Postprint of: Justyna Signerska-Rynkowska, Michał Lipiński, Paweł Dłotko (2024). Testing Topological Conjugacy of Time Series. SIAM JOURNAL ON APPLIED DYNAMICAL SYSTEMS, vol. 23, iss. 4. <u>https://dx.doi.org/10.1137/23M1594728</u>

© 2024. This manuscript version is made available under the CC BY 4.0 license https://creativecommons.org/licenses/by/4.0/

Testing topological conjugacy of time series * 1 Paweł Dłotko[†], Michał Lipiński^{†‡}, and Justyna Signerska-Rynkowska^{†§} 2 3 4 Abstract. This paper considers a problem of testing, from a finite sample, a topological conjugacy of two trajectories coming from dynamical systems (X, f) and (Y, g). More precisely, given $x_1, \ldots, x_n \subset X$ 56 and $y_1, \ldots, y_n \subset Y$ such that $x_{i+1} = f(x_i)$ and $y_{i+1} = g(y_i)$ as well as $h: X \to Y$, we deliver a 7 number of tests to check if f and g are topologically conjugated via h. The values of the tests are 8 close to zero for systems conjugate by h and large for systems that are not. Convergence of the test 9 values, in case when sample size goes to infinity, is established. We provide a number of numerical 10 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' 11 12embedding theorem. Our methods also apply to the situation when we are given two observables 13 of deterministic processes, of a form of one or higher dimensional time-series. In this case, their 14 similarity can be accessed by comparing the dynamics of their Takens' reconstructions. Finally, we 15include a proof-of-concept study using the presented methods to search for an approximation of the homeomorphism conjugating given systems. 16

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

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

1. Introduction. Understanding sampled dynamics is of primal importance in multiple 20 branches of science where there is a lack of solid theoretical models of the underlying phe-2122 nomena [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 23 the considered case we start with two, potentially different, phase spaces X and Y and a 24 map $h: X \to Y$. Given two sampled trajectories, referred to in this paper by time series, $x_1, \ldots, x_n \subset X$ and $y_1, \ldots, y_n \subset Y$ we assume that they are both generated by a continuous 26maps $f: X \to X$ and $g: Y \to Y^1$ in a way that $x_{i+1} = f(x_i)$ and $y_{i+1} = g(y_i)$. In what 27 follows, we build a number of tests that allow to distinguish trajectories that are conjugated 28 by the given map h from those that are not. It should be noted that the problem of finding 29

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.

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

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

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

Thirdly, those methods can be used to estimate the optimal parameters of the dynamics 43 reconstruction. A basic way to achieve such a reconstruction is via time delay embedding, a 44 technique that depends on parameters including the embedding dimension and the time lag (or 45*delay*). When the parameters of the method are appropriately set up and the assumptions of 46 Takens' Embedding Theorem hold (see [29, 7]), then (generically) a reconstruction is obtained, 47meaning that for *generic* dynamical system and *generic* observable, the delay-coordinate map 48 produces a *conjugacy* (dynamical equivalence) between reconstructed dynamics and the orig-49 inal (unknown) dynamics (cut to the limit set of a given trajectory)². However, without the 50prior knowledge of the underlying dynamics (e.g. dimensions of the attractor), the values of 51those parameters have to be determined experimentally from the data. It is typically achieved 52by implicitly testing for a conjugacy of the time delay embeddings to spaces of constitutive dimensions. Specifically, it is assumed that the optimal dimension of reconstruction d is achieved 5455when there is no conjugacy of the reconstruction in dimension d to the reconstruction in the dimension d', where d' < d, while there is a conjugacy between reconstruction in dimension 56 d and reconstruction in dimension d'', where d < d''. Those conditions can be tested with 57methods presented in this paper. 58

The main contributions of this paper include:

• We propose a generalization of the FNN (*False Nearest Neighbor*) method [13] so that it can be applied to test for topological conjugacy of time series³. Moreover, we present its further modification called KNN method.

• We propose two entirely new methods: ConjTest and ConjTest⁺. Instead of providing an almost binary answer to a question if two sampled dynamical systems are conjugate (which happens for the generalized FNN and the KNN method), their result is a continuous variable that can serve as a scale of similarity of two dynamics. This

59

60

61

62

63

64

65

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.

TESTING TOPOLOGICAL CONJUGACY OF TIME SERIES

property makes the two new methods appropriate for noisy data.

- We present a number of benchmark experiments to test the presented methods. In particular we analyze how different methods are robust for the type of testing (e.g. noise, determinism, alignment of a time series).
- $71 \\ 72$

73

67

• Additionally, in one dimensional setting, we propose a heuristic method for approximating the possible conjugating homeomorphism between two dynamical systems given by time series.

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

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

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

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

115 systems, but they will not allow to prove, in any rigorous sense, the conjugacy between them.

116 **2.** Preliminaries.

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

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

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

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

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

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

132
$$d_{\mathrm{H}}(A,B) = \max\{\sup_{a \in A} d(a,B), \sup_{b \in B} d(b,A)\}$$

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

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

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

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

Embedding Theorem allows to reconstruct both the topology of the original attractor and thedynamics.

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

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

151 $\,$ From the theoretical point of view, the Takens' theorem holds for an arbitrary lag. However

in practice a proper choice of l may strongly affect numerical reconstructions (see [14, Chapter 3]).

