

Corollary 3.3. *Let $f : X \to X$ and $g : Y \to Y$, where $X \subset \mathbb{R}^{d_X}$ and $Y \subset \mathbb{R}^{d_Y}$, be continuous maps. Let $x_1 \in X$ and $y_1 \in Y$. Define $\mathcal{A}_n := \left\{a_i := f^{i-1}(x_1) \mid i \in \{1,\ldots,n\}\right\}$ and $\mathcal{B}_m := \left\{b_i := g^{i-1}(y_1) \mid i \in \{1,\ldots,m\}\right\}$. Suppose that $\{h(a_1),\ldots h(a_n)\} \subset \hat{Y}$ for some compact set $\hat{Y} \subset Y$ such that the set $\hat{Y} \cap \mathcal{B}$ is dense in $\hat{Y}$, where $\mathcal{B} = \{b_1,\ldots,b_m,\ldots\}$.*

*If g is **semiconjugate** to f with h as a semiconjugacy, then for every t and k*

$$\lim_{m\to\infty} \operatorname{ConjTest}(\mathcal{A}_n, \mathcal{B}_m; k, t, h) = 0.$$

Remark 3.4. In the above corollary the assumption on the existence of the set $\hat{Y}$ means just that the trajectory of the point y_1 contains points $g^{j_i}(y_1)$ which, respectively, "well-approximate" points $h(a_i)$, $i = 1, 2, \ldots, n$.

Note also that we do not need the compactness of the space X nor the density of $\mathcal{A} = \{a_1, a_2, \ldots, a_n, \ldots\}$ in X.

The following statement is an easy consequence of the statements above

Theorem 3.5. *Let $X \subset \mathbb{R}^{d_X}$ and $Y \subset \mathbb{R}^{d_Y}$ be compact sets and $f : X \to X$ and $g : Y \to Y$ be continuous maps which are **conjugate** by a homeomorphism $h : X \to Y$. Let $x_1 \in X$, $y_1 \in Y$ and $\mathcal{A}_n$ and $\mathcal{B}_m$ be defined as before. Suppose that $\mathcal{A}_n$ and $\mathcal{B}_m$ are dense, respectively, in X and Y as $n \to \infty$ and $m \to \infty$. Then for every t and k*

$$\lim_{m\to\infty} \operatorname{ConjTest}(\mathcal{A}_n, \mathcal{B}_m; k, t, h) = \lim_{n\to\infty} \operatorname{ConjTest}(\mathcal{B}_m, \mathcal{A}_m; k, t, h) = 0.$$

The assumptions on the compactness of the spaces and density of the trajectories can be slightly relaxed in the similar vein as before.

The above results concern ConjTest. Note that in case of ConjTest^+ the neighborhoods $\tilde{h}^+(U_i^k)$, thus also $(g^t \circ \tilde{h}^+)(U_i^k)$, can be significantly enlarged by adding additional points to $\tilde{h}(U_i^k)$ and thus increasing the Hausdorff distance between corresponding sets. In order to still control this distance and formally prove desired convergence additional assumptions concerning space X and the sequence $\mathcal{A}$ are needed:

Theorem 3.6. *Let $f : X \to X$ and $g : Y \to Y$, where $X \subset \mathbb{R}^{d_X}$ and $Y \subset \mathbb{R}^{d_Y}$ be continuous functions. For $x_1 \in X$ and $n \in \mathbb{N}$ define $\mathcal{A}_n := \left\{a_i := f^{i-1}(x_1) \mid i \in \{1,2,\ldots,n\}\right\}$. Similarly, for $y_1 \in Y$ and $m \in \mathbb{N}$ define $\mathcal{B}_m := \{b_i := g^{i-1}(y_1) \mid i \in \{1,2,\ldots,m\}\}$. Assume that X and Y are compact and that the set $\mathcal{A}_n$ becomes dense in X as $n \to \infty$, and $\mathcal{B}_m$ becomes dense in Y as $m \to \infty$. Under those assumptions, if g is semiconjugate to f with $h : X \to Y$ as a semiconjugacy we have that*

$$\lim_{n\to\infty} \lim_{m\to\infty} \operatorname{ConjTest}^+(\mathcal{A}_n, \mathcal{B}_m; k, t, h) = 0 \tag{3.10}$$

for any $k \in \mathbb{N}$ and $t \in \mathbb{N}$.

Proof. Since g is semiconjugate to f via h, for every $t \in \mathbb{N}$ we have $h \circ f^t = g^t \circ h$, where

$h : X \to Y$ is a continuous surjection. Expanding (3.10) yields

$$\lim_{n\to\infty}\lim_{m\to\infty} \frac{\sum_{i=1}^{n} \mathrm{d_H}\left((h\circ f^t)(U_i^k),\ (g^t\circ \tilde{h}^+)(U_i^k)\right)}{n\ \mathrm{diam}(\mathcal{B}_m)} \leq$$
$$\lim_{n\to\infty}\lim_{m\to\infty} \frac{\sum_{i=1}^{n} \mathrm{d_H}\left((h\circ f^t)(U_i^k),\ (g^t\circ \tilde{h})(U_i^k)\right)}{n\ \mathrm{diam}(\mathcal{B}_m)} + \tag{3.11}$$
$$\lim_{n\to\infty}\lim_{m\to\infty} \frac{\sum_{i=1}^{n} \mathrm{d_H}\left((h\circ f^t)(U_i^k),\ (g^t(\tilde{h}^+(U_i^k)\setminus \tilde{h}(U_i^k)))\right)}{n\ \mathrm{diam}(\mathcal{B}_m)}.$$

Recall that $U_i^k := \kappa(a_i, k, \mathcal{A}_n)$, $\tilde{h}(a_i) := \kappa(h(a_i), \mathcal{B}_m)$, $\tilde{h}(U_i^k) := \{\tilde{h}(a_j) : a_j \in U_i^k\}$ and $\tilde{h}^+(U_i^k) := \kappa(h(a_i), k_i, \mathcal{B}_m)$, where k_i is the smallest integer k_i such that $\tilde{h}(U_i^k) \subset \kappa(h(a_i), k_i, \mathcal{B}_m)$. Thus in particular, $\tilde{h}(U_i^k) \subset \tilde{h}^+(U_i^k)$. Obviously all these neighborhoods U_i^k, $\tilde{h}(U_i^k)$ and $\tilde{h}^+(U_i^k)$ depend on n and m (since they are taken with respect to $\mathcal{A}_n$ and $\mathcal{B}_m$).

Note that from Theorem 3.2 already follows that the first of the two terms in the sum in (3.11) vanishes. Thus we will only show that the second double limit vanishes as well.

Let $\varepsilon > 0$, $k \in \mathbb{N}$ and $t \in \mathbb{N}$. Since $g^t : Y \to Y$ is a continuous function on a compact metric space Y, there exists δ such that $|g^t(x) - g^t(y)| < \frac{\varepsilon}{2}$ whenever $x, y \in Y$ are such that $|x - y| < \delta$. Similarly, since X is compact and $h : X \to Y$ is continuous, there exists δ_1 such that $|h(x) - h(y)| < \frac{\delta}{2}$ whenever $x, y \in X$ such that $|x - y| < \delta_1$.

Since $\mathcal{B}$ is dense in Y, there exists $M \in \mathbb{N}$ such that for $m > M$ and every $y \in Y$, there exists $\tilde{b} \in \mathcal{B}_m$ such that $|\tilde{b} - y| < \frac{\delta}{4}$. Moreover, from the density of $\mathcal{A}$, there exists $N \in \mathbb{N}$ such that for every $n > N$ and every $i \in \{1, 2, \ldots, n\}$ we have $\mathrm{diam}(U_i^k) < \delta_1$, i.e. if $a_j \in U_i^k = \kappa(a_i, k, \mathcal{A}_n)$ then $|a_j - a_i| < \delta_1$ and consequently

$$|g^t(h(a_j)) - g^t(h(a_i))| < \frac{\varepsilon}{2}. \tag{3.12}$$

Assume thus $n > N$. Then for $m > M$ and every $i \in \{1, 2, \ldots n\}$ we have $\mathrm{diam}(U_i^k) < \delta_1$ which also implies $\mathrm{diam}(h(U_i^k)) < \frac{\delta}{2}$. As $m > M$, every point of $h(U_i^k)$ can be approximated by some point of $\mathcal{B}_m$ with the accuracy better than $\frac{\delta}{4}$. Consequently, $\mathrm{diam}(\tilde{h}(U_i^k)) < \delta$ for every $i \in \{1, 2, \ldots, n\}$.

Suppose that $\tilde{b} \in \tilde{h}^+(U_i^k) \setminus \tilde{h}(U_i^k)$ for some $\tilde{b} \in \mathcal{B}_m$. Then, by definition of $\tilde{h}^+$,

$$|\tilde{b} - h(a_i)| \leq \mathrm{diam}(\tilde{h}(U_i^k)) < \delta. \tag{3.13}$$

Thus for any $a_j \in U_i^k = \kappa(a_i, k, \mathcal{A}_n)$ and any $\tilde{b} \in \tilde{h}^+(U_i^k) \setminus \tilde{h}(U_i^k)$ we obtain

$$\begin{aligned} &|h(f^t(a_j)) - g^t(\tilde{b})| \\ &\qquad \leq |h(f^t(a_j) - g^t(h(a_j))| + |g^t(h(a_j)) - g^t(h(a_i))| + |g^t(h(a_i)) - g^t(\tilde{b})| \end{aligned}$$

where

- $|h(f^t(a_j) - g^t(h(a_j))| = 0$ by semiconjugacy assumption
- $|g^t(h(a_j)) - g^t(h(a_i))| < \frac{\varepsilon}{2}$ as follows from (3.12)
- $|g^t(h(a_i)) - g^t(\tilde{b})| < \frac{\varepsilon}{2}$ as follows from (3.13).

Finally for every $i \in \{1, 2, \dots n\}$, every $a_j \in U_i^k$ and every $\tilde{b} \in (\tilde{h}^+(U_i^k) \setminus \tilde{h}(U_i^k))$ we have $|h(f^t(a_j)) - g^t(\tilde{b})| < \varepsilon$ meaning that

$$\frac{\sum_{i=1}^n \mathrm{d_H}\left((h \circ f^t)(U_i^k),\ g^t(\tilde{h}^+(U_i^k) \setminus \tilde{h}(U_i^k))\right)}{n\ \mathrm{diam}(\mathcal{B}_m)} < \frac{\varepsilon}{\mathrm{diam}(\mathcal{B}_m)}$$

if only $n > N$ and $m > M$.

This shows that the value of $\mathrm{ConjTest}^+(\mathcal{A}_n, \mathcal{B}_m; k, t, h)$ can be arbitrarily small if n and m are sufficiently large and ends the proof. ■

Finally, let us mention that conjugacy tests described in this Section are not tests in the statistical sense. They should be rather considered as methods of assessing dynamical similarity of the two unknown systems when only small finite samples of their trajectories are available. Trajectories related by topological conjugacy will give values of the tests close to 0, and those coming from not conjugate systems are expected to result with higher values of the tests.

The discussed task is already, to a certain extent, considered in the literature. The paper [23] develops sets of statistics which are intended to characterize, in terms of probabilities and confidence levels, whether time delay embeddings of the two time series are connected by a continuous, injected or differentiable map. The work [23] presents method to assess (generalized) synchronization of time series, coupling in complex population dynamics (see [22]) or detecting damage in some material structures (see [21]). The statistics proposed in those papers are inspired by notions of continuity, differentiability etc., typically involving quantities like ϵ's and δ's. These values need to be fixed and enforce the user to understand how δ's scale with ϵ which is typically hard. It seems to be possible to adopt ConjTest's methods to the framework of statistical tests and will be considered in the future.

4. Conjugacy experiments.

In this section the behavior of the described methods is experimentally studied. For that purpose a benchmark set of a number of time series originating from (non-)conjugate dynamical systems is generated. A time series of length N generated by a map $f : X \to X$ with a starting point $x_1 \in X$ is denoted by

$$\varrho(f, x_1, N) := \left\{ f^{j-1}(x_1) \in X \mid j \in \{1, 2, \dots, N\} \right\}.$$

All the experiments were computed in Python using floating number precision. The implementations of the methods presented in this paper as well as the notebooks recreating the presented experiments are available at https://github.com/dioscuri-tda/conjtest.

4.1. Irrational rotation on a circle.

The first example involves a dynamics generated by rotation on a circle by an irrational angle. Let us define a 1-dimensional circle as a quotient space $\mathbb{S} := \mathbb{R}/\mathbb{Z}$. Denote the operation of taking a decimal part of a number (modulo 1) by $x_{\mathbb{1}} := x - \lfloor x \rfloor$. Then, for a parameter $\phi \in [0, 1)$ we define a rigid rotation on a circle, $f_{[\phi]} : \mathbb{S} \to \mathbb{S}$, as

$$f_{[\phi]}(x) := (x + \phi)_{\mathbb{1}}.$$

We consider the following metric on $\mathbb{S}$

$$\mathbf{d}_{\mathbb{S}} : \mathbb{S} \times \mathbb{S} \ni (x, y) \mapsto \min\left((x-y)_{\mathbb{1}}, (y-x)_{\mathbb{1}}\right) \in [0, 1). \tag{4.1}$$

In this case $\mathbf{d}_{\mathbb{S}}(x, y)$ can be interpreted as the length of the shorter arc joining points x and y on $\mathbb{S}$. It is known that two rigid rotations, $f_{[\phi]}$ and $f_{[\psi]}$, are topologically conjugate if and only if $\phi = \psi$ or $\phi + \psi = 1$ (see e.g. Theorem 2.4.3 and Corollary 2.4.1 in [28]). In the first case the conjugating circle homeomorphism h preserves the orientation i.e. the lift $H : \mathbb{R} \to \mathbb{R}$ of $h : \mathbb{S} \to \mathbb{S}$ satisfies $H(x + 1) = H(x) + 1$ for every $x \in \mathbb{R}$ and in the second case h reverses the orientation $H(x + 1) = H(x) - 1$ and the two rotations $f_{[\phi]}$ and $f_{[\psi]}$ are just mutually inverse.

Moreover, for a map $f_{[\phi]}$ we introduce a family of topologically conjugate maps given by

$$f_{[\phi],s}(x) := \left((x_{\mathbb{1}}^s + \phi)_{\mathbb{1}}\right)^{1/s}, \ x \in \mathbb{R}$$

with $s > 0$. In particular, $f_{[\phi]} = f_{[\phi],1}$. It is easy to check that by putting $h_s(x) := x_{\mathbb{1}}^s$ we get $f_{[\phi],s} = h_s^{-1} \circ f_{[\phi]} \circ h_s$.

4.1.1. Experiment 1A.

Setup. Let $\alpha = \frac{\sqrt{2}}{10}$. In the first experiment we compare the following time series:

$$\mathcal{R}_1 = \varrho(f_{[\alpha]}, 0.0, 2000), \qquad \mathcal{R}_2 = \varrho(f_{[\alpha]}, 0.25, 2000),$$
$$\mathcal{R}_3 = \varrho(f_{[\alpha+0.02]}, 0.0, 2000), \qquad \mathcal{R}_4 = \varrho(f_{[2\alpha]}, 0.0, 2000),$$
$$\mathcal{R}_5 = \varrho(f_{[\alpha],2}, 0.0, 2000), \qquad \mathcal{R}_6 = \mathcal{R}_5 + \mathrm{err}(0.05),$$

where $\mathrm{err}(\epsilon)$ denotes a uniform noise sampled from the interval $[-\epsilon, \epsilon]$.

As follows from Poincaré Classification Theorem, $f_{[\alpha]}$ and $f_{[2\alpha]}$ are not conjugate nor semiconjugate whereas $f_{[\alpha]}$ and $f_{[\alpha],2}$ are conjugate via h_2. Thus the expectation is to confirm conjugacy of $\mathcal{R}_1$ and $\mathcal{R}_2$ and of $\mathcal{R}_1$ and $\mathcal{R}_5$ and indicate deviations from conjugacy in all the remaining cases.

In case of ConjTest the comparison $\mathcal{R}_1$ versus $\mathcal{R}_2$, $\mathcal{R}_1$ versus $\mathcal{R}_3$ and $\mathcal{R}_1$ versus $\mathcal{R}_4$ was done with $h \equiv \mathrm{id}_{\mathbb{S}}$. As we already mentioned, there is no conjugacy between $\mathcal{R}_1$ and $\mathcal{R}_3$, nor between $\mathcal{R}_1$ and $\mathcal{R}_4$, as the angles of these rotations are different. Thus there is no true connecting homeomorphism between $\mathcal{R}_1$ and $\mathcal{R}_3$ and between $\mathcal{R}_1$ and $\mathcal{R}_4$. However, in order to apply ConjTests we need to pick some candidate for a matching map between two point clouds and as the first choice one can always start with the identity map, especially for comparing point clouds generated by trajectories starting at the same or close initial points. Therefore in this experiment we use $h \equiv \mathrm{id}_{\mathbb{S}}$ for comparing $\mathcal{R}_1$ both with $\mathcal{R}_3$ and $\mathcal{R}_4$. When comparing $\mathcal{R}_1$ versus $\mathcal{R}_4$ and $\mathcal{R}_1$ versus $\mathcal{R}_5$ we use homeomorphism $h_2(x) := x_{\mathbb{1}}^2$. Let us recall that for FNN and KNN methods we always use $h(x_i) = y_i$, a connecting homeomorphism based on the indices correspondence.

