

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 $x_1, \dots, x_n \subset X$ and $y_1, \dots, y_n \subset Y$ such that $x_{i+1} = f(x_i)$ and $y_{i+1} = g(y_i)$ as well as $h : X \rightarrow Y$, 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 $h : X \rightarrow Y$. Given two sampled trajectories, referred to in this paper by *time series*, $x_1, \dots, x_n \subset X$ and $y_1, \dots, y_n \subset Y$ we assume that they are both generated by a continuous maps $f : X \rightarrow X$ and $g : Y \rightarrow Y^1$ 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

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

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

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

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

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

59 The main contributions of this paper include:

- 60 • We propose a generalization of the FNN (*False Nearest Neighbor*) method [13] so that
 61 it can be applied to test for topological conjugacy of time series³. Moreover, we present
 62 its further modification called KNN method.
- 63 • We propose two entirely new methods: ConjTest and ConjTest⁺. Instead of providing
 64 an almost binary answer to a question if two sampled dynamical systems are conjugate
 65 (which happens for the generalized FNN and the KNN method), their result is a
 66 continuous variable that can serve as a scale of similarity of two dynamics. This

²One 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.

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

67 property makes the two new methods appropriate for noisy data.

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

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

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

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

108 Finally, it should be noted that in the continuous setting, topological conjugacy is very
109 fragile; it may be destroyed by an infinitesimal change of parameters of the system once that
110 causes bifurcation. However, two finite sample of the trajectories obtained from the system

111 before and after bifurcation are very close and it would require much large change of param-
 112 eters to detect problems with conjugacy. It is a consequence of the fact that the techniques
 113 proposed in this paper operates on finite data. Therefore, they can provide evidences that
 114 the proposed connecting homeomorphism is not a topological conjugacy of the two considered
 115 systems, but they will not allow to prove, in any rigorous sense, the conjugacy between them.

116 2. Preliminaries.

117 **2.1. Topological conjugacy.** We start with a pair of metric spaces X and Y and a pair of
 118 dynamical systems: $\varphi : X \times \mathbb{T} \rightarrow X$ and $\psi : Y \times \mathbb{T} \rightarrow Y$, where $\mathbb{T} \in \{\mathbb{Z}, \mathbb{R}\}$. Fixing $t_X, t_Y \in \mathbb{T}$
 119 define $f : X \ni x \rightarrow \varphi(x, t_X)$ and $g : Y \ni y \rightarrow \psi(y, t_Y)$. We say that f and g are *topologically*
 120 *conjugate* if there exists a homeomorphism $h : X \rightarrow Y$ such that the diagram

$$121 \quad (2.1) \quad \begin{array}{ccc} X & \xrightarrow{f} & X \\ h \downarrow & & \downarrow h \\ Y & \xrightarrow{g} & Y \end{array}$$

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

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

127 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
 128 $x_{i+1} = f^i(x_1)$ and $y_{i+1} = g^i(y_1)$ for $i \in \{1, 2, \dots, n-1\}$, $x_1 \in X$ and $y_1 \in Y$ and derive
 129 criteria to test (semi)topological conjugacy of f and g via h based on those samples and the
 130 given possible (semi)conjugacy h .

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

$$132 \quad d_H(A, B) = \max\{\sup_{a \in A} d(a, B), \sup_{b \in B} d(b, A)\}$$

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

134 **2.2. Takens' Embedding Theorem.** Our work is related to the problem of reconstruction
 135 of dynamics from one dimensional time series. For a fixed map $f : X \rightarrow X$ and $x_1 \in X$ take a
 136 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)
 137 dimension m . Take $s : X \rightarrow \mathbb{R}$, a generic measurement function of observable states of the
 138 system, and one dimensional time series $\mathcal{S} = \{s(x_i)\}_{x_i \in \mathcal{A}}$, associated to \mathcal{A} . The celebrated
 139 Takens' Embedding Theorem [29] states that given \mathcal{S} it is possible to reconstruct the original
 140 system with delay vectors, for instance $(s(x_i), s(x_{i+1}), \dots, s(x_{i+d-1}))$, for sufficiently large
 141 *embedding dimension* $d \geq 2m + 1$ (the bound is often not optimal). The Takens' theorem
 142 implies that, under certain generic assumptions, an embedding of the attractor Ω into \mathbb{R}^d
 143 given by

$$144 \quad (2.2) \quad F_{s,f} : \Omega \ni x \mapsto (s(x), s(f(x)), \dots, s(f^{d-1}(x))) \in \mathbb{R}^d$$

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

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

$$150 \quad \Pi(\mathcal{A}, d, l) := \{(s(x_i), s(x_{i+l}), \dots, s(x_{i+(d-1)l})) \mid i \in \{1, 2, \dots, n - dl\}\}.$$

151 From the theoretical point of view, the Takens' theorem holds for an arbitrary lag. However
 152 in practice a proper choice of l may strongly affect numerical reconstructions (see [14, Chapter
 153 3]).

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

156 **2.3. Search for an optimal dimension for reconstruction.** In practice, the bound in
 157 Takens' theorem is often not sharp and an embedding dimension less than $2m + 1$ is already
 158 sufficient to reconstruct the original dynamics (see [3, 4]). Moreover, for time series encoun-
 159 tered in practice, the attractor's dimension m is almost always unknown. To discover the
 160 sufficient dimension of reconstruction, the False Nearest Neighbor (FNN) method [13, 15], a
 161 heuristic technique for estimating the optimal dimension using a finite time series, is typi-
 162 cally used. It is based on an idea to compare the embeddings of a time series into a couple
 163 of consecutive dimensions and to check if the introduction of an additional $d + 1$ dimension
 164 separates some points that were close in d -dimensional embedding. Hence, it tests whether d -
 165 dimensional neighbors are (false) neighbors just because of the tightness of the d -dimensional
 166 space. The dimension where the value of the test stabilizes and no more false neighbors can
 167 be detected is proclaimed to be the optimal embedding dimension.

168 **2.4. False Nearest Neighbor and beyond.** The False Nearest Neighbor method implic-
 169 itly tests semiconjugacy of d and $d + 1$ dimensional Takens' embedding by checking if the
 170 neighborhood of d -embedded points are preserved in $d + 1$ dimension. This technique was an
 171 inspiration for stating a more general question: given two time series, can we test if they were
 172 generated from conjugate dynamical systems? The positive answer could suggest that the
 173 two observed signals were actually generated by the same dynamics, but obtained by a differ-
 174 ent measurement function. In what follows, a number of tests inspired by these observations
 175 concerning False Nearest Neighbor method and Takens' Embedding Theorem, are presented.

176 **3. Conjugacy testing methods.** In this section we introduce a number of new methods
 177 for quantifying the dynamical similarity of two time series. Before digging into them let us
 178 introduce some basic pieces of notation used throughout the section. From now on we assume
 179 that X is a metric space. Let $\mathcal{A} = \{x_i\}_{i=1}^n$ be a finite time series in space X . For $k \in \mathbb{N}$,
 180 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} .
 181 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
 182 clearly $\kappa(x, \mathcal{A}) = \{x\}$. Hence, it is handfull to consider also $\bar{\kappa}(x, k, \mathcal{A}) := \kappa(x, k, \mathcal{A} \setminus \{x\})$ and
 183 $\bar{\kappa}(x, \mathcal{A}) := \kappa(x, 1, \mathcal{A} \setminus \{x\})$ ⁴.

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



184 **3.1. False Nearest Neighbor method.** The first proposed method is an extension of the
 185 already mentioned FNN technique for estimating the optimal embedding dimension of time
 186 series. The idea of the classical FNN method relies on counting the number of so-called false
 187 nearest neighbors depending on the threshold parameter r . This is based on the observation
 188 that if the two reconstructed points

$$189 \quad \mathbf{s}_d^1 := (s(x_{k_1}), s(x_{k_1+l}), \dots, s(x_{k_1+(d-1)l}))$$

190 and

$$191 \quad \mathbf{s}_d^2 := (s(x_{k_2}), s(x_{k_2+l}), \dots, s(x_{k_2+(d-1)l}))$$

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

$$194 \quad \mathbf{s}_{d+1}^1 := (s(x_{k_1}), \dots, s(x_{k_1+(d-1)l}), s(x_{k_1+dl}))$$

195 and

$$196 \quad \mathbf{s}_{d+1}^2 := (s(x_{k_2}), \dots, s(x_{k_2+(d-1)l}), s(x_{k_2+dl}))$$

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

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

$$219 \quad (3.1) \quad \text{FNN}(\mathcal{A}, \mathcal{B}; r) := \frac{\sum_{i=1}^n \Theta \left(\frac{\mathbf{d}_Y(b_i, \xi(\bar{\kappa}(a_i, \mathcal{A})))}{\mathbf{d}_X(a_i, \bar{\kappa}(a_i, \mathcal{A}))} - r \right) \Theta \left(\frac{\sigma}{r} - \mathbf{d}_X(a_i, \bar{\kappa}(a_i, \mathcal{A})) \right)}{\sum_{i=1}^n \Theta \left(\frac{\sigma}{r} - \mathbf{d}_X(a_i, \bar{\kappa}(a_i, \mathcal{A})) \right)}$$

220 where \mathbf{d}_X and \mathbf{d}_Y denote the distance function respectively in X and Y , σ is the standard
 221 deviation of the data (i.e. the standard deviation of the elements of the sequence \mathcal{A}), r is the

222 parameter of the method and Θ is the usual Heaviside step function, i.e. $\Theta(x) = 1$ if $x > 0$
 223 and 0 otherwise. Note that the distance \mathbf{d} (i.e. \mathbf{d}_X or \mathbf{d}_Y) might be defined in various ways
 224 however, as elements of time-series are usually elements of \mathbb{R}^k (for some k), then $\mathbf{d}(x, y)$ is
 225 often simply the Euclidean norm $|x - y|$.

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

229 (3.2)
$$\text{FNN}(\mathcal{A}; r, d) := \text{FNN}(\Pi_d(\mathcal{A}), \Pi_{d+1}(\mathcal{A}); r),$$

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

235 **3.2. K-Nearest Neighbors.** The key to the method presented in this section is an attempt
 236 to weaken and simplify the condition posed by FNN by considering a larger neighborhood of
 237 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
 238 same length. Let $\xi : \mathcal{A} \rightarrow \mathcal{B}$ be a bijection defined $\xi(a_i) := b_i$. The proposed statistics, taking
 239 into account k nearest neighbors of each point, is given by the following formula:

240 (3.3)
$$\text{KNN}(\mathcal{A}, \mathcal{B}; k) := \frac{\sum_{i=1}^n \min \{e \in \mathbb{N} \mid \xi(\bar{\kappa}(a_i, k, \mathcal{A})) \subseteq \bar{\kappa}(\xi(a_i), e + k, \mathcal{B})\}}{n^2},$$

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

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

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

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

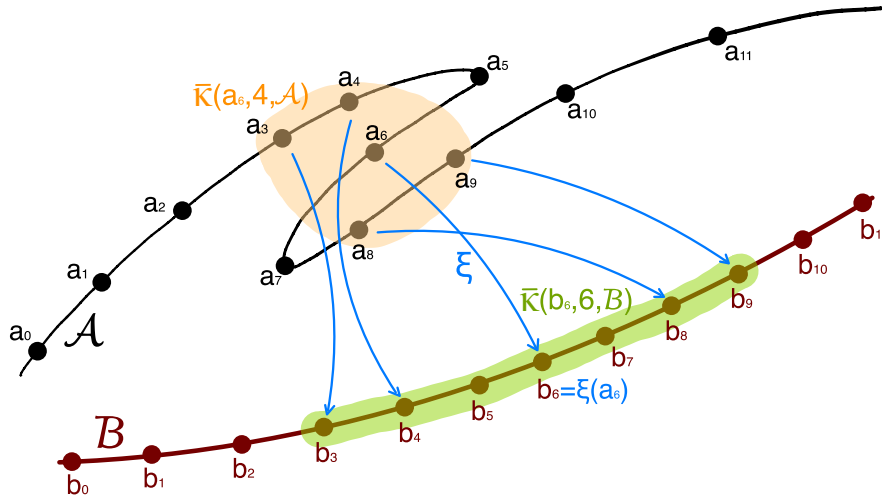


Figure 1. Top (continuous) black line represents trajectory from which \mathcal{A} is sampled (black dots). Bottom (continuous) trajectory is sampled to obtain \mathcal{B} (burgundy dots). Set $U := \bar{\kappa}(a_6, 4, \mathcal{A})$ highlighted with orange color, represents 4-nearest neighbors of $a_6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

260 **3.3. Conjugacy test.** The third method tests the conjugacy of two time series by directly
 261 checking the commutativity of the diagram (2.1) which is tested in a more direct way compared
 262 to the methods presented so far. We no longer assume that both time series are of the
 263 same size, however, the method requires a *connecting map* $h : X \rightarrow Y$, a candidate for a
 264 (semi)conjugating map. Unlike the map ξ in FNN and KNN method map h may transform
 265 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}
 266 are crucial because they carry the information about the dynamics $g : Y \rightarrow Y$. Thus, in order to
 267 follow trajectories of points in Y we introduce $\tilde{h} : \mathcal{A} \rightarrow \mathcal{B}$, a discrete approximation of h :

$$268 \quad (3.4) \quad \tilde{h}(a_i) := \kappa(h(a_i), \mathcal{B}).$$

269 The map \tilde{h} simply assigns to a_i the closest element(s) of $h(a_i)$ from the time series \mathcal{B} . For
 270 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
 271 that it may happen that $\tilde{h}(A)$ has less elements than A .

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

$$274 \quad (3.5) \quad \text{ConjTest}(\mathcal{A}, \mathcal{B}; k, t, h) := \frac{\sum_{i=1}^n d_H \left((h \circ f^t)(U_i^k), (g^t \circ \tilde{h})(U_i^k) \right)}{n \text{diam}(\mathcal{B})},$$

275 where d_H is the Hausdorff distance between two discrete sets and $\text{diam}(\mathcal{B})$ is the diameter
 276 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
 277 series together with map h are close to satisfy diagram (2.1) defining topological conjugacy.
 278 First, we approximate the neighborhood of $a_i \in \mathcal{A}$ with U_i^k and then we try to traverse the

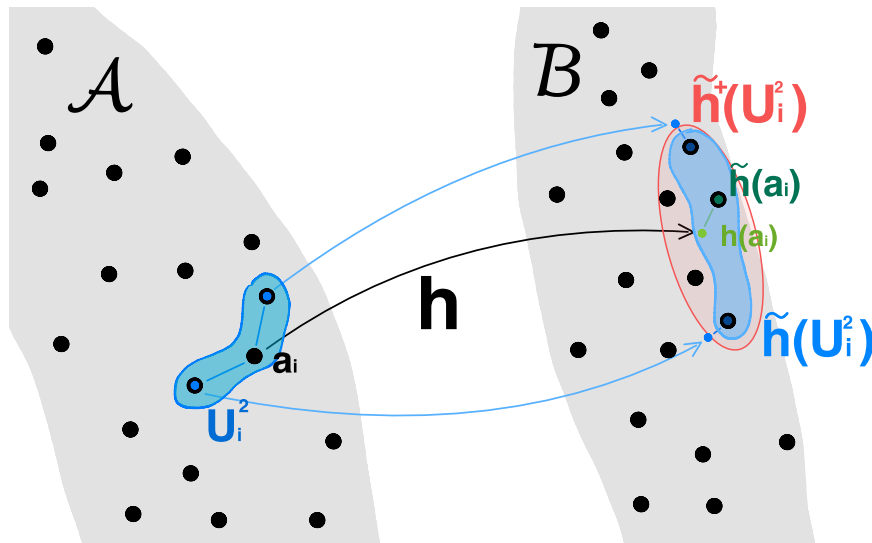


Figure 2. A pictorial visualization of a difference between h , \tilde{h} and \tilde{h}^+ . Map h transforms a point $a \in \mathcal{A}$ into a point $h(a) \in Y$. Map \tilde{h} approximates the value of the map h by finding the closest point in \mathcal{B} for $h(a)$. The discrete neighborhood $U_i^2 \subset \mathcal{A}$ 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 $\tilde{h}(a_i)$ than points in $\tilde{h}(U_i^2)$.

