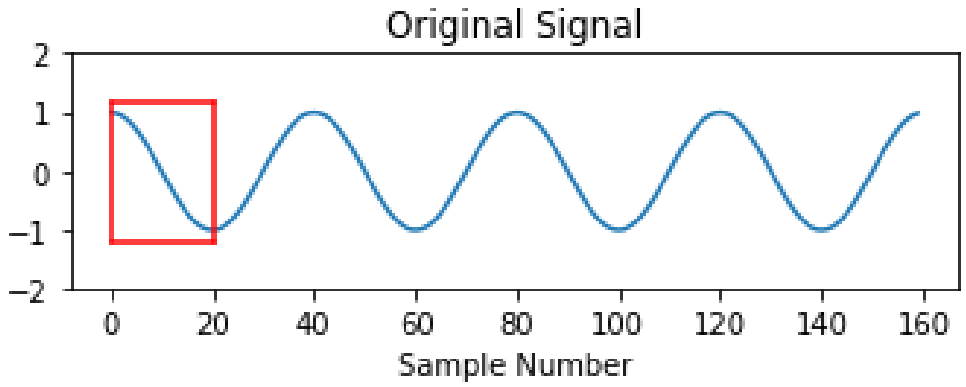
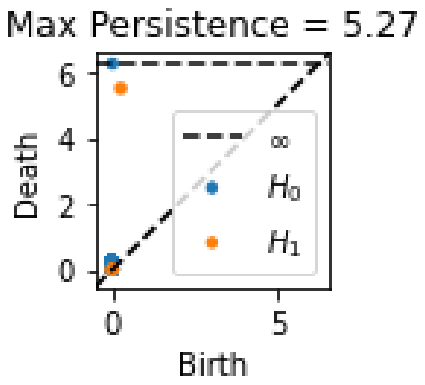
EMALCA 2021 Peru – Análisis Topológico de Datos

Clase 3: Topological Time Series Analysis





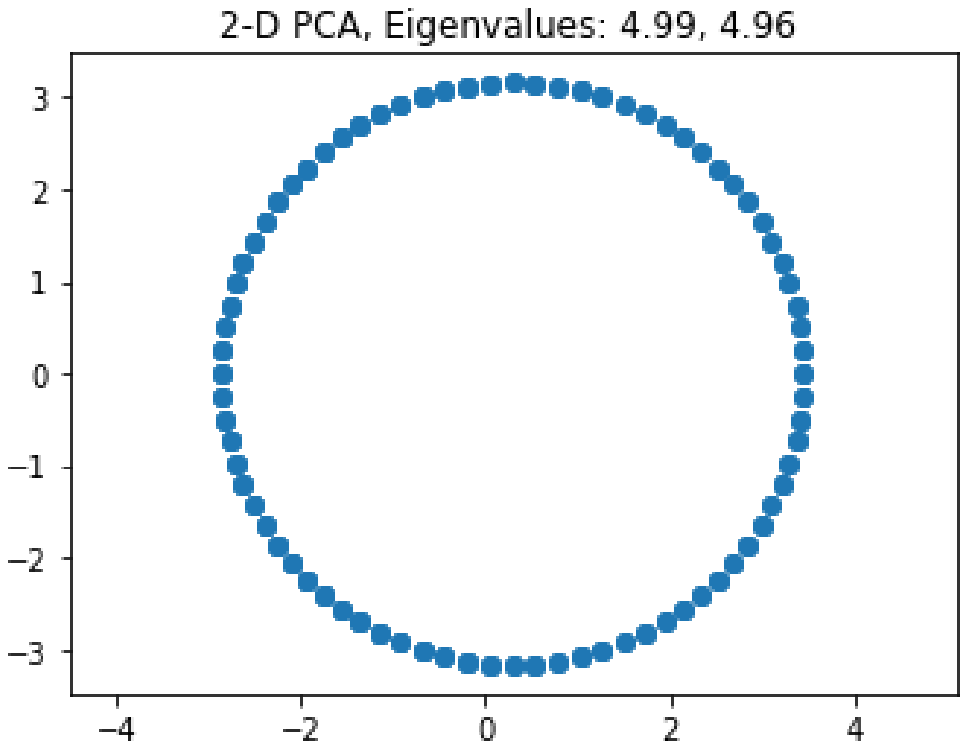
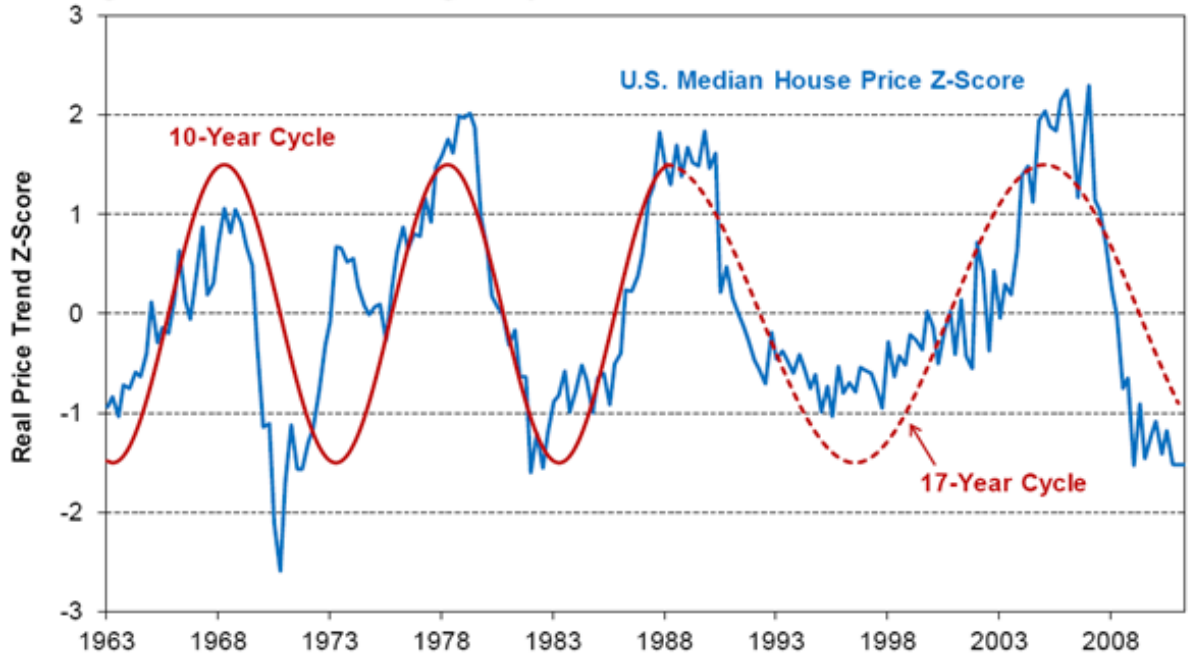
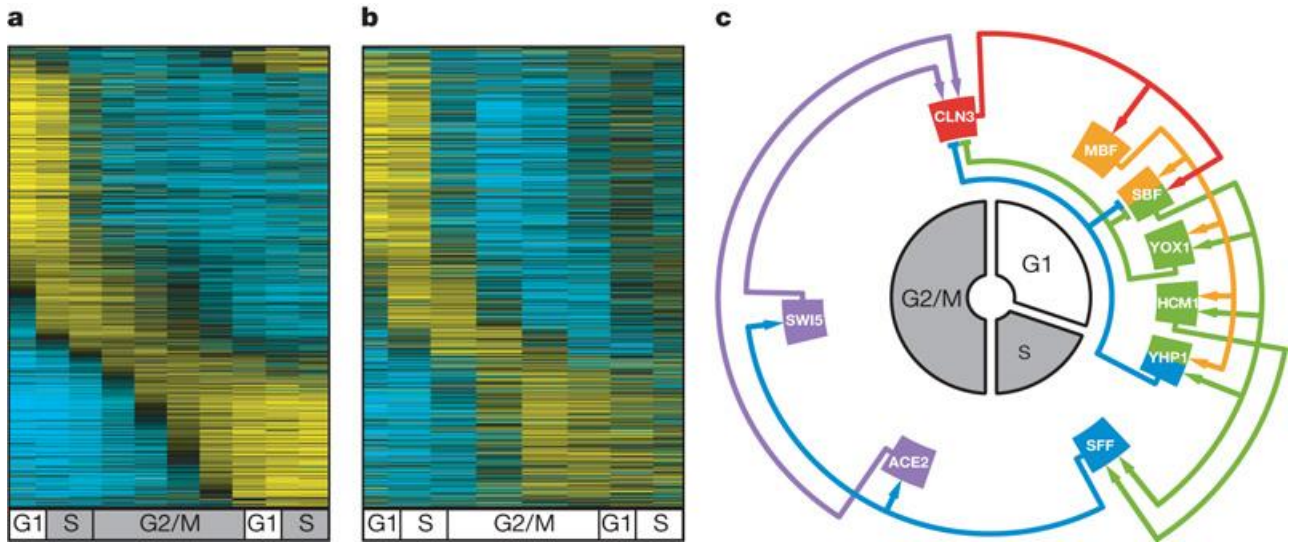


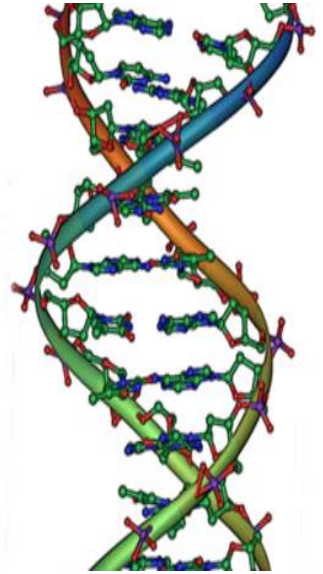
Exhibit 1

U.S. Housing Follows a More or Less Regular Cycle



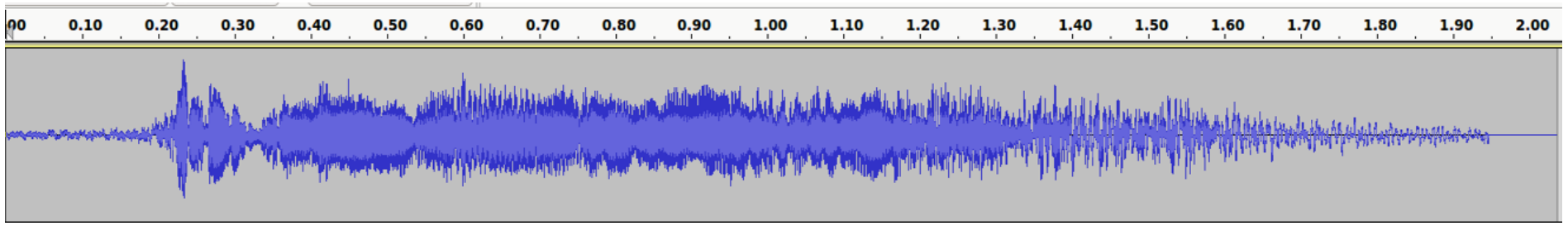
Source: Bureau of the Census, GMO As of 6/30/11





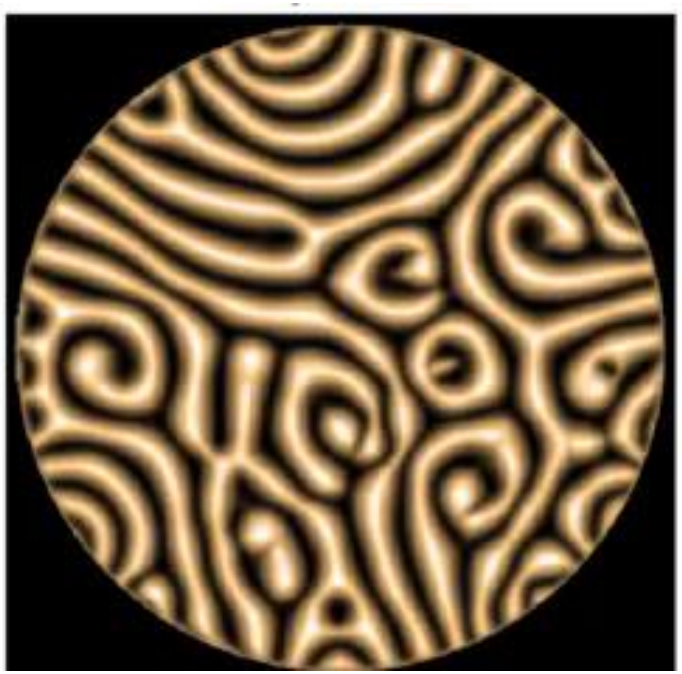
Global control of cell-cycle transcription by coupled SDK and network oscillators, D. Orlando et. al., Nature, 2008