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

2.3. Search for an optimal dimension for reconstruction. In practice, the bound in 156Takens' theorem is often not sharp and an embedding dimension less than 2m + 1 is already 157158sufficient to reconstruct the original dynamics (see [3, 4]). Moreover, for time series encountered in practice, the attractor's dimension m is almost always unknown. To discover the 159sufficient dimension of reconstruction, the False Nearest Neighbor (FNN) method [13, 15], a 160 heuristic technique for estimating the optimal dimension using a finite time series, is typi-161 162cally 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 separates some points that were close in d-dimensional embedding. Hence, it tests whether d-164dimensional neighbors are (false) neighbors just because of the tightness of the d-dimensional 165166space. The dimension where the value of the test stabilizes and no more false neighbors can be detected is proclaimed to be the optimal embedding dimension. 167

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

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

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

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

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

190 and

189

191

194

196

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

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

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

195 and

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

in (d+1)-dimensional embedding differs too much, then \mathbf{s}_d^1 and \mathbf{s}_d^2 were d-dimensional neighbors 197 only due to folding of the space. In this case, we will refer to them as "false nearest neighbors". 198 Precisely, the ordered pair $(\mathbf{s}_d^1, \mathbf{s}_d^2)$ of *d*-dimensional points is counted as false nearest neighbor, 199 if the following conditions are satisfied: (I.) the point \mathbf{s}_d^2 is the closest point to \mathbf{s}_d^1 among all 200points in the d-dimensional embedding, (II.) the distance $|\mathbf{s}_d^1 - \mathbf{s}_d^2|$ between the points \mathbf{s}_d^1 and 201 \mathbf{s}_d^2 is less than σ/r , where σ is the standard deviation of d-dimensional points formed from 202 delay-embedding of the time series and (III.) the ratio between the distance $|\mathbf{s}_{d+1}^1 - \mathbf{s}_{d+1}^2|$ of d + 1-dimensional counterparts of these points, \mathbf{s}_{d+1}^1 and \mathbf{s}_{d+1}^2 , and the distance $|\mathbf{s}_d^1 - \mathbf{s}_d^2|$ 203 204is greater than the threshold r. The condition (III.) is motivated by the fact that under 205continuous evolution, even if the original dynamics is chaotic, the position of two close points 206should not deviate too much in the nearest future (we assume that the system is deterministic, 207even if subjected to some noise, which is the main assumption of all the nonlinear analysis 208time series methods). On the other hand, the condition (II.) means that we consider only pairs 209of points which are originally not too far away since applying the condition (III.) to points 210which are already outliers in d dimensions does not make sense. Next, the statistic FNN(r)211counts the relative number of such false nearest neighbors i.e. after normalizing with respect 212213to the number of all the ordered pairs of points which satisfy (I.) and (II.). For discussion and some examples see e.g. [14]. 214

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

9 (3.1)
$$\operatorname{FNN}(\mathcal{A},\mathcal{B};r) := \frac{\sum_{i=1}^{n} \Theta\left(\frac{\mathbf{d}_{Y}(b_{i},\xi(\overline{\kappa}(a_{i},\mathcal{A})))}{\mathbf{d}_{X}(a_{i},\overline{\kappa}(a_{i},\mathcal{A}))} - r\right) \Theta\left(\frac{\sigma}{r} - \mathbf{d}_{X}(a_{i},\overline{\kappa}(a_{i},\mathcal{A}))\right)}{\sum_{i=1}^{n} \Theta\left(\frac{\sigma}{r} - \mathbf{d}_{X}(a_{i},\overline{\kappa}(a_{i},\mathcal{A}))\right)}$$

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

This manuscript is for review purposes only.

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

225 often simply the Euclidean norm |x - y|.

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

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

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

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

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

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

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.

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

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



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

3.3. Conjugacy test. The third method tests the conjugacy of two time series by directly 260checking the commutativity of the diagram (2.1) which is tested in a more direct way compared 261to the methods presented so far. We no longer assume that both time series are of the 262same size, however, the method requires a connecting map $h: X \to Y$, a candidate for a 263(semi)conjugating map. Unlike the map ξ in FNN and KNN method map h may transform 264a point $a_i \in \mathcal{A}$ into a point in Y that doesn't belong to \mathcal{B} . Nevertheless, the points in \mathcal{B} are 265crucial because they carry the information about the dynamics $q: Y \to Y$. Thus, in order to 266follow trajectories of points in Y we introduce $h : \mathcal{A} \to \mathcal{B}$, a discrete approximation of h: 267

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

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

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

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



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

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

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

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

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

286 (3.7)
$$\operatorname{ConjTest}^{+}(\mathcal{A}, \mathcal{B}; k, t, h) := \frac{\sum_{i=1}^{n} \mathrm{d}_{\mathrm{H}}\left((h \circ f^{t})(U_{i}^{k}), g^{t}\left(\tilde{h}^{+}(U_{i}^{k})\right)\right)}{n \operatorname{diam}(\mathcal{B})}$$

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

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

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

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

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 $m \to \infty$. If g is semiconjugate to f with h as a semiconjugacy map, 299 then for every fixed n, t and k

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

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, $h : X \to Y$ 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 $\mathcal{A}_n := \{a_i := f^{i-1}(x_1) \mid i \in \{1, \ldots, n\}\} \subset X$ of length n (where $x_1 \in X$ is some point in X) 308 it holds that

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

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

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

where
$$a_i = f^{i-1}(x_1) \in \mathcal{A}_n$$
.
Thus for $m > M$, we always (independently of the point $a_i \in X$) have

319
$$|h(f^{t}(a_{i})) - g^{t}(\tilde{h}(a_{i}))| = |g^{t}(h(a_{i})) - g^{t}(\tilde{h}(a_{i}))| < \varepsilon |b_{2} - b_{1}|$$

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

321
$$d_{\mathrm{H}}\left((h \circ f^{t})(U_{i}^{k}), \ (g^{t} \circ \tilde{h})(U_{i}^{k})\right) < \varepsilon |b_{2} - b_{1}|,$$

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

$$\frac{\sum_{i=1}^{n} \mathrm{d}_{\mathrm{H}}\left((h \circ f^{t})(U_{i}^{k}), (g^{t} \circ \tilde{h})(U_{i}^{k})\right)}{n \operatorname{diam}(\mathcal{B}_{m})} < \frac{n \varepsilon |b_{2} - b_{1}|}{n \operatorname{diam}(\mathcal{B}_{m})} \leq \varepsilon$$