279 diagram in two possible ways. Thus, we end up with two sets in Y , that is $(h \circ f^t)(U_i^k)$ and
 280 $(g^t \circ \tilde{h})(U_i^k)$. We measure how those two sets diverge using the Hausdorff distance.

281 The extended version of the test presented above considers a larger approximation of
 282 $\tilde{h}(U_i^k)$. 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
 283 superset defines the enriched approximation (see Figure 2):

284 (3.6)
$$\tilde{h}^+(U_i^k) := \kappa(h(a_i), k_i, \mathcal{B}).$$

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

286 (3.7)
$$\text{ConjTest}^+(\mathcal{A}, \mathcal{B}; k, t, h) := \frac{\sum_{i=1}^n d_H \left((h \circ f^t)(U_i^k), g^t \left(\tilde{h}^+(U_i^k) \right) \right)}{n \text{diam}(\mathcal{B})}.$$

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

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

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

294 **Theorem 3.2.** Let $f : X \rightarrow X$ and $g : Y \rightarrow Y$, where $X \subset \mathbb{R}^{d_X}$ and $Y \subset \mathbb{R}^{d_Y}$, be
 295 continuous maps (d_X and d_Y denote dimensions of the spaces). For $y_1 \in Y$ define $\mathcal{B}_m :=$
 296 $\{b_i := g^{i-1}(y_1) \mid i \in \{1, \dots, m\}\}$.

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

$$300 \quad (3.8) \quad \lim_{m \rightarrow \infty} \text{ConjTest}(\mathcal{A}_n, \mathcal{B}_m; k, t, h) = 0,$$

301 where $\mathcal{A}_n := \{a_i := f^{i-1}(x_1) \mid i \in \{1, \dots, n\}\}$, $x_1 \in X$, is any time-series in X of a length n .
 302 Moreover, the convergence is uniform with respect to n and with respect to the choice of
 303 the starting point x_1 (i.e. the “rate” of convergence does not depend on the time-series \mathcal{A}_n).

304 *Proof.* Since g is semiconjugate to f via h , $h : X \rightarrow Y$ is a continuous surjection such
 305 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
 306 show that there exists M such that for all $m > M$, all $n \in \mathbb{N}$ and every finite time-series
 307 $\mathcal{A}_n := \{a_i := f^{i-1}(x_1) \mid i \in \{1, \dots, n\}\} \subset X$ of length n (where $x_1 \in X$ is some point in X)
 308 it holds that

$$309 \quad (3.9) \quad \text{ConjTest}(\mathcal{A}_n, \mathcal{B}_m; k, t, h) < \varepsilon.$$

310 Note that $|b_2 - b_1| \leq |\mathcal{B}_m|$ for any $m \geq 2$ (with $|\mathcal{B}_m|$ denoting cardinality of the set
 311 \mathcal{B}_m), which we will use at the end of the proof. As g is continuous and Y is compact,
 312 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 $|y - \tilde{y}| < \delta$.
 313 As $\mathcal{B} = \{y_1, g(y_1), \dots, g^m(y_1), \dots\} = \{b_1, \dots, b_m, \dots\}$ is dense in Y , there exists M such
 314 that if $m > M$ then for every $n \in \mathbb{N}$, every $x_1 \in X$ and every $i \in \{1, 2, \dots\}$ there exists
 315 $j_m(i) \in \{1, 2, \dots, m\}$ such that

$$316 \quad |b_{j_m(i)} - h(a_i)| < \delta,$$

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

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

$$319 \quad |h(f^t(a_i)) - g^t(\tilde{h}(a_i))| = |g^t(h(a_i)) - g^t(\tilde{h}(a_i))| < \varepsilon |b_2 - b_1|$$

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

$$321 \quad d_H\left((h \circ f^t)(U_i^k), (g^t \circ \tilde{h})(U_i^k)\right) < \varepsilon |b_2 - b_1|,$$

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

$$323 \quad \frac{\sum_{i=1}^n d_H\left((h \circ f^t)(U_i^k), (g^t \circ \tilde{h})(U_i^k)\right)}{n \text{diam}(\mathcal{B}_m)} < \frac{n \varepsilon |b_2 - b_1|}{n \text{diam}(\mathcal{B}_m)} \leq \varepsilon$$

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

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

328 **Corollary 3.3.** Let $f : X \rightarrow X$ and $g : Y \rightarrow Y$, where $X \subset \mathbb{R}^{d_x}$ and $Y \subset \mathbb{R}^{d_y}$, be
 329 continuous maps. Let $x_1 \in X$ and $y_1 \in Y$. Define $\mathcal{A}_n := \{a_i := f^{i-1}(x_1) \mid i \in \{1, \dots, n\}\}$
 330 and $\mathcal{B}_m := \{b_i := g^{i-1}(y_1) \mid i \in \{1, \dots, m\}\}$. Suppose that $\{h(a_1), \dots, h(a_n)\} \subset \hat{Y}$ for some
 331 compact set $\hat{Y} \subset Y$ such that the set $\hat{Y} \cap \mathcal{B}$ is dense in \hat{Y} , where $\mathcal{B} = \{b_1, \dots, b_m, \dots\}$.
 332 If g is **semiconjugate** to f with h as a semiconjugacy, then for every t and k

$$333 \quad \lim_{m \rightarrow \infty} \text{ConjTest}(\mathcal{A}_n, \mathcal{B}_m; k, t, h) = 0.$$

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

337 Note also that we do not need the compactness of the space X nor the density of $\mathcal{A} =$
 338 $\{a_1, a_2, \dots, a_n, \dots\}$ in X .

339 The following statement is an easy consequence of the statements above

340 **Theorem 3.5.** Let $X \subset \mathbb{R}^{d_x}$ and $Y \subset \mathbb{R}^{d_y}$ be compact sets and $f : X \rightarrow X$ and $g : Y \rightarrow Y$
 341 be continuous maps which are **conjugate** by a homeomorphism $h : X \rightarrow Y$. Let $x_1 \in X$,
 342 $y_1 \in Y$ and \mathcal{A}_n and \mathcal{B}_m be defined as before. Suppose that \mathcal{A}_n and \mathcal{B}_m are dense, respectively,
 343 in X and Y as $n \rightarrow \infty$ and $m \rightarrow \infty$. Then for every t and k

$$344 \quad \lim_{m \rightarrow \infty} \text{ConjTest}(\mathcal{A}_n, \mathcal{B}_m; k, t, h) = \lim_{n \rightarrow \infty} \text{ConjTest}(\mathcal{B}_m, \mathcal{A}_m; k, t, h) = 0.$$

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

347 The above results concern ConjTest . Note that in case of ConjTest^+ the neighborhoods
 348 $\tilde{h}^+(U_i^k)$, thus also $(g^t \circ \tilde{h}^+)(U_i^k)$, can be significantly enlarged by adding additional points
 349 to $\tilde{h}(U_i^k)$ and thus increasing the Hausdorff distance between corresponding sets. In order
 350 to still control this distance and formally prove desired convergence additional assumptions
 351 concerning space X and the sequence \mathcal{A} are needed:

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

$$358 \quad (3.10) \quad \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \text{ConjTest}^+(\mathcal{A}_n, \mathcal{B}_m; k, t, h) = 0$$

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

360 **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

361 $h : X \rightarrow Y$ is a continuous surjection. Expanding (3.10) yields

$$\begin{aligned}
 & \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \frac{\sum_{i=1}^n d_H \left((h \circ f^t)(U_i^k), (g^t \circ \tilde{h}^+)(U_i^k) \right)}{n \operatorname{diam}(\mathcal{B}_m)} \leq \\
 362 \quad (3.11) \quad & \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \frac{\sum_{i=1}^n d_H \left((h \circ f^t)(U_i^k), (g^t \circ \tilde{h})(U_i^k) \right)}{n \operatorname{diam}(\mathcal{B}_m)} + \\
 & \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \frac{\sum_{i=1}^n d_H \left((h \circ f^t)(U_i^k), (g^t(\tilde{h}^+(U_i^k) \setminus \tilde{h}(U_i^k))) \right)}{n \operatorname{diam}(\mathcal{B}_m)}.
 \end{aligned}$$

363 Recall that $U_i^k := \kappa(a_i, k, \mathcal{A}_n)$, $\tilde{h}(a_i) := \kappa(h(a_i), \mathcal{B}_m)$, $\tilde{h}(U_i^k) := \{\tilde{h}(a_j) : a_j \in U_i^k\}$
 364 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 $\tilde{h}(U_i^k) \subset$
 365 $\kappa(h(a_i), k_i, \mathcal{B}_m)$. Thus in particular, $\tilde{h}(U_i^k) \subset \tilde{h}^+(U_i^k)$. Obviously all these neighborhoods
 366 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).

367 Note that from Theorem 3.2 already follows that the first of the two terms in the sum in
 368 (3.11) vanishes. Thus we will only show that the second double limit vanishes as well.

369 Let $\varepsilon > 0$, $k \in \mathbb{N}$ and $t \in \mathbb{N}$. Since $g^t : Y \rightarrow Y$ is a continuous function on a compact
 370 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
 371 $|x - y| < \delta$. Similarly, since X is compact and $h : X \rightarrow Y$ is continuous, there exists δ_1 such
 372 that $|h(x) - h(y)| < \frac{\delta}{2}$ whenever $x, y \in X$ such that $|x - y| < \delta_1$.

373 Since \mathcal{B} is dense in Y , there exists $M \in \mathbb{N}$ such that for $m > M$ and every $y \in Y$,
 374 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
 375 $N \in \mathbb{N}$ such that for every $n > N$ and every $i \in \{1, 2, \dots, n\}$ we have $\operatorname{diam}(U_i^k) < \delta_1$, i.e. if
 376 $a_j \in U_i^k = \kappa(a_i, k, \mathcal{A}_n)$ then $|a_j - a_i| < \delta_1$ and consequently

$$377 \quad (3.12) \quad |g^t(h(a_j)) - g^t(h(a_i))| < \frac{\varepsilon}{2}.$$

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

382 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}^+ ,

$$383 \quad (3.13) \quad |\tilde{b} - h(a_i)| \leq \operatorname{diam}(\tilde{h}(U_i^k)) < \delta.$$

384 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

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

386 where

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

390 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
 391 $|h(f^t(a_j)) - g^t(\tilde{b})| < \varepsilon$ meaning that

$$392 \frac{\sum_{i=1}^n d_H \left((h \circ f^t)(U_i^k), g^t(\tilde{h}^+(U_i^k) \setminus \tilde{h}(U_i^k)) \right)}{n \operatorname{diam}(\mathcal{B}_m)} < \frac{\varepsilon}{\operatorname{diam}(\mathcal{B}_m)}$$

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

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

396 Finally, let us mention that conjugacy tests described in this Section are not tests in
 397 the statistical sense. They should be rather considered as methods of assessing dynamical
 398 similarity of the two unknown systems when only small finite samples of their trajectories are
 399 available. Trajectories related by topological conjugacy will give values of the tests close to
 400 0, and those coming from not conjugate systems are expected to result with higher values of
 401 the tests.

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

412 **4. Conjugacy experiments.** In this section the behavior of the described methods is ex-
 413 perimentally studied. For that purpose a benchmark set of a number of time series originating
 414 from (non-)conjugate dynamical systems is generated. A time series of length N generated
 415 by a map $f : X \rightarrow X$ with a starting point $x_1 \in X$ is denoted by

$$416 \varrho(f, x_1, N) := \{f^{j-1}(x_1) \in X \mid j \in \{1, 2, \dots, N\}\}.$$

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

420 **4.1. Irrational rotation on a circle.** The first example involves a dynamics generated by
 421 rotation on a circle by an irrational angle. Let us define a 1-dimensional circle as a quotient
 422 space $\mathbb{S} := \mathbb{R}/\mathbb{Z}$. Denote the operation of taking a decimal part of a number (modulo 1)
 423 by $x_{\mathbb{1}} := x - \lfloor x \rfloor$. Then, for a parameter $\phi \in [0, 1)$ we define a rigid rotation on a circle,
 424 $f_{[\phi]} : \mathbb{S} \rightarrow \mathbb{S}$, as

$$425 f_{[\phi]}(x) := (x + \phi)_{\mathbb{1}}.$$

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

$$427 \quad (4.1) \quad \mathbf{d}_{\mathbb{S}} : \mathbb{S} \times \mathbb{S} \ni (x, y) \mapsto \min((x - y)_{\mathbb{1}}, (y - x)_{\mathbb{1}}) \in [0, 1).$$

428 In this case $\mathbf{d}_{\mathbb{S}}(x, y)$ can be interpreted as the length of the shorter arc joining points x
 429 and y on \mathbb{S} . It is known that two rigid rotations, $f_{[\phi]}$ and $f_{[\psi]}$, are topologically conjugate
 430 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
 431 the first case the conjugating circle homeomorphism h preserves the orientation i.e. the lift
 432 $H : \mathbb{R} \rightarrow \mathbb{R}$ of $h : \mathbb{S} \rightarrow \mathbb{S}$ satisfies $H(x + 1) = H(x) + 1$ for every $x \in \mathbb{R}$ and in the second case
 433 h reverses the orientation $H(x + 1) = H(x) - 1$ and the two rotations $f_{[\phi]}$ and $f_{[\psi]}$ are just
 434 mutually inverse.

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

$$436 \quad f_{[\phi],s}(x) := \left((x_{\mathbb{1}}^s + \phi)_{\mathbb{1}} \right)^{1/s}, \quad x \in \mathbb{R}$$

437 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
 438 $f_{[\phi],s} = h_s^{-1} \circ f_{[\phi]} \circ h_s$.

439 4.1.1. Experiment 1A.

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

$$\begin{aligned} 441 \quad \mathcal{R}_1 &= \varrho(f_{[\alpha]}, 0.0, 2000), & \mathcal{R}_2 &= \varrho(f_{[\alpha]}, 0.25, 2000), \\ 442 \quad \mathcal{R}_3 &= \varrho(f_{[\alpha+0.02]}, 0.0, 2000), & \mathcal{R}_4 &= \varrho(f_{[2\alpha]}, 0.0, 2000), \\ 443 \quad \mathcal{R}_5 &= \varrho(f_{[\alpha],2}, 0.0, 2000), & \mathcal{R}_6 &= \mathcal{R}_5 + \text{err}(0.05), \end{aligned}$$

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

445 As follows from Poincaré Classification Theorem, $f_{[\alpha]}$ and $f_{[2\alpha]}$ are not conjugate nor
 446 semiconjugate whereas $f_{[\alpha]}$ and $f_{[\alpha],2}$ are conjugate via h_2 . Thus the expectation is to confirm
 447 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
 448 remaining cases.

449 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
 450 done with $h \equiv \text{id}_{\mathbb{S}}$. As we already mentioned, there is no conjugacy between \mathcal{R}_1 and \mathcal{R}_3 ,
 451 nor between \mathcal{R}_1 and \mathcal{R}_4 , as the angles of these rotations are different. Thus there is no
 452 true connecting homeomorphism between \mathcal{R}_1 and \mathcal{R}_3 and between \mathcal{R}_1 and \mathcal{R}_4 . However, in
 453 order to apply ConjTests we need to pick some candidate for a matching map between two
 454 point clouds and as the first choice one can always start with the identity map, especially for
 455 comparing point clouds generated by trajectories starting at the same or close initial points.
 456 Therefore in this experiment we use $h \equiv \text{id}_{\mathbb{S}}$ for comparing \mathcal{R}_1 both with \mathcal{R}_3 and \mathcal{R}_4 . When
 457 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
 458 that for FNN and KNN methods we always use $h(x_i) = y_i$, a connecting homeomorphism
 459 based on the indices correspondence.

