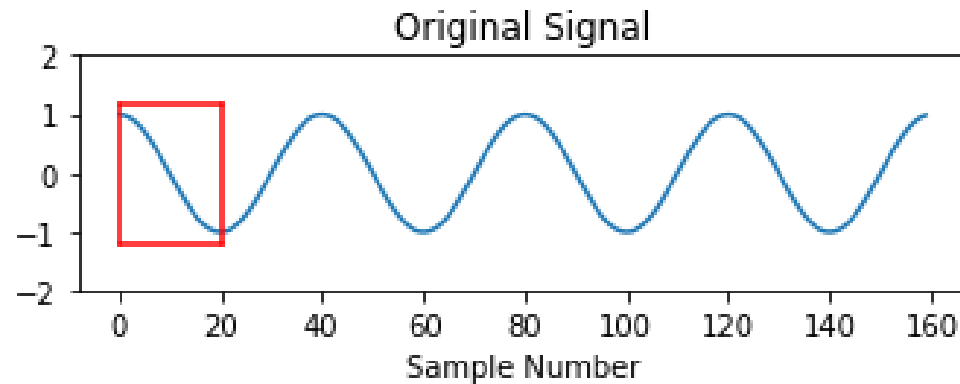


EMALCA 2021 Peru – Análisis Topológico de Datos

Clase 3: Topological Time Series Analysis



Max Persistence = 5.27

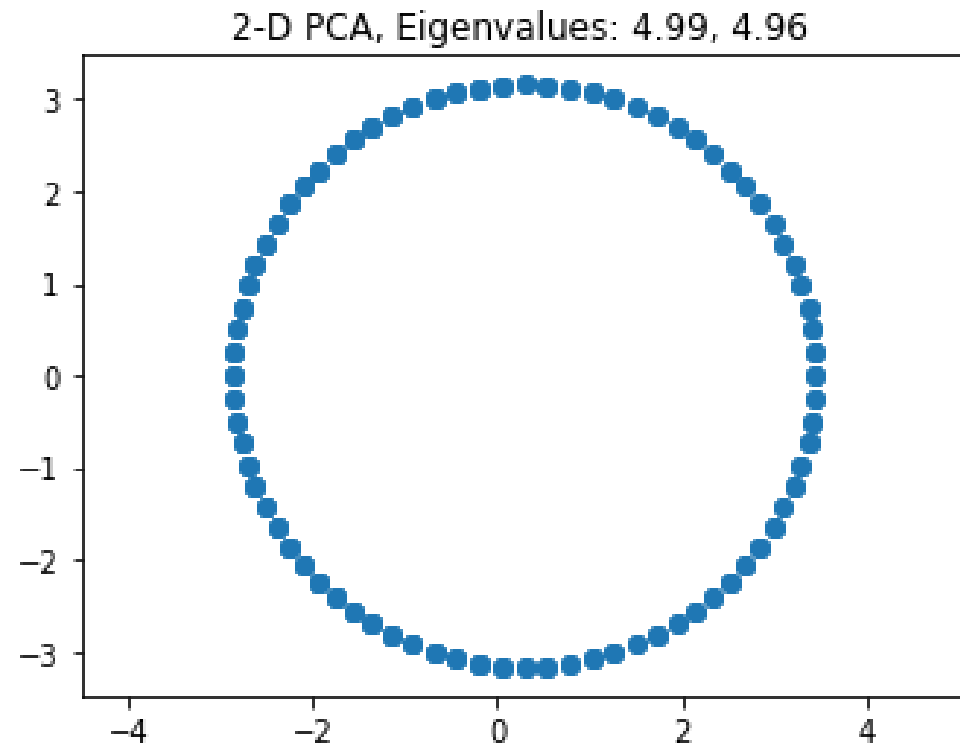