Results. The results are presented in Table 1. Since the presented methods are not symmetric, order of input time series matters. To accommodate this information, every cell contains two values, above and below the diagonal. For the column with header "$\mathcal{R}_i$ vs. $\mathcal{R}_j$", the cells upper value corresponds to the outcome of $\mathrm{FNN}(\mathcal{R}_i, \mathcal{R}_j; r)$, $\mathrm{KNN}(\mathcal{R}_i, \mathcal{R}_j; k)$,

method \ test	$\mathcal{R}_1$ vs. $\mathcal{R}_2$ (starting point perturbation)	$\mathcal{R}_1$ vs. $\mathcal{R}_3$ (angle perturbation)	$\mathcal{R}_1$ vs. $\mathcal{R}_4$ (angle doubled)	$\mathcal{R}_1$ vs. $\mathcal{R}_5$ (nonlinear rotation)	$\mathcal{R}_1$ vs. $\mathcal{R}_6$ (noise)
FNN (r=2)	0.0 / 0.0	1.0 / 1.0	.393 / .393	.063 / 0.0	.987 / .986
KNN (k=5)	0.0 / 0.0	.257 / .756	.003 / .997	0.0 / 0.0	.150 / .152
ConjTest (k=5, t=5)	.001 / .001	.201 / .201	.586 / .586	0.0 / 0.0	.142 / .181
ConjTest$^+$ (k=5, t=5)	.001 / .001	.201 / .201	.586 / .586	0.0 / 0.0	.162 / .181

Table 1

Comparison of conjugacy measures for time series generated by the rotation on a circle. The number in the upper left part of the cell corresponds to a comparison of the first time series vs. the second one, while the lower right corresponds to the inverse comparison. As follows from formulas at the beginning of Section 4.1.1 the considered trajectories have length $N = 2000$, other corresponding parameters are stated in the table.

$\mathrm{ConjTest}(\mathcal{R}_i, \mathcal{R}_j; k, t, h)$ and $\mathrm{ConjTest}^+(\mathcal{R}_i, \mathcal{R}_j; k, t, h)$, respectively to the row. The lower values corresponds to $\mathrm{FNN}(\mathcal{R}_j, \mathcal{R}_i; r)$, $\mathrm{KNN}(\mathcal{R}_j, \mathcal{R}_i; k)$, $\mathrm{ConjTest}(\mathcal{R}_j, \mathcal{R}_i; k, t, h)$ and $\mathrm{ConjTest}^+(\mathcal{R}_j, \mathcal{R}_i; k, t, h)$, respectively.

As we can see from Table 1 the starting point does not affect results of methods ($\mathcal{R}_1$ vs. $\mathcal{R}_2$) since all the values in the first column are close to 0. It is expected due to the symmetry of the considered system. A nonlinearity introduced in time series $\mathcal{R}_5$ also does not affect the results. Despite the fact that $f_{[\alpha],2}$ is nonlinear, it is conjugate to the rotation $f_{[\alpha]}$ which is reflected by tests' values. However, when we change the rotation parameter we can see an increase of measured values ($\mathcal{R}_1$ vs. $\mathcal{R}_3$ and $\mathcal{R}_1$ vs. $\mathcal{R}_4$). It is particularly visible in the case of FNN and KNN. Interestingly, a small perturbation of the angle ($\mathcal{R}_3$) can cause a bigger change in a value then a large one ($\mathcal{R}_4$). We investigate how the perturbation of the rotation parameter affects values of examined methods in Experiment 1B. Moreover, the last column ($\mathcal{R}_1$ vs. $\mathcal{R}_6$) shows that FNN is very sensitive to noise, while KNN and ConjTest methods present some robustness. The influence of noise on the value of the test statistics is further studied in Experiment 1C.

Note also that additional summary comments concerning Table 1, as well as results of other forthcoming experiments, will be also presented at the end of the article.

4.1.2. Experiment 1B.

In this experiment we test how the difference of the system parameter affects tested methods.

Setup. Let $\alpha := \frac{\sqrt{2}}{10} \approx 0.141$. We consider a family of time series parameterized by β.

$$\left\{ \mathcal{R}_\beta := \varrho(f_{[\beta]}, 0.0, 2000) \mid \beta = \alpha + \frac{i\alpha}{100},\ i \in [-50, -49, \ldots, 125] \right\}. \tag{4.2}$$

Thus, the tested interval of values of β is approximately $[0.07, 0.32]$. As a reference value we chose $\alpha = \frac{\sqrt{2}}{10} \approx 0.141$. We denote the corresponding time series as $\mathcal{R}_\alpha$. We compare all time series from the family (4.2) with $\mathcal{R}_\alpha$. In the case of ConjTest methods we use $h = \mathrm{id}$.

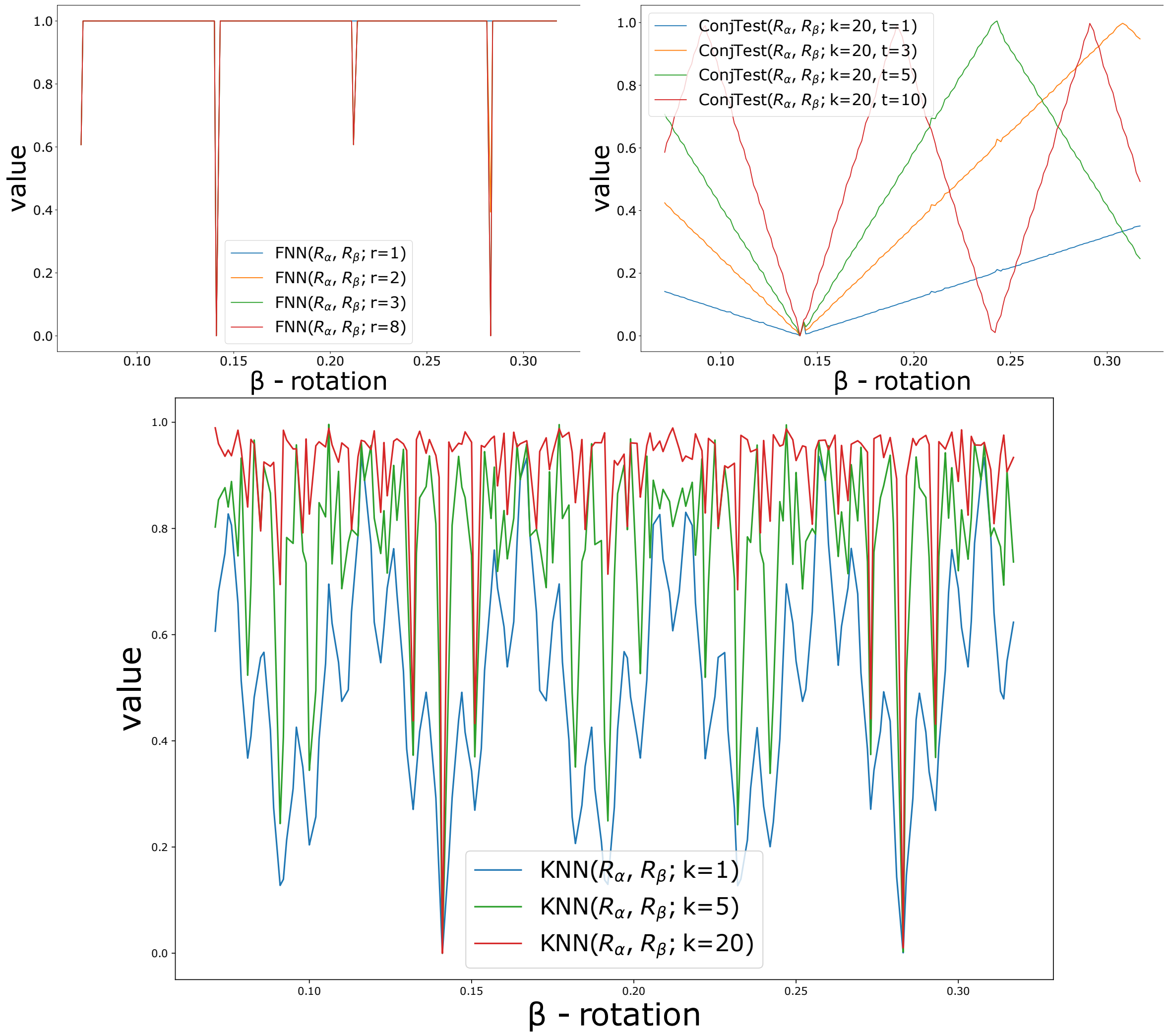

Figure 3. *Dependence of the conjugacy measures on the perturbation of rotation angle. Top left:* FNN *method. Top right:* ConjTest *method. Bottom:* KNN *method.*

Results. The outcome of the experiment is plotted in Figure 3. We can see that all methods give values close to 0 when comparing $\mathcal{R}_\alpha$ with itself. For different values of parameter r of FNN plots (Figure 3 top left looks almost identically. Even a small perturbation of the rotation parameter causes an immediate jump of FNN value from 0 to 1, making it extremely sensitive to any changes in the system. Obviously, unless $\beta = \alpha$, $\mathcal{R}_\alpha$ and $\mathcal{R}_\beta$ are not conjugate. However, sometimes it might be convenient to have a somehow smoother relation of the test value to the infinitesimal change of the rotation angle. KNN method seems to behave inconsistently, but we can see that the higher parameter k gets the closer we get to a shape resembling the curve obtained with FNN. On the other hand, ConjTest shows a linear dependence on β parameter. Moreover, different values of ConjTest's parameter t result in a different slope of this dependency.

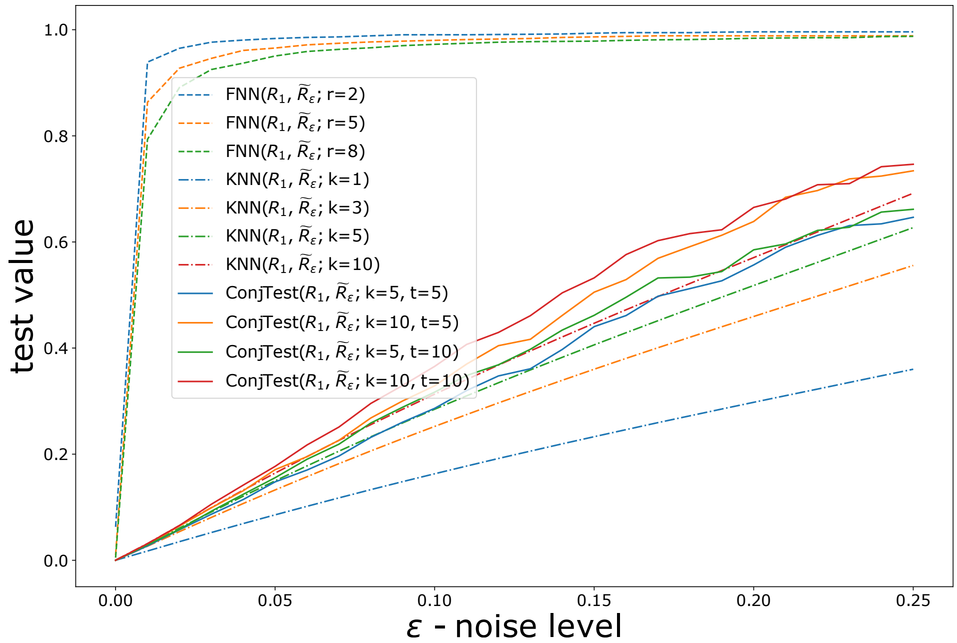

Figure 4. *Dependence of conjugacy measures on the perturbation of time series.*

Both, FNN and KNN exhibit an interesting drop of the value when $\beta \approx 0.283$ that is $\beta = 2\alpha$. Formally, we know that $f_{[\alpha]}$ and $f_{[2\alpha]}$ are not conjugate systems. However, we can explain this outcome by analyzing the methods. Let $a_i \in \mathcal{R}_\alpha$ and let $\tau \in \mathbb{Z}$ such that $a_j := a_{i+\tau} \in \mathcal{R}_\alpha$ be the nearest neighbor of a_i. In particular, $\tau \in \mathbb{Z}$. By (4.1) we get $\mathbf{d}_{\mathbb{S}}(a_i, a_j) = (\alpha\tau)_{\mathbb{1}}$ or $(-\alpha\tau)_{\mathbb{1}}$. There is an $N \in \mathbb{Z}$ and a $\delta \in [0, 1)$ such that $\alpha\tau = N + \delta$. Since $\mathbf{d}_{\mathbb{S}}(a_i, a_j) \approx 0$, it follows that $\delta_{\mathbb{1}} \approx 0$. To get FNN we also need to know $\mathbf{d}_{\mathbb{S}}(b_i, b_j)$. Let $\beta = z\alpha$. Then, $b_i = (z\alpha i)_{\mathbb{1}}$, $b_j = (z\alpha i + z\alpha\tau)_{\mathbb{1}}$ and $\mathbf{d}_{\mathbb{S}}(b_i, b_j) = (z\alpha\tau)_{\mathbb{1}}$ or $(-z\alpha\tau)_{\mathbb{1}}$. Thus, $z\alpha\tau = zN + z\delta$. We assume that $z\delta \in [0, 1)$, because $\delta_{\mathbb{1}} \approx 0$ and z is not very large . Again, there exists an $M \in \mathbb{Z}$ and $\epsilon \in [0, 1)$ such that $zN = M + \epsilon$. Now, if $zN \in \mathbb{Z}$, then $\epsilon = 0$, $\mathbf{d}_{\mathbb{S}}(b_i, b_j) = (z\delta)_{\mathbb{1}} = z\,\mathbf{d}_{\mathbb{S}}(a_i, a_j)$ (last equality given by $\delta_{\mathbb{1}} \approx 0$) and $\frac{\mathbf{d}_{\mathbb{S}}(b_i,b_j)}{\mathbf{d}_{\mathbb{S}}(a_i,a_j)} = z$. If $zN \notin \mathbb{Z}$ then $\epsilon \neq 0$ and $\frac{\mathbf{d}_{\mathbb{S}}(b_i,b_j)}{\mathbf{d}_{\mathbb{S}}(a_i,a_j)} = \frac{\gg 0}{\sim 0}$. Hence, the fraction gives a large number and the numerator of FNN will count most of the points, unless $zN \in \mathbb{Z}$ which is always satisfied when $z \in \mathbb{Z}$. Moreover, for the irrational rotation τ might be large. In our experiments we usually get $|\tau| > 1000$. Thus, N is large and ϵ is basically a random number. In the case of KNN there is a chance that at least for some of the k-nearest neighbors $zN \in \mathbb{Z}$. Hence, the more rugged shape of the curve.

In the case of ConjTest we observe a clear impact of ConjTest's parameter t on the shape of the curve. The method takes k-nearest neighbors of a point x_i (U_i^k in the formula (2.1)) and moves them t times about angle α. At the same time the corresponding image of those points in the system $\mathcal{R}_\beta$ ($\tilde{h}(U_i^k)$ in the formula (2.1)) is rotated t times about β angle. Thus, the discrepancy of the position of those two sets of points is proportional to $t\beta$. In particular, when $(t\beta)_{\mathbb{1}} = \alpha$, these two sets are in the same position.

4.1.3. Experiment 1C. In this experiment, instead of perturbing the parameter of the system we perturb the time series itself by applying a noise to every point of the series.

Setup. Set $\alpha = \frac{\sqrt{2}}{10}$. We compare a time series $\mathcal{R}_1 := \varrho(f_{[\alpha]}, 0.0, 2000)$ with a family of time series:

$$\left\{ \widetilde{\mathcal{R}}_\epsilon := \varrho(f_{[\alpha],2}, 0.0, 2000) + \mathrm{err}(\epsilon) \mid \epsilon \in [0.00, 0.25] \right\}, \tag{4.3}$$

where $\mathrm{err}(\epsilon)$ is a uniform noise sampled from the interval $[-\epsilon, \epsilon]$ applied to every point of the time series. In the case of ConjTest we again use $h(x) = x^2{}_{\mathbb{1}}$.

Results. Results are presented in Figure 4. Again, FNN presents a very high sensitivity on any disruption of a time series and even a small amount of noise gives a conclusion that two systems are not conjugate. On the other hand, KNN and ConjTest present an almost linear dependence on noise level. Note that higher values of parameters k and t make methods more sensitive to the noise.