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

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.

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

467 As we can see from Table 1 the starting point does not affect results of methods (\mathcal{R}_1 vs.
468 \mathcal{R}_2) since all the values in the first column are close to 0. It is expected due to the symmetry
469 of the considered system. A nonlinearity introduced in time series \mathcal{R}_5 also does not affect the
470 results. Despite the fact that $f_{[\alpha],2}$ is nonlinear, it is conjugate to the rotation $f_{[\alpha]}$ which is
471 reflected by tests' values. However, when we change the rotation parameter we can see an
472 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
473 of FNN and KNN. Interestingly, a small perturbation of the angle (\mathcal{R}_3) can cause a bigger
474 change in a value than a large one (\mathcal{R}_4). We investigate how the perturbation of the rotation
475 parameter affects values of examined methods in Experiment 1B. Moreover, the last column
476 (\mathcal{R}_1 vs. \mathcal{R}_6) shows that FNN is very sensitive to noise, while KNN and ConjTest methods
477 present some robustness. The influence of noise on the value of the test statistics is further
478 studied in Experiment 1C.

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

481 **4.1.2. Experiment 1B.** In this experiment we test how the difference of the system pa-
482 rameter affects tested methods.

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

$$484 \quad (4.2) \quad \left\{ \mathcal{R}_\beta := \varrho(f_{[\beta]}, 0.0, 2000) \mid \beta = \alpha + \frac{i\alpha}{100}, i \in [-50, -49, \dots, 125] \right\}.$$

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

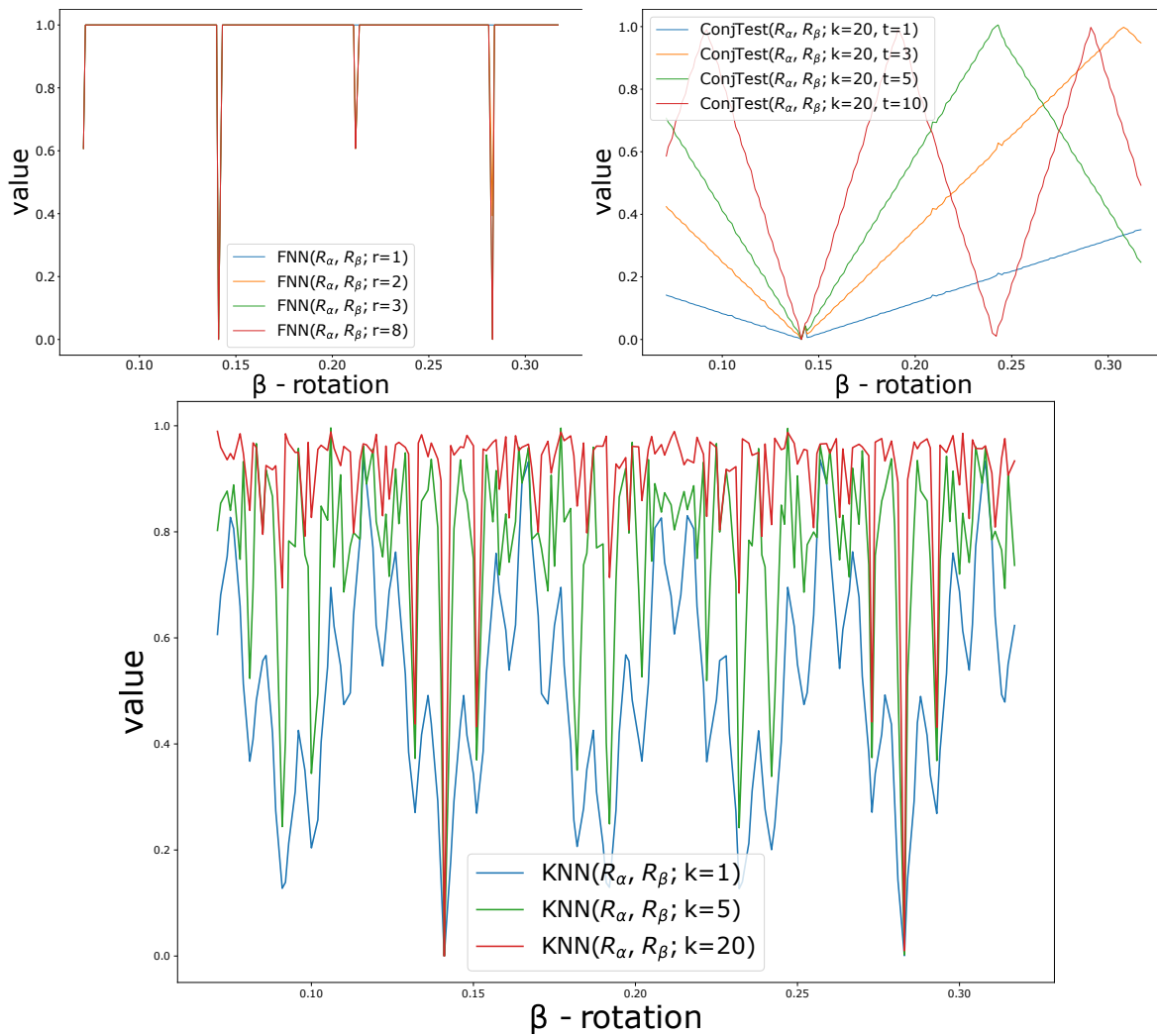


Figure 3. Dependence of the conjugacy measures on the perturbation of rotation angle. Top left: FNN method. Top right: ConjTest method. Bottom: KNN method.

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

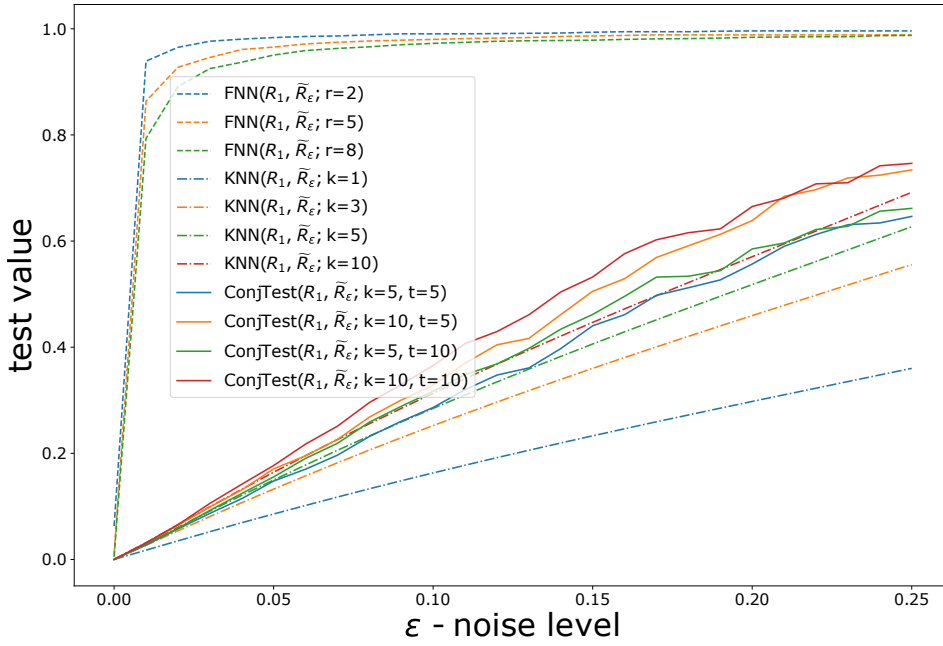


Figure 4. Dependence of conjugacy measures on the perturbation of time series.

499 Both, FNN and KNN exhibit an interesting drop of the value when $\beta \approx 0.283$ that is
 500 $\beta = 2\alpha$. Formally, we know that $f_{[\alpha]}$ and $f_{[2\alpha]}$ are not conjugate systems. However, we
 501 can explain this outcome by analyzing the methods. Let $a_i \in \mathcal{R}_\alpha$ and let $\tau \in \mathbb{Z}$ such that
 502 $a_j := a_{i+\tau} \in \mathcal{R}_\alpha$ be the nearest neighbor of a_i . In particular, $\tau \in \mathbb{Z}$. By (4.1) we get
 503 $\mathbf{d}_S(a_i, a_j) = (\alpha\tau)_\perp$ or $(-\alpha\tau)_\perp$. There is an $N \in \mathbb{Z}$ and a $\delta \in [0, 1)$ such that $\alpha\tau = N + \delta$.
 504 Since $\mathbf{d}_S(a_i, a_j) \approx 0$, it follows that $\delta_\perp \approx 0$. To get FNN we also need to know $\mathbf{d}_S(b_i, b_j)$. Let
 505 $\beta = z\alpha$. Then, $b_i = (z\alpha i)_\perp$, $b_j = (z\alpha i + z\alpha\tau)_\perp$ and $\mathbf{d}_S(b_i, b_j) = (z\alpha\tau)_\perp$ or $(-z\alpha\tau)_\perp$. Thus,
 506 $z\alpha\tau = zN + z\delta$. We assume that $z\delta \in [0, 1)$, because $\delta_\perp \approx 0$ and z is not very large. Again,
 507 there exists an $M \in \mathbb{Z}$ and $\epsilon \in [0, 1)$ such that $zN = M + \epsilon$. Now, if $zN \in \mathbb{Z}$, then $\epsilon = 0$,
 508 $\mathbf{d}_S(b_i, b_j) = (z\delta)_\perp = z\mathbf{d}_S(a_i, a_j)$ (last equality given by $\delta_\perp \approx 0$) and $\frac{\mathbf{d}_S(b_i, b_j)}{\mathbf{d}_S(a_i, a_j)} = z$. If $zN \notin \mathbb{Z}$
 509 then $\epsilon \neq 0$ and $\frac{\mathbf{d}_S(b_i, b_j)}{\mathbf{d}_S(a_i, a_j)} = \frac{\geq 0}{\approx 0}$. Hence, the fraction gives a large number and the numerator
 510 of FNN will count most of the points, unless $zN \in \mathbb{Z}$ which is always satisfied when $z \in \mathbb{Z}$.
 511 Moreover, for the irrational rotation τ might be large. In our experiments we usually get
 512 $|\tau| > 1000$. Thus, N is large and ϵ is basically a random number. In the case of KNN there is
 513 a chance that at least for some of the k -nearest neighbors $zN \in \mathbb{Z}$. Hence, the more rugged
 514 shape of the curve.

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

521 **4.1.3. Experiment 1C.** In this experiment, instead of perturbing the parameter of the
 522 system we perturb the time series itself by applying a noise to every point of the series.

523 *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
 524 time series:

$$525 \quad (4.3) \quad \left\{ \tilde{\mathcal{R}}_\epsilon := \varrho(f_{[\alpha],2}, 0.0, 2000) + \text{err}(\epsilon) \mid \epsilon \in [0.00, 0.25] \right\},$$

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

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

533 **4.2. Example: irrational rotation on a torus.** Let us consider a simple extension of the
 534 previous rotation example to a rotation on a torus. With a torus defined as $\mathbb{T} := \mathbb{S} \times \mathbb{S}$, where
 535 $\mathbb{S} = \mathbb{R}/\mathbb{Z}$, we can introduce map $f_{[\phi_1, \phi_2]} : \mathbb{T} \rightarrow \mathbb{T}$ defined as

$$536 \quad f_{[\phi_1, \phi_2]}(x^{(1)}, x^{(2)}) = ((x^{(1)} + \phi_1)_{\mathbb{1}}, (x^{(2)} + \phi_2)_{\mathbb{1}}),$$

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

$$538 \quad \mathbf{d}_{\mathbb{T}} : \mathbb{T} \times \mathbb{T} \ni ((x_1, y_1), (x_2, y_2)) \mapsto \max(\mathbf{d}_{\mathbb{S}}(x_1, x_2), \mathbf{d}_{\mathbb{S}}(y_1, y_2)) \in [0, 1),$$

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

540 Note that rotation on a torus described above and rotation on a circle $f_{[\phi_i]} : \mathbb{S} \rightarrow \mathbb{S}$ studied
 541 in Section 4.1 give a simple example of semiconjugate systems. Namely, let $h : \mathbb{T} \rightarrow \mathbb{S}$ be a
 542 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
 543 $i \in \{1, 2\}$.

544 **4.2.1. Experiment 2A.**

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

$$\begin{aligned} 546 \quad \mathcal{T}_1 &= \varrho(f_{[\alpha, \beta]}, (0.0, 0.0), 2000), & \mathcal{S}_1 &= \mathcal{T}_1^{(1)}, \\ 547 \quad \mathcal{T}_2 &= \varrho(f_{[1.1\alpha, \beta]}, (0.1, 0.0), 2000), & \mathcal{S}_2 &= \mathcal{T}_2^{(1)}, \\ 548 \quad \mathcal{T}_3 &= \varrho(f_{[\beta, \beta]}, (0.1, 0.0), 2000), & \mathcal{S}_3 &= \mathcal{T}_3^{(1)}, \end{aligned}$$

549 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
 550 projection of the elements of \mathcal{T}_i onto the first coordinate. When comparing \mathcal{T}_i with \mathcal{T}_j for
 551 $i, j \in \{1, 2, 3\}$ we use $h \equiv \text{id}$. When we compare \mathcal{T}_i versus \mathcal{S}_j we use $h(x, y) = x$, and for \mathcal{S}_i
 552 versus \mathcal{T}_j we get $h(x) = (x, 0)$.

553 *Results.* The asymmetry of results in the first column (\mathcal{T}_1 vs. \mathcal{S}_1) in Table 2 shows that
 554 all methods detect a semiconjugacy between \mathcal{T}_1 and \mathcal{S}_1 , i.e. that $f_{[\alpha]}$ is semiconjugate to $f_{[\alpha, \beta]}$
 555 via h_1 . An embedding of a torus into a 1-sphere preserves a neighborhood of a point. The
 556 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.

557 The rest of the results confirm conclusions from the previous experiment. The second
 558 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
 559 perturbation of the system parameters. The fourth and the fifth column (\mathcal{T}_1 vs. \mathcal{T}_3 and \mathcal{T}_1
 560 vs. \mathcal{S}_3) present another example where those two methods produce a false positive answer
 561 suggesting a semiconjugacy. This time the problematic case is not due to a doubling of the
 562 rotation parameter, but because of coinciding rotation angles. Again, the behavior of the
 563 ConjTest method exhibits a response that is relative to the level of perturbation.

564 **4.3. Example: the logistic map and the tent map.** Our next experiment examines two
 565 broadly studied chaotic maps defined on a real line. The logistic map and the tent map,
 566 $f_l, g_\mu : [0, 1] \rightarrow [0, 1]$, respectively defined as:

$$567 \quad (4.4) \quad f_l(x) := lx(1-x) \quad \text{and} \quad g_\mu(x) := \mu \min\{x, 1-x\},$$

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

$$570 \quad (4.5) \quad h(x) := \frac{2 \arcsin(\sqrt{x})}{\pi},$$

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

572 4.3.1. Experiment 3A.

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

$$574 \quad \mathcal{A} = \varrho(f_4, 0.2, 2000), \quad \mathcal{B}_2 = \varrho(f_4, 0.21, 2000),$$

$$575 \quad \mathcal{B}_1 = \varrho(g_2, h(0.2), 2000), \quad \mathcal{B}_3 = \varrho(f_{3.99}, 0.2, 2000),$$

$$576 \quad \mathcal{B}_4 = \varrho(f_{3.99}, 0.21, 2000).$$

577 Time series \mathcal{A} is conjugate to \mathcal{B}_1 through the homeomorphism h . Time series \mathcal{A} and \mathcal{B}_2
 578 come from the same system – f_4 , but are generated using different starting points. Sequences
 579 \mathcal{B}_3 and \mathcal{B}_4 are both generated by the logistic map but with different parameter value ($l = 3.99$)
 580 than \mathcal{A} ; thus, they are not conjugate with \mathcal{A} . For ConjTest methods we use (4.5) to compare
 581 \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.