323

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\lim_{n \to \infty} \lim_{m \to \infty} \frac{\sum_{i=1}^{n} d_{\mathrm{H}} \left((h \circ f^{t})(U_{i}^{k}), (g^{t} \circ \tilde{h}^{+})(U_{i}^{k}) \right)}{n \operatorname{diam}(\mathcal{B}_{m})} \leq \\
362 \quad (3.11) \qquad \qquad \lim_{n \to \infty} \lim_{m \to \infty} \frac{\sum_{i=1}^{n} d_{\mathrm{H}} \left((h \circ f^{t})(U_{i}^{k}), (g^{t} \circ \tilde{h})(U_{i}^{k}) \right)}{n \operatorname{diam}(\mathcal{B}_{m})} + \\
\qquad \qquad \qquad \qquad \lim_{n \to \infty} \lim_{m \to \infty} \frac{\sum_{i=1}^{n} d_{\mathrm{H}} \left((h \circ f^{t})(U_{i}^{k}), (g^{t}(\tilde{h}^{+}(U_{i}^{k}) \setminus \tilde{h}(U_{i}^{k}))) \right)}{n \operatorname{diam}(\mathcal{B}_{m})}.$$

Recall that $U_i^k := \kappa(a_i, k, \mathcal{A}_n)$, $\tilde{h}(a_i) := \kappa(h(a_i), \mathcal{B}_m)$, $\tilde{h}(U_i^k) := \{\tilde{h}(a_j) : a_j \in U_i^k\}$ and $\tilde{h}^+(U_i^k) := \kappa(h(a_i), k_i, \mathcal{B}_m)$, where k_i is the smallest integer k_i such that $\tilde{h}(U_i^k) \subset$ $\kappa(h(a_i), k_i, \mathcal{B}_m)$. Thus in particular, $\tilde{h}(U_i^k) \subset \tilde{h}^+(U_i^k)$. Obviously all these neighborhoods $U_i^k, \tilde{h}(U_i^k)$ and $\tilde{h}^+(U_i^k)$ depend on n and m (since they are taken with respect to \mathcal{A}_n and \mathcal{B}_m). Note that from Theorem 3.2 already follows that the first of the two terms in the sum in (3.11) vanishes. Thus we will only show that the second double limit vanishes as well.

1369 Let $\varepsilon > 0$, $k \in \mathbb{N}$ and $t \in \mathbb{N}$. Since $g^t : Y \to Y$ is a continuous function on a compact 1370 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 1371 $|x - y| < \delta$. Similarly, since X is compact and $h : X \to Y$ is continuous, there exists δ_1 such 1372 that $|h(x) - h(y)| < \frac{\delta}{2}$ whenever $x, y \in X$ such that $|x - y| < \delta_1$.

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

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

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

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

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

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

387

388

389

$$\begin{aligned} |h(f^{t}(a_{j})) - g^{t}(\tilde{b})| \\ &\leq |h(f^{t}(a_{j}) - g^{t}(h(a_{j})))| + |g^{t}(h(a_{j})) - g^{t}(h(a_{i}))| + |g^{t}(h(a_{i})) - g^{t}(\tilde{b})| \end{aligned}$$

386 where

• $|h(f^t(a_i) - g^t(h(a_i)))| = 0$ by semiconjugacy assumption

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

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

$$\frac{\sum_{i=1}^{n} \mathrm{d}_{\mathrm{H}}\left((h \circ f^{t})(U_{i}^{k}), \ g^{t}(\tilde{h}^{+}(U_{i}^{k}) \setminus \tilde{h}(U_{i}^{k}))\right)}{n \operatorname{diam}(\mathcal{B}_{m})} < \frac{\varepsilon}{\operatorname{diam}(\mathcal{B}_{m})}$$

393 if only n > N and m > M.

392

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

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

402 The discussed task is already, to a certain extent, considered in the literature. The paper [23] develops sets of statistics which are intended to characterize, in terms of probabilities 403 and confidence levels, whether time delay embeddings of the two time series are connected by 404 a continuous, injected or differentiable map. The work [23] presents method to assess (gen-405 eralized) synchronization of time series, coupling in complex population dynamics (see [22]) 406or detecting damage in some material structures (see [21]). The statistics proposed in those 407 papers are inspired by notions of continuity, differentiability etc., typically involving quantities 408 like ϵ 's and δ 's. These values need to be fixed and enforce the user to understand how δ 's 409410 scale with ϵ which is typically hard. It seems to be possible to adopt ConjTest's methods to the framework of statistical tests and will be considered in the future. 411

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 $f: X \to X$ with a starting point $x_1 \in X$ is denoted by

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

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

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 $S := \mathbb{R}/\mathbb{Z}$. Denote the operation of taking a decimal part of a number (modulo 1) 423 by $x_{\mathbb{I}} := x - \lfloor x \rfloor$. Then, for a parameter $\phi \in [0, 1)$ we define a rigid rotation on a circle, 424 $f_{[\phi]} : S \to S$, as

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

426 We consider the following metric on S

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

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

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

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

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

441
$$\mathcal{R}_1 = \varrho(f_{[\alpha]}, 0.0, 2000), \qquad \mathcal{R}_2 = \varrho(f_{[\alpha]}, 0.25, 2000),$$

442
$$\mathcal{R}_3 = \varrho(f_{[\alpha+0.02]}, 0.0, 2000), \qquad \mathcal{R}_4 = \varrho(f_{[2\alpha]}, 0.0, 2000)$$

443
$$\mathcal{R}_5 = \varrho(f_{[\alpha],2}, 0.0, 2000), \qquad \mathcal{R}_6 = \mathcal{R}_5 + \operatorname{err}(0.05),$$