Quantifying recurrence in time varying systems

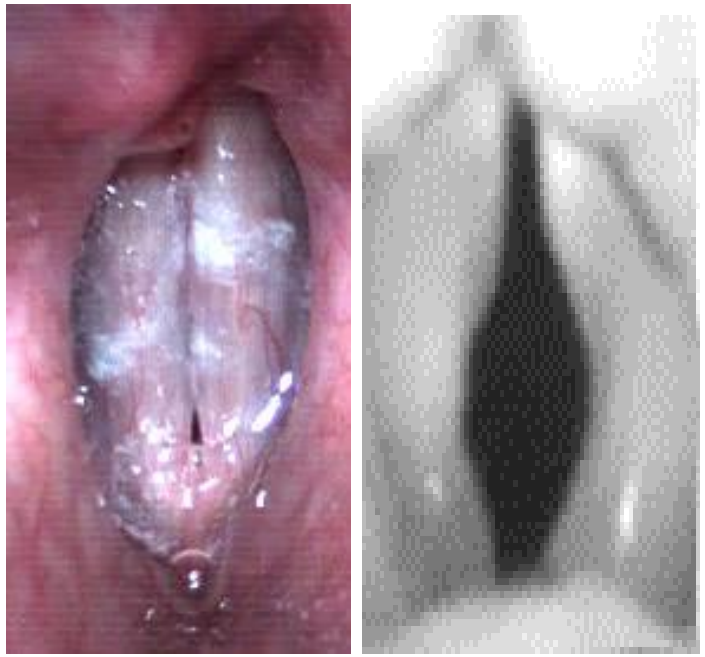


Recording

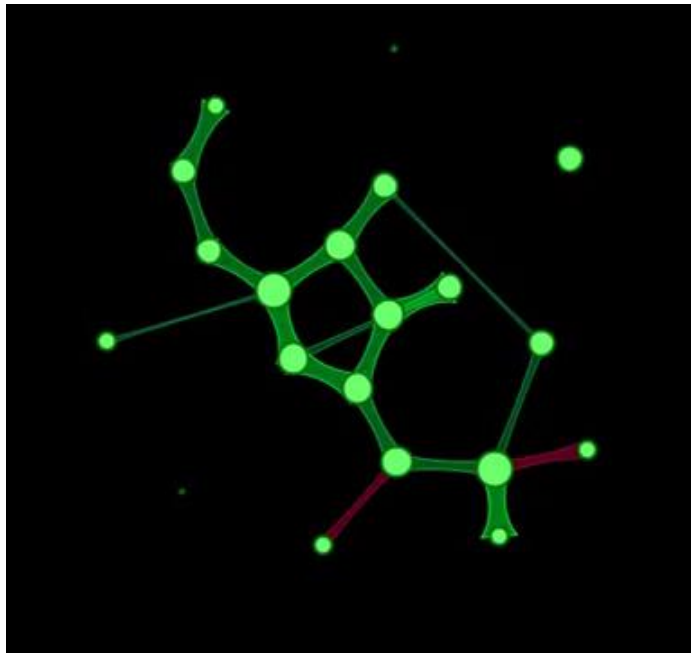
Evolving solution (RB conv.)

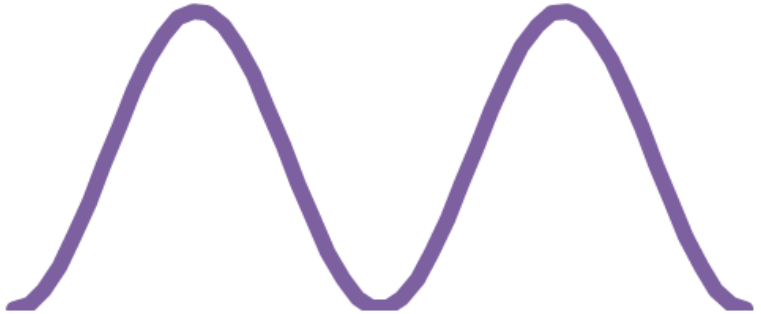


Video data

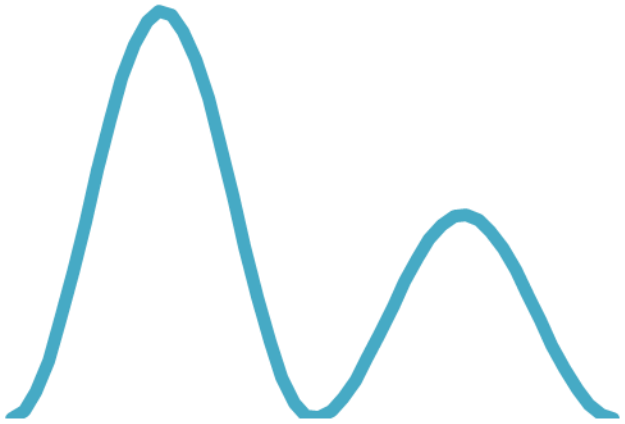


Dynamic network

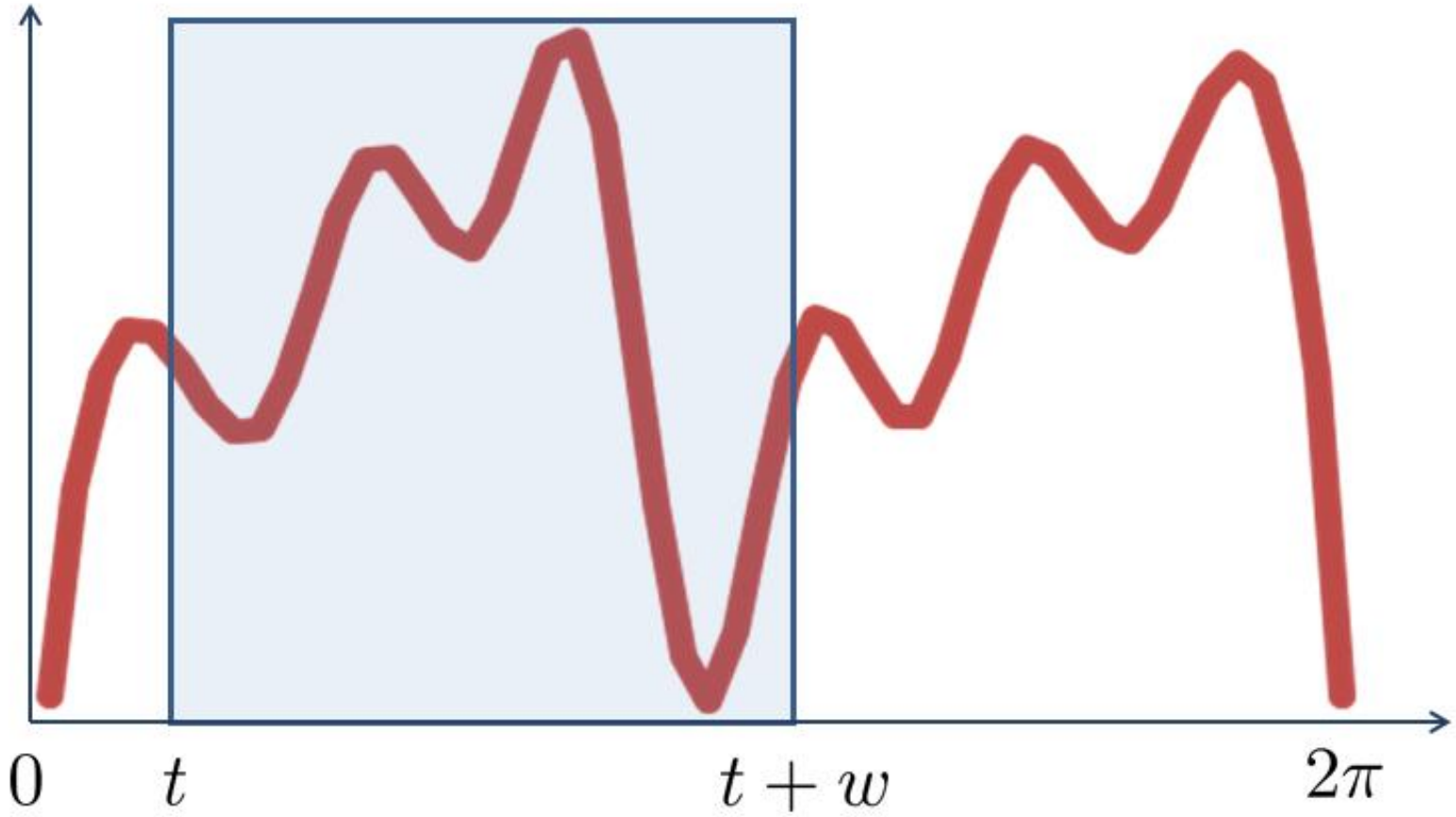




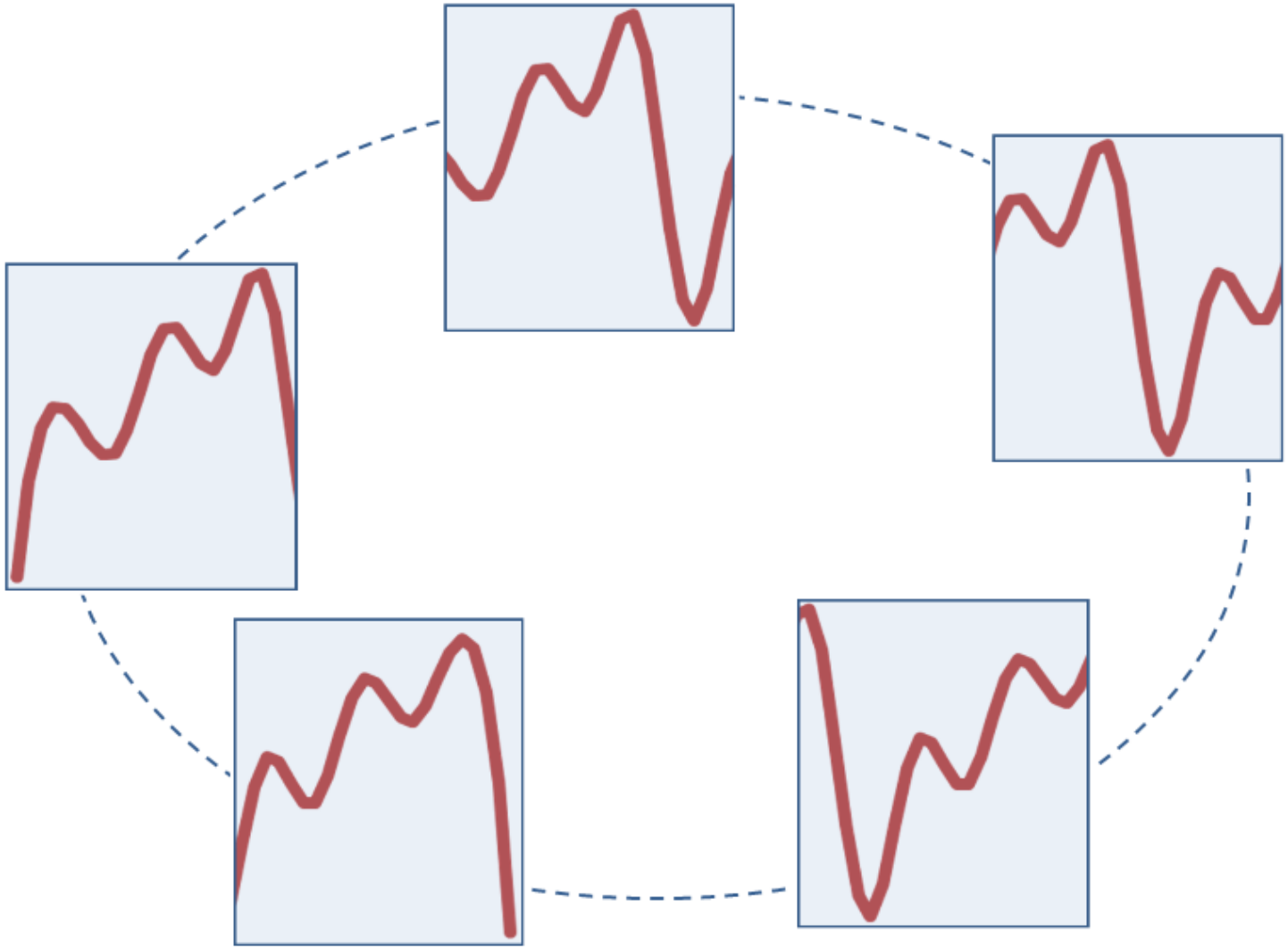
What is recurrence, and
how do we quantify it?