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

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

600 **4.3.2. Experiment 3B.** The logistic map is one of the standard examples of chaotic maps.
601 Thus, we expect that the behavior of the system will change significantly if we modify the
602 parameter l . Here, we examine how the perturbation of l affects the outcome of tested
603 methods.

604 **Setup.** We generated a collection of time series:

$$605 \quad \{\mathcal{B}(l) := \varrho(f_l, 0.2, 2000) \mid l \in \{3.8, 3.805, 3.81, \dots, 4.0\}\}.$$

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

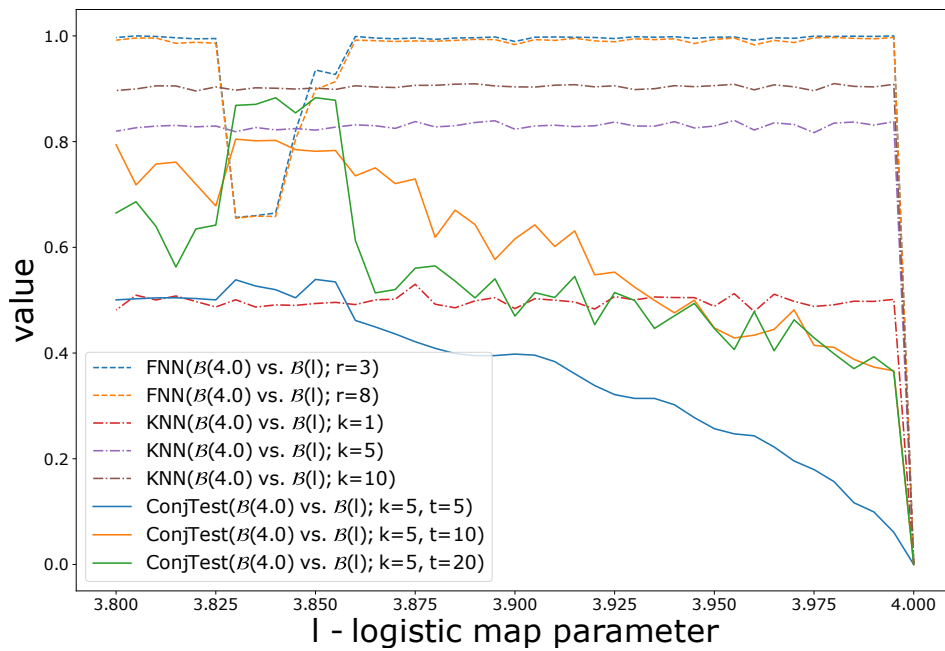


Figure 5. Dependence of the conjugacy measures on a change of the parameter of the logistic map.

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

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

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

624 **4.3.3. Experiment 3C.** As observed in the previous experiment, a change of the parameter
 625 l in the logistic equation may significantly change the dynamical nature of the system. In this
 626 experiment we use the ConjTest to grasp the types of dynamics as a function of l parameter.

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

$$628 \quad (4.6) \quad \{\mathcal{B}(l, p) := \varrho(f_l, f_l^{500}(p), 2000) \mid l \in \{3.4, 3.405, 3.41, \dots, 4.0\}, p \in \{0.11, 0.31\}\}.$$

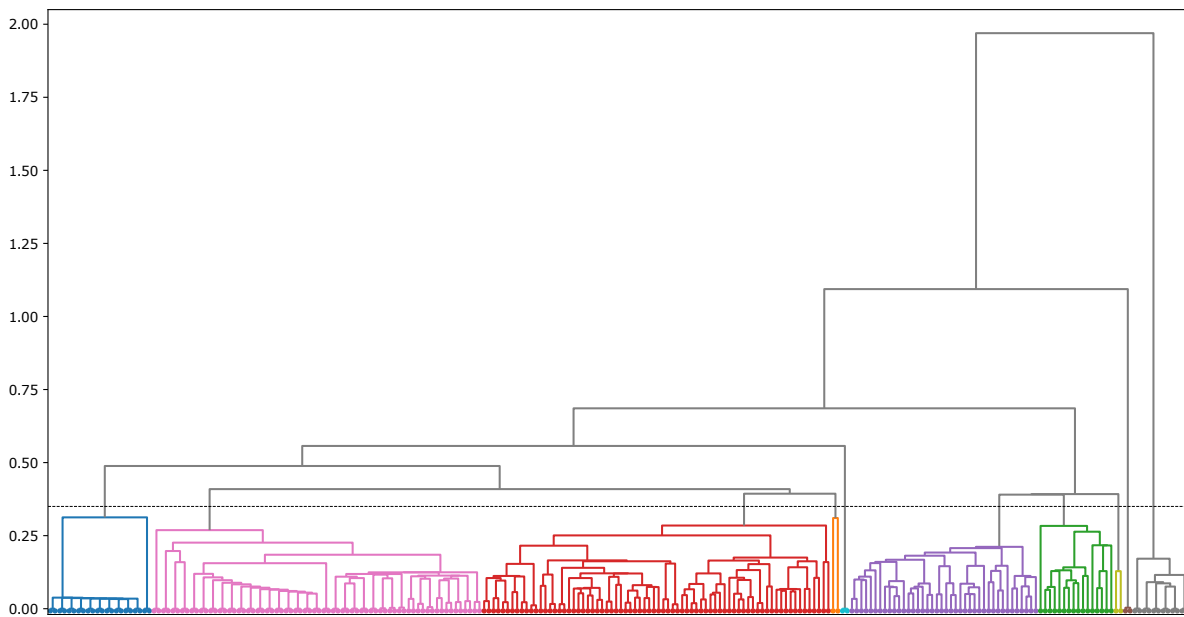


Figure 6. Dendrogram obtained from the single-linkage agglomerative hierarchical clustering of the collection of time series (4.6) generated from the logistic map using similarity score defined (4.7). The horizontal dashed line represents the threshold chosen for the clustering. The distribution of the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

629 Note that each time series starts at 500-th iterate of point p . It is a standard procedure
 630 allowing the trajectory to settle down on an attractor. We compared every two time series
 631 $\mathcal{B}(l, p)$ and $\mathcal{B}(l', p')$ with the formula We assign a similarity for each pair of time series $\mathcal{B}(l, p)$
 632 and $\mathcal{B}(l', p')$ via

$$633 \quad (4.7) \quad \max \{ \text{ConjTest}(\mathcal{B}(l, p), \mathcal{B}(l', p'); k, t, \text{id}), \text{ConjTest}(\mathcal{B}(l', p'), \mathcal{B}(l, p); k, t, \text{id}) \}$$

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

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

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

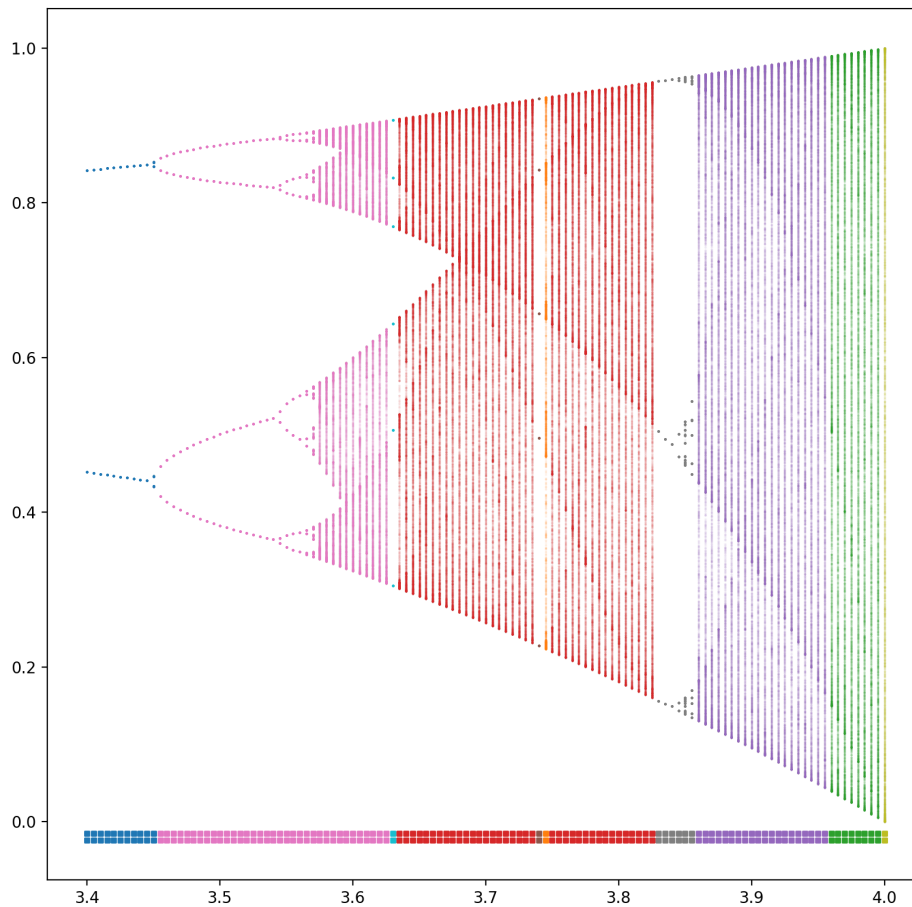


Figure 7. Partition of the bifurcation diagram based on the clustering of time series generated from the logistic map presented in Figure 6. Every square at the bottom of the image represents a time series corresponding to a value of parameter l given by the horizontal axis. Top row corresponds to the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

646 **4.4. Example: Lorenz attractor and its embeddings.** The fourth example is based on
 647 the Lorenz system defined by equations:

$$648 \quad (4.8) \quad \begin{cases} \dot{x} &= \sigma(y - x), \\ \dot{y} &= x(\rho - z) - y, \\ \dot{z} &= xy - \beta z, \end{cases}$$

649 which induces a continuous dynamical system $\varphi : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}^3$. We consider the classical
 650 values of the parameters: $\sigma = 10$, $\rho = 28$, and $\beta = 8/3$. A time series can be generated by
 651 iterates of the map $f(x) := \varphi(x, \tilde{t})$, where $\tilde{t} > 0$ is a fixed value of the time parameter. For
 652 the following experiments we chose $\tilde{t} = 0.02$ and we use the Runge-Kutta method of an order
 653 5(4) to generate the time series.

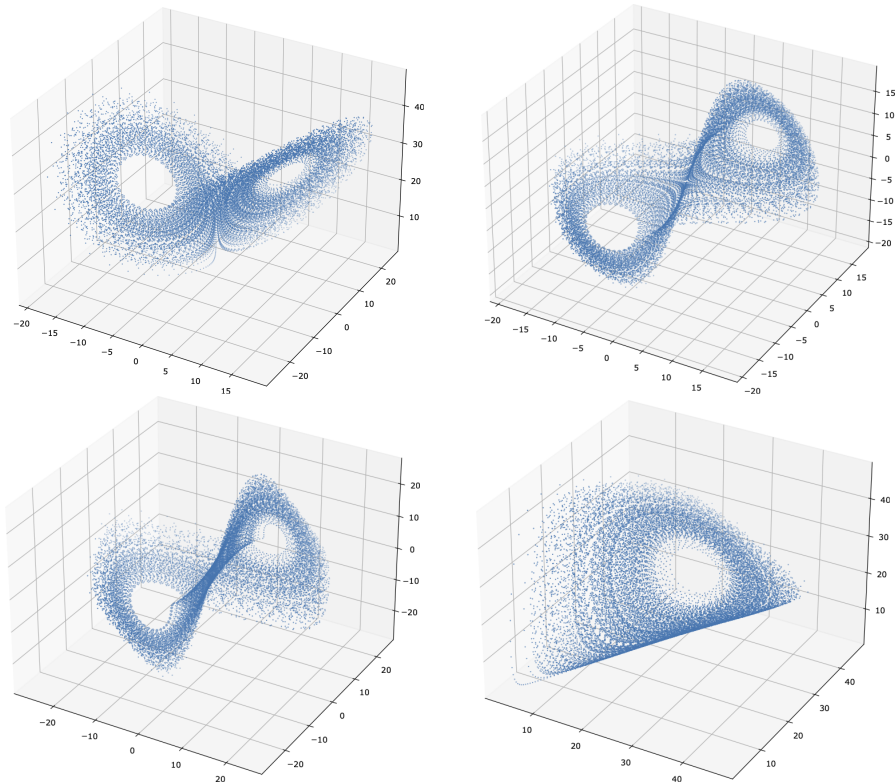


Figure 8. A time series generated from the Lorenz system (top left) and 3-d embeddings of its projections onto the x -coordinate (top right), y -coordinate (bottom left) and z -coordinate (bottom right) with a delay parameter $l = 5$.

654 4.4.1. Experiment 4A.

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

$$657 \quad \mathcal{L}_i = \varrho(f, f^{2000}(p_i), 10000), \quad \mathcal{P}_{x,d}^i = \Pi(\mathcal{L}_i^{(1)}, d, 5), \quad \mathcal{P}_{z,d}^i = \Pi(\mathcal{L}_i^{(3)}, d, 5),$$

658 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
659 projection of time series \mathcal{L}_i onto its j -th coordinate. In all the embeddings we choose the lag
660 $l = 5$. Note that the first point of time series \mathcal{L}_i is equal to the 2000-th iterate of point p_i
661 under map f . It is a standard procedure to cut off some transient part of the time series.

662 Time series $\mathcal{P}_{x,d}^i$ and $\mathcal{P}_{z,d}^i$ are embeddings of the first and third coordinate of \mathcal{L}_i , respec-
663 tively. As Figure 8 (top right) suggests, the embedding of the first coordinate into \mathbb{R}^3 results
664 in a structure topologically similar to the Lorenz attractor. The embedding of the third coord-
665 inate, due to the symmetry of the system, produces a collapsed structure with “wings” of the
666 attractor glued together (Figure 8, right). Thus, we expect time series $\mathcal{P}_{z,d}^i$ to be recognized
667 as non-conjugate to \mathcal{L}_i .

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

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

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

$$675 \quad (4.9) \quad h : \mathcal{L}_i \ni \mathbf{x}_t \mapsto (\mathbf{x}_t^{(1)}, \mathbf{x}_{t+5}^{(1)}, \dots, \mathbf{x}_{t+5d}^{(1)}) \in P_{x,d}^i \subseteq \mathbb{R}^d,$$

676 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 $\mathbf{x}_t^{(1)} = x_t$
 677 denotes its projection onto the x -coordinate. When $j = i$ this formula matches the points of
 678 \mathcal{L}_i to the corresponding points of $P_{x,d}^j = P_{x,d}^i$. However, if $j \neq i$ then the points of \mathcal{L}_i are
 679 mapped onto the points of $P_{x,d}^i$, not to $P_{x,d}^j$, thus in fact in our comparison tests we verify
 680 how well $P_{x,d}^j$ approximates the image of \mathcal{L}_i under h and the original dynamics.