4.2. Example: irrational rotation on a torus. Let us consider a simple extension of the previous rotation example to a rotation on a torus. With a torus defined as $\mathbb{T} := \mathbb{S} \times \mathbb{S}$, where $\mathbb{S} = \mathbb{R}/\mathbb{Z}$, we can introduce map $f_{[\phi_1,\phi_2]} : \mathbb{T} \to \mathbb{T}$ defined as

$$f_{[\phi_1,\phi_2]}(x^{(1)}, x^{(2)}) = ((x^{(1)} + \phi_1)_{\mathbb{1}}, (x^{(2)} + \phi_2)_{\mathbb{1}}),$$

where $\phi_1, \phi_2 \in [0, 1)$. We equip the space with the maximum metric $\mathbf{d}_{\mathbb{T}}$:

$$\mathbf{d}_{\mathbb{T}} : \mathbb{T} \times \mathbb{T} \ni ((x_1, y_1), (x_2, y_2)) \mapsto \max\left(\mathbf{d}_{\mathbb{S}}(x_1, x_2), \mathbf{d}_{\mathbb{S}}(y_1, y_2)\right) \in [0, 1),$$

where $\mathbf{d}_{\mathbb{S}}$ is the sphere metric (see (4.1)).

Note that rotation on a torus described above and rotation on a circle $f_{[\phi_i]} : \mathbb{S} \to \mathbb{S}$ studied in Section 4.1 give a simple example of semiconjugate systems. Namely, let $h : \mathbb{T} \to \mathbb{S}$ be a projection $h_i(x^{(1)}, x^{(2)}) = x^{(i)}$, $i = 1, 2$. Then we get the equality $h_i \circ f_{[\phi_1,\phi_2]} = f_{[\phi_i]} \circ h_i$ for $i \in \{1, 2\}$.

4.2.1. Experiment 2A.

Setup. For this experiment we consider the following time series:

$$\begin{aligned} \mathcal{T}_1 &= \varrho(f_{[\alpha,\beta]}, (0.0, 0.0), 2000), & \mathcal{S}_1 &= \mathcal{T}_1^{(1)}, \\ \mathcal{T}_2 &= \varrho(f_{[1.1\alpha,\beta]}, (0.1, 0.0), 2000), & \mathcal{S}_2 &= \mathcal{T}_2^{(1)}, \\ \mathcal{T}_3 &= \varrho(f_{[\beta,\beta]}, (0.1, 0.0), 2000), & \mathcal{S}_3 &= \mathcal{T}_3^{(1)}, \end{aligned}$$

where $\alpha = \sqrt{2}/10$, $\beta = \sqrt{3}/10$, and $\mathcal{S}_i = \mathcal{T}_i^{(1)}$, $i = 1, 2, 3$, is a time series obtained from the projection of the elements of $\mathcal{T}_i$ onto the first coordinate. When comparing $\mathcal{T}_i$ with $\mathcal{T}_j$ for $i, j \in \{1, 2, 3\}$ we use $h \equiv \mathrm{id}$. When we compare $\mathcal{T}_i$ versus $\mathcal{S}_j$ we use $h(x, y) = x$, and for $\mathcal{S}_i$ versus $\mathcal{T}_j$ we get $h(x) = (x, 0)$.

Results. The asymmetry of results in the first column ($\mathcal{T}_1$ vs. $\mathcal{S}_1$) in Table 2 shows that all methods detect a semiconjugacy between $\mathcal{T}_1$ and $\mathcal{S}_1$, i.e. that $f_{[\alpha]}$ is semiconjugate to $f_{[\alpha,\beta]}$ via h_1. An embedding of a torus into a 1-sphere preserves a neighborhood of a point. The inverse map clearly does not exist.

method \ test	$\mathcal{T}_1$ vs. $\mathcal{S}_1$	$\mathcal{T}_1$ vs. $\mathcal{T}_2$	$\mathcal{T}_1$ vs. $\mathcal{S}_2$	$\mathcal{T}_1$ vs. $\mathcal{T}_3$	$\mathcal{T}_1$ vs. $\mathcal{S}_3$
FNN (r=2)	0.0 / 1.0	1.0 / .978	1.0 / 1.0	0.0 / 1.0	0.0 / 1.0
KNN (k=5)	.042 / .617	.275 / .855	.514 / .690	.041 / .938	.041 / .938
ConjTest (k=5, t=5)	.001 / .270	.149 / .148	.142 / .270	.451 / .322	.318 / .322
ConjTest$^+$ (k=5, t=5)	.018 / .272	.154 / .158	.143 / .271	.458 / .325	.319 / .324

Table 2

Comparison of conjugacy measures for time series generated by the rotation on a torus. The number in the upper left part of the cell corresponds to a comparison of the first time series vs. the second one, while the lower right number corresponds to the inverse comparison.

The rest of the results confirm conclusions from the previous experiment. The second and the third column ($\mathcal{T}_1$ vs. $\mathcal{T}_2$ and $\mathcal{T}_1$ vs. $\mathcal{S}_2$) show that FNN and KNN are sensitive to a perturbation of the system parameters. The fourth and the fifth column ($\mathcal{T}_1$ vs. $\mathcal{T}_3$ and $\mathcal{T}_1$ vs. $\mathcal{S}_3$) present another example where those two methods produce a false positive answer suggesting a semiconjugacy. This time the problematic case is not due to a doubling of the rotation parameter, but because of coinciding rotation angles. Again, the behavior of the ConjTest method exhibits a response that is relative to the level of perturbation.

4.3. Example: the logistic map and the tent map. Our next experiment examines two broadly studied chaotic maps defined on a real line. The logistic map and the tent map, $f_l, g_\mu : [0,1] \to [0,1]$, respectively defined as:

$$f_l(x) := lx(1-x) \quad \text{and} \quad g_\mu(x) := \mu \min\{x, 1-x\}, \tag{4.4}$$

where, typically, $l \in [0,4]$ and $\mu \in [0,2]$. For parameters $l = 4$ and $\mu = 2$ the systems are conjugate via homeomorphism:

$$h(x) := \frac{2\arcsin(\sqrt{x})}{\pi}, \tag{4.5}$$

that is, $h \circ f_4 = g_2 \circ h$. In this example we use the standard metric induced from $\mathbb{R}$.

4.3.1. Experiment 3A.

Setup. In the initial experiment for those systems we compare the following time series:

$$\mathcal{A} = \varrho(f_4, 0.2, 2000), \qquad \mathcal{B}_2 = \varrho(f_4, 0.21, 2000),$$
$$\mathcal{B}_1 = \varrho(g_2, h(0.2), 2000), \qquad \mathcal{B}_3 = \varrho(f_{3.99}, 0.2, 2000),$$
$$\mathcal{B}_4 = \varrho(f_{3.99}, 0.21, 2000).$$

Time series $\mathcal{A}$ is conjugate to $\mathcal{B}_1$ through the homeomorphism h. Time series $\mathcal{A}$ and $\mathcal{B}_2$ come from the same system – f_4, but are generated using different starting points. Sequences $\mathcal{B}_3$ and $\mathcal{B}_4$ are both generated by the logistic map but with different parameter value ($l = 3.99$) than $\mathcal{A}$; thus, they are not conjugate with $\mathcal{A}$. For ConjTest methods we use (4.5) to compare $\mathcal{A}$ with $\mathcal{B}_1$, and the identity map to compare $\mathcal{A}$ with $\mathcal{B}_2$, $\mathcal{B}_3$ and $\mathcal{B}_4$.

method \ test	$\mathcal{A}$ vs. $\mathcal{B}_1$	$\mathcal{A}$ vs. $\mathcal{B}_2$ (starting point perturbation)	$\mathcal{A}$ vs. $\mathcal{B}_3$ (parameter perturbation)	$\mathcal{A}$ vs. $\mathcal{B}_4$ (st.point + param. perturbation)
FNN (r=2)	.205 / 0.0	.998 / 1.0	1.0 / 1.0	1.0 / .999
KNN (k=5)	0.0 / 0.0	.825 / .828	.831 / .832	.835 / .833
ConjTest (k=5, t=5)	0.0 / 0.0	.017 / .017	.099 / .059	.099 / .059
ConjTest$^+$ (k=5, t=5)	0.0 / .001	.027 / .023	.104 / .065	.104 / .064

Table 3

Comparison of conjugacy measures for time series generated by logistic and tent maps. The number in the upper left part of the cell corresponds to a comparison of the first time series vs. the second one, while the lower right number corresponds to the inverse comparison.

Results. The first column of Table 3 shows that all methods properly identify the tent map as a system conjugate to the logistic map (provided that the two time series are generated by dynamically corresponding points, i.e. a_1 and $b_1 := h(a_1)$, respectively). The second column demonstrates that FNN and KNN get confused by a perturbation of the starting point generating time series. This effect was not present in the circle and the torus example (Sections 4.1 and 4.2) due to a full symmetry in those examples. The ConjTest methods are only weakly affected by the perturbation of the starting point. Nevertheless, we expect that higher values of parameter t may significantly affect the outcome of ConjTest due to the chaotic nature of the map. We test it further in the context of Lorenz attractor (Experiment 4C). The third and the fourth column reflect high sensitivity of FNN and KNN to the parameter of the system. On the other hand, ConjTest methods admit rather conservative response to a change of the parameter.

The experiment shows that FNN and KNN are able to detect a change caused by a perturbation of a system immediately. However, in the context of empirical data we may not be able to determine whether the starting point was perturbed, or if the system has actually changed, or whether there was a noise in our measurements. Thus, some robustness with respect to noise might be desirable and the seemingly blurred concept of the conjugacy represented by ConjTest might be helpful.

4.3.2. Experiment 3B.

The logistic map is one of the standard examples of chaotic maps. Thus, we expect that the behavior of the system will change significantly if we modify the parameter l. Here, we examine how the perturbation of l affects the outcome of tested methods.

Setup. We generated a collection of time series:

$$\{\mathcal{B}(l) := \varrho(f_l, 0.2, 2000) \mid l \in \{3.8, 3.805, 3.81, \dots, 4.0\}\}\,.$$

Every time series $\mathcal{B}(l)$ in the collection was compared with a reference time series $\mathcal{B}(4.0)$.

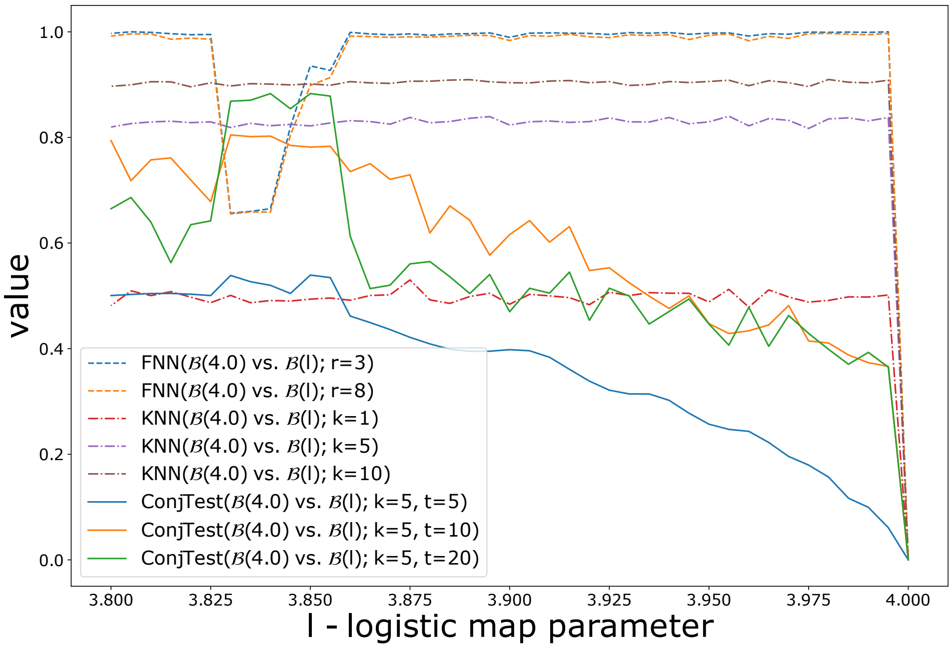

Figure 5. *Dependence of the conjugacy measures on a change of the parameter of the logistic map.*

Results. The results are plotted in Figure 5. As Experiment 3A suggested, FNN and KNN quickly saturate, providing almost "binary" response i.e. the output value is either 0 or a fixed non-zero number depending on parameter k. Similarly to Experiment 1B we observe that with a higher parameter k the curve corresponding to KNN gets more similar to FNN and becomes nearly a step function. ConjTest admits approximately continuous dependence on the value of the parameter of the system. However, higher values of the parameter t of ConjTest make the curve more steep and forms a significant step down in the vicinity of $l = 4$. This makes sense, because the more time-steps forward we take into account the more nonlinearity of the system affects the tested neighborhood.

We presume that the observed drop of FNN values and increase of ConjTest values for the parameter l value approximately in the interval $\{3.83, 3.86\}$ is caused by the collapse of the attractor to the 3-periodic orbit observed for these parameter values (see bifurcation diagram in Figure 7).

Obviously, the logistic map with different parameter values won't be conjugate. However, since we work with only finite samples, it might be not enough to rigorously distinguish them if the difference of the parameter is small. The results can only suggest an empirical similarity of the underlying dynamical systems.

4.3.3. Experiment 3C. As observed in the previous experiment, a change of the parameter l in the logistic equation may significantly change the dynamical nature of the system. In this experiment we use the ConjTest to grasp the types of dynamics as a function of l parameter.

Setup. First, we generated the following collection of time series:

$$\{\mathcal{B}(l,p) := \varrho(f_l, f_l^{500}(p), 2000) \mid l \in \{3.4, 3.405, 3.41, \ldots, 4.0\},\ p \in \{0.11, 0.31\}\}. \tag{4.6}$$

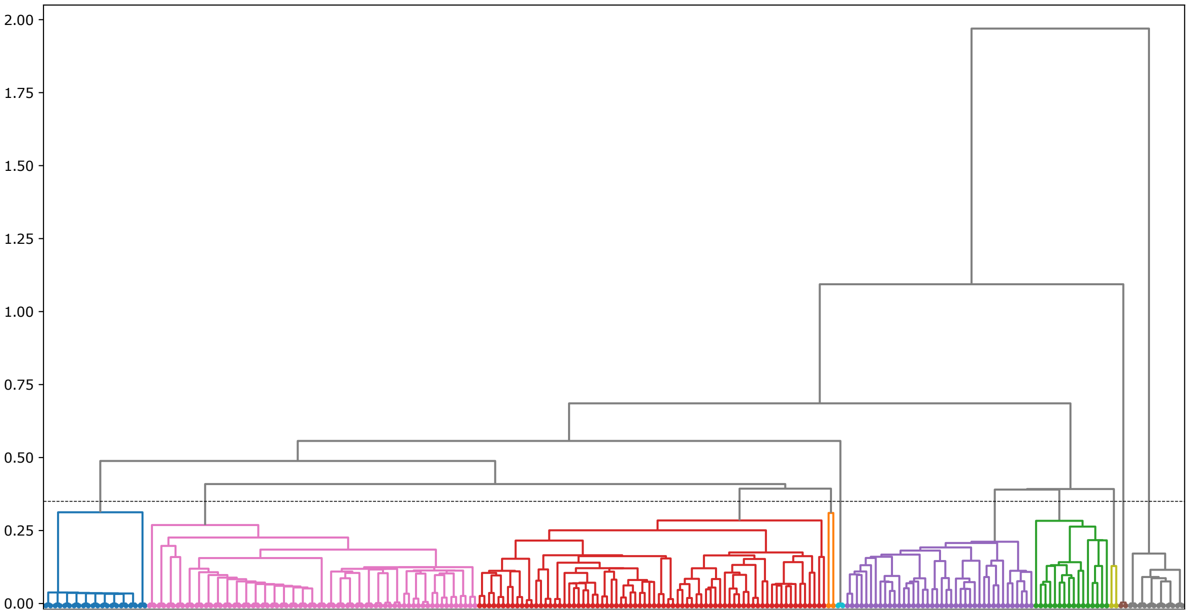

Figure 6. *Dendrogram obtained from the single-linkage agglomerative hierarchical clustering of the collection of time series (4.6) generated from the logistic map using similarity score defined (4.7). The horizontal dashed line represents the threshold chosen for the clustering. The distribution of the clustered time series on bifurcation diagram is presented in Figure 7.*

Note that each time series starts at 500-th iterate of point p. It is a standard procedure allowing the trajectory to settle down on an attractor. We compared every two time series $\mathcal{B}(l,p)$ and $\mathcal{B}(l',p')$ with the formula We assign a similarity for each pair of time series $\mathcal{B}(l,p)$ and $\mathcal{B}(l',p')$ via

$$(4.7) \qquad \max\left\{\mathrm{ConjTest}(\mathcal{B}(l,p),\mathcal{B}(l',p');k,t,\mathrm{id}),\mathrm{ConjTest}(\mathcal{B}(l',p'),\mathcal{B}(l,p);k,t,\mathrm{id})\right\}$$

with fixed $k=5$ and $t=2$. The obtained similarity matrix was then applied to the single-linkage agglomerative hierarchical clustering.