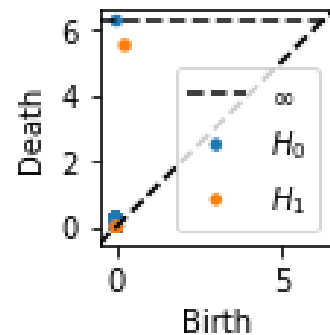
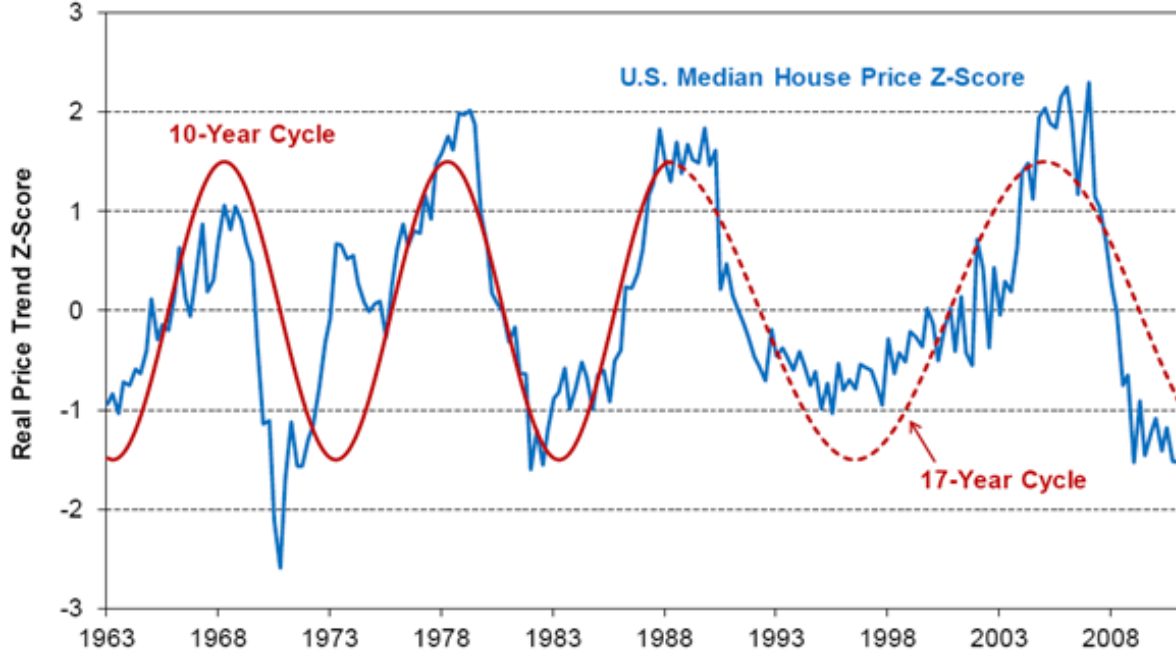
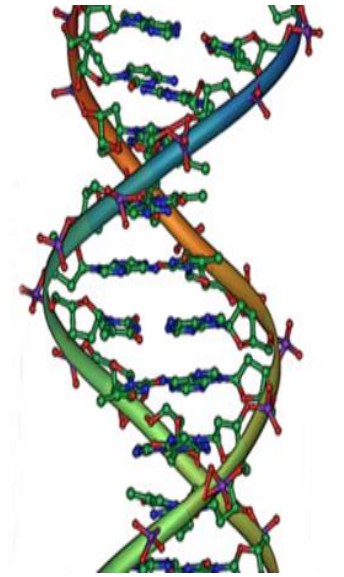
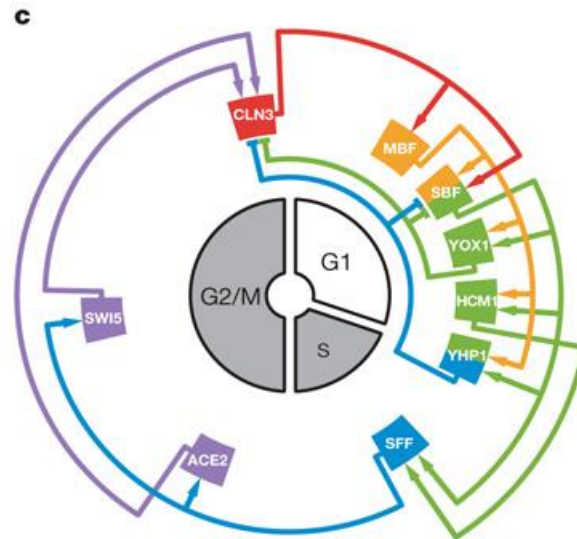
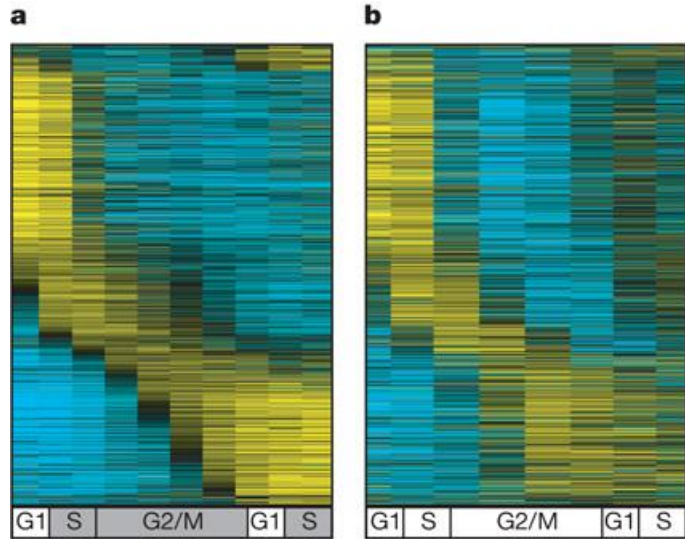