Sliding Windows

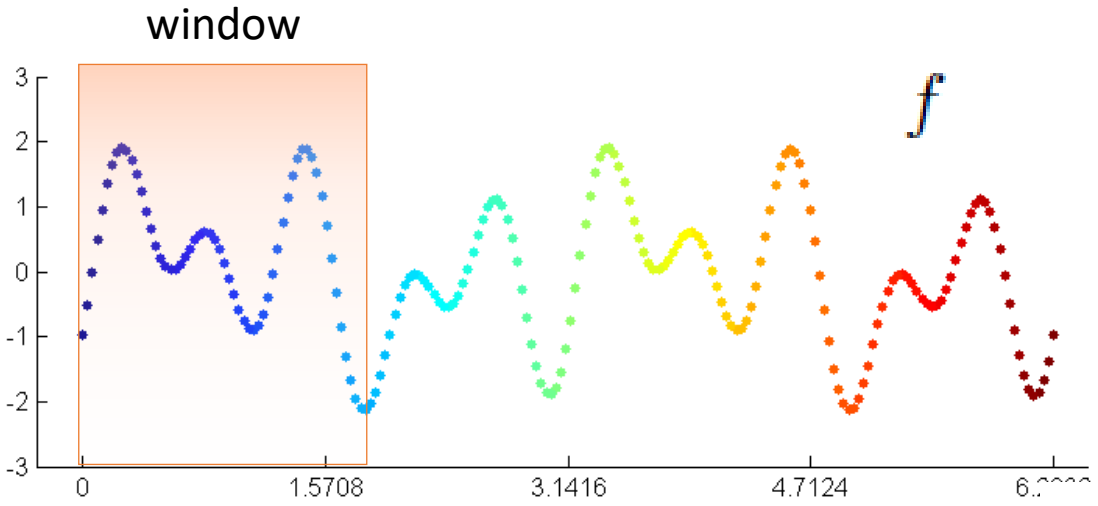


Sliding Windows



SW1PerS: Sliding Windows and 1-Persistence Scoring

Sliding window embedding

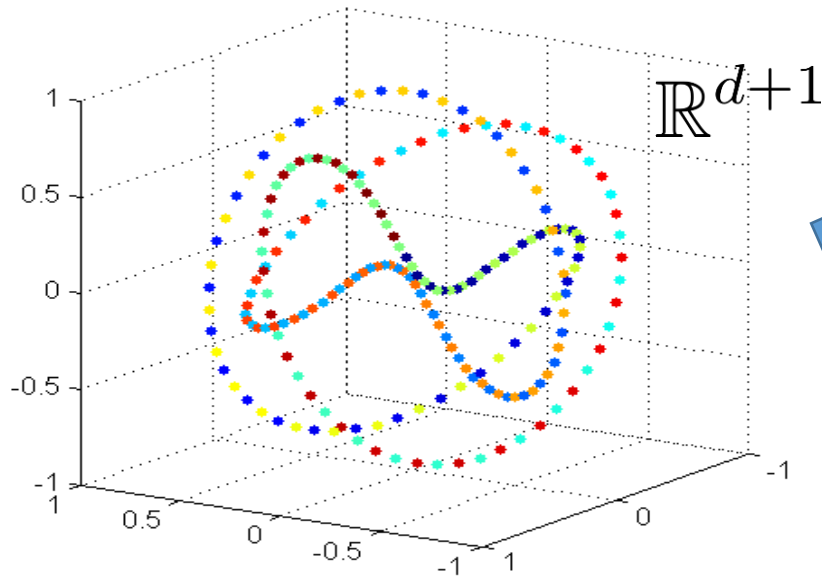


τ ⟶ Step/delay

$d\tau$ ⟶ Window size

$d+1$ ⟶ Dimension

$$SW_{d,\tau} f(t) = \begin{bmatrix} f(t) \\ f(t+\tau) \\ \vdots \\ f(t+d\tau) \end{bmatrix}$$

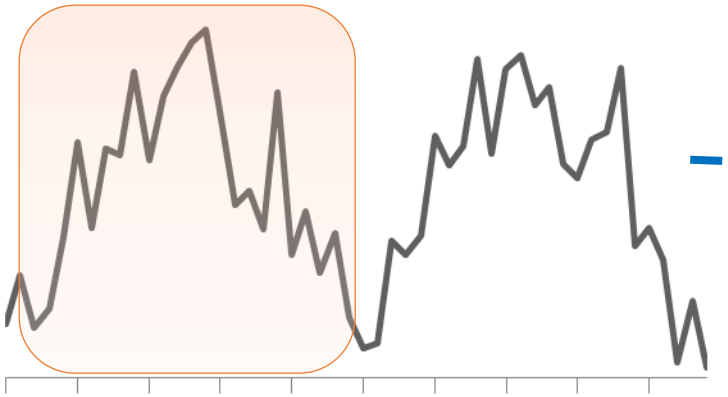


Sliding window point-cloud

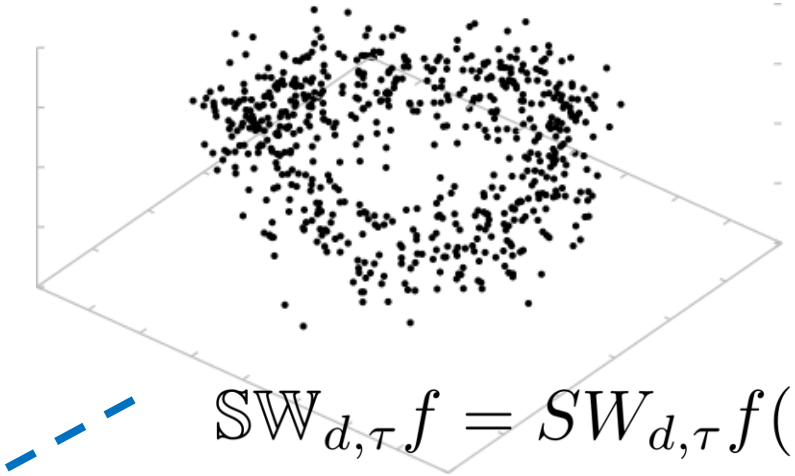
$$\mathbb{SW}_{d,\tau} f$$
$$\| $$
$$SW_{d,\tau} f(I)$$

$I \subset \mathbb{R}$

Sliding Windows and Persistence: An application of topology to signal analysis, J. Perea and J. Harer, 2015

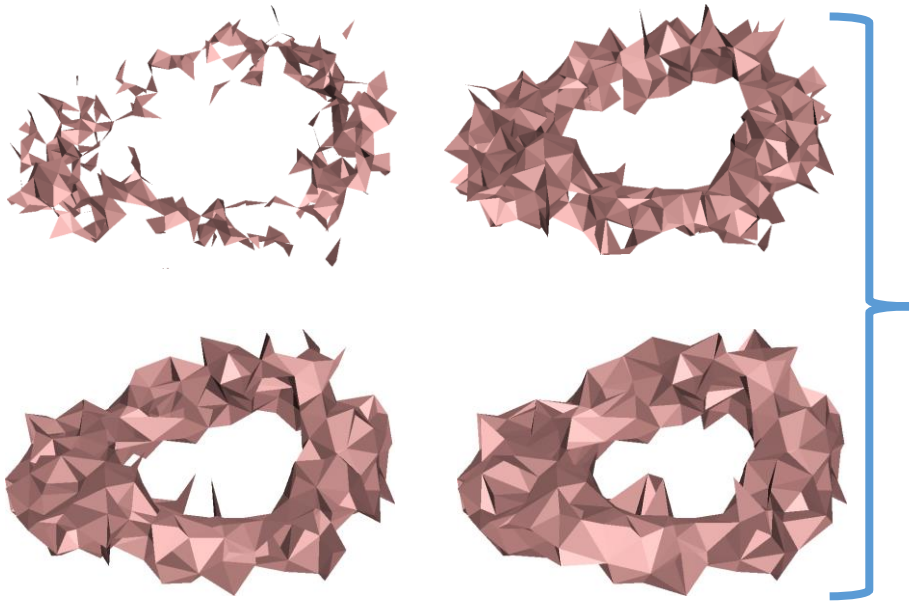


Sliding window point-cloud



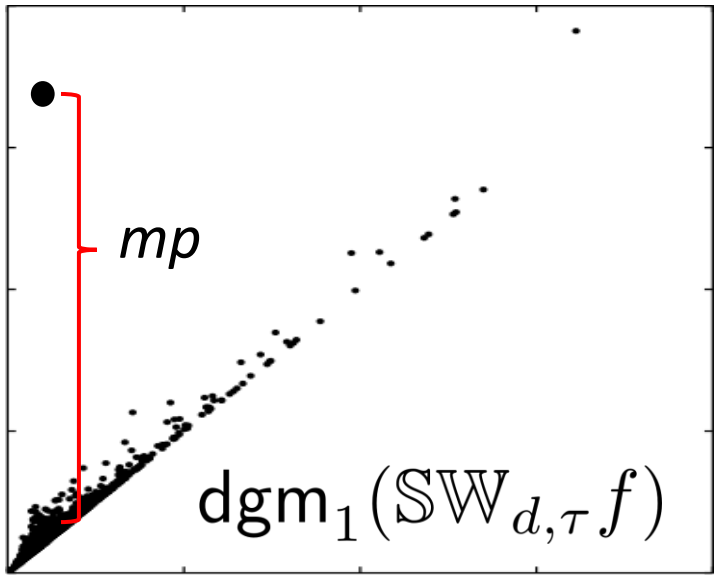
$$\mathbb{SW}_{d,\tau} f = SW_{d,\tau} f(I)^*$$

$\{R_\epsilon(X)\}_{\epsilon \geq 0}$ Rips complexes

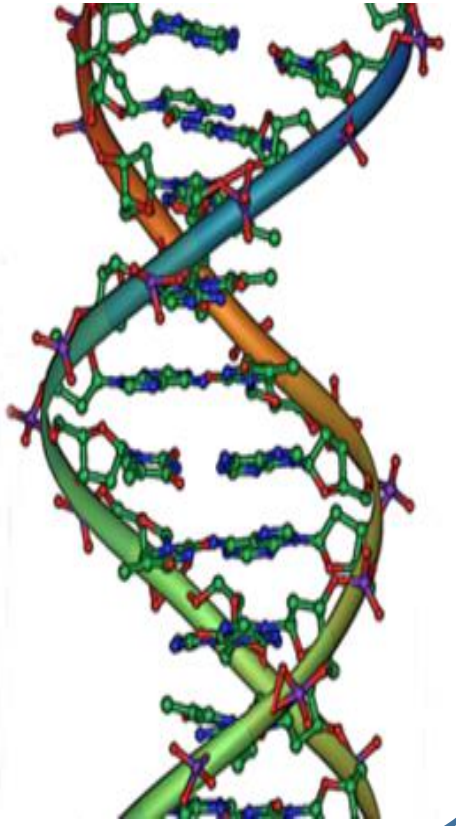


dgm

1D- Persistence Diagram



Biological Clocks



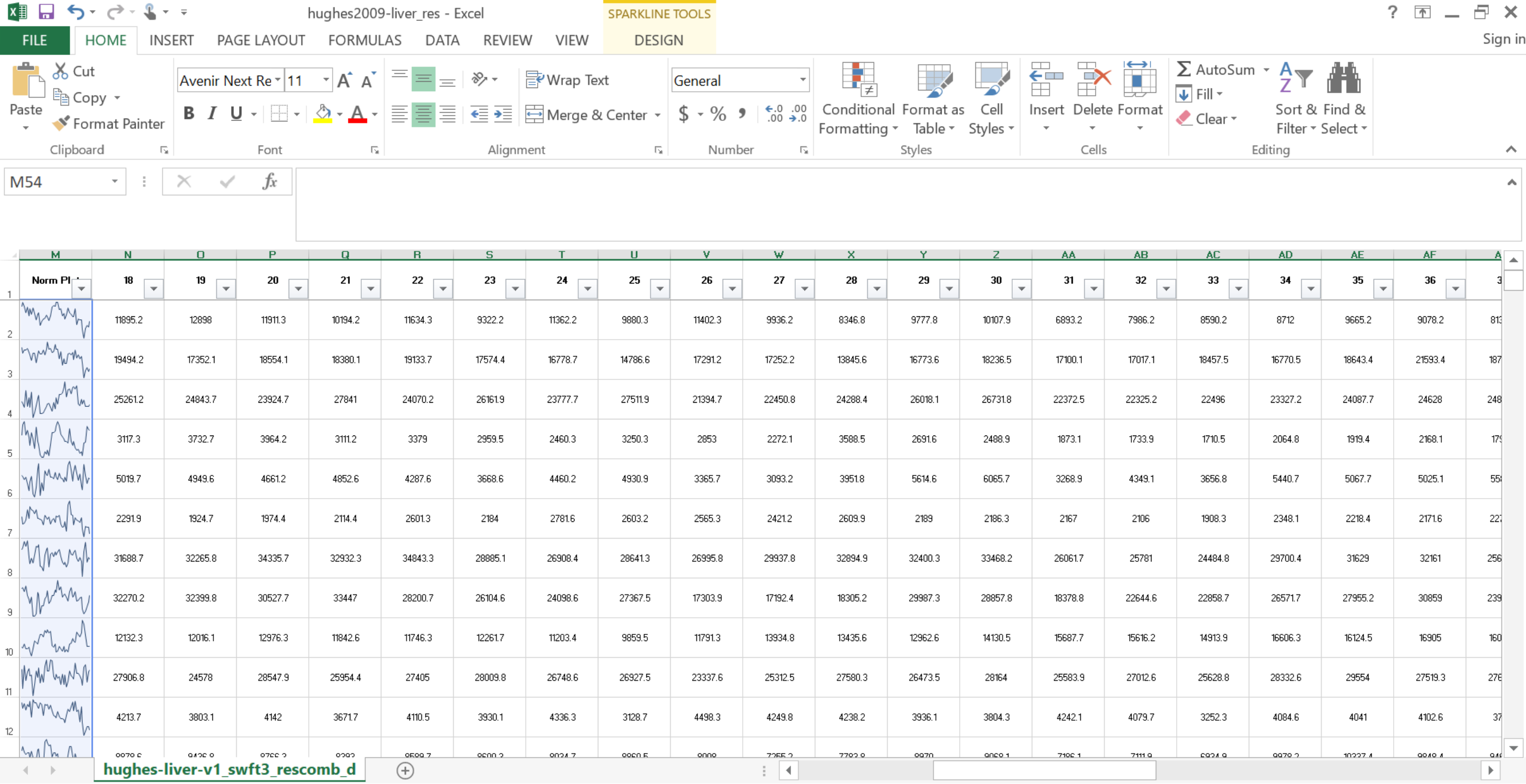
Perea *et al. BMC Bioinformatics* (2015) 16:257
DOI 10.1186/s12859-015-0645-6

BMC Bioinformatics

METHODOLOGY ARTICLE Open Access

SW1PerS: Sliding windows and 1-persistence scoring; discovering periodicity in gene expression time series data

Jose A. Perea[1,2*], Anastasia Deckard[3], Steve B. Haase[4,5] and John Harer[1,4,6]



SW1PerS: Sliding Windows and 1-Persistence Scoring, J. Perea et. al., 2015

hughes2009-liver_res - Excel

FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW Sign in

Paste Cut Copy Format Painter | Avenir Next Re 11 | Wrap Text | General | Conditional Formatting | Format as Table | Cell Styles | Insert Delete Format | AutoSum Fill Clear | Sort & Filter Find & Select