681 For the symmetric comparison of $P_{x,d}^j$ with \mathcal{L}_i the local approximation of $h^{-1} : \mathbb{R}^d \rightarrow \mathbb{R}^3$
 682 will take a form:

$$683 \quad (4.10) \quad h^{-1} : P_{x,d}^j \ni (\mathbf{x}_t^{(1)}, \mathbf{x}_{t+5}^{(1)}, \dots, \mathbf{x}_{t+5d}^{(1)}) \mapsto \mathbf{x}_t \in \mathcal{L}_j \subset \mathbb{R}^3.$$

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

689 In the experiments below we use the maximum metric.

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

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

703 Interestingly, when we use ConjTest to compare \mathcal{L}_1 with embedding time series generated
 704 from \mathcal{L}_1 we always get values 0.0. The connecting maps used in this experiment, defined by
 705 (4.9) and (4.10), establish a direct correspondence between points in two time series. As a
 706 result we get $\tilde{h} = h$ in the definition of ConjTest, and consequently, every pair of sets in the
 707 numerator of equation (3.5) is the same. If the embedded time series comes from another
 708 trajectory then $\tilde{h} \neq h$ and ConjTest gives the expected results, as visible in Table 5. On the
 709 other hand, computationally more demanding ConjTest⁺ exhibits virtually the same results

method \ test	\mathcal{L}_1 vs. $\mathcal{P}_{x,1}^1$	\mathcal{L}_1 vs. $\mathcal{P}_{x,2}^1$	\mathcal{L}_1 vs. $\mathcal{P}_{x,3}^1$	\mathcal{L}_1 vs. $\mathcal{P}_{z,1}^1$	\mathcal{L}_1 vs. $\mathcal{P}_{z,3}^1$
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. $\mathcal{P}_{x,1}^2$	\mathcal{L}_1 vs. $\mathcal{P}_{x,2}^2$	\mathcal{L}_1 vs. $\mathcal{P}_{x,3}^2$	\mathcal{L}_1 vs. $\mathcal{P}_{z,1}^2$	\mathcal{L}_1 vs. $\mathcal{P}_{z,3}^2$
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.

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

712 **4.4.2. Experiment 4B.** This experiment is proceeded according to the standard use of
713 FNN for estimating optimal embedding dimension without an explicit knowledge about the
714 original system.

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

$$716 \quad \mathcal{L} = \varrho(f, f^{2000}(p), 10000), \quad \mathcal{P}_d = \Pi(\mathcal{L}^{(1)}, d, 5),$$

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

723 left to be separated.

724 **Results.** Results are presented in Figure 9. In general, the outputs of all methods are
 725 consistent. When the one-dimensional embedding, \mathcal{P}_1 , is compared with two-dimensional
 726 embedding, \mathcal{P}_2 , we get large comparison values for the entire range of every parameter. When
 727 we compare \mathcal{P}_2 with \mathcal{P}_3 the estimation of dissimilarity drops significantly, i.e. we conclude
 728 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
 729 with \mathcal{P}_4 still decreases the values, suggesting that the third dimension improves the quality of
 730 our embedding. The curve corresponding to \mathcal{P}_4 vs. \mathcal{P}_5 essentially overlaps \mathcal{P}_3 vs. \mathcal{P}_4 curve.
 731 Thus, the third dimension seems to be a reasonable choice.

732 We can see that FNN (Figure 9 top left), originally designed for this test, gives a clear
 733 answer. However, in the case of KNN (Figure 9 top right) the difference between the yellow
 734 and the green curve is rather subtle. Thus, the outcome could be alternatively interpreted
 735 with a claim that two dimensions are enough for this embedding. In the case of ConjTest⁺
 736 we have two parameters. For the fixed value of $t = 10$ we manipulated the value of k (Figure
 737 9 bottom left) and the outcome matched up with the FNN result. However, the situation is
 738 slightly different when we fix $k = 5$ and vary the t (Figure 9 bottom right). For $t < 30$ the
 739 results suggest dimension 3 to be optimal for the embedding, but for $t > 40$ the green and
 740 the red curve split. Moreover, for $t > 70$, we can observe the beginning of another split of the
 741 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.
 742 We attribute this effect to the chaotic nature of the attractor. The higher the value of t the
 743 higher the effect. We investigate it further in the following experiment.

744 **4.4.3. Experiment 4C.** In this experiment we investigate the dependence of ConjTest⁺
 745 on the choice of value of parameter t . Parameter t of ConjTest⁺ controls how far we push
 746 the approximation of a neighborhood of a point x_i (U_k^i in (3.7)) through the dynamics. In
 747 the case of systems with a sensitive dependence on initial conditions (e.g., the Lorenz system)
 748 we could expect that higher values of t spread the neighborhood over the attractor. As a
 749 consequence, we obtain higher values of ConjTest⁺.

750 **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
 751 we study the following time series:

$$752 \quad \mathcal{L}_i = \varrho(f, f^{2000}(p_i), 10000), \quad \mathcal{P}_{x,d}^i = \Pi(\mathcal{L}_i^{(1)}, d, 5), \quad \mathcal{P}_{y,d}^i = \Pi(\mathcal{L}_i^{(2)}, d, 5),$$

753 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
 754 others using ConjTest⁺ method with the range of parameter

$$755 \quad t \in \{1, 5, 9, 13, 17, 21, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80\}.$$

756 **Results.** The top plot of Figure 10 presents the results of comparing \mathcal{L}_1 to the time series
 757 \mathcal{L}_i and $\mathcal{P}_{x,d}^i$ with $i \in \{2, 3, 4\}$ and $d \in \{1, 2, 3, 4\}$. Red curves correspond to \mathcal{L}_1 vs. $\mathcal{P}_{x,1}^i$,
 758 green curves to \mathcal{L}_1 vs. $\mathcal{P}_{x,2}^i$, blue curves to \mathcal{L}_1 vs. $\mathcal{P}_{x,3}^i$, and dark yellow curves to \mathcal{L}_1 vs.
 759 $\mathcal{P}_{x,4}^i$. There are three curves of every color, each one corresponds to a different starting point
 760 p_i , $i \in \{2, 3, 4\}$. The bottom part shows results for comparison of \mathcal{L}_1 to $\mathcal{P}_{y,d}^i$ (we embed
 761 the y -coordinate time series instead of x -coordinate). The color of the curves is interpreted
 762 analogously. Black curves on both plots are the same and correspond to the comparison of
 763 \mathcal{L}_1 with \mathcal{L}_j for $j \in \{2, 3, 4\}$.

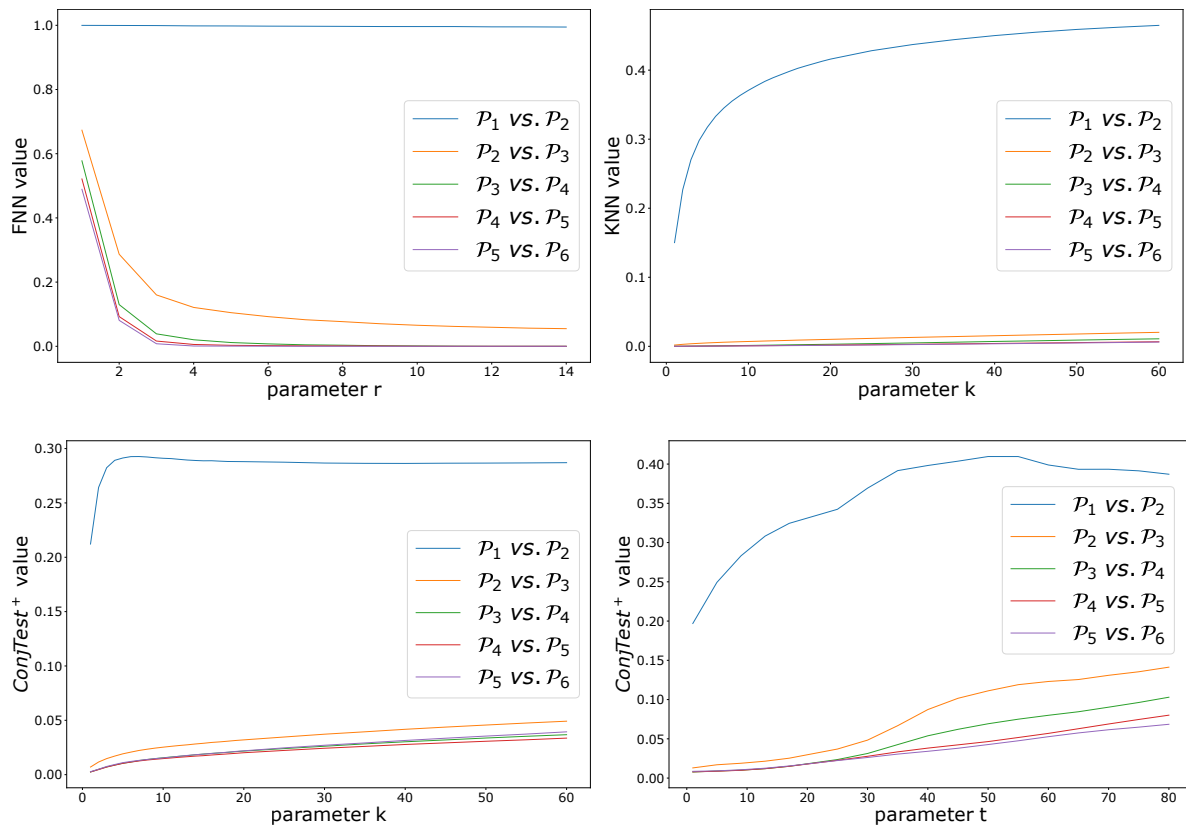


Figure 9. A comparison of embeddings for consecutive dimensions. Top left: FNN with respect to parameter r . Top right: KNN with respect to parameter k . Bottom left: ConjTest^+ with respect to parameter k . Bottom right: ConjTest^+ with respect to parameter t .

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

776 The y -coordinate embeddings presented in the bottom part of Figure 10 give similar
 777 results. However, we can see that gaps between curves corresponding to different dimensions
 778 are more visible. Moreover, the absolute level of all curves is higher. We interpret this outcome
 779 with a claim that the y -coordinate inherits a bit less information about the original system

780 than the x -coordinate. In Figure 8 we can see that y -embedding is more twisted in the center
 781 of the attractor. Hence, generally values are higher, and more temporal information is needed
 782 (reflected by higher embedding dimension) to compensate.

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

785 **4.5. Example: rotation on the Klein bottle.** In the next example we consider the Klein
 786 bottle, denoted \mathbb{K} and defined as an image $\mathbb{K} := \text{im } \beta$ of the map β :

$$787 \quad (4.11) \quad \beta : [0, 2\pi) \times [0, 2\pi) \ni \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} \cos \frac{x}{2} \cos y - \sin \frac{x}{2} \sin(2y) \\ \sin \frac{x}{2} \cos y + \cos \frac{x}{2} \sin(2y) \\ 8 \cos x (1 + \frac{\sin y}{2}) \\ 8 \sin x (1 + \frac{\sin y}{2}) \end{bmatrix} \in \mathbb{R}^4.$$

788 In particular, the map β is a bijection onto its image and the following “rotation map”
 789 $f_{[\phi_1, \phi_2]} : \mathbb{K} \rightarrow \mathbb{K}$ over the Klein bottle is well-defined:

$$790 \quad f_{[\phi_1, \phi_2]}(x) := \beta \left(\beta^{-1}(x) + \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} \pmod{2\pi} \right).$$

791 **4.5.1. Experiment 5A.** We conduct an experiment analogous to Experiment 4B on esti-
 792 mating the optimal embedding dimension of a projection of the Klein bottle.

793 *Setup.* We generate the following time series

$$794 \quad \mathcal{K} = \varrho(f_{[\phi_1, \phi_2]}, (0, 0, 0, 0), 8000),$$

$$795 \quad \mathcal{P}_d = \Pi \left((\mathcal{K}^{(1)} + \mathcal{K}^{(2)} + \mathcal{K}^{(3)} + \mathcal{K}^{(4)})/4, d, 8 \right),$$

796 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
 797 coordinate. Note that in previous experiments we mostly used a simple observable s which
 798 was a projection onto a given coordinate. However, in general, one can consider any (smooth)
 799 function as an observable. Therefore in the current experiment, in the definition of \mathcal{P}_d , s is a
 800 sum of all the coordinates, not the projection onto a chosen one. Note also that because of
 801 the symmetries (see formula (4.11)) a single coordinate might be not enough to reconstruct
 802 the Klein bottle.

803 *Results.* We can proceed with the interpretation similar to Experiment 4B. The FNN
 804 results (Figure 11 top left) suggests that 4 is a sufficient embedding dimension. The similar
 805 conclusion follows from KNN (Figure 11 top right) and ConjTest⁺ with a fixed parameter
 806 $k = 10$ (Figure 11 bottom right). The bottom left figure of 11 is inconclusive as for the higher
 807 values of k the curves do not stabilize even with high dimension.

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

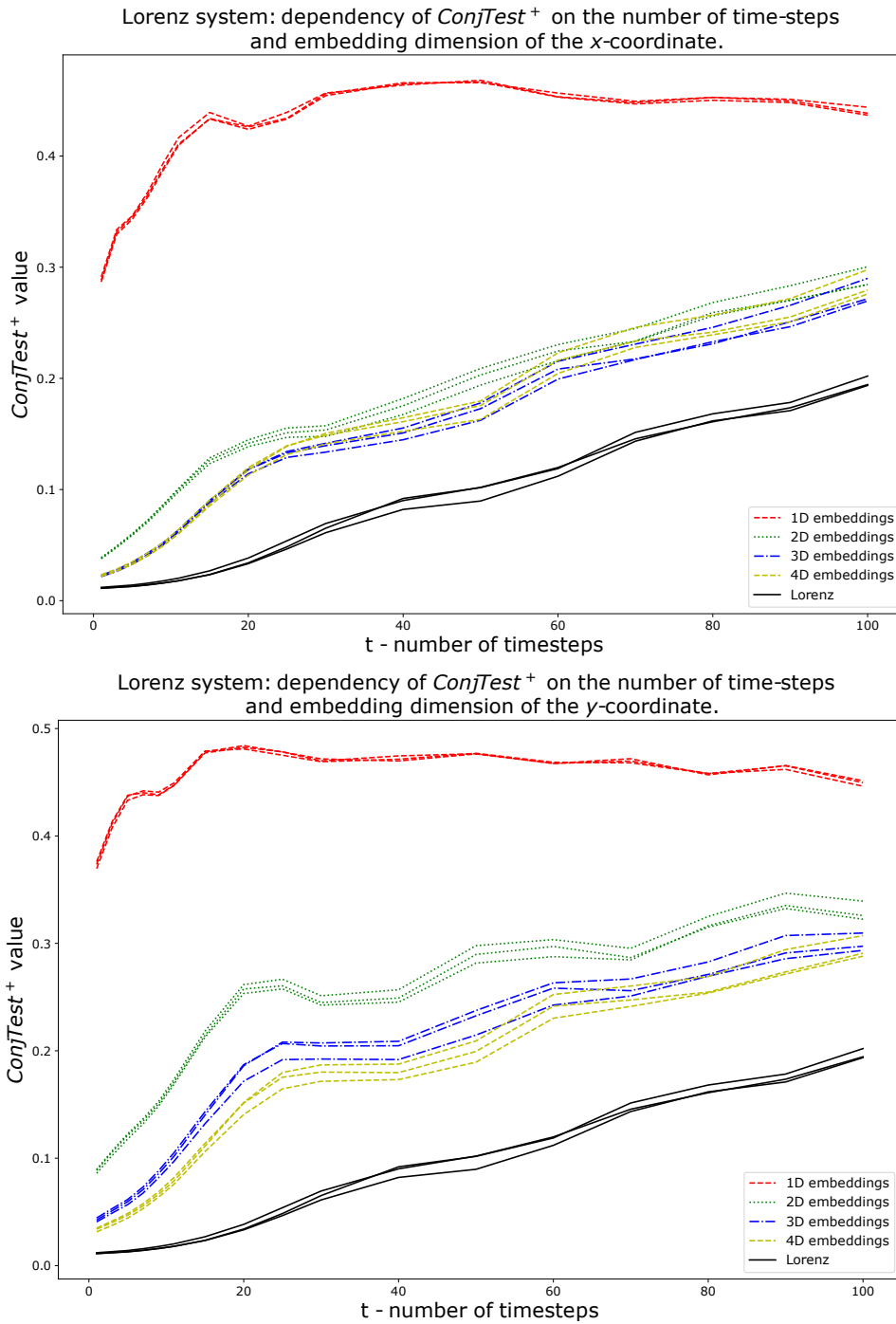


Figure 10. *Dependence of $ConjTest^+$ on the parameter t for Lorenz system. In this experiment multiple time series with different starting points were generated. Each of them was used to produce an embedding. Top: comparison of x -coordinate embedding with \mathcal{L}_1 . Bottom: comparison of y -coordinate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