Results. The dendrogram in Figure 6 presents the output of the experiment. Every leaf represents a single time series corresponding to a pair (l,p) of the parameter value and a starting point. With a threshold value 0.35 we can distinguish 10 clusters. For every value of parameter l, both time series $\mathcal{B}(l,0.11)$ and $\mathcal{B}(l,0.31)$ fall into the same cluster.

We draw the result of the clustering on the bifurcation diagram in Figure 7. As one can expect, the time series grouped according to their dynamics type and their proximity in the parameter space. The dendrogram structure indicates additional substructures within the clusters. For instance, the pink cluster contains two visible subclasses from which one corresponds to a set of 4-periodic orbit, while the second aggregates the attractors after the initial period doubling bifurcations.

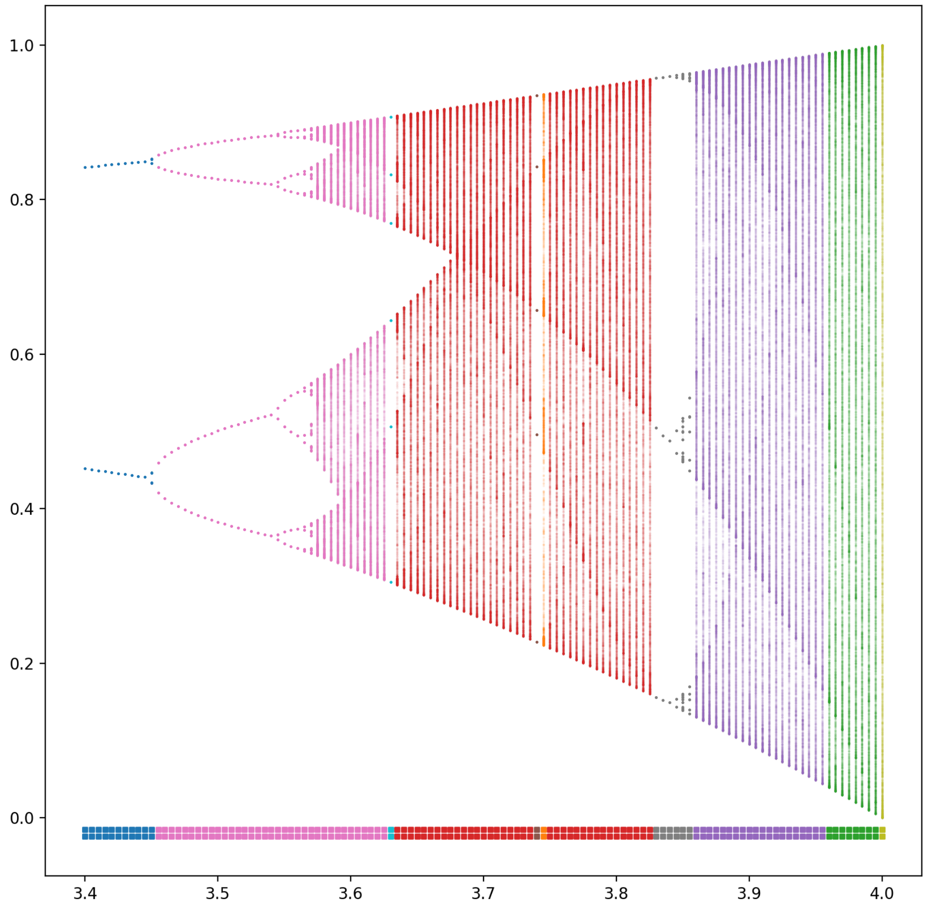

Figure 7. *Partition of the bifurcation diagram based on the clustering of time series generated from the logistic map presented in Figure 6. Every square at the bottom of the image represents a time series corresponding to a value of parameter l given by the horizontal axis. Top row corresponds to the starting point* 0.11*, bottom row to* 0.31*. Color of a square indicates the cluster into which the corresponding trajectory belongs.*

4.4. Example: Lorenz attractor and its embeddings. The fourth example is based on the Lorenz system defined by equations:

$$\begin{cases} \dot{x} &= \sigma(y-x), \\ \dot{y} &= x(\rho - z) - y, \\ \dot{z} &= xy - \beta z, \end{cases} \tag{4.8}$$

which induces a continuous dynamical system $\varphi : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$. We consider the classical values of the parameters: $\sigma = 10$, $\rho = 28$, and $\beta = 8/3$. A time series can be generated by iterates of the map $f(x) := \varphi(x, \tilde{t})$, where $\tilde{t} > 0$ is a fixed value of the time parameter. For the following experiments we chose $\tilde{t} = 0.02$ and we use the Runge-Kutta method of an order 5(4) to generate the time series.

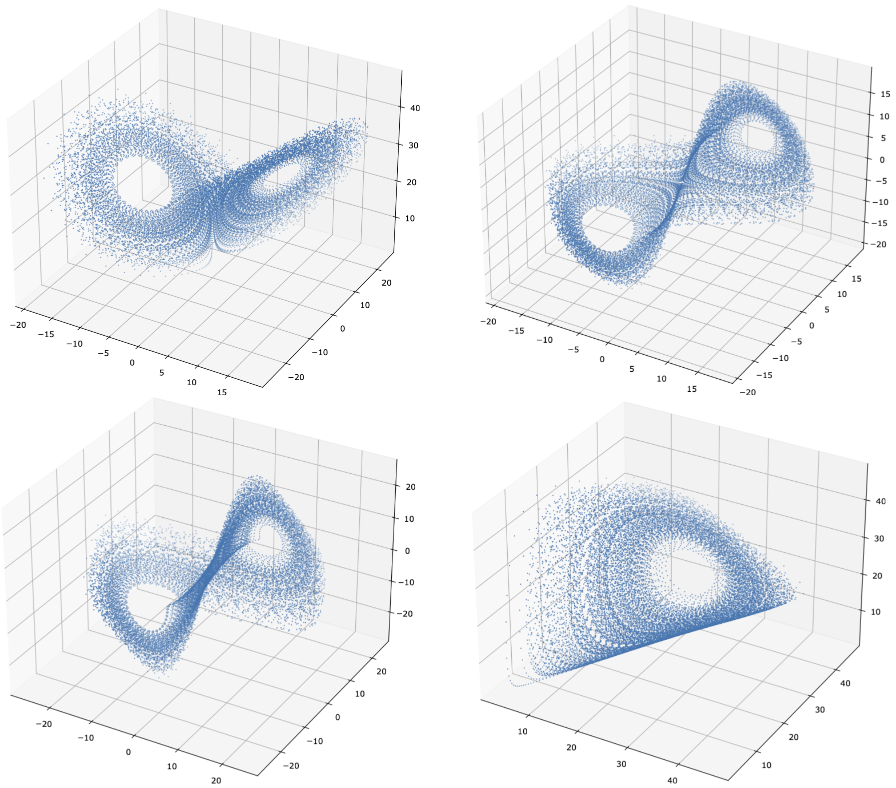

Figure 8. *A time series generated from the Lorenz system (top left) and 3-d embeddings of its projections onto the x-coordinate (top right), y-coordinate (bottom left) and z-coordinate (bottom right) with a delay parameter $l = 5$.*

4.4.1. Experiment 4A.

Setup. Let $p_1 = (1, 1, 1)$ and $p_2 = (2, 1, 1)$. In this experiment we compare the following time series:

$$\mathcal{L}_i = \varrho(f, f^{2000}(p_i), 10000), \quad \mathcal{P}^i_{x,d} = \Pi(\mathcal{L}_i^{(1)}, d, 5), \quad \mathcal{P}^i_{z,d} = \Pi(\mathcal{L}_i^{(3)}, d, 5),$$

where $i \in \{1, 2\}$. Recall that Π denotes the embedding of a time series into $\mathbb{R}^d$ and $\mathcal{L}_i^j$ is a projection of time series $\mathcal{L}_i$ onto its j-th coordinate. In all the embeddings we choose the lag $l = 5$. Note that the first point of time series $\mathcal{L}_i$ is equal to the 2000-th iterate of point p_i under map f. It is a standard procedure to cut off some transient part of the time series.

Time series $\mathcal{P}^i_{x,d}$ and $\mathcal{P}^i_{z,d}$ are embeddings of the first and third coordinate of $\mathcal{L}_i$, respectively. As Figure 8 (top right) suggests, the embedding of the first coordinate into $\mathbb{R}^3$ results in a structure topologically similar to the Lorenz attractor. The embedding of the third coordinate, due to the symmetry of the system, produces a collapsed structure with "wings" of the attractor glued together (Figure 8, right). Thus, we expect time series $\mathcal{P}^i_{z,d}$ to be recognized as non-conjugate to $\mathcal{L}_i$.

In order to compare $\mathcal{L}_i$ and embedded time series with ConjTest we shall find the suitable map h. Ideally, such a map should be a homeomorphism between the Lorenz attractor $L \subset \mathbb{R}^3$

(or precisely the ω-limit set of the corresponding initial condition under the system (4.8)) and its image $h(L)$. However, the construction of the time series allows us to easily define the best candidate for such a correspondence map pointwise for all elements of the time series.

For instance, a local approximation of h when comparing $\mathcal{L}_i \subset \mathbb{R}^3$ and $P^j_{x,d} \subset \mathbb{R}^d$ will be given by:

$$h : \mathcal{L}_i \ni \mathbf{x}_t \mapsto (\mathbf{x}_t^{(1)}, \mathbf{x}_{t+5}^{(1)}, \ldots, \mathbf{x}_{t+5d}^{(1)}) \in P^i_{x,d} \subseteq \mathbb{R}^d, \tag{4.9}$$

where $\mathbf{x}_t := (x_t, y_t, z_t) \in \mathbb{R}^3$ denotes the state of the system (4.8) at time t and $\mathbf{x}_t^{(1)} = x_t$ denotes its projection onto the x-coordinate. When $j = i$ this formula matches the points of $\mathcal{L}_i$ to the corresponding points of $P^j_{x,d} = P^i_{x,d}$. However, if $j \neq i$ then the points of $\mathcal{L}_i$ are mapped onto the points of $P^i_{x,d}$, not to $P^j_{x,d}$, thus in fact in our comparison tests we verify how well $P^j_{x,d}$ approximates the image of $\mathcal{L}_i$ under h and the original dynamics.

For the symmetric comparison of $P^j_{x,d}$ with $\mathcal{L}_i$ the local approximation of $h^{-1} : \mathbb{R}^d \to \mathbb{R}^3$ will take a form:

$$h^{-1} : P^j_{x,d} \ni (\mathbf{x}_t^{(1)}, \mathbf{x}_{t+5}^{(1)}, \ldots, \mathbf{x}_{t+5d}^{(1)}) \mapsto \mathbf{x}_t \in \mathcal{L}_j \subset \mathbb{R}^3. \tag{4.10}$$

These are naive and data driven approximations of the potential connecting map h. In particular the homeomorphism from the Lorenz attractor to its 1D embedding cannot exist, but we still can construct map h using the above receipt, which seems natural and the best candidate for such a comparison. More sophisticated ways of finding the optimal h in general situations will be the subject of our future studies.

In the experiments below we use the maximum metric.

Results. As one can expect, Table 4 shows that embeddings of the first coordinate give in general noticeably lower values then embeddings of the z'th coordinate. Thus, suggesting that $\mathcal{L}_1$ is conjugate to $\mathcal{P}^1_{x,3}$, but not to $\mathcal{P}^1_{z,3}$. Again, Table 5 shows that, in the case of chaotic systems, FNN and KNN are highly sensitive to variation in starting points of the series.

All methods suggest that 2-d embedding of the x-coordinate has structure reasonably compatible with $\mathcal{L}_1$. With the additional dimension values gets only slightly lower. One could expect that 3 dimensions would be necessary for an accurate reconstruction of the attractor. Note that Takens' Embedding Theorem suggests even dimension of 5, as the Hausdorff dimension of the Lorenz attractor is about 2.06 [30]. However, it often turns out that the dynamics can be reconstructed with the embedding dimension less than given by Takens' Embedding Theorem (as implied e.g. by Probabilistic Takens' Embedding Theorem, see [3, 4]). We also attribute our outcome to the observation that the x-coordinate carries a large piece of the system information, which is visually presented in Figure 8.

Interestingly, when we use ConjTest to compare $\mathcal{L}_1$ with embedding time series generated from $\mathcal{L}_1$ we always get values 0.0. The connecting maps used in this experiment, defined by (4.9) and (4.10), establish a direct correspondence between points in two time series. As a result we get $\tilde{h} = h$ in the definition of ConjTest, and consequently, every pair of sets in the numerator of equation (3.5) is the same. If the embedded time series comes from another trajectory then $\tilde{h} \neq h$ and ConjTest gives the expected results, as visible in Table 5. On the other hand, computationally more demanding ConjTest$^+$ exhibits virtually the same results

method \ test	$\mathcal{L}_1$ vs. $\mathcal{P}^1_{x,1}$	$\mathcal{L}_1$ vs. $\mathcal{P}^1_{x,2}$	$\mathcal{L}_1$ vs. $\mathcal{P}^1_{x,3}$	$\mathcal{L}_1$ vs. $\mathcal{P}^1_{z,1}$	$\mathcal{L}_1$ vs. $\mathcal{P}^1_{z,3}$
FNN (r=3)	0.0 / 1.0	0.0 / .362	.05 / .196	0.0 / 1.0	.111 / .541
KNN (k=5)	.019 / .465	.003 / .036	.003 / .007	.024 / .743	.002 / .519
ConjTest (k=5, t=10)	0.0 / 0.0	0.0 / 0.0	0.0 / 0.0	0.0 / 0.0	0.0 / 0.0
ConjTest$^+$ (k=5, t=10)	.330 / .401	.030 / .087	.024 / .051	.406 / .396	.046 / .407

Table 4

Comparison of conjugacy measures for time series generated by the Lorenz system. The number in the upper left part of the cell corresponds to a comparison of $\mathcal{L}_1$ vs. the second time series, while the lower right number corresponds to the inverse comparison.

method \ test	$\mathcal{L}_1$ vs. $\mathcal{L}_2$	$\mathcal{L}_1$ vs. $\mathcal{P}^2_{x,1}$	$\mathcal{L}_1$ vs. $\mathcal{P}^2_{x,2}$	$\mathcal{L}_1$ vs. $\mathcal{P}^2_{x,3}$	$\mathcal{L}_1$ vs. $\mathcal{P}^2_{z,1}$	$\mathcal{L}_1$ vs. $\mathcal{P}^2_{z,3}$
FNN (r=3)	.995 / .996	.955 / 1.0	.987 / .996	.991 / .996	.963 / 1.0	.996 / 997.
KNN (k=5)	.822 / .827	.826 / .832	.829 / .826	.829 / .825	.823 / .828	.820 / .833
ConjTest (k=5, t=10)	.010 / .009	.236 / .012	.016 / .010	.010 / .009	.391 / .012	.017 / .009
ConjTest$^+$ (k=5, t=10)	.020 / .017	.331 / .400	.039 / .092	.033 / .056	.431 / .392	.060 / .404

Table 5

Comparison of conjugacy measures for time series generated by the Lorenz system. The number in the upper left part of the cell corresponds to a comparison of $\mathcal{L}_1$ vs. the second time series, the lower right corresponds to the symmetric comparison.

in both cases, when $\mathcal{L}_1$ is compared with embeddings of its own (Table 4) and when $\mathcal{L}_1$ is compared with embeddings of $\mathcal{L}_2$ (Table 5).

4.4.2. Experiment 4B. This experiment is proceeded according to the standard use of FNN for estimating optimal embedding dimension without an explicit knowledge about the original system.

Setup. Let $p = (1, 1, 1)$, we generate the following collection of time series

$$\mathcal{L} = \varrho(f, f^{2000}(p), 10000), \quad \mathcal{P}_d = \Pi(\mathcal{L}^{(1)}, d, 5),$$

where $d \in \{1, 2, 3, 4, 5, 6\}$. In the experiment we compare pairs of embedded time series corresponding to consecutive dimensions, e.g., $\mathcal{P}_d$ with $\mathcal{P}_{d+1}$, for the entire range of parameter values. We are looking for the minimal value of d such that $\mathcal{P}_{d-1}$ is dynamically different from $\mathcal{P}_d$, but $\mathcal{P}_d$ is similar to $\mathcal{P}_{d+1}$. The interpretation says that d is optimal, because by passing from $d-1$ to d we split some false neighborhoods apart (hence, dissimilarity of dynamics), but by passing from d to $d+1$ there is no difference, because there is no false neighborhood left to be separated.

Results. Results are presented in Figure 9. In general, the outputs of all methods are consistent. When the one-dimensional embedding, $\mathcal{P}_1$, is compared with two-dimensional embedding, $\mathcal{P}_2$, we get large comparison values for the entire range of every parameter. When we compare $\mathcal{P}_2$ with $\mathcal{P}_3$ the estimation of dissimilarity drops significantly, i.e. we conclude that the time series $\mathcal{P}_2$ and $\mathcal{P}_3$ are more "similar" than $\mathcal{P}_1$ and $\mathcal{P}_2$. The comparison of $\mathcal{P}_3$ with $\mathcal{P}_4$ still decreases the values, suggesting that the third dimension improves the quality of our embedding. The curve corresponding to $\mathcal{P}_4$ vs. $\mathcal{P}_5$ essentially overlaps $\mathcal{P}_3$ vs. $\mathcal{P}_4$ curve. Thus, the third dimension seems to be a reasonable choice.