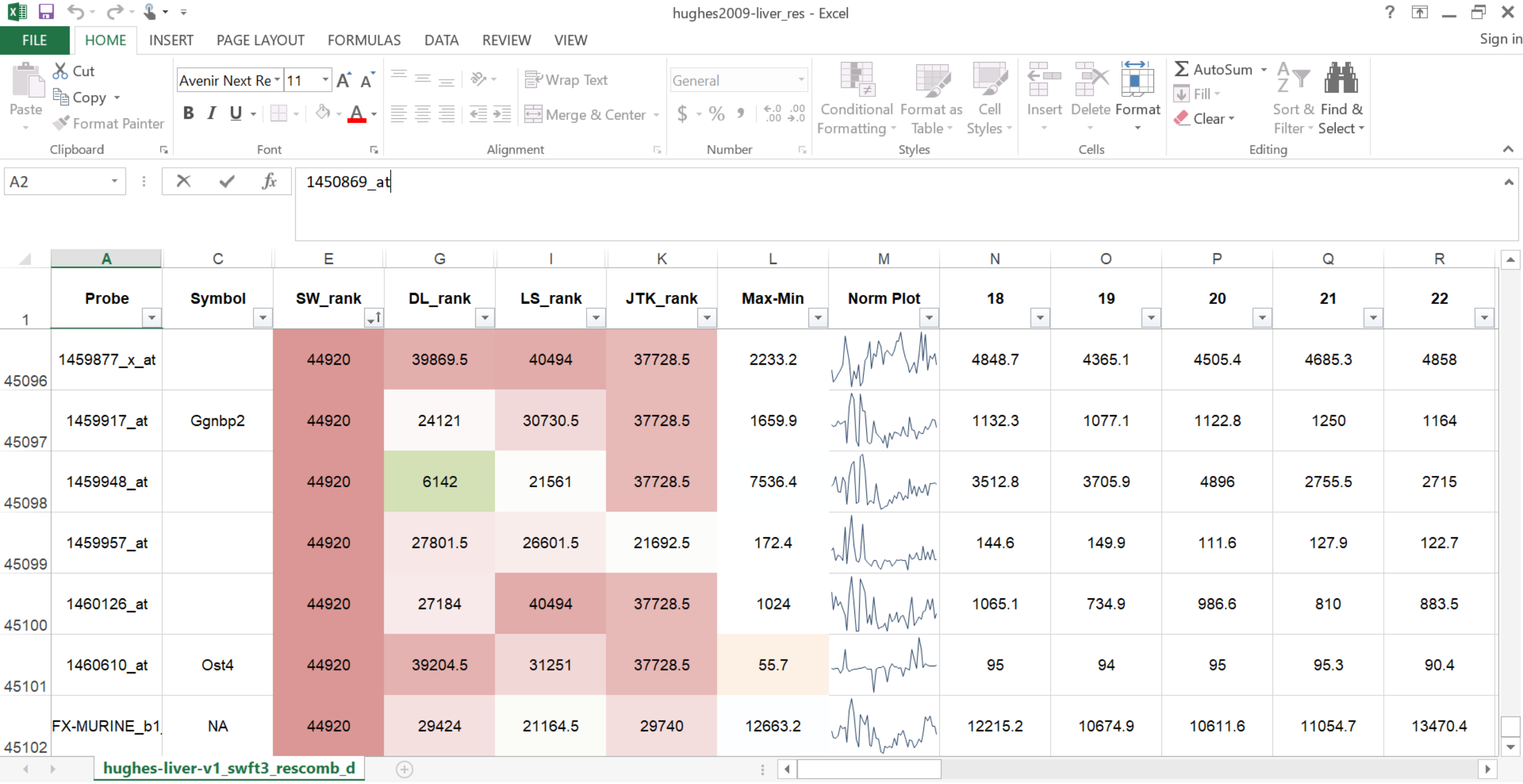
Clipboard Font Alignment Number Styles Cells Editing

A2 1450869_at

	A	C	E	G	I	K	L	M	N	O	P	Q	R
1	Probe	Symbol	SW_rank	DL_rank	LS_rank	JTK_rank	Max-Min	Norm Plot	18	19	20	21	22
2	1450869_at	Fgf1	1	1451.5	39	116	10916.2		15201.2	13204.8	15041.3	14251	11082.3
3	1416958_at	Nr1d2	2	26.5	10	25.5	62708.9		13006.6	12287.7	10224	9298	7393.1
4	1417190_at	Nampt	3	242.5	48	5	17275.4		10043.8	9446.7	9994.4	7548.1	6338.7
5	1450714_at	Azin1	4	4053.5	134	121	8391.4		13494.8	14060.1	13963.6	12193.6	11334.7
6	1436590_at		5	98	86.5	144.5	49573.2		47739.7	38509.4	38855.4	30598.5	35784.8
7	1420722_at	Elovl3	6.5	26.5	1	1	149978.3		76307.7	93712.8	98998.6	121038.2	127449
8	1437250_at	Mreg	6.5	1534.5	38	34.5	30373.3		29040.6	29839.6	31687.1	39536	35259.5

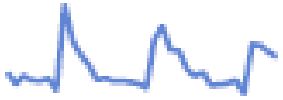
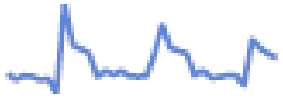
hughes-liver-v1_swft3_rescomb_d

SW1PerS: Sliding Windows and 1-Persistence Scoring, J. Perea et. al., 2015



SW1PerS: Sliding Windows and 1-Persistence Scoring, J. Perea et. al., 2015

Yeast Metabolic Cycle Data

Gene	SW	DL	LS	JTK	Amp	Plot
ECM33	137	1552	1194.5	1492	35.86	
CDC9	291	1494	1993.5	2714.5	2.81	
SAM1,2	628	1133	1723	3289.5	60.82	
MSH6	715	3569	2381	3341.5	5.06	

Rankings of genes in the top 10% (out of 9,330) according to SW, and not in the top 10% for any other algorithm

SW1PerS: Sliding Windows and 1-Persistence Scoring, J. Perea et. al., 2015

Consensus Analysis

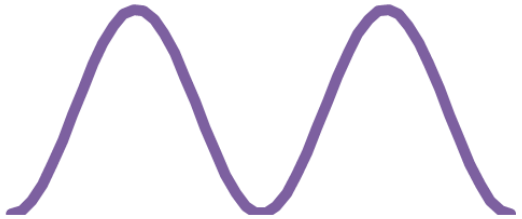
Data Set	Alg	#Consensus	%Consensus
Yeast Cell Cycle	sw	316	0.90
Consensus: 353	dl	289	0.82
	ls	298	0.84
	jtk	311	0.88
Yeast Met. Cycle	sw	553	0.93
Consensus: 596	dl	345	0.58
	ls	563	0.94
	jtk	541	0.91
Mammal Circadian	sw	3090	0.82
Consensus: 3767	dl	3330	0.88
	ls	3640	0.97
	jtk	3636	0.97

Table 3 Number and percentage of probes in the top 10% of rankings from each algorithm that are in a consensus set. That is, those which appear in the top 10% of rankings for at least three algorithms.

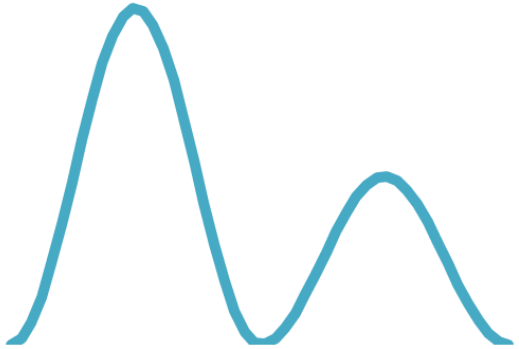
SW1PerS: Sliding Windows and 1-Persistence Scoring, J. Perea et. al., 2015

SW1PerS:
Sliding Windows and 1-Peristence Scoring

Shape agnostic



Resistant to dampening



SW1PerS: Sliding Windows and 1-Persistence Scoring, J. Perea et. al., 2015

Found Comput Math (2015) 15:799–838
DOI 10.1007/s10208-014-9206-z

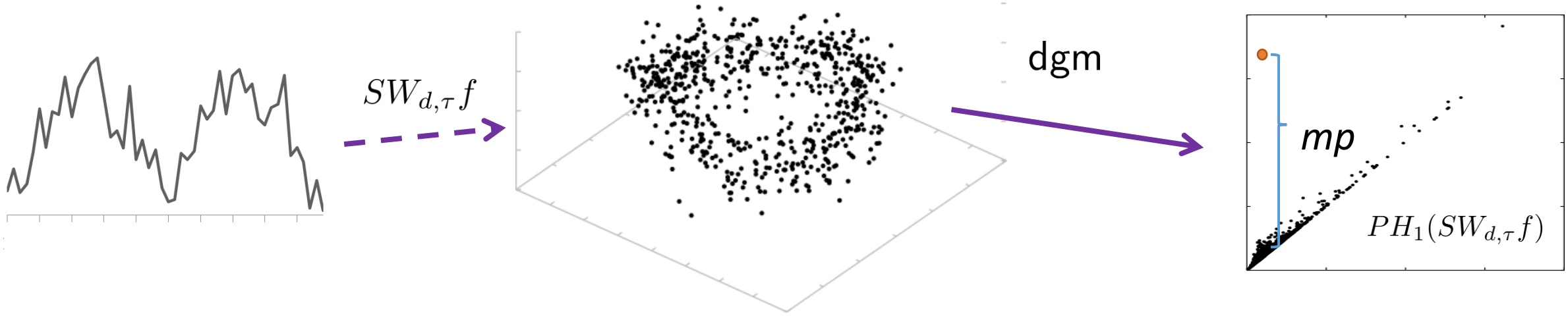
FOUNDATIONS OF COMPUTATIONAL MATHEMATICS

The Journal of the Society for the Foundations of Computational Mathematics

Sliding Windows and Persistence: An Application of Topological Methods to Signal Analysis

Jose A. Perea · John Harer

Recap: SW1PerS (Sliding Windows and 1-Persistence Scoring)



$N \sim$ Sample size

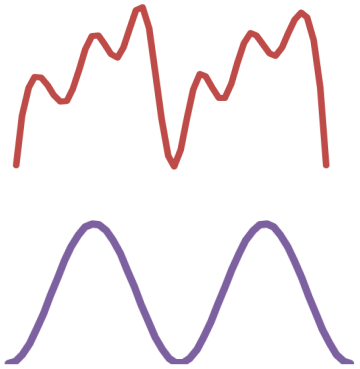
$d = 2N$

$L =$ number of periods

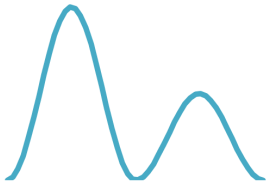
$$\tau = \frac{2\pi}{L(d+1)}$$

mp : maximum persistence (recurrence score)