Exhibit 1

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

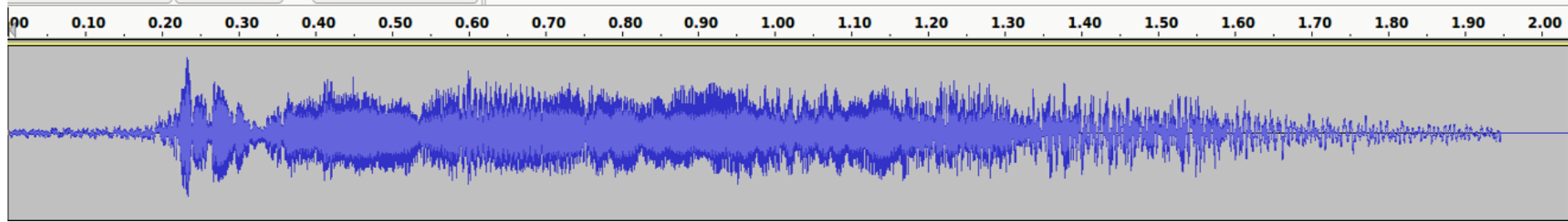


Source: Bureau of the Census, GMO As of 6/30/11



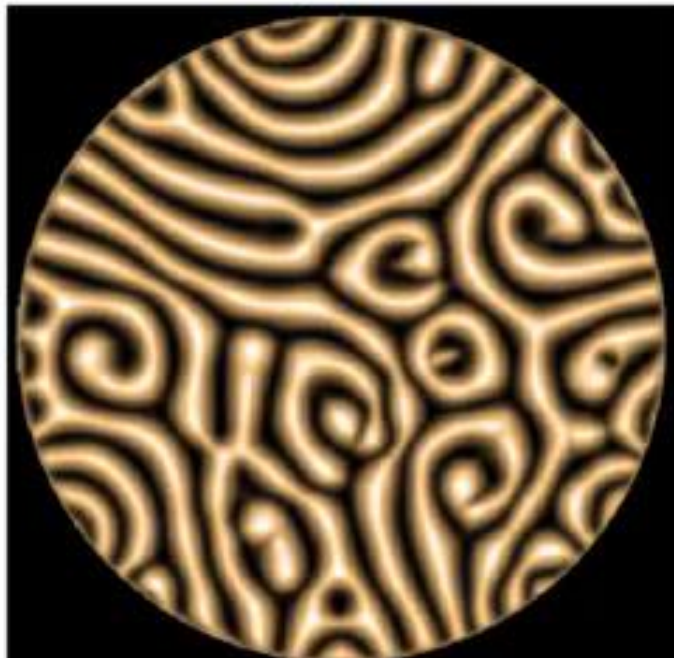
Global control of cell-cycle transcription by coupled SDK and network oscillators, D. Orlando et. al., Nature, 2008