444 where $\operatorname{err}(\epsilon)$ 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 done with $h \equiv id_{\mathbb{S}}$. As we already mentioned, there is no conjugacy between \mathcal{R}_1 and \mathcal{R}_3 , 450nor between \mathcal{R}_1 and \mathcal{R}_4 , as the angles of these rotations are different. Thus there is no 451true connecting homeomorphism between \mathcal{R}_1 and \mathcal{R}_3 and between \mathcal{R}_1 and \mathcal{R}_4 . However, in 452453order to apply ConjTests we need to pick some candidate for a matching map between two point clouds and as the first choice one can always start with the identity map, especially for 454 comparing point clouds generated by trajectories starting at the same or close initial points. 455Therefore in this experiment we use $h \equiv id_{\mathbb{S}}$ for comparing \mathcal{R}_1 both with \mathcal{R}_3 and \mathcal{R}_4 . When 456comparing \mathcal{R}_1 versus \mathcal{R}_4 and \mathcal{R}_1 versus \mathcal{R}_5 we use homeomorphism $h_2(x) := x_1^2$. Let us recall 457that for FNN and KNN methods we always use $h(x_i) = y_i$, a connecting homeomorphism 458 based on the indices correspondence. 459

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 $\text{FNN}(\mathcal{R}_i, \mathcal{R}_j; r)$, $\text{KNN}(\mathcal{R}_i, \mathcal{R}_j; k)$,

TESTING TOPOLOGICAL CONJUGACY OF TIME SERIE

test	\mathcal{R}_1 vs. \mathcal{R}_2 (starting point perturbation)	$\mathcal{R}_1 ext{ vs. } \mathcal{R}_3 \ (ext{angle} \ ext{perturbation})$	\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	.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	.586	0.0 0.0	.142 .181
$\begin{array}{c} \text{ConjTest}^+ \\ (\text{k=5, t=5}) \end{array}$.001 .001	.201	.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 $\text{FNN}(\mathcal{R}_j, \mathcal{R}_i; r)$, $\text{KNN}(\mathcal{R}_j, \mathcal{R}_i; k)$, $\text{ConjTest}(\mathcal{R}_j, \mathcal{R}_i; k, t, h)$ and
- 466 ConjTest⁺($\mathcal{R}_j, \mathcal{R}_i; k, t, h$), respectively.

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

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

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

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

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

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

This manuscript is for review purposes only.



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 give values close to 0 when comparing \mathcal{R}_{α} with itself. For different values of parameter 489 r of FNN plots (Figure 3 top left looks almost identically. Even a small perturbation of 490 the rotation parameter causes an immediate jump of FNN value from 0 to 1, making it 491extremely sensitive to any changes in the system. Obviously, unless $\beta = \alpha$, \mathcal{R}_{α} and \mathcal{R}_{β} 492 are not conjugate. However, sometimes it might be convenient to have a somehow smoother 493 relation of the test value to the infinitesimal change of the rotation angle. KNN method seems 494 to behave inconsistently, but we can see that the higher parameter k gets the closer we get 495 to a shape resembling the curve obtained with FNN. On the other hand, ConjTest shows a 496 linear dependence on β parameter. Moreover, different values of ConjTest's parameter t result 497 498in 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 $\beta = 2\alpha$. Formally, we know that f_{α} and f_{α} are not conjugate systems. However, we 500can explain this outcome by analyzing the methods. Let $a_i \in \mathcal{R}_{\alpha}$ and let $\tau \in \mathbb{Z}$ such that 501 $a_i := a_{i+\tau} \in \mathcal{R}_{\alpha}$ be the nearest neighbor of a_i . In particular, $\tau \in \mathbb{Z}$. By (4.1) we get 502503 $\mathbf{d}_{\mathbb{S}}(a_i, a_j) = (\alpha \tau)_{\mathbb{I}}$ or $(-\alpha \tau)_{\mathbb{I}}$. There is an $N \in \mathbb{Z}$ and a $\delta \in [0, 1)$ such that $\alpha \tau = N + \delta$. Since $\mathbf{d}_{\mathbb{S}}(a_i, a_j) \approx 0$, it follows that $\delta_{\mathbb{I}} \approx 0$. To get FNN we also need to know $\mathbf{d}_{\mathbb{S}}(b_i, b_j)$. Let 504 $\beta = z\alpha$. Then, $b_i = (z\alpha i)_{\mathbb{I}}, b_j = (z\alpha i + z\alpha \tau)_{\mathbb{I}}$ and $\mathbf{d}_{\mathbb{S}}(b_i, b_j) = (z\alpha \tau)_{\mathbb{I}}$ or $(-z\alpha \tau)_{\mathbb{I}}$. Thus, 505 $z\alpha\tau = zN + z\delta$. We assume that $z\delta \in [0,1)$, because $\delta_{\mathbb{I}} \approx 0$ and z is not very large. Again, 506there exists an $M \in \mathbb{Z}$ and $\epsilon \in [0,1)$ such that $zN = M + \epsilon$. Now, if $zN \in \mathbb{Z}$, then $\epsilon = 0$, 507 $\mathbf{d}_{\mathbb{S}}(b_i, b_j) = (z\delta)_{\mathbb{I}} = z \, \mathbf{d}_{\mathbb{S}}(a_i, a_j)$ (last equality given by $\delta_{\mathbb{I}} \approx 0$) and $\frac{\mathbf{d}_{\mathbb{S}}(b_i, b_j)}{\mathbf{d}_{\mathbb{S}}(a_i, a_j)} = z$. If $zN \notin \mathbb{Z}$ 508then $\epsilon \neq 0$ and $\frac{\mathbf{d}_{\mathbb{S}}(b_i, b_j)}{\mathbf{d}_{\mathbb{S}}(a_i, a_j)} = \frac{\geq 0}{\sim 0}$. Hence, the fraction gives a large number and the numerator 509 of FNN will count most of the points, unless $zN \in \mathbb{Z}$ which is always satisfied when $z \in \mathbb{Z}$. 510Moreover, for the irrational rotation τ might be large. In our experiments we usually get 511 $|\tau| > 1000$. Thus, N is large and ϵ is basically a random number. In the case of KNN there is 512a chance that at least for some of the k-nearest neighbors $zN \in \mathbb{Z}$. Hence, the more rugged 513shape of the curve. 514

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

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

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 time series:

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

526 where $\operatorname{err}(\epsilon)$ 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 $h(x) = x^2_{1}$.