Shape agnostic



Resistant to dampening

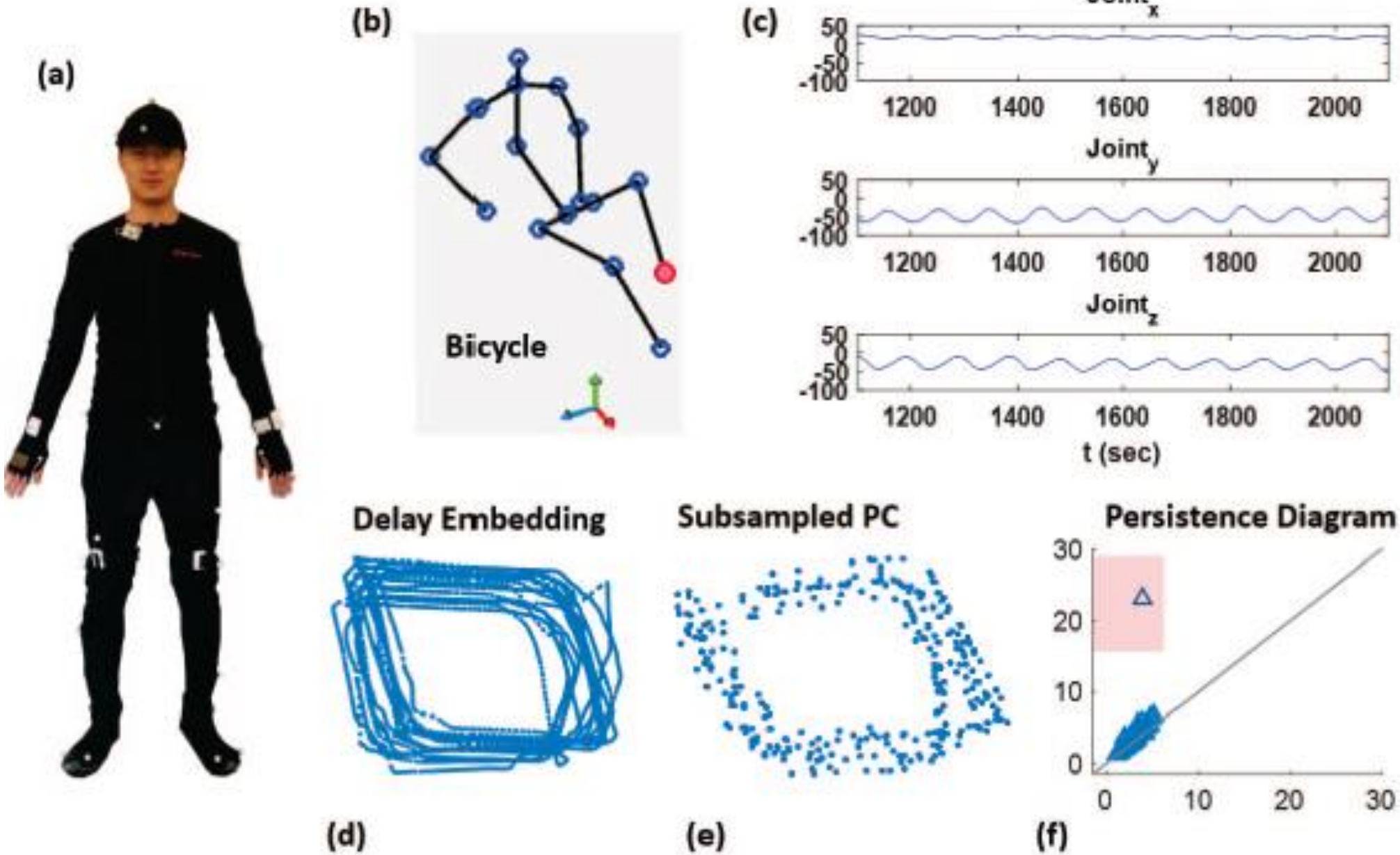


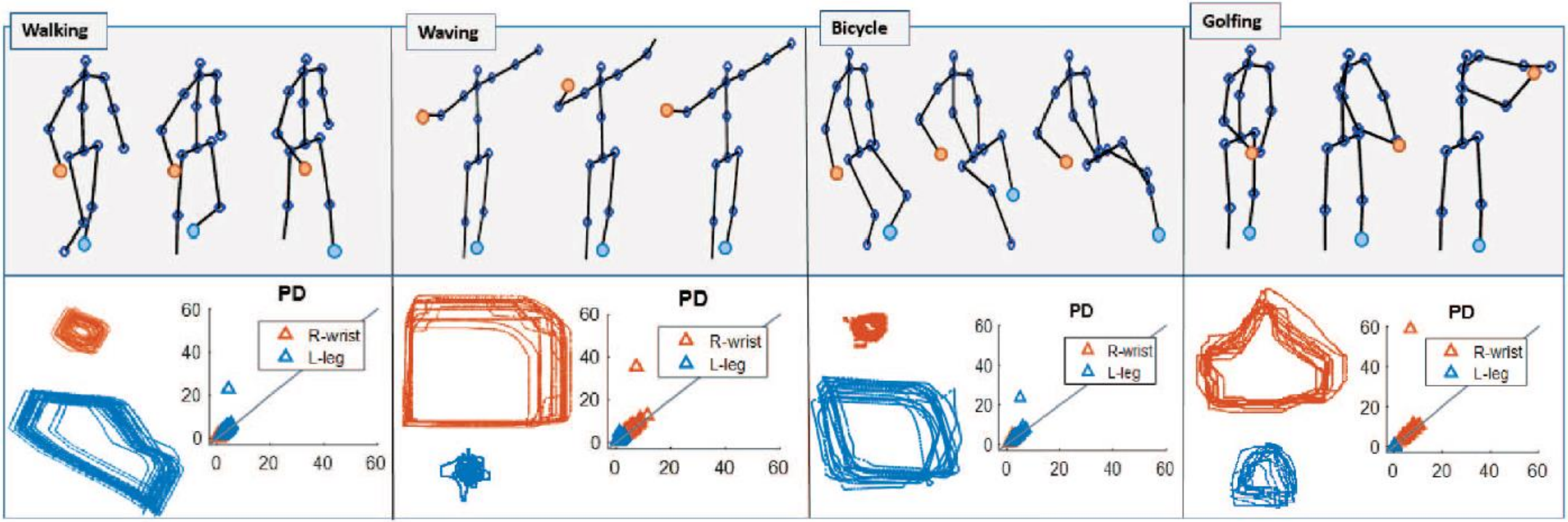
Field of coefficients

$$\mathbb{Z}_p \ , \quad p > d$$

ACTION CLASSIFICATION FROM MOTION CAPTURE DATA USING TOPOLOGICAL DATA ANALYSIS

Alireza Dirafzoon, Namita Lokare and Edgar Lobaton


(a)
(b)
Bicycle
(c)
Joint x
Joint y
Joint z
50
0
-50
-100
1200
1400
1600
1800
2000
t (sec)
Delay Embedding
Subsampled PC
Persistence Diagram
30
20
10
0
(d)
(e)
(f)


Walking
Waving
Bicycle
Golfing
PD
R-wrist
L-leg
0
20
40
60

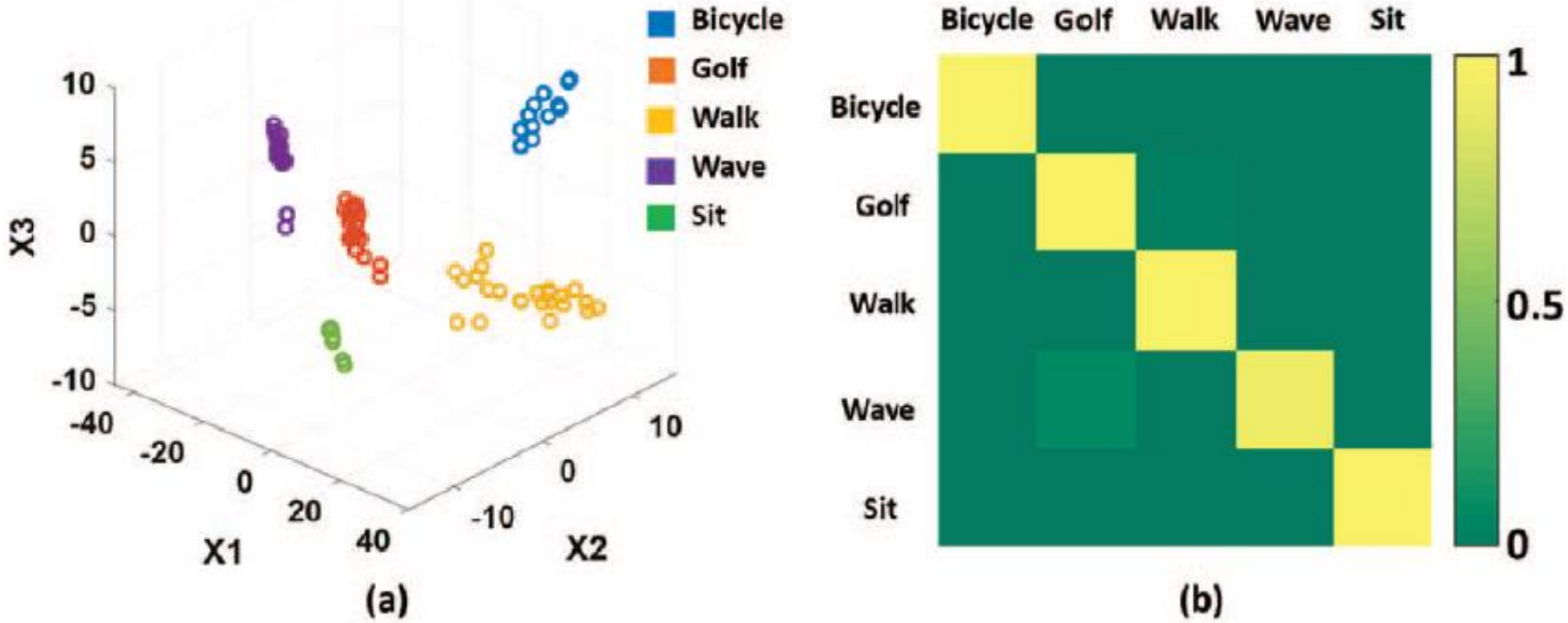


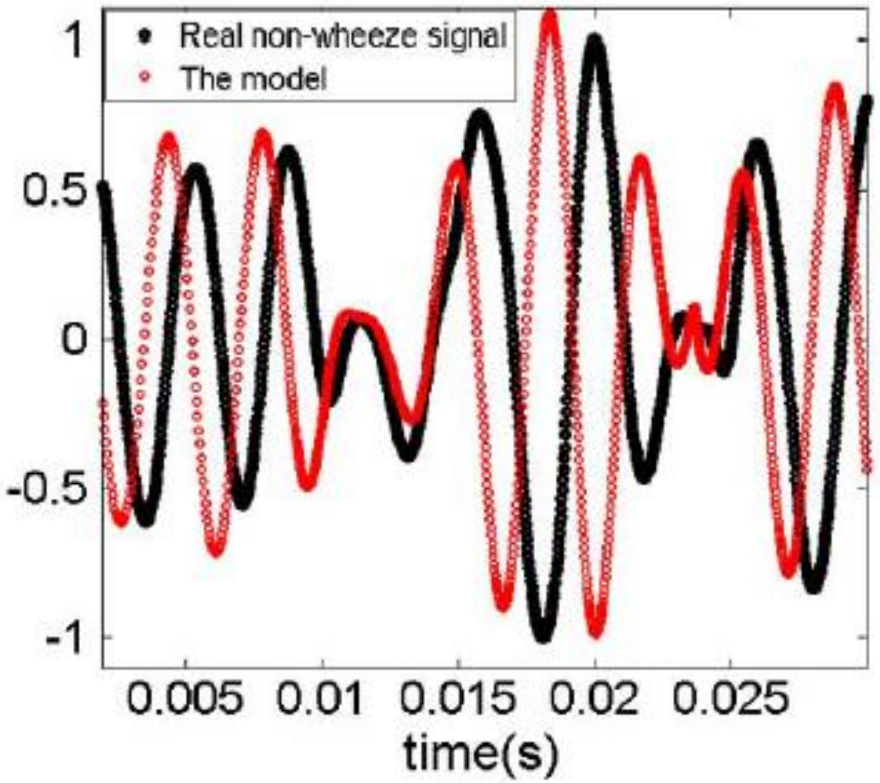
Fig. 4. (a) Separation of classes from the training set, (b) Confusion matrix over the predicted and true classes

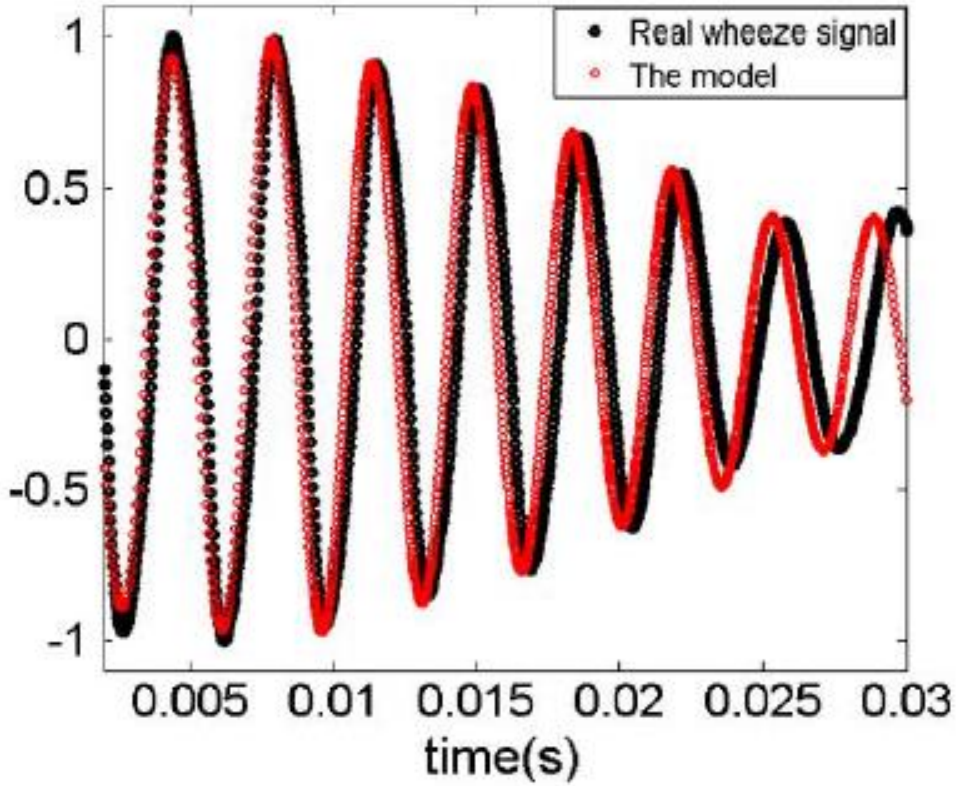
Table 1. Class accuracy results for the activities