Quantifying recurrence in time varying systems



Recording

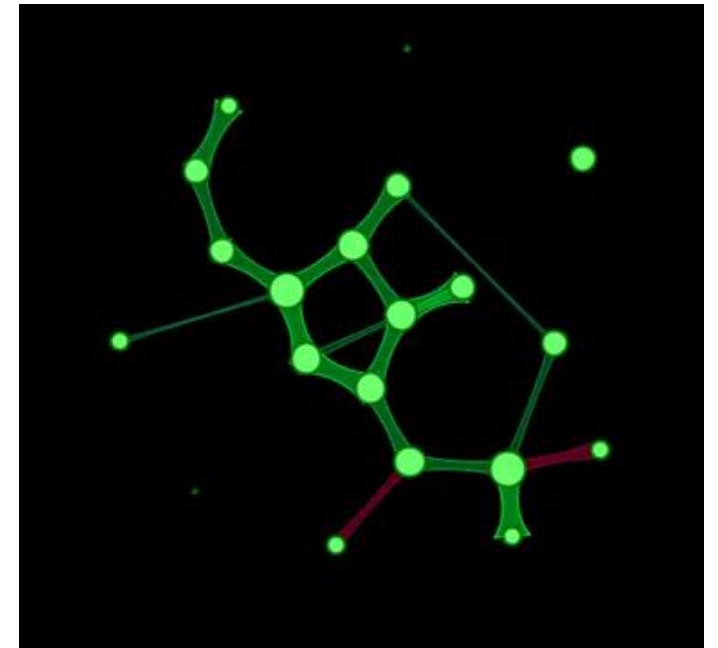
Evolving solution (RB conv.)