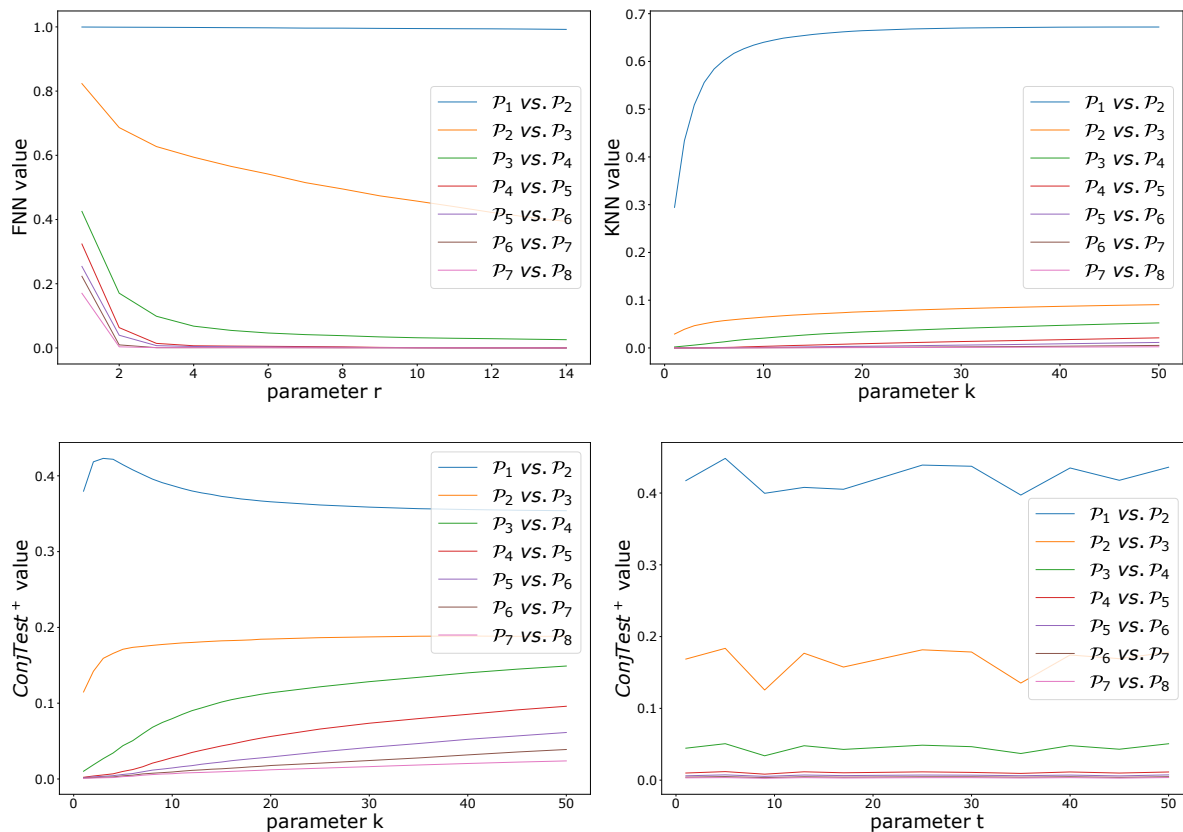


Figure 11. A comparison of the conjugacy measures for embeddings of the Klein bottle for consecutive dimensions. Top left: FNN with respect to parameter r . Top right: KNN with respect to parameter k . Bottom left: ConjTest⁺ with respect to parameter k ($t = 10$ fixed). Bottom right: ConjTest⁺ with respect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

811 **5. Approximation of the connecting homeomorphism.** In whole generality, finding the
 812 connecting homeomorphism between conjugate dynamical systems, is a very difficult task.
 813 Some prior work has been done in this direction but the existing methods still have many
 814 limitations in applying for a broader class of systems. In particular, the works [26, 27, 32]
 815 developed a method to produce conjugacy functions based on a functional fixed-point iteration
 816 scheme that can also be generalized to compare non-conjugate dynamical systems in which
 817 case the limit point of a fixed-point iteration scheme yields a function called a “commuter”.
 818 Quantifying how much the commuter function fails to be a homeomorphism (in various mea-
 819 sures) led to the notion of a “homeomorphic defect” that, as the authors point out, allows one
 820 to quantify the dissimilarity of the two dynamical systems. However, the method has been
 821 illustrated on a very few specific examples and could be rigorously mathematically justified
 822 only under very restrictive assumptions e.g. uniform contraction of at least one of the systems
 823 when comparing systems in one dimension. Significant problems occur in rigorous extension
 824 to systems of higher dimension. Consequently, later work [5] extended this theory to allow for
 825 multivariate transformations and presented ideas on constructing commuter functions differ-

ent 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 commutators 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 $0 = a_1 < a_2 < \dots < a_{n+1} = 1$. Denote $\mathbb{A}_{i,j} := [a_i, a_{i+1}] \times [a_j, a_{j+1}]$ and $\mathbb{A} := \{\mathbb{A}_{i,j} \mid i, j \in \{1, 2, \dots, n\}\}$. Let $h : I \rightarrow I$ 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 $\mathbf{h} = \{\mathbb{A}_{i,j} \in \mathbb{A} \mid \text{int } \mathbb{A}_{i,j} \cap \pi(h) \neq \emptyset\}$ is a *the cubical approximation* of h and we denote it by $[h]$. Equivalently, \mathbf{h} is the minimal subset of \mathbb{A} such that $\pi(h) \subset \bigcup \mathbf{h}$. We refer to

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

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 $\mathbf{h} \in \mathbb{H}$. We denote a *selector* of \mathbf{h} , that is a homeomorphism representing \mathbf{h} , by $f_{\mathbf{h}}$. 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 \mathbb{H} grows exponentially with the resolution of the grid (the explicit

869 formula for the size of H is given in A). For instance, the number of cubical homeomorphisms
 870 for $m = 21$ is about $2.6 \cdot 10^{14}$. Thus, it is hopeless to find the optimal approximation of the
 871 connecting homeomorphism by a brute examination of all elements of H . Instead, we propose
 872 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, $\mathbf{h}_0 \in \mathbf{h}$ – initial approximation of the homeomor-
 phism, `nsteps` – number of steps, `p` – memory size.

Output: `best_h` – approximated connecting homeomorphism

```

1:  $\mathbf{h} \leftarrow \mathbf{h}_0$ 
2:  $\mathbf{q} \leftarrow$  initialize queue of size  $p$  with null's
3:  $\mathbf{best\_h} \leftarrow \mathbf{h}$ 
4: for  $t = 0$  to nsteps do
5:    $\mathbf{c} \leftarrow \{\mathbf{h}' \in \mathbf{h} \mid \mathbf{h}' \in \text{nbhd}(\mathbf{h}) \text{ and } \text{diff}(\mathbf{h}, \mathbf{h}') \notin \mathbf{q}\}$ 
6:   if  $\#\mathbf{c} = 0$  then
7:     break
8:   else
9:      $\mathbf{h} \leftarrow \mathbf{h}' \in \mathbf{c}$  with a minimal value of  $\text{score}(\mathbf{h}', \mathcal{A}, \mathcal{B})$ 
10:    if  $\text{score}(\mathbf{h}, \mathcal{A}, \mathcal{B}) < \text{score}(\mathbf{best\_h}, \mathcal{A}, \mathcal{B})$  then
11:       $\mathbf{best\_h} \leftarrow \mathbf{h}$ 
12:    end if
13:     $\mathbf{q.append}(\text{diff}(\mathbf{h}, \mathbf{h}'))$  {append the unique element differentiating  $\mathbf{h}$  and  $\mathbf{h}'$ }
14:     $\mathbf{q.pop}()$ 
15:  end if
16: end for
17: return  $\mathbf{best\_h}$ 

```

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

$$876 \quad \text{score}(\mathbf{h}, \mathcal{A}, \mathcal{B}) := \max \{ \text{ConjTest}(\mathcal{A}, \mathcal{B}; k, t, f_{\mathbf{h}}), \text{ConjTest}(\mathcal{B}, \mathcal{A}; k, t, f_{\mathbf{h}}^{-1}) \}.$$

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

$$879 \quad \text{diff}(\mathbf{h}, \mathbf{h}') := (\mathbf{h} \setminus \mathbf{h}') \cup (\mathbf{h}' \setminus \mathbf{h}).$$

880 Let $\mathbf{h} \in H$ be an initial guess for the connecting homeomorphism. In each step of the
 881 algorithm an attempt is made to update it in a way that the `score` gets improved. Each
 882 iteration of the main loop considers all neighbors \mathbf{h}' of \mathbf{h} in H such that $\text{nbhd}(\mathbf{h}) := \{\mathbf{h}' \in H \mid$
 883 $\#\text{diff}(\mathbf{h}, \mathbf{h}') = 1\}$ (a unit sphere in a Hamming distance centered in \mathbf{h}). The element \mathbf{h}' of
 884 $\text{nbhd}(\mathbf{h})$ with minimal $\text{score}(\mathbf{h}')$ is chosen for the next iteration of the algorithm. Note that
 885 it might happen that $\text{score}(\mathbf{h}') < \text{score}(\mathbf{h})$. This prevents the algorithm from being stuck at
 886 a local minimum. In addition, to avoid orbiting around them we exclude elements of $\text{nbhd}(\mathbf{h})$

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

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

$$893 \quad \mathcal{A} = \varrho(f_4, 0.02, 500) \quad \text{and} \quad \mathcal{B} = \varrho(g_2, 0.87, 500),$$

894 where f_4 and g_2 are respectively a logistic and a tent map as in Section 4.3. We use `score`
 895 with parameters $k = 5$ and $t = 1$. As an initial guess of the connecting homeomorphism h_0
 896 we naively took a cubical approximation of a $h(x) = x^5$ with resolution $m = 21$, as presented
 897 in the top-left panel of Figure 12. We run Algorithm 5.1 for time series \mathcal{A} and \mathcal{B} for 1000
 898 steps with the memory parameter $p = m = 21$. Figure 13 shows the values of the `score`
 899 for the consecutive approximations. We can see that the algorithm falls temporarily into
 900 local minima, but eventually, thanks to the memory parameter, it escapes them and settles
 901 down towards the low score values. Figure 12 shows relations corresponding to the 1st, 200th,
 902 400th, and 612nd iteration of the algorithm run. The bottom-right panel, the 612nd iteration
 903 is the relation inducing the lowest score among all iterations. The orange curve, at the
 904 same panel, is the graph of homeomorphism (4.5) – the analytically correct map conjugating
 905 f_4 and g_2 . Clearly, the iterations are converging towards the right value of the connecting
 906 homeomorphism.

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

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

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

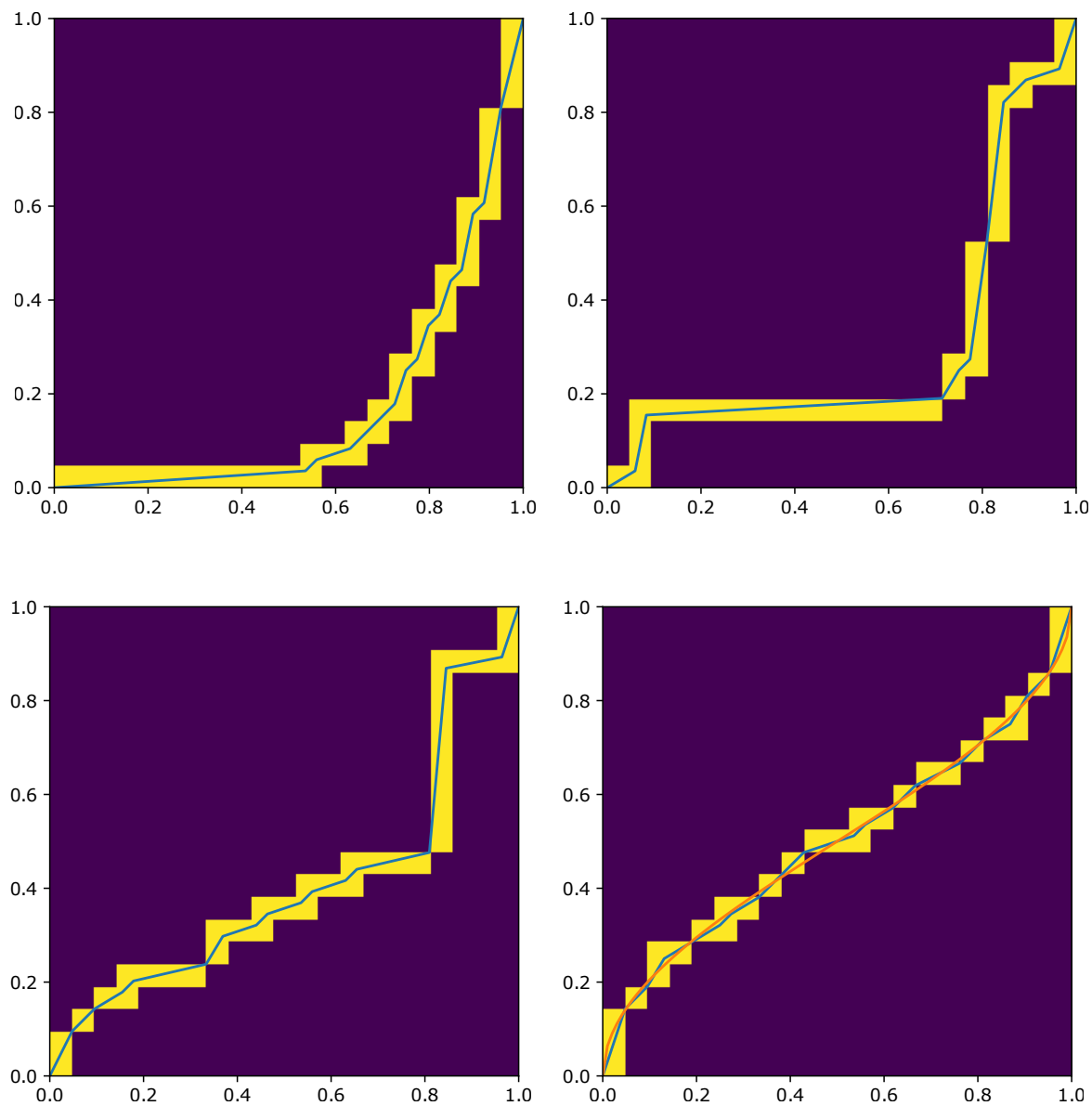


Figure 12. Steps 0, 200, 400, and 612 (top left, top right, bottom left, bottom right, respectively) of the run of Algorithm 5.1 in a search for the connecting homeomorphism between the logistic and the tent map. The blue lines corresponds to a selector of a cubical homeomorphism. The orange curve in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

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

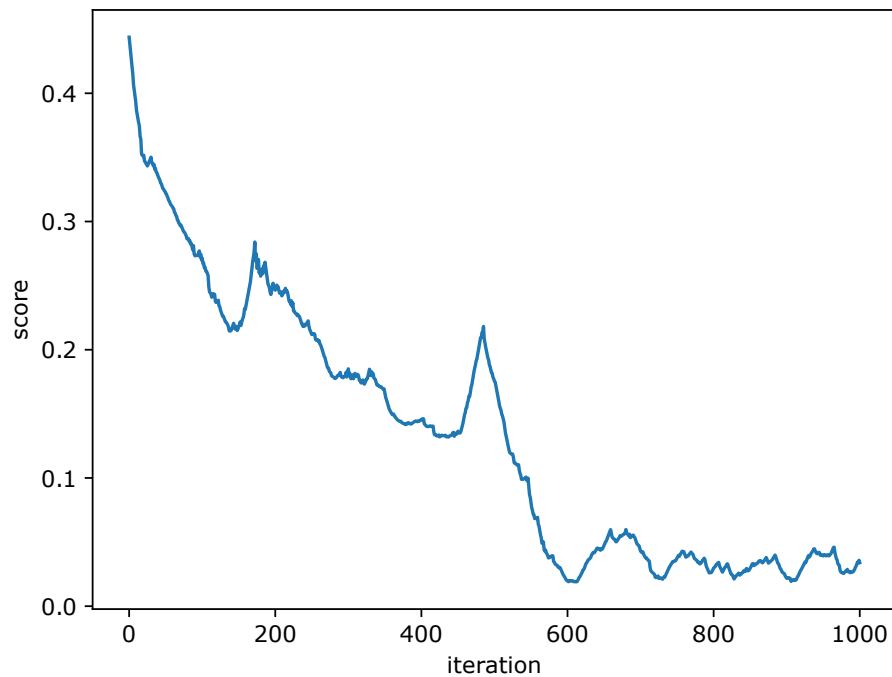


Figure 13. The value of $\text{score}(h, A, B)$ for every step of the experiment approximating the connecting homeomorphism between the logistic and the tent map.