Action	Bicycle	Golf	Walk	Wave	Sit
Accuracy	1.00	0.9787	0.9929	0.9858	1.00

Persistent Homology of Delay Embeddings and its Application to Wheeze Detection

Saba Emrani, Thanos Gentimis, and Hamid Krim, *Fellow, IEEE*

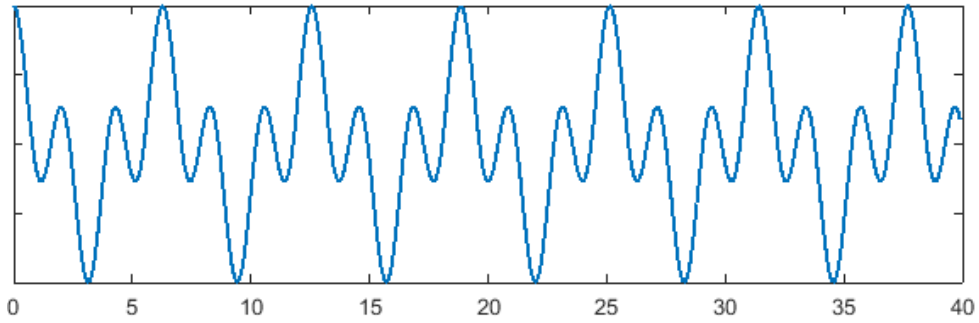


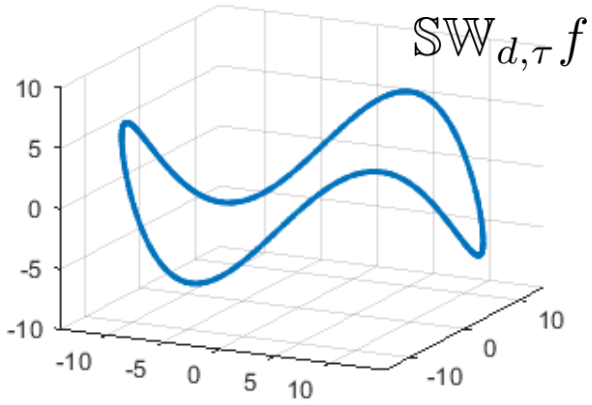


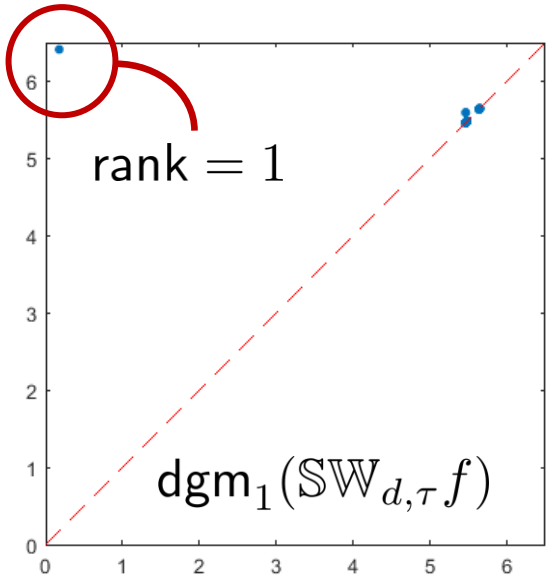
The accuracy of our proposed technique is 98.39% while the accuracy of the techniques proposed in [8] and [9] are 86.2% and 95.5%, respectively.

Geometric QuasiPeriodicity

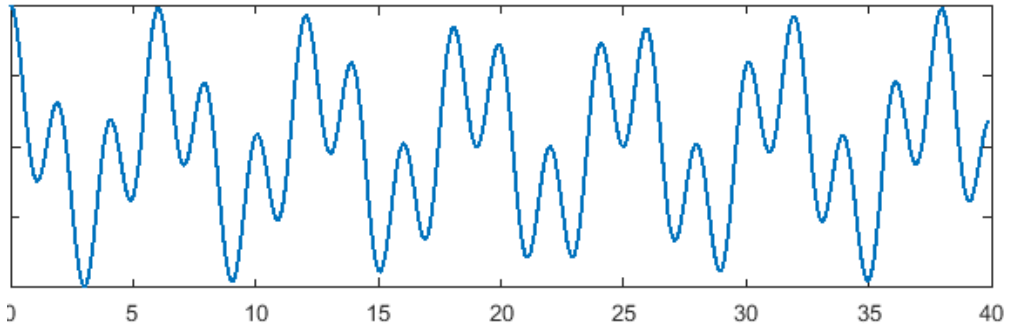
$f(t) = \cos(t) + \cos(3t)$

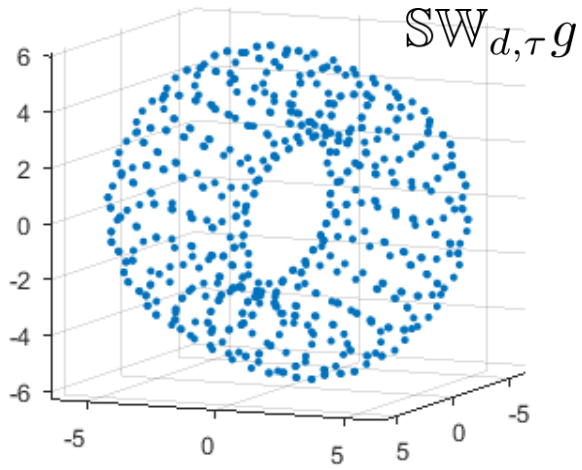


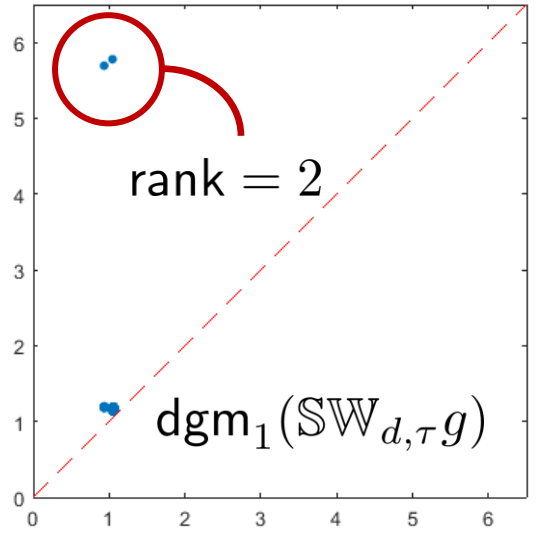




$g(t) = \cos(t) + \cos(\pi t)$







Time Series

Sliding Window Point Cloud

Persistent Homology

(joint with Adam Huston)

Audio for this story is unavailable.

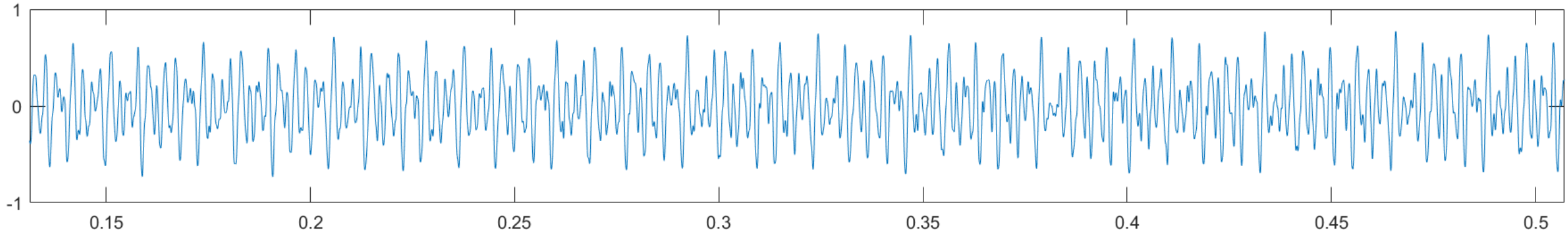
MUSIC LISTS

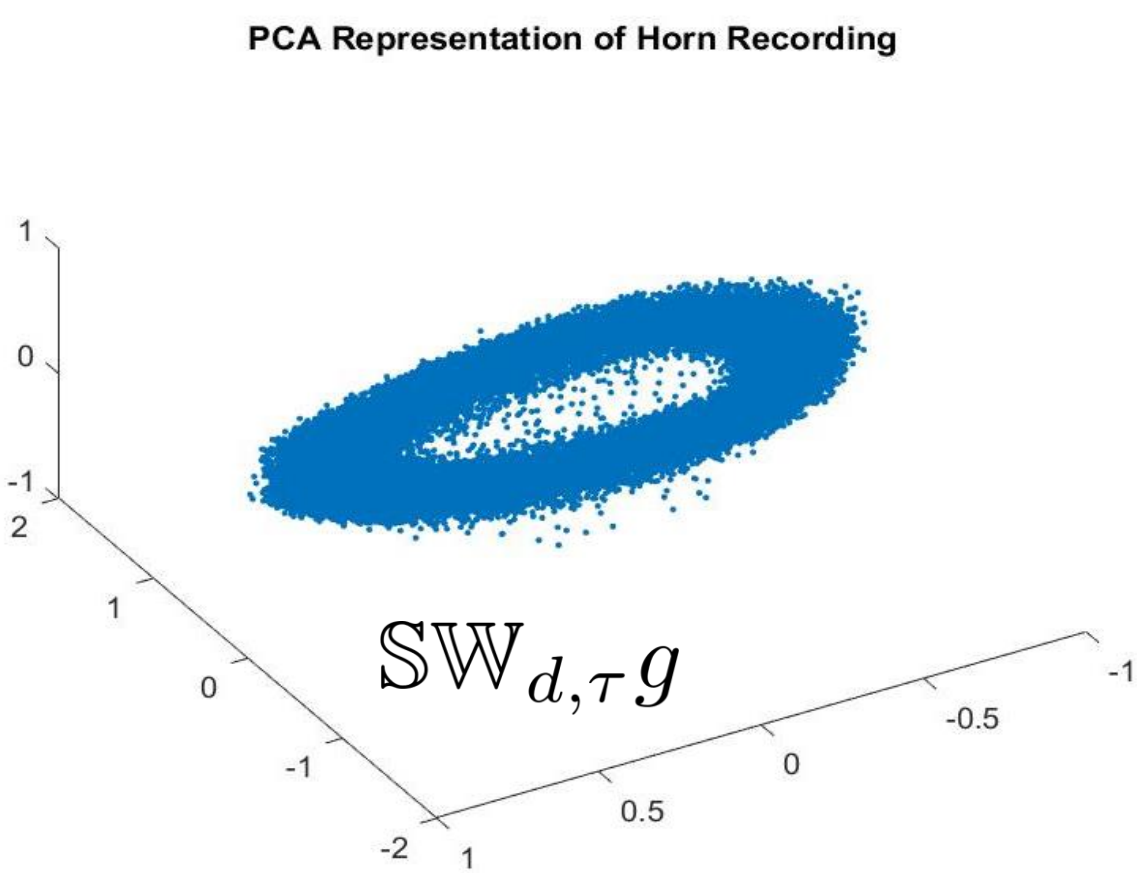
The Unsettling Sound Of Tritones, The Devil's Interval

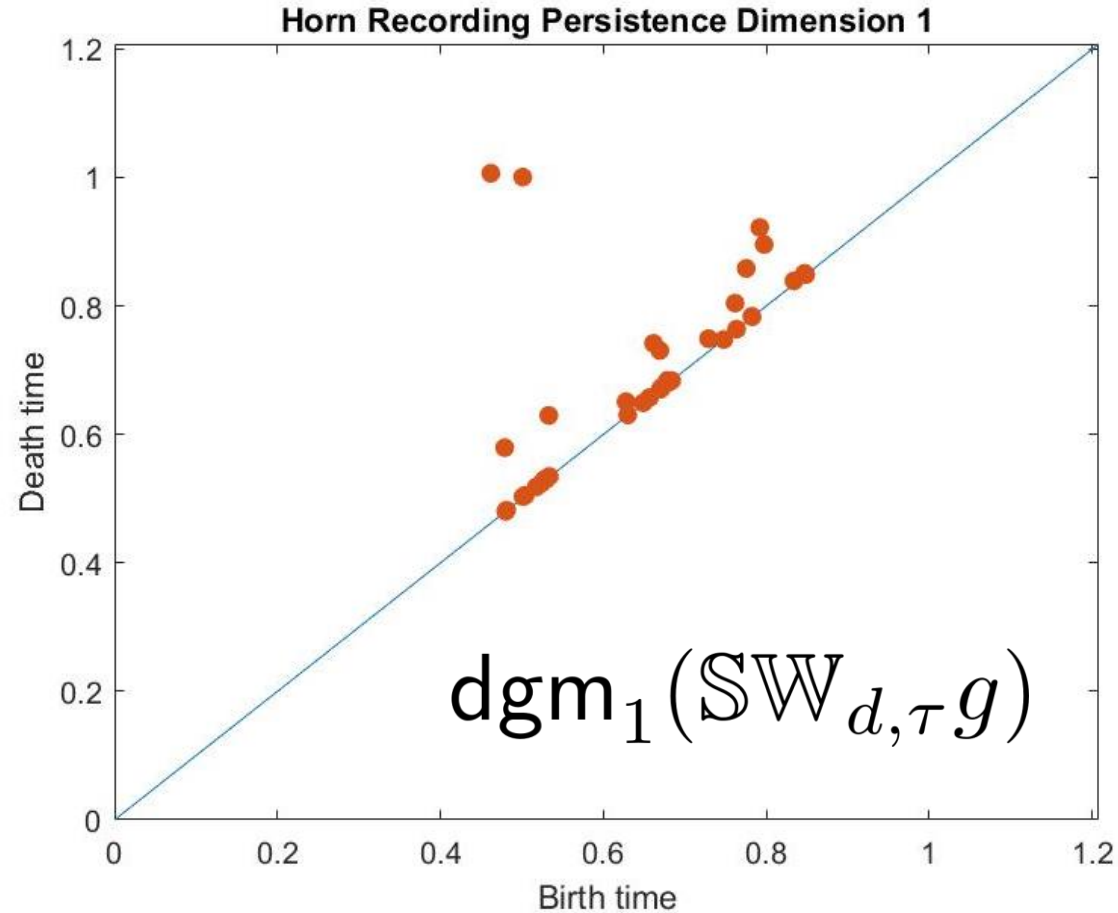
October 31, 2017 · 10:52 AM ET

JUDITH KOGAN

Frans Francken's *Death Playing the Violin*. In music theory, the tritone came to be known as the devil's interval.