We can see that FNN (Figure 9 top left), originally designed for this test, gives a clear answer. However, in the case of KNN (Figure 9 top right) the difference between the yellow and the green curve is rather subtle. Thus, the outcome could be alternatively interpreted with a claim that two dimensions are enough for this embedding. In the case of ConjTest$^+$ we have two parameters. For the fixed value of $t = 10$ we manipulated the value of k (Figure 9 bottom left) and the outcome matched up with the FNN result. However, the situation is slightly different when we fix $k = 5$ and vary the t (Figure 9 bottom right). For $t < 30$ the results suggest dimension 3 to be optimal for the embedding, but for $t > 40$ the green and the red curve split. Moreover, for $t > 70$, we can observe the beginning of another split of the red ($\mathcal{P}_4$ vs. $\mathcal{P}_5$) and the violet ($\mathcal{P}_5$ vs. $\mathcal{P}_6$) curves. Hence, the answer is not fully conclusive. We attribute this effect to the chaotic nature of the attractor. The higher the value of t the higher the effect. We investigate it further in the following experiment.

4.4.3. Experiment 4C. In this experiment we investigate the dependence of ConjTest$^+$ on the choice of value of parameter t. Parameter t of ConjTest$^+$ controls how far we push the approximation of a neighborhood of a point x_i (U_k^i in (3.7)) through the dynamics. In the case of systems with a sensitive dependence on initial conditions (e.g., the Lorenz system) we could expect that higher values of t spread the neighborhood over the attractor. As a consequence, we obtain higher values of ConjTest$^+$.

Setup. Let $p_1 = (1, 1, 1)$, $p_2 = (2, 1, 1)$, $p_3 = (1, 2, 1)$, and $p_4 = (1, 1, 2)$. In this experiment we study the following time series:

$$\mathcal{L}_i = \varrho(f, f^{2000}(p_i), 10000), \quad \mathcal{P}^i_{x,d} = \Pi(\mathcal{L}^{(1)}_i, d, 5), \quad \mathcal{P}^i_{y,d} = \Pi(\mathcal{L}^{(2)}_i, d, 5),$$

where $i \in \{1, 2, 3, 4\}$ and $d \in \{1, 2, 3, 4\}$. We compare the reference time series $\mathcal{L}_1$ with all the others using ConjTest$^+$ method with the range of parameter

$$t \in \{1, 5, 9, 13, 17, 21, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80\}.$$

Results. The top plot of Figure 10 presents the results of comparing $\mathcal{L}_1$ to the time series $\mathcal{L}_i$ and $\mathcal{P}^i_{x,d}$ with $i \in \{2, 3, 4\}$ and $d \in \{1, 2, 3, 4\}$. Red curves correspond to $\mathcal{L}_1$ vs. $\mathcal{P}^i_{x,1}$, green curves to $\mathcal{L}_1$ vs. $\mathcal{P}^i_{x,2}$, blue curves to $\mathcal{L}_1$ vs. $\mathcal{P}^i_{x,3}$, and dark yellow curves to $\mathcal{L}_1$ vs. $\mathcal{P}^i_{x,4}$. There are three curves of every color, each one corresponds to a different starting point p_i, $i \in \{2, 3, 4\}$. The bottom part shows results for comparison of $\mathcal{L}_1$ to $\mathcal{P}^i_{y,d}$ (we embed the y-coordinate time series instead of x-coordinate). The color of the curves is interpreted analogously. Black curves on both plots are the same and correspond to the comparison of $\mathcal{L}_1$ with $\mathcal{L}_j$ for $j \in \{2, 3, 4\}$.

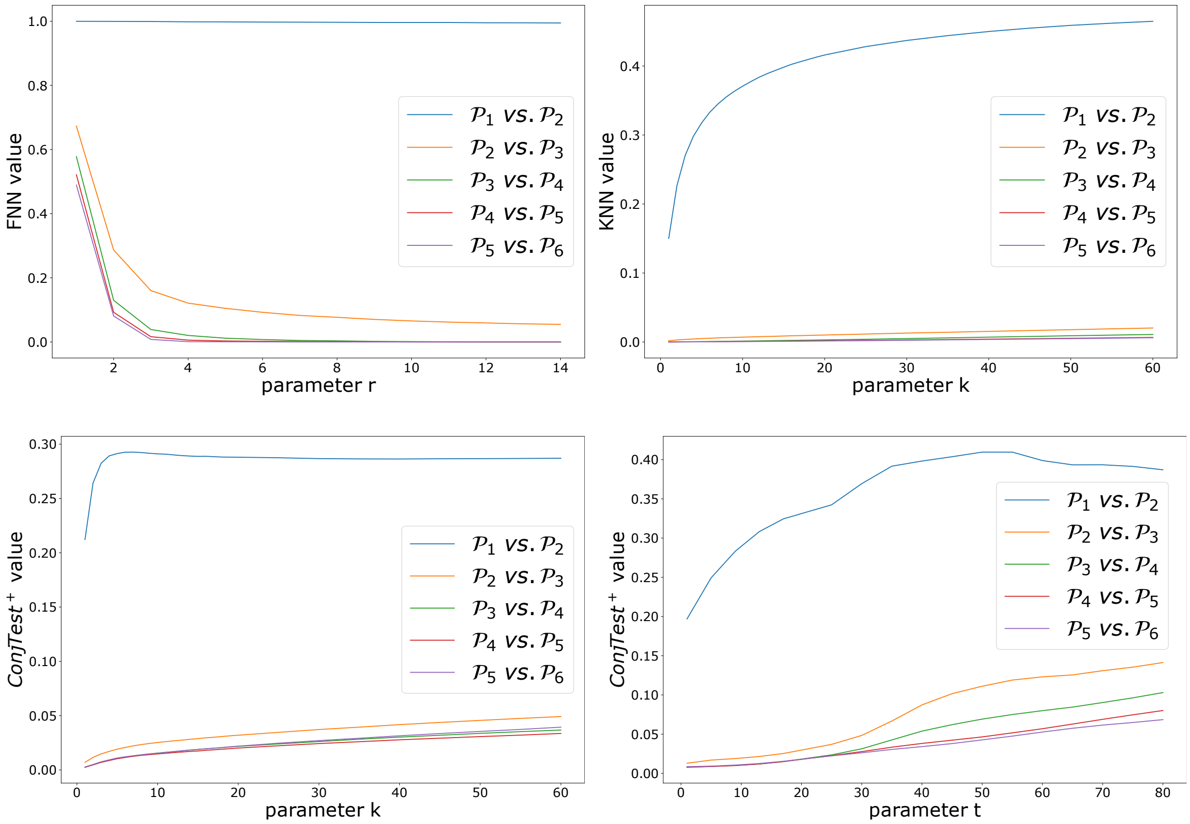

Figure 9. *A comparison of embeddings for consecutive dimensions. Top left:* FNN *with respect to parameter* r*. Top right:* KNN *with respect to parameter* k*. Bottom left:* ConjTest$^+$ *with respect to parameter* k*. Bottom right:* ConjTest$^+$ *with respect to parameter* t*.*

As expected, we can observe a drift toward higher values of ConjTest$^+$ as the value of parameter t increases. Let us recall that U_i^k in (3.7) is a k-element approximation of a neighborhood of a point x_i. The curve reflects how the image of U_i^k under f^t gets spread across the attractor with more iterations. In consequence, a 2D embedding with $t = 10$ might get lower value than 3D embedding with $t = 40$. Nevertheless, Figure 10 (top) shows consistency of the results across the tested range of values of parameter t. Red curves corresponding to 1D embeddings give significantly higher values then the others. We observe the strongest drop of values for 2D embeddings (green curves). The third dimension (blue curves) does not improve the situation essentially, except for $t \in [1, 25]$. The curves corresponding to 4D embeddings (yellow curves) overlap those of 3D embeddings. Thus, the 4D embedded system does not resemble the Lorenz attractor essentially better than the 3D embedding. It agrees with the analysis in the Experiment 4B.

The y-coordinate embeddings presented in the bottom part of Figure 10 give similar results. However, we can see that gaps between curves corresponding to different dimensions are more visible. Moreover, the absolute level of all curves is higher. We interpret this outcome with a claim that the y-coordinate inherits a bit less information about the original system than the x-coordinate. In Figure 8 we can see that y-embedding is more twisted in the center of the attractor. Hence, generally values are higher, and more temporal information is needed (reflected by higher embedding dimension) to compensate.

Note that the comparison of $\mathcal{L}_1$ to any embedding $\mathcal{P}^i_{x,d}$ is always significantly worse than comparison of $\mathcal{L}_1$ to any $\mathcal{L}_j$. This may suggest that any embedding is not perfect.

4.5. Example: rotation on the Klein bottle. In the next example we consider the Klein bottle, denoted $\mathbb{K}$ and defined as an image $\mathbb{K} := \operatorname{im}\beta$ of the map β:

$$\beta : [0, 2\pi) \times [0, 2\pi) \ni \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} \cos\frac{x}{2}\cos y - \sin\frac{x}{2}\sin(2y) \\ \sin\frac{x}{2}\cos y + \cos\frac{x}{2}\sin(2y) \\ 8\cos x(1 + \frac{\sin y}{2}) \\ 8\sin x(1 + \frac{\sin y}{2}) \end{bmatrix} \in \mathbb{R}^4. \tag{4.11}$$

In particular, the map β is a bijection onto its image and the following "rotation map" $f_{[\phi_1,\phi_2]} : \mathbb{K} \to \mathbb{K}$ over the Klein bottle is well-defined:

$$f_{[\phi_1,\phi_2]}(x) := \beta\left(\beta^{-1}(x) + \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} \mod 2\pi\right).$$

4.5.1. Experiment 5A. We conduct an experiment analogous to Experiment 4B on estimating the optimal embedding dimension of a projection of the Klein bottle.

Setup. We generate the following time series

$$\mathcal{K} = \varrho(f_{[\phi_1,\phi_2]}, (0, 0, 0, 0), 8000),$$
$$\mathcal{P}_d = \Pi\left((\mathcal{K}^{(1)} + \mathcal{K}^{(2)} + \mathcal{K}^{(3)} + \mathcal{K}^{(4)})/4, d, 8\right),$$

where $\phi_1 = \frac{\sqrt{2}}{10}$, $\phi_2 = \frac{\sqrt{3}}{10}$, $d \in \{2, 3, 4, 5\}$ and $\mathcal{K}^{(i)}$ denotes the projection onto the i-th coordinate. Note that in previous experiments we mostly used a simple observable s which was a projection onto a given coordinate. However, in general, one can consider any (smooth) function as an observable. Therefore in the current experiment, in the definition of $\mathcal{P}_d$, s is a sum of all the coordinates, not the projection onto a chosen one. Note also that because of the symmetries (see formula (4.11)) a single coordinate might be not enough to reconstruct the Klein bottle.

Results. We can proceed with the interpretation similar to Experiment 4B. The FNN results (Figure 11 top left) suggests that 4 is a sufficient embedding dimension. The similar conclusion follows from KNN (Figure 11 top right) and ConjTest$^+$ with a fixed parameter $k = 10$ (Figure 11 bottom right). The bottom left figure of 11 is inconclusive as for the higher values of k the curves do not stabilize even with high dimension.

Note that the increase of parameter t in ConjTest$^+$ (Figure 11 bottom right) does not result in drift of values as in Figure 9 (bottom right). In contrast to the Lorenz system studied in Experiment 4B the rotation on the Klein bottle is not sensitive to the initial conditions.

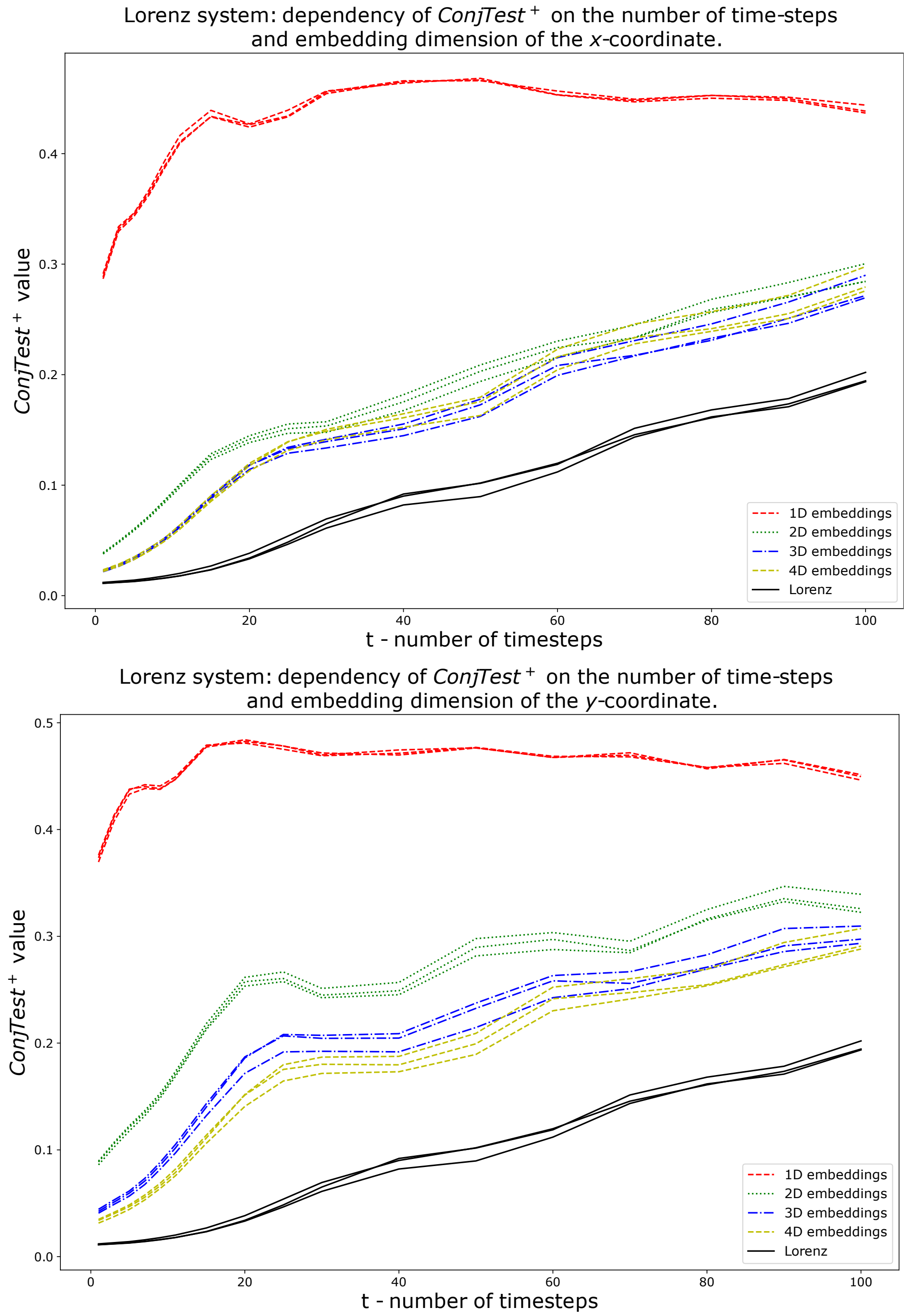

Figure 10. *Dependence of* ConjTest$^+$ *on the parameter t for Lorenz system. In this experiment multiple time series with different starting points were generated. Each of them was used to produce an embedding. Top: comparison of x-coordinate embedding with $\mathcal{L}_1$. Bottom: comparison of y-coordinate embedding with $\mathcal{L}_1$. For more explanation see text.*