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

937 Both ConjTest and ConjTest⁺ (collectively called ConjTest methods) are directly inspired
 938 by the definition and properties of topological conjugacy. They are more flexible in all con-
 939 sidered experiments and can be applied to time series of different lengths and generated by
 940 different initial conditions (the first point of the series). In contrast to FNN and KNN, they
 941 admit more robust behavior with respect to any kind of perturbation, be it measurement
 942 noise (Experiment 1C), perturbation of the initial condition (Experiments 1A, 2A, 3A, and
 943 4A), t parameter (Experiment 4C), or a parameter of a system (Experiment 1B). In most
 944 experiments, we can observe a continuous-like dependence of the test value on the level of per-
 945 turbations. We see this effect as softening the concept of topological conjugacy by ConjTest
 946 methods. A downside of this weakening is a lack of definite response whether two time series
 947 come from conjugate dynamical systems. Hence the ConjTest methods should be considered
 948 as a means for a quantification of a dynamical similarity of two processes. Experiments 1A,
 949 2A, and 3A show that both methods, ConjTest and ConjTest⁺, capture essentially the same
 950 information from data. In general, ConjTest is simpler and, thus, computationally more ef-
 951 ficient. However, Experiment 4A shows that ConjTest (in contrast to ConjTest⁺) does not
 952 work well in the context of embedded time series, especially when the compared embeddings
 953 are constructed from the same time series. Experiments 4B and 5A show that the varia-
 954 tion of ConjTest methods with respect to the t parameter can also be used for estimating

Property \ Method	FNN	KNN	ConjTest	ConjTest ⁺
Requirements	identical matching between indexes of the elements (in particular the series must be of the same length)		<ul style="list-style-type: none"> • 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 	
Parameters	only one parameter: r (but one should examine large interval of r values)	only one parameter: k (but is recommend to check a couple of different k values)	involve two parameters: k and t	
Robustness	<ul style="list-style-type: none"> • 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 		<ul style="list-style-type: none"> • 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 	
Recurrent properties	takes into account only the one closest return of a series (trajectory) to each neighborhood	takes into account k -closest returns		
Further properties			more likely to give false positive answer than ConjTest ⁺	more computationally demanding than ConjTest but usually more reliable

Table 6

Comparison of the properties of discussed conjugacy measures.

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

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

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

973 Appendix A. Cubical homeomorphisms.

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

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

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

978 *Proof.* Since h is an increasing homeomorphism it follows that $(0,0), (1,1) \in \pi(h)$. In
 979 consequence, $\mathbf{A}_{1,1}, \mathbf{A}_{n,n} \in [h]$, because these are the only elements of \mathbb{A} containing $(0,0)$ and
 980 $(1,1)$. ■

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

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

- 983 (1) $\text{int } \mathbf{A}_{i+1,j} \cap \pi(h) \neq \emptyset$ and $\text{int } \mathbf{A}_{i,j+1} \cap \pi(h) = \emptyset$, when $h(a_{i+1}) \in (a_j, a_{j+1})$,
 984 (2) $\text{int } \mathbf{A}_{i,j+1} \cap \pi(h) \neq \emptyset$ and $\text{int } \mathbf{A}_{i+1,j} \cap \pi(h) = \emptyset$, when $h^{-1}(a_{i+1}) \in (a_i, a_{i+1})$,
 985 (3) $\text{int } \mathbf{A}_{i+1,j+1} \cap \pi(h) \neq \emptyset$, $\text{int } \mathbf{A}_{i+1,j} \cap \pi(h) = \emptyset$ and $\text{int } \mathbf{A}_{i,j+1} \cap \pi(h) = \emptyset$, when $h(a_{i+1}) =$
 986 a_{j+1} .

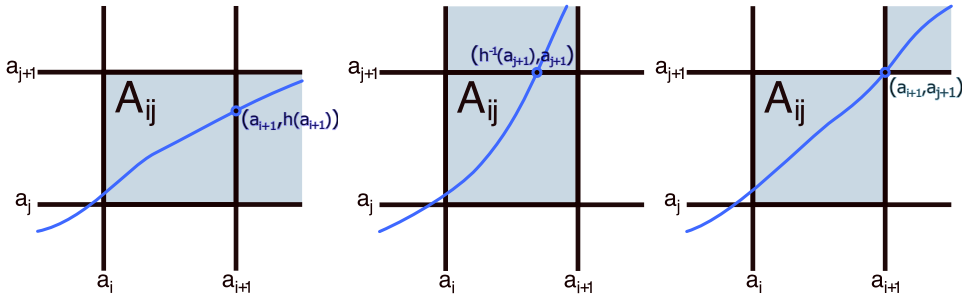


Figure 14. From left to right, cases (1), (2) and (3) of Proposition A.2.

987 Again, the proof for the next proposition is a consequence of basic properties of homeo-
 988 morphism h such that $[h] = \mathbf{h}$.

989 **Proposition A.3.** *Let $\mathbf{h} \in \mathbb{H}$. Then*

- 990 (i) for every $i \in \{1, \dots, n\}$ set $\bigcup \{ \mathbf{A}_{i,j} \in \mathbf{h} \mid j \in \{0, 1, \dots, n\} \}$ is nonempty and connected,
 991 (ii) for every $j \in \{1, \dots, n\}$ set $\bigcup \{ \mathbf{A}_{i,j} \in \mathbf{h} \mid i \in \{0, 1, \dots, n\} \}$ is nonempty and connected.
 992 (iii) if $\mathbf{A}_{i,j} \in \mathbf{h}$ then for every $i' > i$ and $j' < j$ we have $\mathbf{A}_{i',j'} \notin \mathbf{h}$,
 993 (iv) if $\mathbf{A}_{i,j} \in \mathbf{h}$ then for every $j' > j$ and $i' < i$ we have $\mathbf{A}_{i',j'} \notin \mathbf{h}$,

994 The above propositions implies that every element $\mathbf{h} \in \mathbb{H}$ can be seen as a path starting
 995 from element $A_{1,1}$ and ending at $A_{n,n}$. In particular, \mathbf{h} can be represented as a vector of symbols
 996 R (right, incrementation of index i , case (1)), U (up, incrementation of index j , case (2)), and
 997 D (diagonal, incrementation of both indices, case (3)). The vectors do not have to be of the
 998 same length. In particular, a single symbol D can replace a pair of symbols R and U . Denote
 999 by n_R , n_U and n_D the number of corresponding symbols in the vector. As indices i and j
 1000 have to be incremented from 1 to n we have the following properties:

$$1001 \quad (\text{A.1}) \quad 0 \leq n_R, n_U, n_D \leq n - 1, \quad n_R + n_U + 2n_D = 2(n - 1) \quad \text{and} \quad n_R = n_U.$$

1002 Actually, any vector \mathbf{V} of symbols $\{R, U, D\}$ satisfying the above conditions (A.1) corre-
 1003 sponds to a cubical homeomorphism. We show it by constructing a piecewise-linear homeo-
 1004 morphism h such that $[h] = \mathbf{h}$ for any \mathbf{h} represented by \mathbf{V} . We refer to the constructed h as
 1005 a *selector* of \mathbf{h} . In particular, Algorithm A.1 produces a sequence of points corresponding
 1006 to points of non-differentiability of the homeomorphism. We have five type of points, the
 1007 starting point $(0, 0)$ (type B), the ending point $(1, 1)$ (type E), and points corresponding to
 1008 subsequences UR (type UR), RU (type RU) and D (type D). Proposition A.4 shows that
 1009 the map generated by the algorithm is an actual homeomorphism.

Algorithm A.1 FindSelector

Input: \mathbf{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

```

1:  $L \leftarrow \{(0, 0)\}$ 
2:  $\text{prev} \leftarrow \text{Null}$ 
3:  $i, j \leftarrow 0$ 
4: for  $\mathbf{s} \in \mathbf{V}$  do
5:   if  $\text{prev} = U$  and  $\mathbf{s} = R$  then
6:      $L = L \cup \{(p a_i + (1 - p) a_{i+1}, (1 - p) a_j + p a_{j+1})\}$ 
7:   else if  $\text{prev} = R$  and  $\mathbf{s} = U$  then
8:      $L = L \cup \{(1 - p) a_i + p a_{i+1}, p a_j + (1 - p) a_{j+1}\}$ 
9:   else if  $\mathbf{s} = D$  then
10:     $L = L \cup \{(a_{i+1}, a_{j+1})\}$ 
11:   end if
12:    $\text{prev} \leftarrow \mathbf{s}$ 
13:   if  $\mathbf{s} = R$  or  $\mathbf{s} = D$  then
14:      $i \leftarrow i + 1$ 
15:   end if
16:   if  $\mathbf{s} = U$  or  $\mathbf{s} = D$  then
17:      $j \leftarrow j + 1$ 
18:   end if
19: end for
20:  $L \leftarrow \{(1, 1)\}$ 
21: return  $L$ 

```

1010 **Proposition A.4.** Let \mathbf{V} be a vector of symbols $\{R, U, D\}$ satisfying (A.1) and \mathbf{h} the corre-
 1011 sponding cubical set. Let $L = \{(x_1, y_1), (x_2, y_2), \dots, (x_K, y_K)\}$ be a sequence of points gener-
 1012 ated by Algorithm A.1 for \mathbf{V} . Then, for every $k \in \{1, 2, \dots, K-1\}$ following properties are
 1013 satisfied:

- 1014 (i) $x_k < x_{k+1}$ and $y_k < y_{k+1}$,
 1015 (ii) if $(x_k, y_k) \in L$ then $(x_k, y_k) + (1-t)(x_k, y_{k+1}) \in \bigcup \mathbf{h}$ for $t \in [0, 1]$.

1016 *Proof.* First, note that if (x_k, y_k) is of type UR or RU then we have $(x_k, y_k) \in \text{int } A_{i,j}$.
 1017 If (x_k, y_k) is of type D we have $(x_k, y_k) = A_{i,j} \cap A_{i+1,j+1}$. In particular, in that case $x_k, y_k \in$
 1018 $\{a_2, a_3, \dots, a_n\}$.

1019 Sequence L always begins $(x_0, y_0) = (0, 0)$ and ends with $(x_K, y_K) = (1, 1)$. By Proposition
 1020 A.1 we have $A_{0,0}, A_{n,n} \in \mathbf{h}$. The above observations shows that for every (x_k, y_k) with $k \in$
 1021 $\{1, 2, \dots, K-1\}$ we have $0 < x_k, y_k < 1$. Thus, two first and two last points of L satisfies
 1022 (i). If (x_1, y_1) is of type UR then it follows that \mathbf{V} begins with a sequence of U 's. We have
 1023 $(x_1, y_1) \in \text{int } A_{0,j}$ for some $j > 0$. By Proposition A.3(ii) all $A_{0,j'} \in \mathbf{h}$ for $0 \leq j' \leq j$. Hence,
 1024 the interval spanned by (x_0, y_0) and (x_1, y_1) is contained in $\bigcup \mathbf{h}$ proving (ii). The cases when
 1025 (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
 1026 similar argument.

1027 Let (x, y) and (x', y') be two consecutive points of L . Suppose that (x, y) is of type UR
 1028 and (x', y') of type RU . This situation arises when two symbols U are separated by a positive
 1029 number of symbols R (see Figure 15 top). It follows that

$$1030 \quad (x, y) = (p a_i + (1-p) a_{i+1}, (1-p) a_j + p a_{j+1}) \in \text{int } A_{i,j},$$

$$1031 \quad (x', y') = ((1-p) a_{i'} + p a_{i'+1}, p a_j + (1-p) a_{j+1}) \in \text{int } A_{i',j},$$

1032 where $i < i'$. Thus,

$$1033 \quad y' - y = p a_j + (1-p) a_{j+1} - ((1-p) a_j + p a_{j+1})$$

$$1034 \quad = 2p a_j + a_{j+1} - a_j > a_{j+1} - a_j > 0.$$

1035 Consequently, we get $x < x'$ and $y < y'$ proving (i) for this case. By Proposition A.3(ii) we
 1036 get that $A_{i'',j} \in \mathbf{h}$ for all $i \leq i'' \leq i'$. Hence, the interval spanned by (x, y) and (x', y') is
 1037 contained in $\bigcup \mathbf{h}$ proving (ii).

1038 The case when (x, y) is of type RU and (x', y') of type UR is analogous.

1039 Suppose that (x, y) is of type UR and (x', y') of type D . This situation arises when
 1040 symbols U and D are separated by a positive number of symbols R (see Figure 15 middle).
 1041 It follows that

$$1042 \quad (x, y) = (p a_i + (1-p) a_{i+1}, (1-p) a_j + p a_{j+1}) \in \text{int } A_{i,j},$$

$$1043 \quad (x', y') = (a_{i'+1}, a_{j+1}) = A_{i',j} \cap A_{i'+1,j+1},$$

1044 where $i < i'$. It follows that $x < x'$ and $y < y'$ proving (i) for this case. Again, by Proposition
 1045 A.3(ii) we can prove (ii).

1046 The case when (x, y) is of type RU and (x', y') of type D is analogous.

1047 Now, suppose that (x, y) is of type D and (x', y') of type UR . This situation arises when
 1048 symbols D and R are separated by a positive number of symbols U (see Figure 15 bottom).
 1049 It follows that

$$1050 \quad (x, y) = (a_{i+1}, a_{j+1}) = \mathbf{A}_{i,j} \cap \mathbf{A}_{i+1,j+1},$$

$$1051 \quad (x', y') = (p a_{i+1} + (1 - p) a_{i+2}, (1 - p) a_{j'} + p a_{j'+1}) \in \text{int } \mathbf{A}_{i+1,j'+1},$$

1052 where $j < j'$. It follows that $x < x'$ and $y < y'$ proving (i) for this case. By Proposition A.3(i)
 1053 follows property (ii).