1
0
-1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5


PCA Representation of Horn Recording
$\mathbb{SW}_{d,\tau} g$


Horn Recording Persistence Dimension 1
Death time
Birth time
$\mathrm{dgm}_1(\mathbb{SW}_{d,\tau} g)$

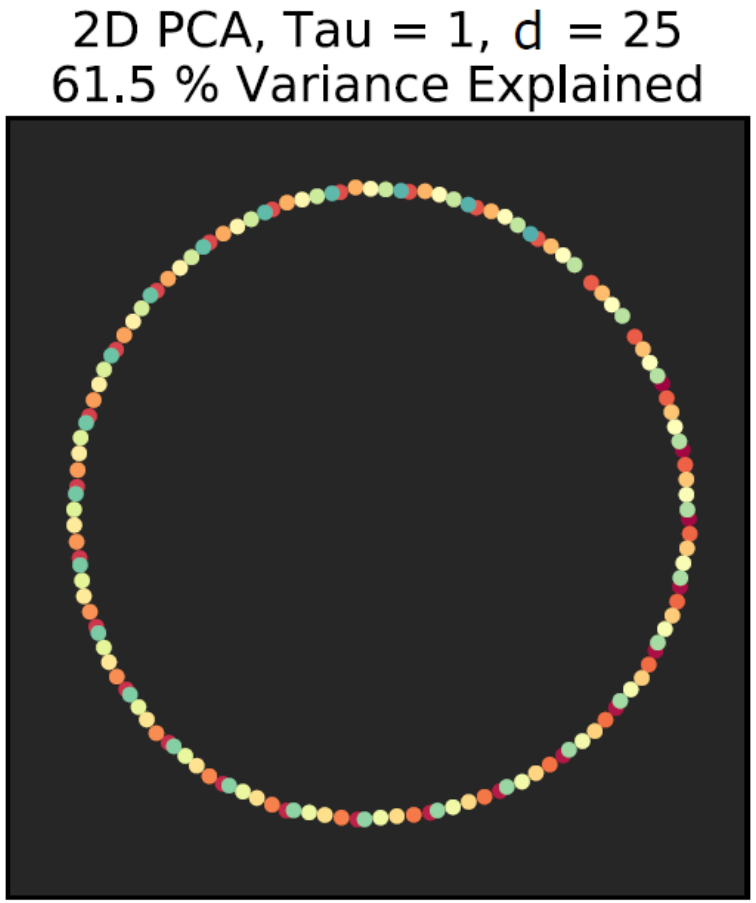
SIAM J. IMAGING SCIENCES
Vol. 11, No. 2, pp. 1049–1077

© 2018 Society for Industrial and Applied Mathematics

(Quasi)Periodicity Quantification in Video Data, Using Topology*

Christopher J. Tralie† and Jose A. Perea‡

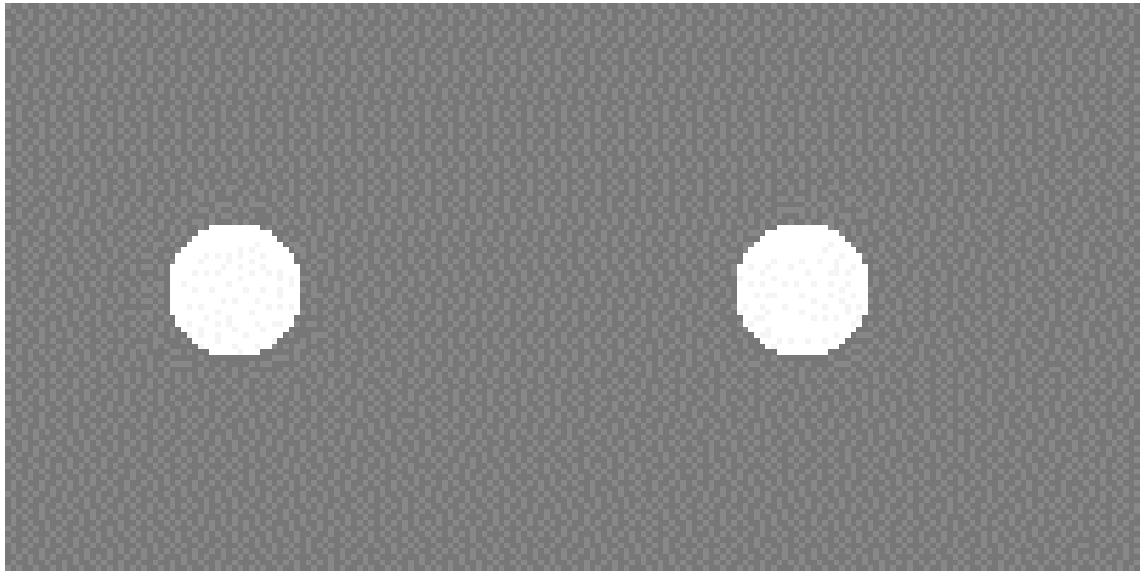
Recurrence in video data (SW1PerS-video)

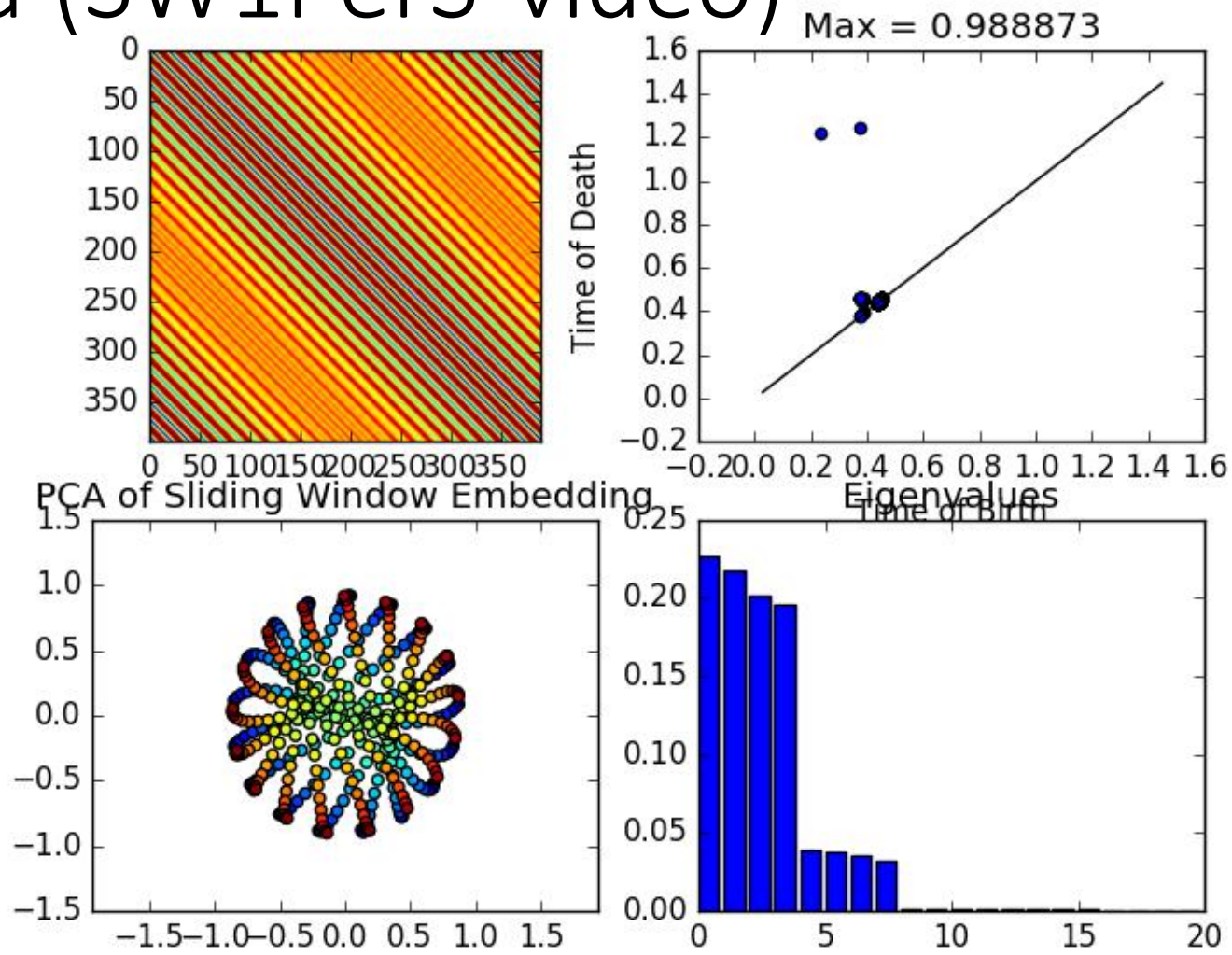


Sliding window embedding

C. Tralie and J. Perea, *(Quasi)Periodicity Quantification in Video Data, Using Topology,* 2018.

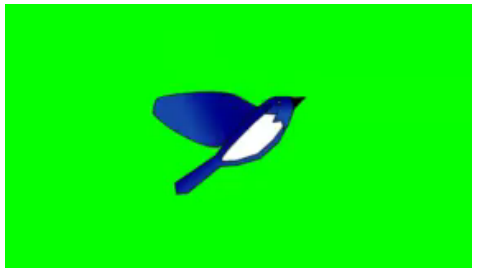
Recurrence in video data (SW1PerS-video)





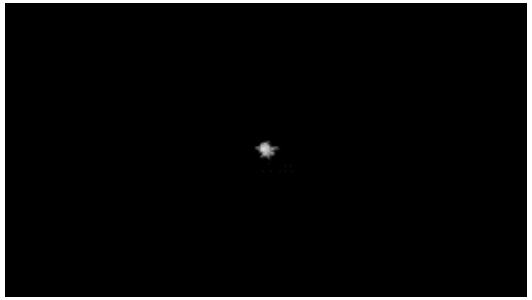
C. Tralie and J. Perea, *(Quasi)Periodicity Quantification in Video Data, Using Topology,* 2018.

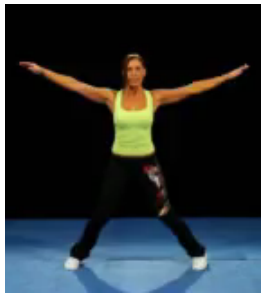
Experiments

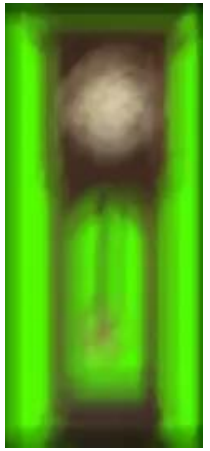




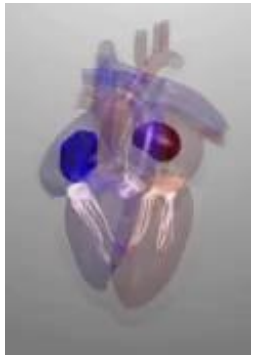
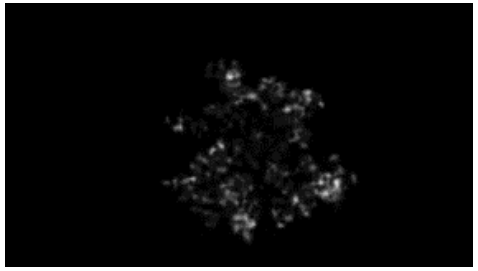