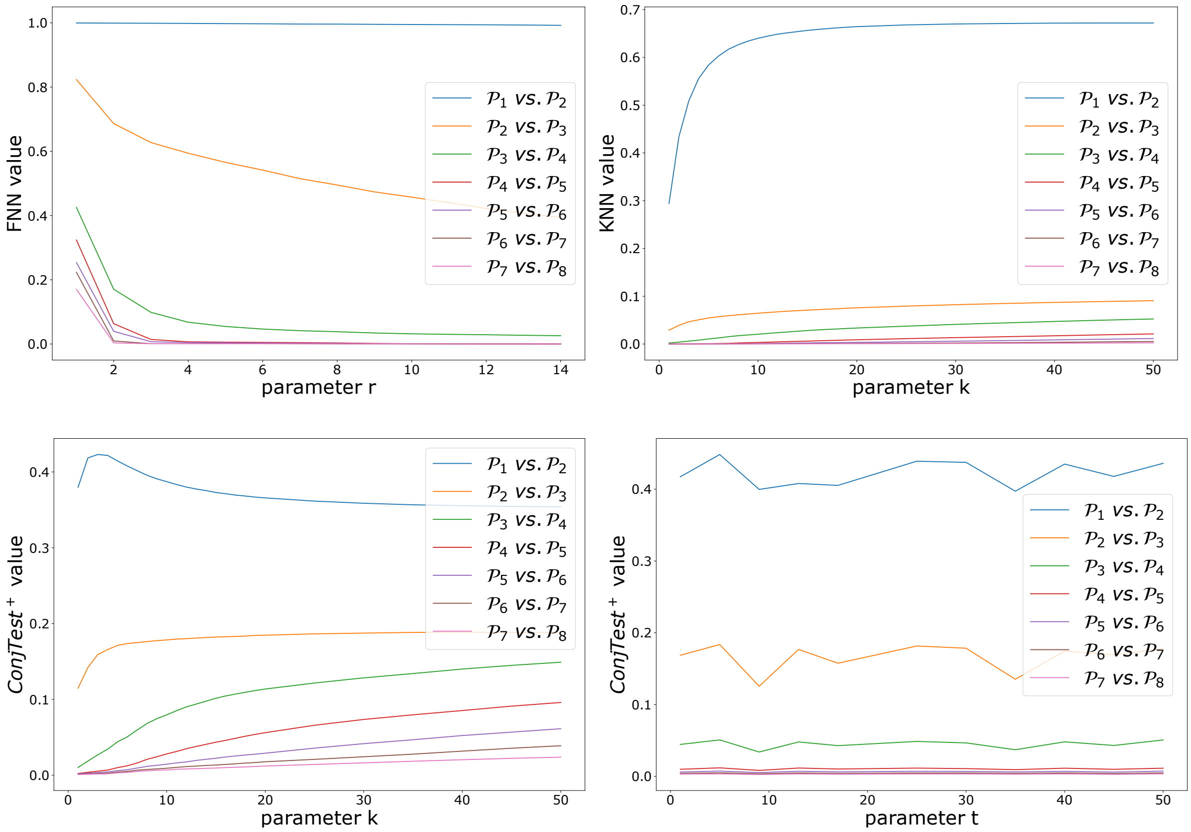

Figure 11. *A comparison of the conjugacy measures for embeddings of the Klein bottle for consecutive dimensions. Top left:* FNN *with respect to parameter* r*. Top right:* KNN *with respect to parameter* k*. Bottom left:* ConjTest$^+$ *with respect to parameter* k *(*$t = 10$ *fixed). Bottom right:* ConjTest$^+$ *with respect to parameter* t *(*$k = 5$ *fixed).*

5. Approximation of the connecting homeomorphism.

In whole generality, finding the connecting homeomorphism between conjugate dynamical systems, is a very difficult task. Some prior work has been done in this direction but the existing methods still have many limitations in applying for a broader class of systems. In particular, the works [26, 27, 32] developed a method to produce conjugacy functions based on a functional fixed-point iteration scheme that can also be generalized to compare non-conjugate dynamical systems in which case the limit point of a fixed-point iteration scheme yields a function called a "commuter". Quantifying how much the commuter function fails to be a homeomorphism (in various measures) led to the notion of a "homeomorphic defect" that, as the authors point out, allows one to quantify the dissimilarity of the two dynamical systems. However, the method has been illustrated on a very few specific examples and could be rigorously mathematically justified only under very restrictive assumptions e.g. uniform contraction of at least one of the systems when comparing systems in one dimension. Significant problems occur in rigorous extension to systems of higher dimension. Consequently, later work [5] extended this theory to allow for multivariate transformations and presented ideas on constructing commuter functions different than fixed-point iteration scheme. This method is based on symbolic dynamics approach which, however, requires existence and finding a general partition for systems being compared. Although the approach of finding a (semi-)conjugacy to a symbolic shift space through generating or Markov partition seems very natural from the point of view of dynamical systems theory, in practice finding a *reasonable* partition even for systems given by explicit equations (not to mention time series of real data) is often not feasible. Moreover, one can face an explosion of computational complexity as the number of symbols increases. Later work [33] employs a method of graph matching between the graphs representing the underlying symbolic dynamics or, alternatively, between the graphs approximating the action of the systems on some eligible partition. Interestingly, the authors show that the permutation matrices that relate the adjacency matrices of the merging graphs coincide with the solution of Monge's mass transport problem.

The above earlier works contain valuable ideas on finding to-be-conjugacies or commuter functions and the defect measures of the arising commuters might serve as a quantification of the dynamical similarity between two given systems. In turn, our proposed tools, ConjTest and ConjTest$^+$ can be applied, among others, to explicitly assess the quality of the matching between the two systems through the commuting functions obtained by the above mentioned methods and these matching functions can be candidates for testing (semi-)conjugacy of given systems by ConjTests. Note also that, contrary to the previous works, ConjTest methods can be applied directly to the time-series since we work on the point clouds and do not need a priori the formulas for systems which generated them - these are only used as benchmark tests.

However, as our contribution and small step forward towards effective algorithms of finding conjugating maps, in this section we present a proof-of-concept gradient-descent algorithm, utilizing the ConjTest as a cost function, to approximate such a connecting homeomorphism. More precisely, as an example, we use it to discover an approximation of the map (4.5) that constitutes a topological conjugacy between the tent and the logistic map (see Section 4.3). Instead of finding an analytical formula approximating the connecting homeomorphism our strategy aims to construct a cubical set representing the map. Further development and generalization of the presented procedure will be a subject of forthcoming studies.

Consider the following sequence $0 = a_1 < a_2 < \ldots < a_{n+1} = 1$. Denote $\mathtt{A}_{i,j} := [a_i, a_{i+1}] \times [a_j, a_{j+1}]$ and $\mathbb{A} := \{\mathtt{A}_{i,j} \mid i,j \in \{1,2,\ldots n\}\}$. Let $h : I \to I$ be an increasing homeomorphism from the unit interval I to itself and by $\pi(h) := \{(x,y) \in I \times I \mid y = h(x)\}$ denote the graph of h. We say that a collection $\mathtt{h} = \{\mathtt{A}_{i,j} \in \mathbb{A} \mid \operatorname{int} \mathtt{A}_{i,j} \cap \pi(h) \neq \emptyset\}$ is a *the cubical approximation* of h and we denote it by $[h]$. Equivalently, $\mathtt{h}$ is the minimal subset of $\mathbb{A}$ such that $\pi(h) \subset \bigcup \mathtt{h}$. We refer to

$$\mathtt{H} := \{[h] \subset \mathbb{A} \mid h : I \to I \text{ - an increasing homeomorphism}\}$$

as a family of all *cubical homeomorphisms* of $\mathbb{A}$.

In A we show how to construct a class of piecewise linear homeomorphisms for any $\mathtt{h} \in \mathtt{H}$. We denote a *selector* of $\mathtt{h}$, that is a homeomorphism representing $\mathtt{h}$, by $f_{\mathtt{h}}$. Take, as an example, cubical sets marked with yellow cubes in Figure 12. At every panel, the blue curve corresponds to the graph of the selector.

The size of family $\mathtt{H}$ grows exponentially with the resolution of the grid (the explicit formula for the size of $\mathtt{H}$ is given in A). For instance, the number of cubical homeomorphisms for $m = 21$ is about $2.6 \cdot 10^{14}$. Thus, it is hopeless to find the optimal approximation of the connecting homeomorphism by a brute examination of all elements of $\mathtt{H}$. Instead, we propose an algorithm based on the gradient descent strategy using ConjTest as a cost function.

Algorithm 5.1 ApproximateH

Input: $\mathcal{A}, \mathcal{B}$ – time series on an interval, $\mathtt{h}_0 \in \mathtt{h}$ – initial approximation of the homeomorphism, nsteps – number of steps, p – memory size.

Output: best_h – approximated connecting homeomorphism

$\mathtt{h} \leftarrow \mathtt{h}_0$
$\mathtt{q} \leftarrow$ initialize queue of size p with null's
$\mathtt{best\_h} \leftarrow \mathtt{h}$
for $\mathtt{t} = 0$ to nsteps **do**
  $\mathtt{c} \leftarrow \{\mathtt{h}' \in \mathtt{h} \mid \mathtt{h}' \in \mathtt{nbhd}(\mathtt{h})$ and $\mathtt{diff}(\mathtt{h}, \mathtt{h}') \not\subset \mathtt{q}\}$
  if $\#\mathtt{c} = 0$ **then**
    break
  else
    $\mathtt{h} \leftarrow \mathtt{h}' \in \mathtt{c}$ with a minimal value of $\mathtt{score}(\mathtt{h}', \mathcal{A}, \mathcal{B})$
    if $\mathtt{score}(\mathtt{h}, \mathcal{A}, \mathcal{B}) < \mathtt{score}(\mathtt{best\_h}, \mathcal{A}, \mathcal{B})$ **then**
      $\mathtt{best\_h} \leftarrow \mathtt{h}$
    end if
    $\mathtt{q}.\mathrm{append}(\mathtt{diff}(\mathtt{h}, \mathtt{h}'))$ {append the unique element differentiating h and h′}
    $\mathtt{q}.\mathrm{pop}()$
  end if
end for
return best_h

Let $\mathcal{A}$ and $\mathcal{B}$ be time series on a unit interval. Algorithm 5.1 attempts to find an element of $\mathtt{H}$ with as small value of the ConjTest as possible. For that purpose, each element $\mathtt{h} \in \mathtt{H}$ can be assigned with the following score:

$$\mathtt{score}(\mathtt{h}, \mathcal{A}, \mathcal{B}) := \max\left\{\mathrm{ConjTest}(\mathcal{A}, \mathcal{B}; k, t, f_{\mathtt{h}}), \mathrm{ConjTest}(\mathcal{B}, \mathcal{A}; k, t, f_{\mathtt{h}}^{-1})\right\}.$$

Since elements $\mathtt{h}, \mathtt{h}' \in \mathtt{H}$ are collections of sets, the symmetric difference gives a set of cubes differing $\mathtt{h}$ and $\mathtt{h}'$. We denote it by

$$\mathtt{diff}(\mathtt{h}, \mathtt{h}') := (\mathtt{h} \setminus \mathtt{h}') \cup (\mathtt{h}' \setminus \mathtt{h}).$$

Let $\mathtt{h} \in \mathtt{H}$ be an initial guess for the connecting homeomorphism. In each step of the algorithm an attempt is made to update it in a way that the score gets improved. Each iteration of the main loop considers all neighbors $\mathtt{h}'$ of $\mathtt{h}$ in $\mathtt{H}$ such that $\mathtt{nbhd}(\mathtt{h}) := \{\mathtt{h}' \in \mathtt{H} \mid \#\,\mathtt{diff}(\mathtt{h}, \mathtt{h}') = 1\}$ (a unit sphere in a Hamming distance centered in $\mathtt{h}$). The element $\mathtt{h}'$ of $\mathtt{nbhd}(\mathtt{h})$ with minimal $\mathtt{score}(\mathtt{h}')$ is chosen for the next iteration of the algorithm. Note that it might happen that $\mathtt{score}(\mathtt{h}') < \mathtt{score}(\mathtt{h})$. This prevents the algorithm from being stuck at a local minimum. In addition, to avoid orbiting around them we exclude elements of $\mathtt{nbhd}(\mathtt{h})$

for which element $\texttt{diff}(\texttt{h}, \texttt{h}')$ is an element added or removed in previous `p` iterations of the algorithm. This strategy is more restrictive from just avoiding assigning to `h` the same cubical homeomorphism twice within `p` consecutive steps which still could result in oscillatory changes of some cubes.

We conducted an experiment for the problem studied in Section 4.3, that is, a comparison of time series generated by the logistic and the tent map. Take the following time series

$$\mathcal{A} = \varrho(f_4, 0.02, 500) \quad \text{and} \quad \mathcal{B} = \varrho(g_2, 0.87, 500),$$

where f_4 and g_2 are respectively a logistic and a tent map as in Section 4.3. We use `score` with parameters $k = 5$ and $t = 1$. As an initial guess of the connecting homeomorphism $\texttt{h}_0$ we naively took a cubical approximation of a $h(x) = x^5$ with resolution $m = 21$, as presented in the top-left panel of Figure 12. We run Algorithm 5.1 for time series $\mathcal{A}$ and $\mathcal{B}$ for 1000 steps with the memory parameter $\texttt{p} = m = 21$. Figure 13 shows the values of the `score` for the consecutive approximations. We can see that the algorithm falls temporarily into local minima, but eventually, thanks to the memory parameter, it escapes them and settles down towards the low score values. Figure 12 shows relations corresponding to the 1st, 200th, 400th, and 612nd iteration of the algorithm run. The bottom-right panel, the 612nd iteration is the relation inducing the lowest score among all iterations. The orange curve, at the same panel, is the graph of homeomorphism (4.5) – the analytically correct map conjugating f_4 and g_2. Clearly, the iterations are converging towards the right value of the connecting homeomorphism.

Clearly, the presented approach can be applied to any one-dimensional time series. A generalization of the algorithm will be a subject of further studies.

6. Discussion and Conclusions.

There is a considerable gap between theory and practice when working with dynamical systems; In theoretical consideration, the exact formulas describing the considered system is usually known. Yet in biology, economy, medicine, and many other disciplines, those formulas are unknown; only a finite sample of dynamics is given. This sample contains either sequence of points in the phase space, or one-dimensional time series obtained by applying an observable function to the trajectory of the unknown dynamics. This paper provides tools, FNN, KNN, ConjTest, and ConjTest$^+$, which can be used to test how similar two dynamical systems are, knowing them only through a finite sample. Proof of consistency of some of the presented methods is given.

The first method, FNN distance, is a modification of the classical False Nearest Neighbor technique designed to estimate the embedding dimension of a time series. The second one, KNN distance, has been proposed as an alternative to FNN that takes into account larger neighborhood of a point, not only the nearest neighbor. The conducted experiments show a strong similarity of FNN and KNN methods. Additionally, both methods admit similar requirements with respect to the time series being compared: they should have the same length and their points should be in the exact correspondence, i.e., we imply that an i-th point of the first time series is a dynamical counterpart of the i-th point of the second time series. An approximately binary response characterizes both methods in the sense that they return either a value close to 0 when the compared time series come from conjugate systems, or a significantly higher, non-zero value in the other case. This rigidness might be advantageous

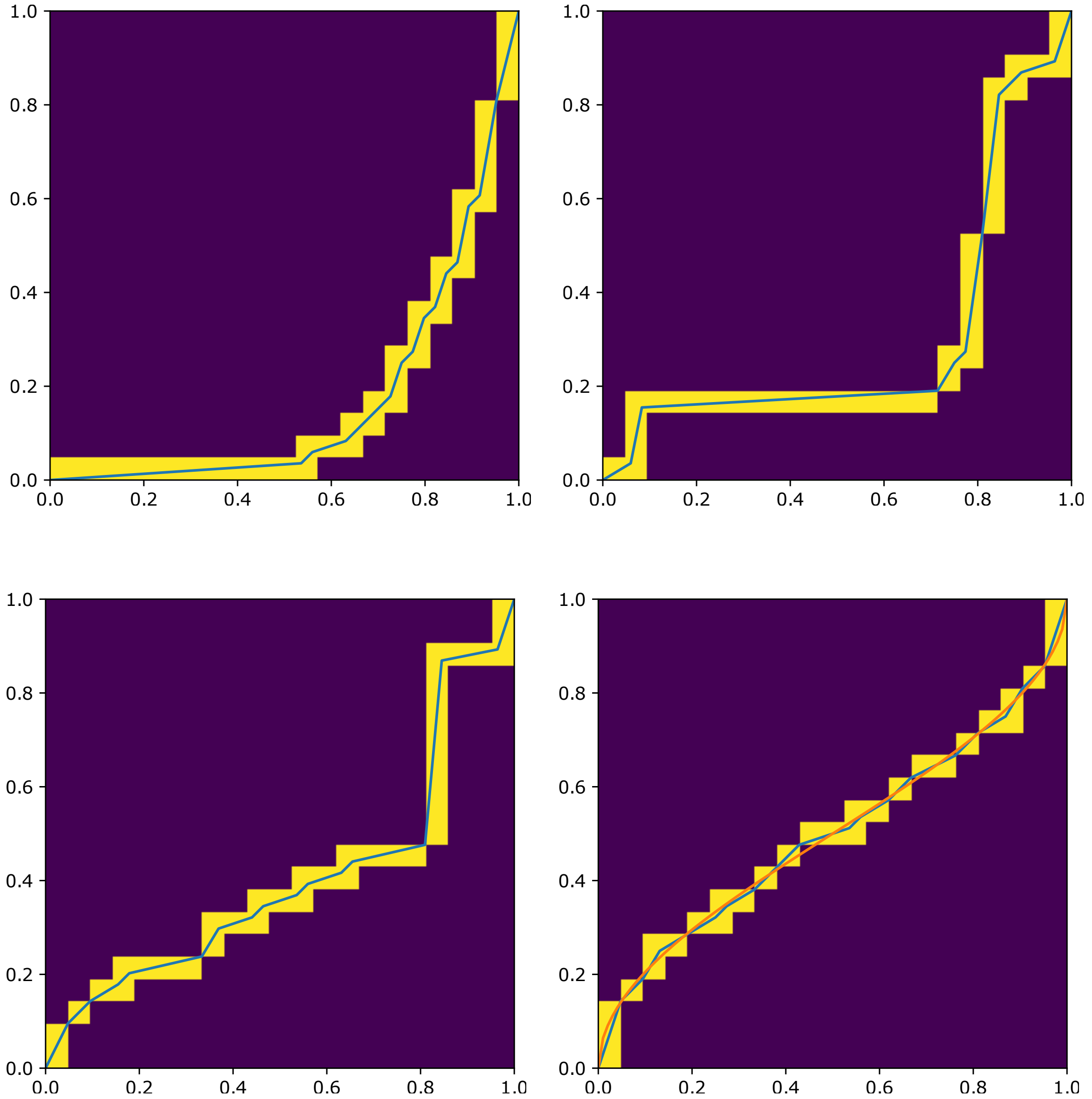

Figure 12. *Steps* 0, 200, 400, *and* 612 *(top left, top right, bottom left, bottom right, respectively) of the run of Algorithm* 5.1 *in a search for the connecting homeomorphism between the logistic and the tent map. The blue lines corresponds to a selector of a cubical homeomorphism. The orange curve in bottom right panel is a graph of the actual connecting homeomorphism* (4.5) *between the logistic and the tent map.*

in some cases. However, for most empirical settings, due to the presence of various kind of noise, FNN and KNN may fail to recognize similarities between time series. Consequently, these two methods are very sensitive to any perturbation of the initial condition of time series as well as the parameters of the considered systems. However, KNN, in contrast to FNN, admits robustness on a measurements noise as presented in Experiment 1C. On the