1054 The case when (x, y) is of type D and (x', y') of type RU is analogous.

1055 Finally, if both (x, y) and (x', y') are of type D it follows that

$$1056 \quad (x, y) = (a_{i+1}, a_{j+1}) = \mathbf{A}_{i,j} \cap \mathbf{A}_{i+1,j+1},$$

$$1057 \quad (x', y') = (a_{i+2}, a_{j+2}) = \mathbf{A}_{i+1,j+1} \cap \mathbf{A}_{i+2,j+2}.$$

1058 Thus, (x, y) and (x', y') are the opposite corners of cube $\mathbf{A}_{i+1,j+1}$ which immediately gives
 1059 both properties (i) and (ii). ■

1060 By counting all possible vectors of symbols $\{R, U, D\}$ satisfying (A.1) we obtain an exact
 1061 size of family \mathbf{H} . In case of $n_D = 0$, the vector has size $2(n - 1)$ and, therefore, we get $\binom{2(n-1)}{(n-1)}$
 1062 ways of ordering symbols R and U . If $n_D = 1$ then the vector size is $2(n - 1) - 1 = 2n - 3$
 1063 and $n_R = n - 2$. Hence, we have $\binom{2n-3}{n-2}$ choices of slots for symbols R and we have choose
 1064 a place for D symbol among the remaining $2n - 3 - (n - 2) = n - 1$ slots. Thus, the total
 1065 number of ordering for $n_D = 1$ is $\binom{2n-3}{n-2}(n - 1)$. In the general case, we get $\binom{2n-2-n_D}{n-1-n_D} \binom{n-1}{n_D}$.
 1066 Finally, the total number of vectors of \mathbf{H} is given by the following formula:

$$1067 \quad \sum_{n_D=0}^{n-1} \binom{2n-2-n_D}{n-1-n_D} \binom{n-1}{n_D}.$$

1068 REFERENCES

1069 [1] A. ALBANO, A. PASSAMANTE, AND M. E. FARRELL, *Using higher-order correlations to define an em-*
 1070 *bedding window*, Physica D: Nonlinear Phenomena, 54 (1991), pp. 85–97, [https://doi.org/10.1016/](https://doi.org/10.1016/0167-2789(91)90110-U)
 1071 [0167-2789\(91\)90110-U](https://doi.org/10.1016/0167-2789(91)90110-U).
 1072 [2] A. M. ALBANO, J. MUENCH, C. SCHWARTZ, A. I. MEES, AND P. E. RAPP, *Singular-value decomposition*
 1073 *and the Grassberger-Procaccia algorithm*, Phys. Rev. A, 38 (1988), pp. 3017–3026, [https://link.aps.](https://link.aps.org/doi/10.1103/PhysRevA.38.3017)
 1074 [org/doi/10.1103/PhysRevA.38.3017](https://link.aps.org/doi/10.1103/PhysRevA.38.3017).
 1075 [3] K. BARAŃSKI, Y. GUTMAN, AND A. ŚPIEWAK, *A probabilistic Takens theorem*, Nonlinearity, 33 (2020),
 1076 p. 4940, <https://dx.doi.org/10.1088/1361-6544/ab8fb8>.
 1077 [4] K. BARAŃSKI, Y. GUTMAN, AND A. ŚPIEWAK, *On the Shroer–Sauer–Ott–Yorke predictability conjecture*
 1078 *for time-delay embeddings*, Commun. Math. Phys., 391 (2022), pp. 609–641, [https://doi.org/10.1007/](https://doi.org/10.1007/s00220-022-04323-y)
 1079 [s00220-022-04323-y](https://doi.org/10.1007/s00220-022-04323-y).
 1080 [5] E. M. BOLLT AND J. D. SKUFCA, *On comparing dynamical systems by defective conjugacy: A sym-*
 1081 *bolic dynamics interpretation of commuter functions*, Physica D: Nonlinear Phenomena, 239 (2010),
 1082 pp. 579–590, <https://doi.org/https://doi.org/10.1016/j.physd.2009.12.007>.
 1083 [6] J. J. BRAMBURGER, S. L. BRUNTON, AND J. N. KUTZ, *Deep learning of conjugate mappings.*, Physica
 1084 D: Nonlinear Phenomena, 427 (2021), p. 133008.

- 1085 [7] H. BROER AND F. TAKENS, *Reconstruction and time series analysis*, Springer New York, New York, NY,
1086 2011, pp. 205–242, https://doi.org/10.1007/978-1-4419-6870-8_6.
- 1087 [8] T. BUZUG AND G. PFISTER, *Comparison of algorithms calculating optimal embedding parameters for*
1088 *delay time coordinates*, *Physica D: Nonlinear Phenomena*, 58 (1992), pp. 127–137, [https://www.](https://www.sciencedirect.com/science/article/pii/016727899290104U)
1089 [sciencedirect.com/science/article/pii/016727899290104U](https://www.sciencedirect.com/science/article/pii/016727899290104U).
- 1090 [9] K. CHAMPION, B. LUSCH, J. N. KUTZ, AND S. L. BRUNTON, *Data-driven discovery of coordinates and*
1091 *governing equations*, *Proceedings of the National Academy of Sciences*, 116 (2019), pp. 22445–22451.
- 1092 [10] D. L. DEANGELIS AND S. YUREK, *Equation-free modeling unravels the behavior of complex ecological*
1093 *systems.*, *Proc Natl Acad Sci U S A*, 112 (2015), pp. 3856–3857.
- 1094 [11] V. DESHMUKH, E. BRADLEY, J. GARLAND, AND J. D. MEISS, *Using curvature to select the time lag for*
1095 *delay reconstruction*, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 30 (2020), p. 063143,
1096 <https://doi.org/10.1063/5.0005890>.
- 1097 [12] A. M. FRASER AND H. L. SWINNEY, *Independent coordinates for strange attractors from mutual informa-*
1098 *tion*, *Phys. Rev. A*, 33 (1986), pp. 1134–1140, <https://link.aps.org/doi/10.1103/PhysRevA.33.1134>.
- 1099 [13] R. HEGGER AND H. KANTZ, *Improved false nearest neighbor method to detect determinism in time series*
1100 *data*, *Phys. Rev. E*, 60 (1999), pp. 4970–4973, <https://link.aps.org/doi/10.1103/PhysRevE.60.4970>.
- 1101 [14] H. KANTZ AND T. SCHREIBER, *Nonlinear Time Series Analysis*, Cambridge University Press, 2 ed., 2003,
1102 <https://doi.org/10.1017/CBO9780511755798>.
- 1103 [15] M. B. KENNEL, R. BROWN, AND H. D. I. ABARBANEL, *Determining embedding dimension for phase-*
1104 *space reconstruction using a geometrical construction*, *Phys. Rev. A*, 45 (1992), pp. 3403–3411, <https://link.aps.org/doi/10.1103/PhysRevA.45.3403>.
- 1106 [16] A. KIANIMAJD, M. RUANO, P. CARVALHO, J. HENRIQUES, T. ROCHA, S. PAREDES, AND A. RU-
1107 *ANO*, *Comparison of different methods of measuring similarity in physiologic time series*, *IFAC-*
1108 *PapersOnLine*, 50 (2017), pp. 11005–11010, [https://www.sciencedirect.com/science/article/pii/](https://www.sciencedirect.com/science/article/pii/S2405896317333967)
1109 [S2405896317333967](https://www.sciencedirect.com/science/article/pii/S2405896317333967). 20th IFAC World Congress.
- 1110 [17] H. KIM, R. EYKHOLT, AND J. SALAS, *Nonlinear dynamics, delay times, and embedding windows*, *Physica*
1111 *D: Nonlinear Phenomena*, 127 (1999), pp. 48–60, [https://www.sciencedirect.com/science/article/pii/](https://www.sciencedirect.com/science/article/pii/S0167278998002401)
1112 [S0167278998002401](https://www.sciencedirect.com/science/article/pii/S0167278998002401).
- 1113 [18] F. LEJARZA AND M. BALDEA, *Data-driven discovery of the governing equations of dynamical systems via*
1114 *moving horizon optimization*, *Scientific Reports*, 12 (2022), p. 11836.
- 1115 [19] M. MATILLA-GARCÍA, I. MORALES, J. M. RODRÍGUEZ, AND M. RUIZ MARÍN, *Selection of embedding*
1116 *dimension and delay time in phase space reconstruction via symbolic dynamics*, *Entropy*, 23 (2021),
1117 <https://doi.org/10.3390/e23020221>, <https://www.mdpi.com/1099-4300/23/2/221>.
- 1118 [20] K. MISCHAIKOW, M. MROZEK, J. REISS, AND A. SZYMCZAK, *Construction of symbolic dynamics from*
1119 *experimental time series*, *Phys. Rev. Lett.*, 82 (1999), pp. 1144–1147.
- 1120 [21] L. MONIZ, L. PECORA, J. NICHOLS, M. TODD, AND J. R. WAIT, *Dynamical assessment of structural*
1121 *damage using the continuity statistic*, *Structural Health Monitoring*, 3 (2004), pp. 199–212, <https://doi.org/10.1177/1475921704042681>.
- 1123 [22] J. NICHOLS, L. MONIZ, J. NICHOLS, L. PECORA, AND E. COOCH, *Assessing spatial coupling in complex*
1124 *population dynamics using mutual prediction and continuity statistics*, *Theoretical Population Biology*,
1125 67 (2005), pp. 9–21, <https://doi.org/https://doi.org/10.1016/j.tpb.2004.08.004>.
- 1126 [23] L. M. PECORA, T. L. CARROLL, AND J. F. HEAGY, *Statistics for mathematical properties of maps*
1127 *between time series embeddings*, *Phys. Rev. E*, 52 (1995), pp. 3420–3439, [https://doi.org/10.1103/](https://doi.org/10.1103/PhysRevE.52.3420)
1128 [PhysRevE.52.3420](https://doi.org/10.1103/PhysRevE.52.3420).
- 1129 [24] L. M. PECORA, L. MONIZ, J. NICHOLS, AND T. L. CARROLL, *A unified approach to attractor re-*
1130 *construction*, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 17 (2007), p. 013110,
1131 <https://doi.org/10.1063/1.2430294>, <https://arxiv.org/abs/https://doi.org/10.1063/1.2430294>.
- 1132 [25] T. SAUER, J. A. YORKE, AND M. CASDAGLI, *Embedology*, *Journal of Statistical Physics*, 65 (1991),
1133 pp. 579–616.
- 1134 [26] J. D. SKUFCA AND E. M. BOLLT, *Relaxing conjugacy to fit modeling in dynamical systems*, *Phys. Rev.*
1135 *E*, 76 (2007), p. 026220, <https://doi.org/10.1103/PhysRevE.76.026220>.
- 1136 [27] J. D. SKUFCA AND E. M. BOLLT, *A concept of homeomorphic defect for defining mostly conjugate*
1137 *dynamical systems*, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 18 (2008), p. 013118,
1138 <https://doi.org/10.1063/1.2837397>.

- 1139 [28] W. SZLENK, *An Introduction to the Theory of Smooth Dynamical Systems*, Państwowe Wydawnictwo
1140 Naukowe (PWN), 1984.
- 1141 [29] F. TAKENS, *Detecting strange attractors in turbulence*, in *Dynamical Systems and Turbulence*, Warwick
1142 1980, D. Rand and L.-S. Young, eds., Berlin, Heidelberg, 1981, Springer Berlin Heidelberg, pp. 366–
1143 381.
- 1144 [30] D. VISWANATH, *The fractal property of the Lorenz attractor*, *Physica D: Nonlinear Phenomena*, 190
1145 (2004), pp. 115–128, <https://www.sciencedirect.com/science/article/pii/S0167278903004093>.
- 1146 [31] Y. YUAN, X. TANG, W. ZHOU, W. PAN, X. LI, H.-T. ZHANG, H. DING, AND J. GONCALVES, *Data
1147 driven discovery of cyber physical systems*, *Nature Communications*, 10 (2019), p. 4894.
- 1148 [32] J. ZHENG, E. BOLLT, AND J. SKUFCA, *Regularity of commuter functions for homeomorphic defect measure
1149 in dynamical systems model comparison*, *Dynamics of Continuous, Discrete & Impulsive Systems.
1150 Series A: Mathematical Analysis*, 18 (2011).
- 1151 [33] J. ZHENG, J. D. SKUFCA, AND E. M. BOLLT, *Comparing dynamical systems by a graph matching method*,
1152 *Physica D: Nonlinear Phenomena*, 255 (2013), pp. 12–21, [https://doi.org/https://doi.org/10.1016/j.
1153 physd.2013.03.012](https://doi.org/https://doi.org/10.1016/j.physd.2013.03.012).

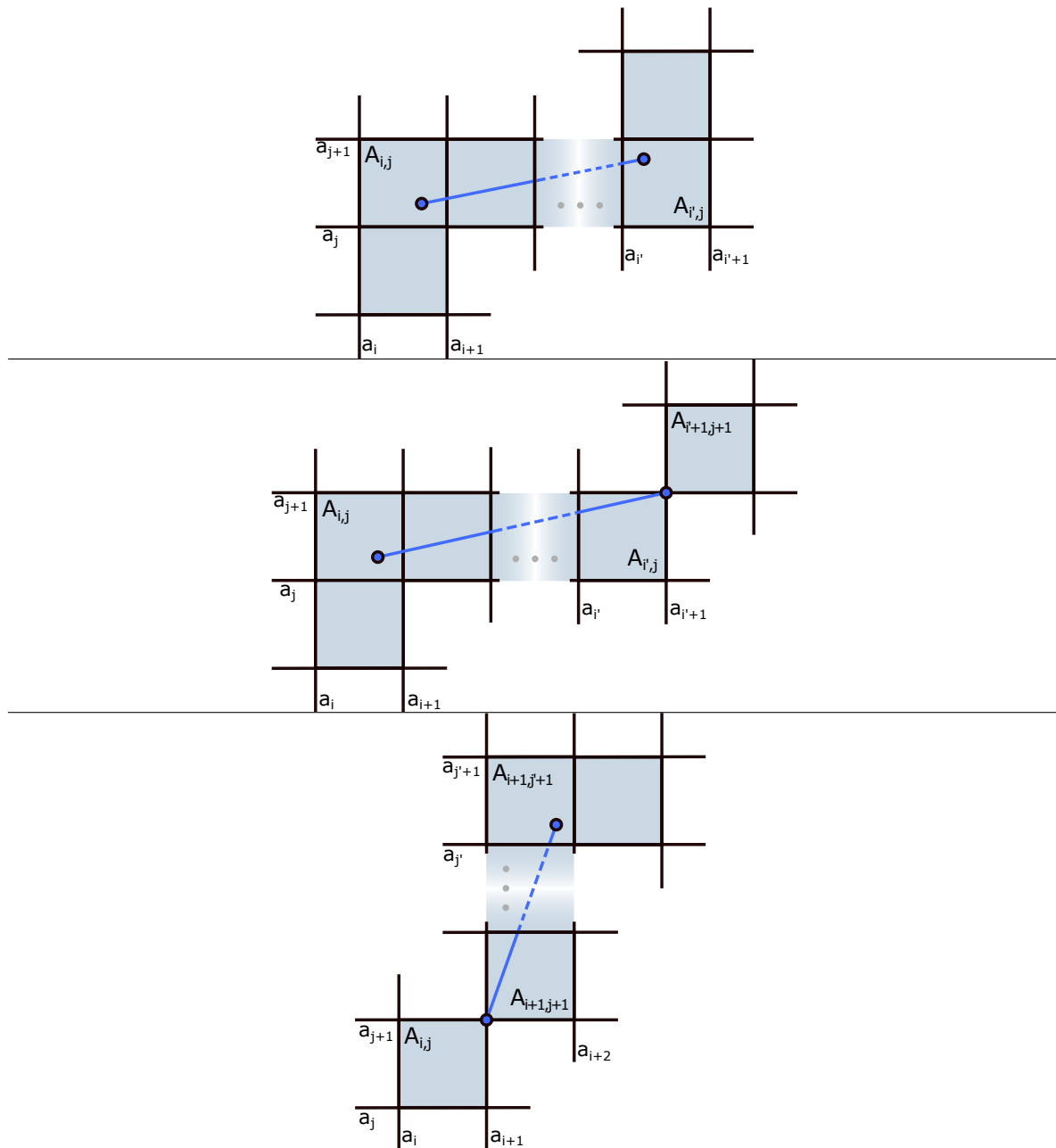


Figure 15. Segments of the piecewise linear homeomorphism being a selector constructed by Algorithm A.1 for a certain cubical homeomorphism h . Three panels corresponds to cases when: point of type UR is followed by point of type UR (top), point of type UR is followed by point of type D (middle), point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