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

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

536
$$f_{[\phi_1,\phi_2]}(x^{(1)},x^{(2)}) = ((x^{(1)}+\phi_1)_1,(x^{(2)}+\phi_2)_1),$$

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

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

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

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

544 **4.2.1. Experiment 2A.**

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

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

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

546

TESTING TOPOLOGICAL CONJUGACY OF TIME SERIES

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	0.0 1.0	0.0 1.0
KNN $(k=5)$.042 .617	.275 .855	.514 .690	.041 .938	.041 .938
$\begin{array}{c} \text{ConjTest} \\ (k=5, t=5) \end{array}$.001 .270	.149 .148	.142 .270	.451 .322	.318 .322
$\begin{array}{c} \text{ConjTest}^+ \\ (\text{k=5, t=5}) \end{array}$.018 .272	.154 .158	.143 .271	.458 .325	.319 .324
Table 2					

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

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

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

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

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

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

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

4.3.1. Experiment 3A.

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

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

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

This manuscript is for review purposes only.

572

P. DŁOTKO, M. LIPIŃSKI, AND J. SIGNERSKA-RYNKOWSKA

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 .999
KNN (k=5)	0.0 0.0	.825 .828	.831 .832	.835 .833
$\begin{array}{c} \text{ConjTest} \\ (k=5, t=5) \end{array}$	0.0 0.0	.017	.099 .059	.099 .059
$\begin{array}{c} ConjTest^+ \\ (k=5, t=5) \end{array}$	0.0 .001	.027	.104 .065	.104 .064

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

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

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

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

Setup. We generated a collection of time series:

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

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

604



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

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

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

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.

4.3.3. Experiment 3C. As observed in the previous experiment, a change of the parameter *l* in the logistic equation may significantly change the dynamical nature of the system. In this experiment we use the ConjTest to grasp the types of dynamics as a function of *l* parameter. *Setup.* First, we generated the following collection of time series:

8 (4.6)
$$\left\{ \mathcal{B}(l,p) := \varrho(f_l, f_l^{500}(p), 2000) \mid l \in \{3.4, 3.405, 3.41, \dots, 4.0\}, \ p \in \{0.11, 0.31\} \right\}.$$



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 allowing the trajectory to settle down on an attractor. We compared every two time series 630 $\mathcal{B}(l,p)$ and $\mathcal{B}(l',p')$ with the formula We assign a similarity for each pair of time series $\mathcal{B}(l,p)$ 631 and $\mathcal{B}(l', p')$ via 632

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

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

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

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

548 (4.8)
$$\begin{cases} \dot{x} = \sigma(y - x), \\ \dot{y} = x(\rho - z) - y, \\ \dot{z} = xy - \beta z, \end{cases}$$

649 which induces a continuous dynamical system $\varphi : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$. 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:

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

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

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

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

TESTING TOPOLOGICAL CONJUGACY OF TIME SERIES

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.

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 given by:

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

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

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

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

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

In the experiments below we use the maximum metric.

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.

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

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

This manuscript is for review purposes only.

P. DŁOTKO, M. LIPIŃSKI, AND J. SIGNERSKA-RYNKOWSKA

test	\mathcal{L}_1 vs. $\mathcal{P}^1_{x,1}$	\mathcal{L}_1 vs. $\mathcal{P}^1_{x,2}$	\mathcal{L}_1 vs. $\mathcal{P}^1_{x,3}$	\mathcal{L}_1 vs. $\mathcal{P}_{z,1}^1$	\mathcal{L}_1 vs. $\mathcal{P}^1_{z,3}$
FNN (r=3)	0.0 1.0	0.0 .362	.05 .196	0.0 1.0	.111 .541
KNN (k=5)	.019 .465	.003 .036	.003 .007	.024 .743	.002 .519
$\begin{array}{c} ConjTest \\ (k=5, t=10) \end{array}$	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0
$\begin{array}{c} \text{ConjTest}^+ \\ \text{(k=5, t=10)} \end{array}$.330 .401	.030	.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.

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

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

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

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

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

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

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.

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

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

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

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:

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

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

755

752

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

Results. The top plot of Figure 10 presents the results of comparing \mathcal{L}_1 to the time series 756 \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. $\mathcal{P}_{x,1}^i$, 757 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. 758 \mathcal{P}_{x4}^{i} . There are three curves of every color, each one corresponds to a different starting point 759 $p_i, i \in \{2, 3, 4\}$. The bottom part shows results for comparison of \mathcal{L}_1 to $\mathcal{P}_{y,d}^i$ (we embed 760761 the y-coordinate time series instead of x-coordinate). The color of the curves is interpreted analogously. Black curves on both plots are the same and correspond to the comparison of 762 \mathcal{L}_1 with \mathcal{L}_j for $j \in \{2, 3, 4\}$. 763



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

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

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

than the x-coordinate. In Figure 8 we can see that y-embedding is more twisted in the center of the attractor. Hence, generally values are higher, and more temporal information is needed

782 (reflected by higher embedding dimension) to compensate.

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

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

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

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

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

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

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

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

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

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

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



Lorenz system: dependency of *ConjTest*⁺ on the number of time-steps and embedding dimension of the *x*-coordinate.