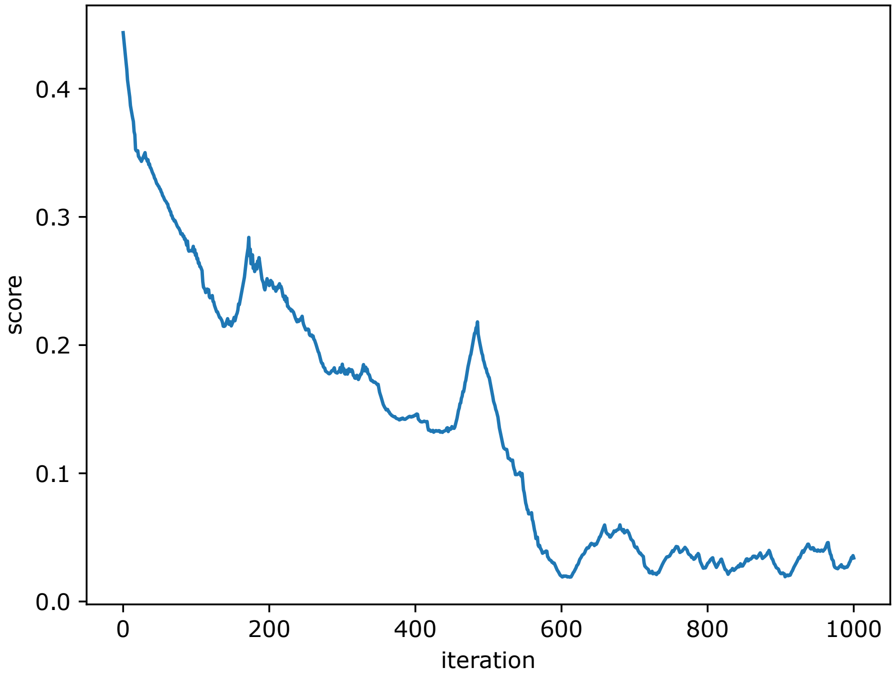

Figure 13. *The value of* $\texttt{score}(\texttt{h}, \mathcal{A}, \mathcal{B})$ *for every step of the experiment approximating the connecting homeomorphism between the logistic and the tent map.*

other hand, FNN performs better than KNN in estimating the sufficient embedding dimension (Experiments 4B, 5A). Moreover, the apparently clear response given by FNN and KNN tests might not be correct (see Experiment $1A$, $\mathcal{R}_1$ vs. $\mathcal{R}_4$).

Both ConjTest and ConjTest$^+$ (collectively called ConjTest methods) are directly inspired by the definition and properties of topological conjugacy. They are more flexible in all considered experiments and can be applied to time series of different lengths and generated by different initial conditions (the first point of the series). In contrast to FNN and KNN, they admit more robust behavior with respect to any kind of perturbation, be it measurement noise (Experiment $1C$), perturbation of the initial condition (Experiments 1A, 2A, 3A, and 4A), t parameter (Experiment $4C$), or a parameter of a system (Experiment $1B$). In most experiments, we can observe a continuous-like dependence of the test value on the level of perturbations. We see this effect as softening the concept of topological conjugacy by ConjTest methods. A downside of this weakening is a lack of definite response whether two time series come from conjugate dynamical systems. Hence the ConjTest methods should be considered as a means for a quantification of a dynamical similarity of two processes. Experiments 1A, 2A, and 3A show that both methods, ConjTest and ConjTest$^+$, capture essentially the same information from data. In general, ConjTest is simpler and, thus, computationally more efficient. However, Experiment 4A shows that ConjTest (in contrast to ConjTest$^+$) does not work well in the context of embedded time series, especially when the compared embeddings are constructed from the same time series. Experiments 4B and 5A show that the variation of ConjTest methods with respect to the t parameter can also be used for estimating

<table>
<tr><th>Property \ Method</th><th>FNN</th><th>KNN</th><th>ConjTest</th><th>ConjTest$^+$</th></tr>
<tr><td>Requirements</td><td colspan="2">identical matching between indexes of the elements (in particular the series must be of the same length)</td><td colspan="2">• an exact correspondence between the two time series is not needed
• allow examining arbitrary, even very complicated potential relations between the series
• can be used for comparison of time series of different length
• require defining the possible (semi)conjugacy h at least locally i.e. giving the corresponding relation between indexes of the elements of the two series</td></tr>
<tr><td>Parameters</td><td>only one parameter: r (but one should examine large interval of r values)</td><td>only one parameter: k (but is recommend to check a couple of different k values)</td><td colspan="2">involve two parameters: k and t</td></tr>
<tr><td>Robustness</td><td colspan="2">• less robust to noise and perturbation than ConjTest methods
• give nearly a binary output
• KNN seems to admit robustness with respect to the measurement noise</td><td colspan="2">• more robust to noise and perturbation
• the returned answer depends continuously on the level of perturbation and noise compared to the binary response given by FNN or KNN</td></tr>
<tr><td>Recurrent properties</td><td>takes into account only the one closest return of a series (trajectory) to each neighborhood</td><td colspan="3">takes into account k-closest returns</td></tr>
<tr><td>Further properties</td><td></td><td></td><td>more likely to give false positive answer than ConjTest$^+$</td><td>more computationally demanding than ConjTest but usually more reliable</td></tr>
</table>

Table 6

Comparison of the properties of discussed conjugacy measures.

a good embedding dimension. Further comparison between ConjTest and ConjTest$^+$ reveals that ConjTest$^+$ is more computationally demanding than ConjTest, but also more reliable. Indeed, in our examples with rotations on the circle and torus and with the logistic map, both these tests gave nearly identical results, but the examples with the Lorenz system show that ConjTest is more likely to give a false positive answer. This is due to the fact that ConjTest works well if the map h connecting time series $\mathcal{A}$ and $\mathcal{B}$ is a reasonably good approximation of the true conjugating homeomorphism, but in case of embeddings and naive, point-wise connection map, as in some of our examples with Lorenz system, the Hausdorff distance in formula (3.5) might vanish resulting in false positive.

The advantages of ConjTest and ConjTest$^+$ methods come with the price of finding a connecting map relating two time series. When it is unknown, in the simplest case, one can try the map h which is defined only locally i.e. on points of the time series and provide an order- preserving matching of indexes of corresponding points in the time series. The simplest example of such a map is an identity map between indices. The question of finding an optimal matching is, however, much more challenging and will be a subject of a further study. Nonetheless, in Section 5 and Appendix A we present preliminary results approaching this challenge.

A convenient summary of the presented methods is gathered in Table 6.

Appendix A. Cubical homeomorphisms.

This appendix offers additional characterization of the family of cubical homeomorphisms introduced in Section 5.

At first, observe that elements of $\mathtt{H}$ have the following straightforward observations.

Proposition A.1. *Let* $[h] \in \mathtt{H}$. *Then,* $\mathtt{A}_{1,1}, \mathtt{A}_{n,n} \in [h]$.

Proof. Since h is an increasing homeomorphism it follows that $(0,0),(1,1) \in \pi(h)$. In consequence, $\mathtt{A}_{1,1}, \mathtt{A}_{n,n} \in [h]$, because these are the only elements of $\mathbb{A}$ containing $(0,0)$ and $(1,1)$. ■

We picture the idea of the following simple proposition with Figure 14.

Proposition A.2. *Let* $\pi(h) \cap \operatorname{int} \mathtt{A}_{i,j} \neq \emptyset$ *then exactly one of the following holds*

(1) $\operatorname{int} \mathtt{A}_{i+1,j} \cap \pi(h) \neq \emptyset$ *and* $\operatorname{int} \mathtt{A}_{i,j+1} \cap \pi(h) = \emptyset$*, when* $h(a_{i+1}) \in (a_j, a_{j+1})$*,*

(2) $\operatorname{int} \mathtt{A}_{i,j+1} \cap \pi(h) \neq \emptyset$ *and* $\operatorname{int} \mathtt{A}_{i+1,j} \cap \pi(h) = \emptyset$*, when* $h^{-1}(a_{i+1}) \in (a_i, a_{i+1})$*,*

(3) $\operatorname{int} \mathtt{A}_{i+1,j+1} \cap \pi(h) \neq \emptyset$, $\operatorname{int} \mathtt{A}_{i+1,j} \cap \pi(h) = \emptyset$ *and* $\operatorname{int} \mathtt{A}_{i,j+1} \cap \pi(h) = \emptyset$*, when* $h(a_{i+1}) = a_{j+1}$.

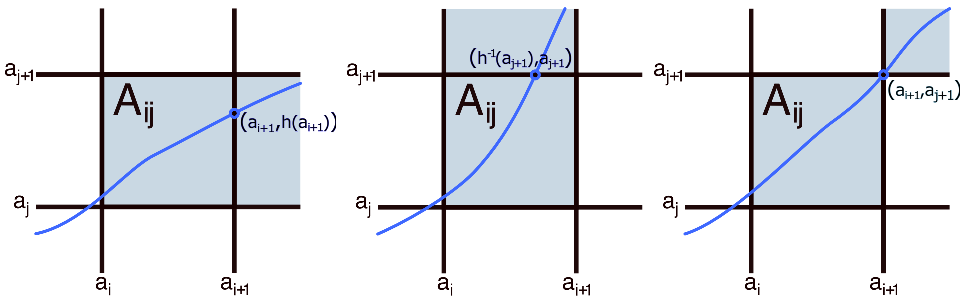

Figure 14. *From left to right, cases (1), (2) and (3) of Proposition A.2.*

Again, the proof for the next proposition is a consequence of basic properties of homeomorphism h such that $[h] = \mathtt{h}$.

Proposition A.3. *Let* $\mathtt{h} \in \mathtt{H}$. *Then*

(i) for every $i \in \{1, \dots, n\}$ *set* $\bigcup\{\mathtt{A}_{i,j} \in \mathtt{h} \mid j \in \{0,1,\dots,n\}\}$ *is nonempty and connected,*

(ii) for every $j \in \{1, \dots, n\}$ *set* $\bigcup\{\mathtt{A}_{i,j} \in \mathtt{h} \mid i \in \{0,1,\dots,n\}\}$ *is nonempty and connected.*

(iii) if $\mathtt{A}_{i,j} \in \mathtt{h}$ *then for every* $i' > i$ *and* $j' < j$ *we have* $\mathtt{A}_{i',j'} \notin \mathtt{h}$*,*

(iv) if $\mathtt{A}_{i,j} \in \mathtt{h}$ *then for every* $j' > j$ *and* $i' < i$ *we have* $\mathtt{A}_{i',j'} \notin \mathtt{h}$*,*

The above propositions implies that every element $\mathtt{h} \in \mathtt{H}$ can be seen as a path starting from element $\mathtt{A}_{1,1}$ and ending at $\mathtt{A}_{n,n}$. In particular, $\mathtt{h}$ can be represented as a vector of symbols R (right, incrementation of index i, case (1)), U (up, incrementation of index j, case (2)), and D (diagonal, incrementation of both indices, case (3)). The vectors do not have to be of the same length. In particular, a single symbol D can replace a pair of symbols R and U. Denote by n_R, n_U and n_D the number of corresponding symbols in the vector. As indices i and j have to be incremented from 1 to n we have the following properties:

(A.1) $$0 \leq n_R, n_U, n_D \leq n-1, \;\; n_R + n_U + 2n_D = 2(n-1) \;\; \text{and} \;\; n_R = n_U.$$

Actually, any vector $\mathtt{V}$ of symbols $\{R, U, D\}$ satisfying the above conditions (A.1) corresponds to a cubical homeomorphism. We show it by constructing a piecewise-linear homeomorphism h such that $[h] = \mathtt{h}$ for any $\mathtt{h}$ represented by $\mathtt{V}$. We refer to the constructed h as a *selector* of $\mathtt{h}$. In particular, Algorithm A.1 produces a sequence of points corresponding to points of non-differentiability of the homeomorphism. We have five type of points, the starting point $(0,0)$ (type B), the ending point $(1,1)$ (type E), and points corresponding to subsequences UR (type UR), RU (type RU) and D (type D). Proposition A.4 shows that the map generated by the algorithm is an actual homeomorphism.

Algorithm A.1 `FindSelector`

Input: $\mathtt{V}$ – a vector of symbols $\{R, U, D\}$ satisfying (A.1), p – a parameter for breaking points selection

Output: L – a sequence encoding the selector

$L \leftarrow \{(0,0)\}$
$\mathtt{prev} \leftarrow \mathtt{Null}$
$i, j \leftarrow 0$
for $\mathtt{s} \in \mathtt{V}$ **do**
  if $\mathtt{prev} = U$ and $\mathtt{s} = R$ **then**
    $L = L \cup \{(p\, a_i + (1-p)\, a_{i+1}, (1-p)\, a_j + p\, a_{j+1})\}$
  else if $\mathtt{prev} = R$ and $\mathtt{s} = U$ **then**
    $L = L \cup \{((1-p)\, a_i + p\, a_{i+1}, p\, a_j + (1-p)\, a_{j+1})\}$
  else if $\mathtt{s} = D$ **then**
    $L = L \cup \{(a_{i+1}, a_{j+1})\}$
  end if
  $\mathtt{prev} \leftarrow \mathtt{s}$
  if $\mathtt{s} = R$ or $\mathtt{s} = D$ **then**
    $i \leftarrow i + 1$
  end if
  if $\mathtt{s} = U$ or $\mathtt{s} = D$ **then**
    $j \leftarrow j + 1$
  end if
end for
$L \leftarrow \{(1,1)\}$
return L

Proposition A.4. *Let* $\mathtt{V}$ *be a vector of symbols* $\{R, U, D\}$ *satisfying* (A.1) *and* $\mathtt{h}$ *the corresponding cubical set. Let* $L = \{(x_1, y_1), (x_2, y_2), \ldots, (x_K, y_K)\}$ *be a sequence of points generated by Algorithm A.1 for* $\mathtt{V}$*. Then, for every* $k \in \{1, 2, \ldots K-1\}$ *following properties are satisfied:*

(i) $x_k < x_{k+1}$ *and* $y_k < y_{k+1}$*,*

(ii) if $(x_k, y_k) \in L$ *then* $(x_k, y_k) + (1-t)(x_k, y_{k+1}) \in \bigcup \mathtt{h}$ *for* $t \in [0, 1]$*.*

Proof. First, note that if (x_k, y_k) is of type UR or RU then we have $(x_k, y_k) \in \operatorname{int} A_{i,j}$. If (x_k, y_k) is of type D we have $(x_k, y_k) = \mathtt{A}_{i,j} \cap \mathtt{A}_{i+1,j+1}$. In particular, in that case $x_k, y_k \in \{a_2, a_3, \ldots, a_n\}$.

Sequence L always begins $(x_0, y_0) = (0, 0)$ and ends with $(x_K, y_K) = (1, 1)$. By Proposition A.1 we have $\mathtt{A}_{0,0}, \mathtt{A}_{n,n} \in \mathtt{h}$. The above observations shows that for every (x_k, y_k) with $k \in \{1, 2 \ldots, K-1\}$ we have $0 < x_k, y_k < 1$. Thus, two first and two last points of L satisfies (i). If (x_1, y_1) is of type UR then it follows that $\mathtt{V}$ begins with a sequence of U's. We have $(x_1, y_1) \in \operatorname{int} \mathtt{A}_{0,j}$ for some $j > 0$. By Proposition A.3(ii) all $\mathtt{A}_{0,j'} \in \mathtt{h}$ for $0 \le j' \le j$. Hence, the interval spanned by (x_0, y_0) and (x_1, y_1) is contained in $\bigcup \mathtt{h}$ proving (ii). The cases when (x_1, y_1) is of type RU or D as well as analysis of points (x_{n-1}, y_{n-1}) and (x_n, y_n) follows by similar argument.