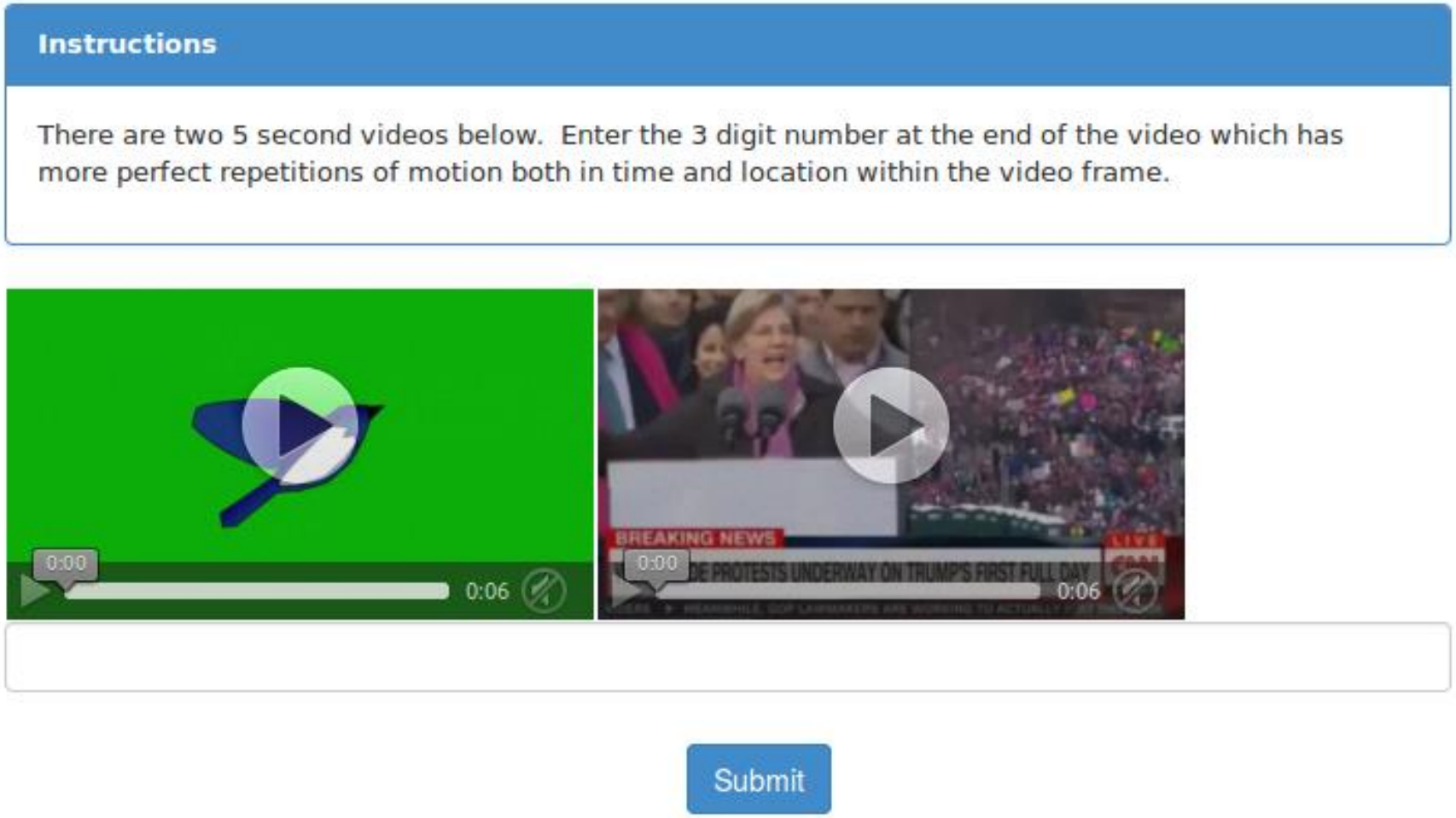




C. Tralie and J. Perea, *(Quasi)Periodicity Quantification in Video Data, Using Topology,* 2018.

Experiment: amazon Mechanical Turk



C. Tralie and J. Perea, *(Quasi)Periodicity Quantification in Video Data, Using Topology,* 2018.

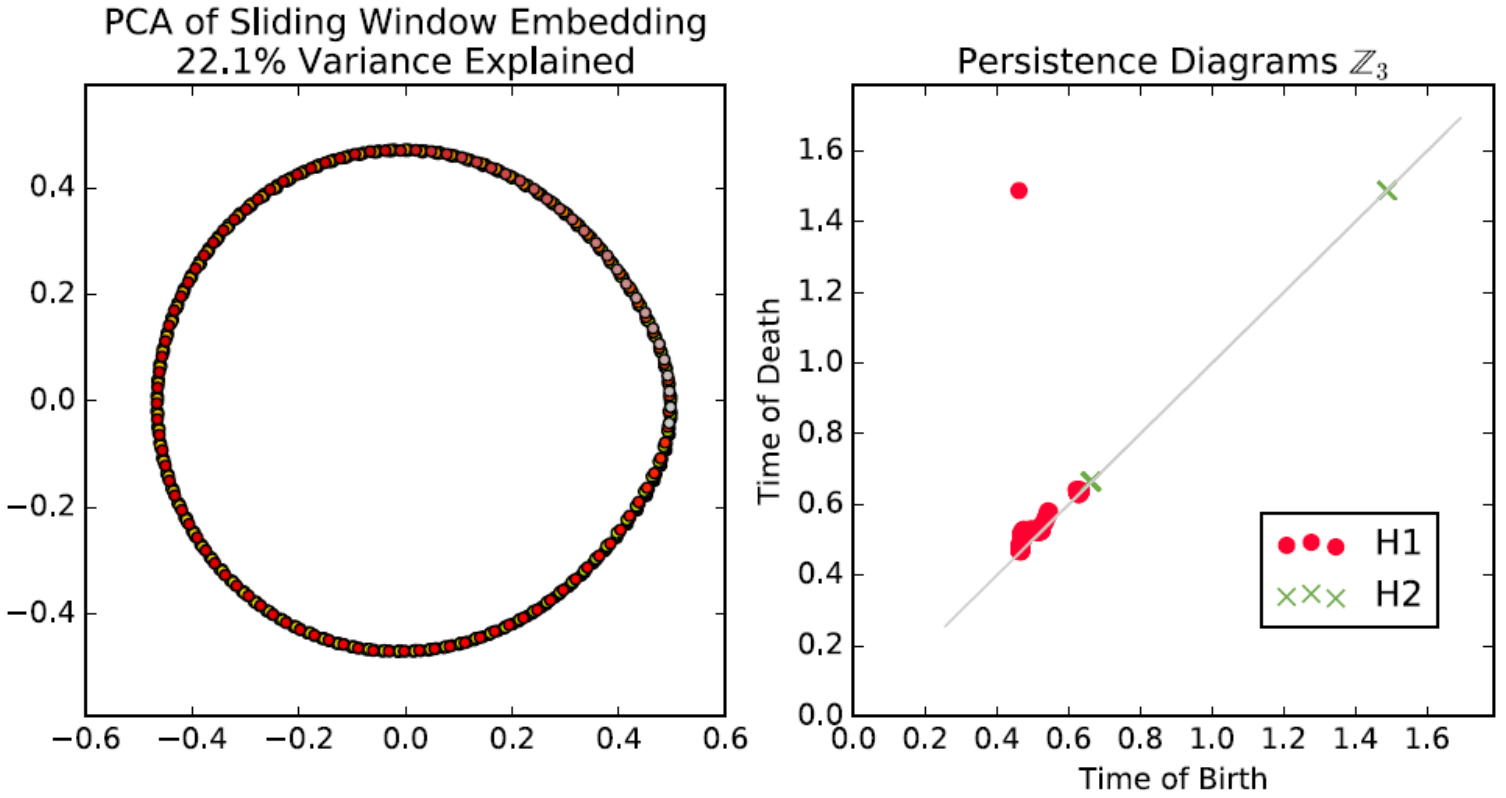
Results: Humans (Amazon turk) vs Computers

Correlation of rakings
(from most periodic to least periodic)
across 20 videos

Kendall's Tau	SW	CutlerDavis Freq	CutlerDavis Lattice	Humans
SW	1	-0.315	0.221	**0.663**
CutlerDavis Freq		1	-0.0842	-0.294
CutlerDavis Lattice			1	0.347
Humans				1

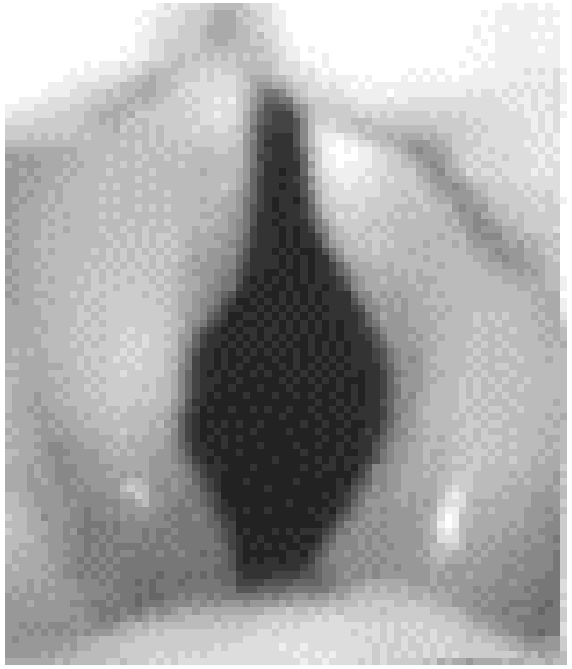
Recurrence in video data (SW1PerS-video)

normal

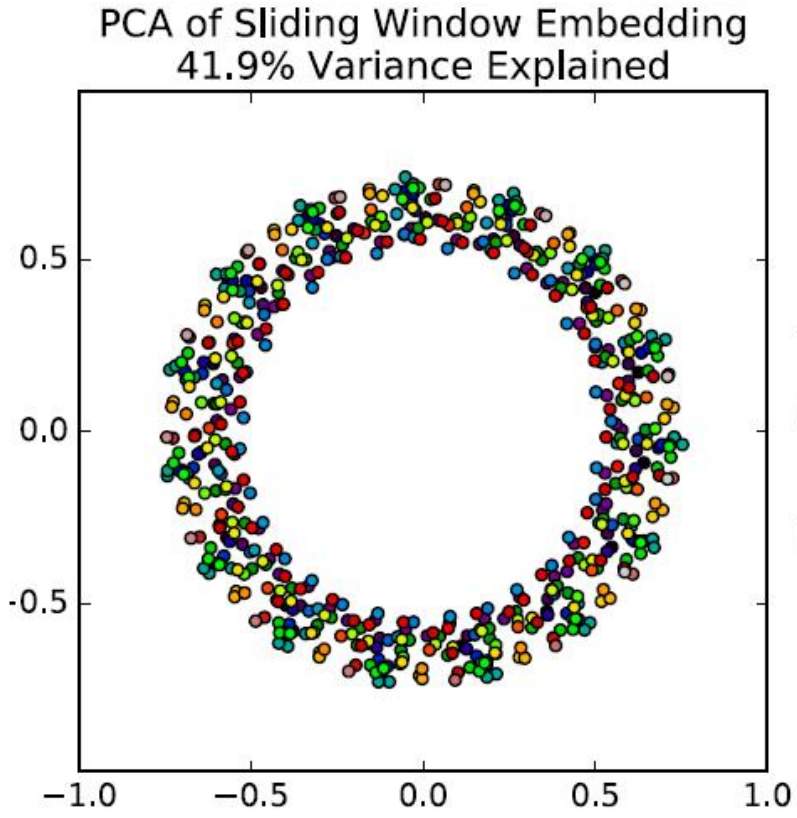


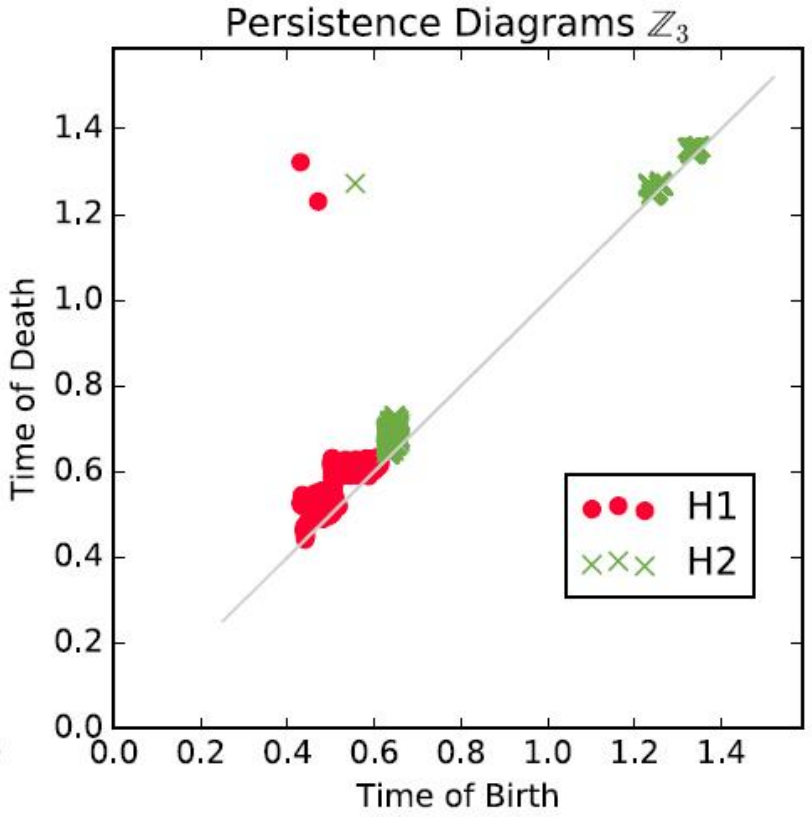
C. Tralie and J. Perea, *(Quasi)Periodicity Quantification in Video Data, Using Topology,* 2018.

Recurrence in video data (SW1PerS-video)



Clinical asymmetry



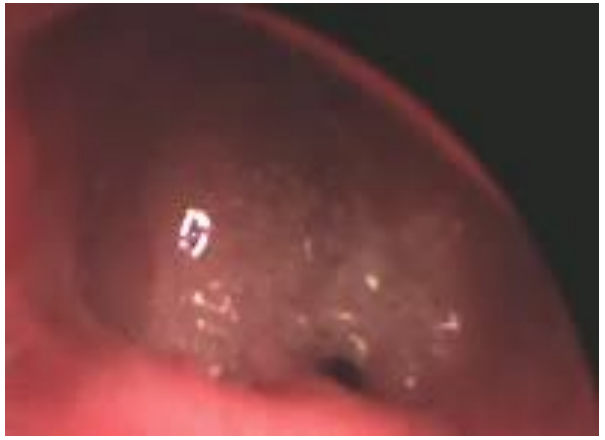


C. Tralie and J. Perea, *(Quasi)Periodicity Quantification in Video Data, Using Topology,* 2018.

Laryngeal video-endoscopy

normal

Clinical asymmetry



Mucus irregular

AP-Biphonation

C. Tralie and J. Perea, *(Quasi)Periodicity Quantification in Video Data, Using Topology,* 2018.

Video Name	Periodicity Score	Quasiperiodicity Score
HerbstPeriodic	**0.794**	0.005
NormalPeriodic	**0.561**	0.010
APBiphonation	0.296	**0.298**
APBiphonation2	0.486	0.133
ClinicalAsymmetry	0.072	**0.423**
MucusPerturbedPeriodic	0.019	0.005
HerbstIrregular	0.019	0.037

C. Tralie and J. Perea, *(Quasi)Periodicity Quantification in Video Data, Using Topology,* 2018.