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

ent than fixed-point iteration scheme. This method is based on symbolic dynamics approach 826 which, however, requires existence and finding a general partition for systems being compared. 827 Although the approach of finding a (semi-)conjugacy to a symbolic shift space through gen-828 erating or Markov partition seems very natural from the point of view of dynamical systems 829 830 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 831 explosion of computational complexity as the number of symbols increases. Later work [33] 832 employs a method of graph matching between the graphs representing the underlying sym-833 bolic dynamics or, alternatively, between the graphs approximating the action of the systems 834 on some eligible partition. Interestingly, the authors show that the permutation matrices that 835 relate the adjacency matrices of the merging graphs coincide with the solution of Monge's 836 mass transport problem. 837

The above earlier works contain valuable ideas on finding to-be-conjugacies or commuter 838 functions and the defect measures of the arising commuters might serve as a quantification of 839 the dynamical similarity between two given systems. In turn, our proposed tools, ConjTest 840 and ConjTest⁺ can be applied, among others, to explicitly assess the quality of the matching 841 between the two systems through the commuting functions obtained by the above mentioned 842 methods and these matching functions can be candidates for testing (semi-)conjugacy of given 843 systems by ConjTests. Note also that, contrary to the previous works, ConjTest methods can 844 be applied directly to the time-series since we work on the point clouds and do not need a 845 priori the formulas for systems which generated them - these are only used as benchmark 846 tests. 847

However, as our contribution and small step forward towards effective algorithms of finding 848 conjugating maps, in this section we present a proof-of-concept gradient-descent algorithm, 849 utilizing the ConjTest as a cost function, to approximate such a connecting homeomorphism. 850 851 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). 852 Instead of finding an analytical formula approximating the connecting homeomorphism our 853 854 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. 855

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

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

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

862

formula for the size of H is given in A). For instance, the number of cubical homeomorphisms for m = 21 is about $2.6 \cdot 10^{14}$. Thus, it is hopeless to find the optimal approximation of the connecting homeomorphism by a brute examination of all elements of H. Instead, we propose

an algorithm based on the gradient descent strategy using ConjTest as a cost function.

Algorithm 5.1 ApproximateH

```
Input: \mathcal{A}, \mathcal{B} – time series on an interval, h_0 \in 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: q \leftarrow initialize queue of size p with null's
 3: best_h \leftarrow h
 4: for t = 0 to nsteps do
       c \leftarrow \{h' \in h \mid h' \in nbhd(h) \text{ and } diff(h, h') \not\subset q\}
 5:
       if \#c = 0 then
 6:
          break
 7:
 8:
       else
          h \leftarrow h' \in c with a minimal value of score(h', \mathcal{A}, \mathcal{B})
 9:
          if score(h, A, B) < score(best_h, A, B) then
10:
             \texttt{best\_h} \gets \texttt{h}
11:
          end if
12:
          q.append(diff(h, h')) {append the unique element differentiating h and h'}
13:
14:
          q.pop()
       end if
15:
16: end for
17: return best_h
```

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

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

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

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

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

This manuscript is for review purposes only.

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

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

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

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

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

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

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

34

TESTING TOPOLOGICAL CONJUGACY OF TIME SERIES



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 score(h, A, B) for every step of the experiment approximating the connecting homeomorphism between the logistic and the tent map.

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

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

TESTING TOPOLOGICAL CONJUGACY OF TIME SERIES

Method	FNN	KNN	ConjTest	$\operatorname{ConjTest}^+$	
			• an exact correspondence between the		
Requirements	identical matching between indexes of the elements (in particular the se- ries must be of the same length)		 two time series is not needed allow examining arbitrary, even complicated potential relations tween the series can be used for comparison of tim ries of different length require defining the po (semi)conjugacy h at least le i.e. giving the corresponding rel between indexes of the elements of two series 		
Parameters	only one param- eter: r (but one should examine large interval of r values)	only one parame- ter: k (but is rec- ommend to check a couple of differ- ent k values)	involve two pa	rameters: k and t	
Robustness	 less robust to noise and perturba- tion than ConjTest methods give nearly a binary output KNN seems to admit robustness with respect to the measurement noise 		 more robust to tion the returned an uously on the 2 and noise com response given 	noise and perturba- swer depends contin- level of perturbation pared to the binary by FNN or KNN	
Recurrent properties	takes into ac- count only the one closest return of a series (tra- jectory) to each neighborhood		at k -closest returns		
Further properties			more likely to give false posi- tive answer than ConjTest ⁺	more computation- ally demanding than ConjTest but usually more reliable	

Table 6

Comparison of the properties of discussed conjugacy measures.

a good embedding dimension. Further comparison between ConjTest and ConjTest⁺ reveals 955 that ConjTest⁺ is more computationally demanding than ConjTest, but also more reliable. 956 Indeed, in our examples with rotations on the circle and torus and with the logistic map, both 957 these tests gave nearly identical results, but the examples with the Lorenz system show that 958 ConjTest is more likely to give a false positive answer. This is due to the fact that ConjTest 959 works well if the map h connecting time series \mathcal{A} and \mathcal{B} is a reasonably good approximation 960 of the true conjugating homeomorphism, but in case of embeddings and naive, point-wise 961 connection map, as in some of our examples with Lorenz system, the Hausdorff distance in 962963 formula (3.5) might vanish resulting in false positive.