Let (x, y) and (x', y') be two consecutive points of L. Suppose that (x, y) is of type UR and (x', y') of type RU. This situation arises when two symbols U are separated by a positive number of symbols R (see Figure 15 top). It follows that

$$(x, y) = (p\, a_i + (1-p)\, a_{i+1}, (1-p)\, a_j + p\, a_{j+1}) \operatorname{int} \mathtt{A}_{i,j},$$
$$(x', y') = ((1-p)\, a_{i'} + p\, a_{i'+1}, p\, a_j + (1-p)\, a_{j+1}) \operatorname{int} \mathtt{A}_{i',j},$$

where $i < i'$. Thus,

$$\begin{aligned} y' - y &= pa_j + (1-p)a_{j+1} - ((1-p)a_j + pa_{j+1}) \\ &= 2pa_j + a_{j+1} - a_j > a_{j+1} - a_j > 0. \end{aligned}$$

Consequently, we get $x < x'$ and $y < y'$ proving (i) for this case. By Proposition A.3(ii) we get that $\mathtt{A}_{i'',j} \in \mathtt{h}$ for all $i \le i'' \le i'$. Hence, the interval spanned by (x, y) and (x', y') is contained in $\bigcup \mathtt{h}$ proving (ii).

The case when (x, y) is of type RU and (x', y') of type UR is analogous.

Suppose that (x, y) is of type UR and (x', y') of type D. This situation arises when symbols U and D are separated by a positive number of symbols R (see Figure 15 middle). It follows that

$$(x, y) = (p\, a_i + (1-p)\, a_{i+1}, (1-p)\, a_j + p\, a_{j+1}) \in \operatorname{int} \mathtt{A}_{i,j},$$
$$(x', y') = (a_{i'+1}, a_{j+1}) = \mathtt{A}_{i',j} \cap \mathtt{A}_{i'+1,j+1},$$

where $i < i'$. It follows that $x < x'$ and $y < y'$ proving (i) for this case. Again, by Proposition A.3(ii) we can prove (ii).

The case when (x, y) is of type RU and (x', y') of type D is analogous.

Now, suppose that (x, y) is of type D and (x', y') of type UR. This situation arises when symbols D and R are separated by a positive number of symbols U (see Figure 15 bottom). It follows that

$$
\begin{aligned}
(x, y) &= (a_{i+1}, a_{j+1}) = \mathtt{A}_{i,j} \cap \mathtt{A}_{i+1,j+1}, \\
(x', y') &= (p\, a_{i+1} + (1-p)\, a_{i+2}, (1-p)\, a_{j'} + p\, a_{j'+1}) \in \operatorname{int} \mathtt{A}_{i+1,j'+1},
\end{aligned}
$$

where $j < j'$. It follows that $x < x'$ and $y < y'$ proving (i) for this case. By Proposition A.3(i) follows property (ii).

The case when (x, y) is of type D and (x', y') of type RU is analogous.

Finally, if both (x, y) and (x', y') are of type D it follows that

$$
\begin{aligned}
(x, y) &= (a_{i+1}, a_{j+1}) = \mathtt{A}_{i,j} \cap \mathtt{A}_{i+1,j+1}, \\
(x', y') &= (a_{i+2}, a_{j+2}) = \mathtt{A}_{i+1,j+1} \cap \mathtt{A}_{i+2,j+2}.
\end{aligned}
$$

Thus, (x, y) and (x', y') are the opposite corners of cube $\mathtt{A}_{i+1,j+1}$ which immediately gives both properties (i) and (ii). ■

By counting all possible vectors of symbols $\{R, U, D\}$ satisfying (A.1) we obtain an exact size of family $\mathtt{H}$. In case of $n_D = 0$, the vector has size $2(n-1)$ and, therefore, we get $\binom{2(n-1)}{(n-1)}$ ways of ordering symbols R and U. If $n_D = 1$ then the vector size is $2(n-1) - 1 = 2n - 3$ and $n_R = n - 2$. Hence, we have $\binom{2n-3}{n-2}$ choices of slots for symbols R and we have choose a place for D symbol among the remaining $2n - 3 - (n - 2) = n - 1$ slots. Thus, the total number of ordering for $n_D = 1$ is $\binom{2n-3}{n-2}(n-1)$. In the general case, we get $\binom{2n-2-n_D}{n-1-n_D}\binom{n-1}{n_D}$. Finally, the total number of vectors of $\mathtt{H}$ is given by the following formula:

$$
\sum_{n_D=0}^{n-1} \binom{2n-2-n_D}{n-1-n_D}\binom{n-1}{n_D}.
$$

REFERENCES

[1] A. Albano, A. Passamante, and M. E. Farrell, *Using higher-order correlations to define an embedding window*, Physica D: Nonlinear Phenomena, 54 (1991), pp. 85–97, https://doi.org/10.1016/0167-2789(91)90110-U.

[2] A. M. Albano, J. Muench, C. Schwartz, A. I. Mees, and P. E. Rapp, *Singular-value decomposition and the Grassberger-Procaccia algorithm*, Phys. Rev. A, 38 (1988), pp. 3017–3026, https://link.aps.org/doi/10.1103/PhysRevA.38.3017.

[3] K. Barański, Y. Gutman, and A. Śpiewak, *A probabilistic Takens theorem*, Nonlinearity, 33 (2020), p. 4940, https://dx.doi.org/10.1088/1361-6544/ab8fb8.

[4] K. Barański, Y. Gutman, and A. Śpiewak, *On the Shroer–Sauer–Ott–Yorke predictability conjecture for time-delay embeddings*, Commun. Math. Phys., 391 (2022), pp. 609–641, https://doi.org/10.1007/s00220-022-04323-y.

[5] E. M. Bollt and J. D. Skufca, *On comparing dynamical systems by defective conjugacy: A symbolic dynamics interpretation of commuter functions*, Physica D: Nonlinear Phenomena, 239 (2010), pp. 579–590, https://doi.org/https://doi.org/10.1016/j.physd.2009.12.007.

[6] J. J. Bramburger, S. L. Brunton, and J. N. Kutz, *Deep learning of conjugate mappings.*, Physica D: Nonlinear Phenomena, 427 (2021), p. 133008.

[7] H. Broer and F. Takens, *Reconstruction and time series analysis*, Springer New York, New York, NY, 2011, pp. 205–242, https://doi.org/10.1007/978-1-4419-6870-8_6.
[8] T. Buzug and G. Pfister, *Comparison of algorithms calculating optimal embedding parameters for delay time coordinates*, Physica D: Nonlinear Phenomena, 58 (1992), pp. 127–137, https://www.sciencedirect.com/science/article/pii/016727899290104U.
[9] K. Champion, B. Lusch, J. N. Kutz, and S. L. Brunton, *Data-driven discovery of coordinates and governing equations*, Proceedings of the National Academy of Sciences, 116 (2019), pp. 22445–22451.
[10] D. L. DeAngelis and S. Yurek, *Equation-free modeling unravels the behavior of complex ecological systems.*, Proc Natl Acad Sci U S A, 112 (2015), pp. 3856–3857.
[11] V. Deshmukh, E. Bradley, J. Garland, and J. D. Meiss, *Using curvature to select the time lag for delay reconstruction*, Chaos: An Interdisciplinary Journal of Nonlinear Science, 30 (2020), p. 063143, https://doi.org/10.1063/5.0005890.
[12] A. M. Fraser and H. L. Swinney, *Independent coordinates for strange attractors from mutual information*, Phys. Rev. A, 33 (1986), pp. 1134–1140, https://link.aps.org/doi/10.1103/PhysRevA.33.1134.
[13] R. Hegger and H. Kantz, *Improved false nearest neighbor method to detect determinism in time series data*, Phys. Rev. E, 60 (1999), pp. 4970–4973, https://link.aps.org/doi/10.1103/PhysRevE.60.4970.
[14] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, 2 ed., 2003, https://doi.org/10.1017/CBO9780511755798.
[15] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, *Determining embedding dimension for phase-space reconstruction using a geometrical construction*, Phys. Rev. A, 45 (1992), pp. 3403–3411, https://link.aps.org/doi/10.1103/PhysRevA.45.3403.
[16] A. Kianimajd, M. Ruano, P. Carvalho, J. Henriques, T. Rocha, S. Paredes, and A. Ruano, *Comparison of different methods of measuring similarity in physiologic time series*, IFAC-PapersOnLine, 50 (2017), pp. 11005–11010, https://www.sciencedirect.com/science/article/pii/S2405896317333967. 20th IFAC World Congress.
[17] H. Kim, R. Eykholt, and J. Salas, *Nonlinear dynamics, delay times, and embedding windows*, Physica D: Nonlinear Phenomena, 127 (1999), pp. 48–60, https://www.sciencedirect.com/science/article/pii/S0167278998002401.
[18] F. Lejarza and M. Baldea, *Data-driven discovery of the governing equations of dynamical systems via moving horizon optimization*, Scientific Reports, 12 (2022), p. 11836.
[19] M. Matilla-García, I. Morales, J. M. Rodríguez, and M. Ruiz Marín, *Selection of embedding dimension and delay time in phase space reconstruction via symbolic dynamics*, Entropy, 23 (2021), https://doi.org/10.3390/e23020221, https://www.mdpi.com/1099-4300/23/2/221.
[20] K. Mischaikow, M. Mrozek, J. Reiss, and A. Szymczak, *Construction of symbolic dynamics from experimental time series*, Phys. Rev. Lett., 82 (1999), pp. 1144–1147.
[21] L. Moniz, L. Pecora, J. Nichols, M. Todd, and J. R. Wait, *Dynamical assessment of structural damage using the continuity statistic*, Structural Health Monitoring, 3 (2004), pp. 199–212, https://doi.org/10.1177/1475921704042681.
[22] J. Nichols, L. Moniz, J. Nichols, L. Pecora, and E. Cooch, *Assessing spatial coupling in complex population dynamics using mutual prediction and continuity statistics*, Theoretical Population Biology, 67 (2005), pp. 9–21, https://doi.org/https://doi.org/10.1016/j.tpb.2004.08.004.
[23] L. M. Pecora, T. L. Carroll, and J. F. Heagy, *Statistics for mathematical properties of maps between time series embeddings*, Phys. Rev. E, 52 (1995), pp. 3420–3439, https://doi.org/10.1103/PhysRevE.52.3420.
[24] L. M. Pecora, L. Moniz, J. Nichols, and T. L. Carroll, *A unified approach to attractor reconstruction*, Chaos: An Interdisciplinary Journal of Nonlinear Science, 17 (2007), p. 013110, https://doi.org/10.1063/1.2430294, https://arxiv.org/abs/https://doi.org/10.1063/1.2430294.
[25] T. Sauer, J. A. Yorke, and M. Casdagli, *Embedology*, Journal of Statistical Physics, 65 (1991), pp. 579–616.
[26] J. D. Skufca and E. M. Bollt, *Relaxing conjugacy to fit modeling in dynamical systems*, Phys. Rev. E, 76 (2007), p. 026220, https://doi.org/10.1103/PhysRevE.76.026220.
[27] J. D. Skufca and E. M. Bollt, *A concept of homeomorphic defect for defining mostly conjugate dynamical systems*, Chaos: An Interdisciplinary Journal of Nonlinear Science, 18 (2008), p. 013118, https://doi.org/10.1063/1.2837397.

[28] W. Szlenk, *An Introduction to the Theory of Smooth Dynamical Systems*, Państwowe Wydawnictwo Naukowe (PWN), 1984.

[29] F. Takens, *Detecting strange attractors in turbulence*, in Dynamical Systems and Turbulence, Warwick 1980, D. Rand and L.-S. Young, eds., Berlin, Heidelberg, 1981, Springer Berlin Heidelberg, pp. 366–381.

[30] D. Viswanath, *The fractal property of the Lorenz attractor*, Physica D: Nonlinear Phenomena, 190 (2004), pp. 115–128, https://www.sciencedirect.com/science/article/pii/S0167278903004093.

[31] Y. Yuan, X. Tang, W. Zhou, W. Pan, X. Li, H.-T. Zhang, H. Ding, and J. Goncalves, *Data driven discovery of cyber physical systems*, Nature Communications, 10 (2019), p. 4894.

[32] J. Zheng, E. Bollt, and J. Skufca, *Regularity of commuter functions for homeomorphic defect measure in dynamical systems model comparison*, Dynamics of Continuous, Discrete & Impulsive Systems. Series A: Mathematical Analysis, 18 (2011).

[33] J. Zheng, J. D. Skufca, and E. M. Bollt, *Comparing dynamical systems by a graph matching method*, Physica D: Nonlinear Phenomena, 255 (2013), pp. 12–21, https://doi.org/https://doi.org/10.1016/j.physd.2013.03.012.

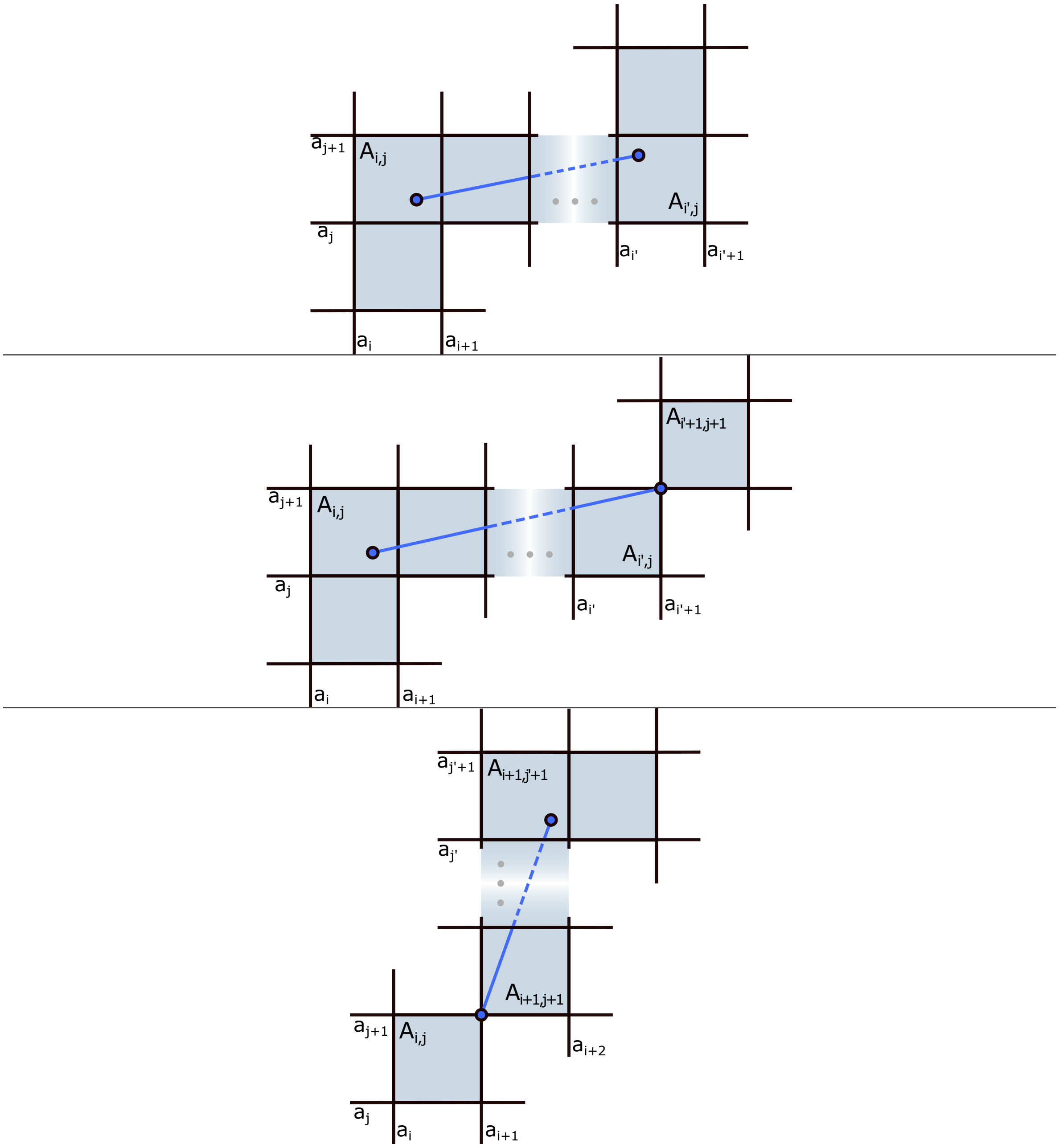

Figure 15. *Segments of the piecewise linear homeomorphism being a selector constructed by Algorithm A.1 for a certain cubical homeomorphism* h. *Three panels corresponds to cases when: point of type UR is followed by point of type UR (top), point of type UR is followed by point of type D (middle), point of type D is followed by point of type UR (bottom).*