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

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

973 Appendix A. Cubical homeomorphisms.

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

976 At first, observe that elements of H have the following straightforward observations.

977 Proposition A.1. Let $[h] \in H$. Then, $A_{1,1}, A_{n,n} \in [h]$.

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

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

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

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

984 (2) int $A_{i,j+1} \cap \pi(h) \neq \emptyset$ and int $A_{i+1,j} \cap \pi(h) = \emptyset$, when $h^{-1}(a_{i+1}) \in (a_i, a_{i+1})$,

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



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

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

989 Proposition A.3. Let $h \in H$. Then

990 (i) for every
$$i \in \{1, ..., n\}$$
 set $\bigcup \{A_{i,j} \in h \mid j \in \{0, 1, ..., n\}\}$ is nonempty and connected,

(ii) for every $j \in \{1, \ldots, n\}$ set $\bigcup \{A_{i,j} \in h \mid i \in \{0, 1, \ldots, n\}\}$ is nonempty and connected.

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

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

This manuscript is for review purposes only.

991

992

TESTING TOPOLOGICAL CONJUGACY OF TIME SERIES

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

1001 (A.1)
$$0 \le n_R, n_U, n_D \le n-1, n_R + n_U + 2n_D = 2(n-1)$$
 and $n_R = n_U$

Actually, any vector V of symbols $\{R, U, D\}$ satisfying the above conditions (A.1) corre-1002 sponds to a cubical homeomorphism. We show it by constructing a piecewise-linear homeo-1003morphism h such that [h] = h for any h represented by V. We refer to the constructed h as 1004 1005a selector of h. In particular, Algorithm A.1 produces a sequence of points corresponding to points of non-differentiability of the homeomorphism. We have five type of points, the 1006 starting point (0,0) (type B), the ending point (1,1) (type E), and points corresponding to 1007subsequences UR (type UR), RU (type RU) and D (type D). Proposition A.4 shows that 10081009 the map generated by the algorithm is an actual homeomorphism.

Algorithm A.1 FindSelector

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

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

1014 (i) $x_k < x_{k+1}$ and $y_k < y_{k+1}$,

(*ii*) if $(x_k, y_k) \in L$ then $(x_k, y_k) + (1 - t)(x_k, y_{k+1}) \in []$ h for $t \in [0, 1]$. 1015

Proof. First, note that if (x_k, y_k) is of type UR or RU then we have $(x_k, y_k) \in \text{int } A_{i,j}$. 1016 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 $x_k, y_k \in$ 1017 $\{a_2, a_3, \ldots, a_n\}.$ 1018

Sequence L always begins $(x_0, y_0) = (0, 0)$ and ends with $(x_K, y_K) = (1, 1)$. By Proposition 1019 1020 A.1 we have $A_{0,0}, A_{n,n} \in h$. The above observations shows that for every (x_k, y_k) with $k \in$ $\{1, 2, \ldots, K-1\}$ we have $0 < x_k, y_k < 1$. Thus, two first and two last points of L satisfies 1021 (i). If (x_1, y_1) is of type UR then it follows that V begins with a sequence of U's. We have 1022 $(x_1, y_1) \in \operatorname{int} A_{0,j}$ for some j > 0. By Proposition A.3(ii) all $A_{0,j'} \in h$ for $0 \leq j' \leq j$. Hence, 1023 the interval spanned by (x_0, y_0) and (x_1, y_1) is contained in $\bigcup h$ proving (ii). The cases when 1024 (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 10251026 similar argument.

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

1030
$$(x,y) = (p a_i + (1-p) a_{i+1}, (1-p) a_j + p a_{j+1}) \operatorname{int} \mathbf{A}_{i,j},$$

1031
$$(x', y') = ((1-p)a_{i'} + pa_{i'+1}, pa_j + (1-p)a_{j+1}) \operatorname{int} \mathbf{A}_{i',j},$$

where i < i'. Thus, 1032

1033

$$y' - y = pa_j + (1 - p)a_{j+1} - ((1 - p)a_j + pa_{j+1})$$
1034

$$= 2pa_j + a_{j+1} - a_j > a_{j+1} - a_j > 0.$$

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

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

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

1042
$$(x,y) = (p a_i + (1-p) a_{i+1}, (1-p) a_j + p a_{j+1}) \in \operatorname{int} \mathbf{A}_{i,j},$$

1043
$$(x',y') = (a_{i'+1},a_{j+1}) = \mathbf{A}_{i',j} \cap \mathbf{A}_{i'+1,j+1},$$

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

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

1051
$$(x',y') = (p a_{i+1} + (1-p) a_{i+2}, (1-p) a_{j'} + p a_{j'+1}) \in \operatorname{int} A_{i+1,j'+1},$$

where j < j'. It follows that x < x' and y < y' proving (i) for this case. By Proposition A.3(i) 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
$$(x,y) = (a_{i+1}, a_{j+1}) = \mathbf{A}_{i,j} \cap \mathbf{A}_{i+1,j+1},$$

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

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

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

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

1	\sim	\sim	0
T	U	0	8

1072 1073

1074

1075

1076

1077

1078

1079 1080

1081

1082

1083

1084

REFERENCES

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

[1]	 T	2011, pp. 205–242, https://doi.org/10.1007/978-1-4419-6870-8-6.
[8]	т.	BUZUG AND G. PFISTER, Comparison of algorithms calculating optimal embedding parameters for delay time coordinates, Physica D: Nonlinear Phenomena, 58 (1992), pp. 127–137, https://www.sciencedirect.com/science/article/pii/016727899290104U
[9]	Κ.	CHAMPION, B. LUSCH, J. N. KUTZ, AND S. L. BRUNTON, <i>Data-driven discovery of coordinates and</i>
)]	D.	<i>governing equations</i> , Proceedings of the National Academy of Sciences, 116 (2019), pp. 22445–22451. L. DEANGELIS AND S. YUREK, <i>Equation-free modeling unravels the behavior of complex ecological systems.</i> , Proc Natl Acad Sci U S A, 112 (2015), pp. 3856–3857.
.1]	V.	DESHMUKH, E. BRADLEY, J. GARLAND, AND J. D. MEISS, Using curvature to select the time lag for delay reconstruction, Chaos: An Interdisciplinary Journal of Nonlinear Science, 30 (2020), p. 063143, https://doi.org/10.1063/5.0005890
12]	А.	M. FRASER AND H. L. SWINNEY, Independent coordinates for strange attractors from mutual informa- tion, Phys. Rev. A, 33 (1986), pp. 1134–1140, https://link.aps.org/doi/10.1103/PhysRevA.33.1134.
[3] [4]	R. H.	HEGGER AND H. KANTZ, Improved false nearest neighbor method to detect determinism in time series data, Phys. Rev. E, 60 (1999), pp. 4970–4973, https://link.aps.org/doi/10.1103/PhysRevE.60.4970. KANTZ AND T. SCHREIBER, Nonlinear Time Series Analysis, Cambridge University Press, 2 ed., 2003.
15]	м	https://doi.org/10.1017/CBO9780511755798.
10]	111.	space reconstruction using a geometrical construction, Phys. Rev. A, 45 (1992), pp. 3403–3411, https: //link.aps.org/doi/10.1103/PhysRevA.45.3403.
[16]	Α.	KIANIMAJD, M. RUANO, P. CARVALHO, J. HENRIQUES, T. ROCHA, S. PAREDES, AND A. RU- ANO, Comparison of different methods of measuring similarity in physiologic time series, IFAC- PapersOnLine, 50 (2017), pp. 11005–11010, https://www.sciencedirect.com/science/article/pii/ S2405296317333067, 20th IFAC World Congress
17]	H.	 S2403890517553907. 20th IFAC world Congress. KIM, R. EYKHOLT, AND J. SALAS, Nonlinear dynamics, delay times, and embedding windows, Physica D: Nonlinear Phenomena, 127 (1999), pp. 48–60, https://www.sciencedirect.com/science/article/pii/ S0167278998002401
[18]	F.	LEJARZA AND M. BALDEA, Data-driven discovery of the governing equations of dynamical systems via moving horizon antimization. Scientific Reports 12 (2022), p. 11836
[19]	М.	MATILLA-GARCÍA, I. MORALES, J. M. RODRÍGUEZ, AND M. RUIZ MARÍN, Selection of embedding dimension and delay time in phase space reconstruction via symbolic dynamics, Entropy, 23 (2021), https://doi.org/10.3390/e23020221, https://www.mdpi.com/1099-4300/23/2/221.
[20]	K.	MISCHAIKOW, M. MROZEK, J. REISS, AND A. SZYMCZAK, Construction of symbolic dynamics from experimental time series, Phys. Rev. Lett., 82 (1999), pp. 1144–1147.
[21]	L.	MONIZ, L. PECORA, J. NICHOLS, M. TODD, AND J. R. WAIT, <i>Dynamical assessment of structural damage using the continuity statistic</i> , Structural Health Monitoring, 3 (2004), pp. 199–212, https://doi.org/10.1177/1475921704042681.
[22]	J.	NICHOLS, L. MONIZ, J. NICHOLS, L. PECORA, AND E. COOCH, Assessing spatial coupling in complex population dynamics using mutual prediction and continuity statistics, Theoretical Population Biology, 67 (2005), pp. 9–21, https://doi.org/https://doi.org/10.1016/j.tpb.2004.08.004.
[23]	L.	M. PECORA, T. L. CARROLL, AND J. F. HEAGY, Statistics for mathematical properties of maps between time series embeddings, Phys. Rev. E, 52 (1995), pp. 3420–3439, https://doi.org/10.1103/ PhysRevE.52.3420.
[24]	L.	M. PECORA, L. MONIZ, J. NICHOLS, AND T. L. CARROLL, A unified approach to attractor re- construction, Chaos: An Interdisciplinary Journal of Nonlinear Science, 17 (2007), p. 013110, https://doi.org/10.1063/1.2430294, https://arxiv.org/abs/https://doi.org/10.1063/1.2430294.
[25]	Т.	SAUER, J. A. YORKE, AND M. CASDAGLI, <i>Embedology</i> , Journal of Statistical Physics, 65 (1991), pp. 579–616.
26]	J.	D. SKUFCA AND E. M. BOLLT, Relaxing conjugacy to fit modeling in dynamical systems, Phys. Rev. E, 76 (2007), p. 026220, https://doi.org/10.1103/PhysRevE.76.026220.
[27]	J.	D. SKUFCA AND E. M. BOLLT, A concept of homeomorphic defect for defining mostly conjugate dynamical systems, Chaos: An Interdisciplinary Journal of Nonlinear Science, 18 (2008), p. 013118, https://doi.org/10.1063/1.2837307

P. DŁOTKO, M. LIPIŃSKI, AND J. SIGNERSKA-RYNKOWSKA

- 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, 1901145(2004), pp. 115–128, https://www.sciencedirect.com/science/article/pii/S0167278903004093.
- [31] Y. YUAN, X. TANG, W. ZHOU, W. PAN, X. LI, H.-T. ZHANG, H. DING, AND J. GONCALVES, Data driven discovery of cyber physical systems, Nature Communications, 10 (2019), p. 4894.
- [32] J. ZHENG, E. BOLLT, AND J. SKUFCA, Regularity of commuter functions for homeomorphic defect measure in dynamical systems model comparison, Dynamics of Continuous, Discrete & Impulsive Systems.
 Series A: Mathematical Analysis, 18 (2011).
- [33] J. ZHENG, J. D. SKUFCA, AND E. M. BOLLT, Comparing dynamical systems by a graph matching method, Physica D: Nonlinear Phenomena, 255 (2013), pp. 12–21, https://doi.org/https://doi.org/10.1016/j.
 physd.2013.03.012.



Figure 15. Segments of the piecewise linear homeomorphism being a selector constructed by Algorithm A.1 for a certain cubical homeomorphism \mathbf{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).