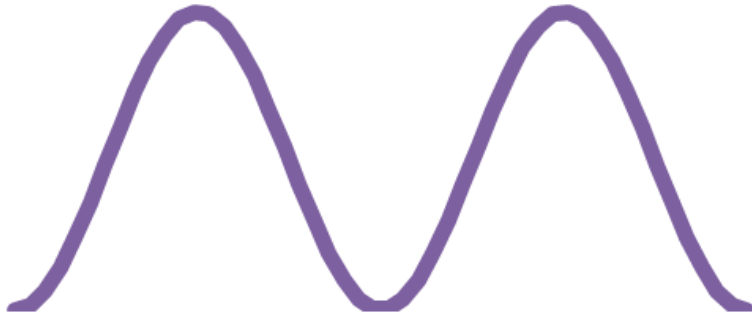


Video data

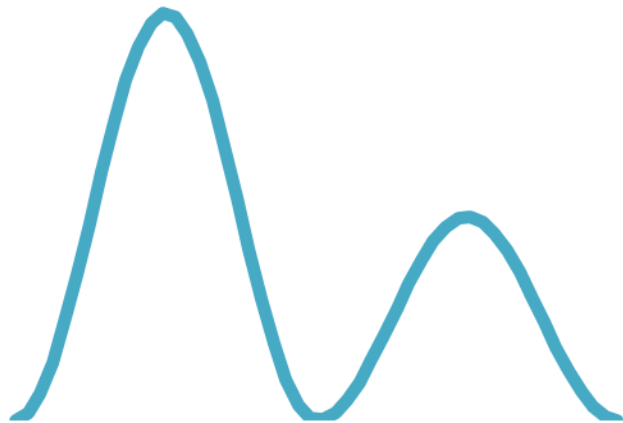


Dynamic network

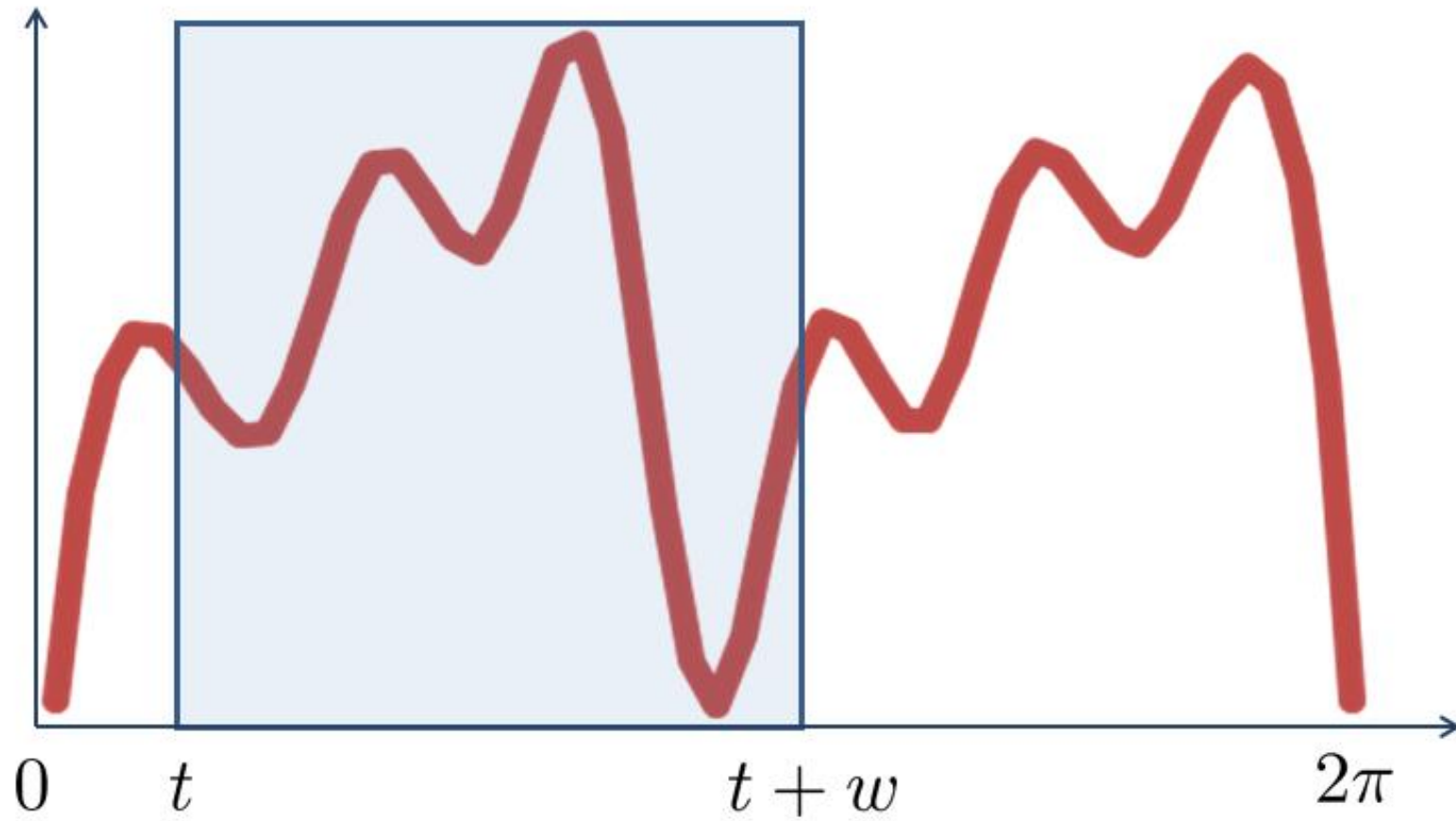




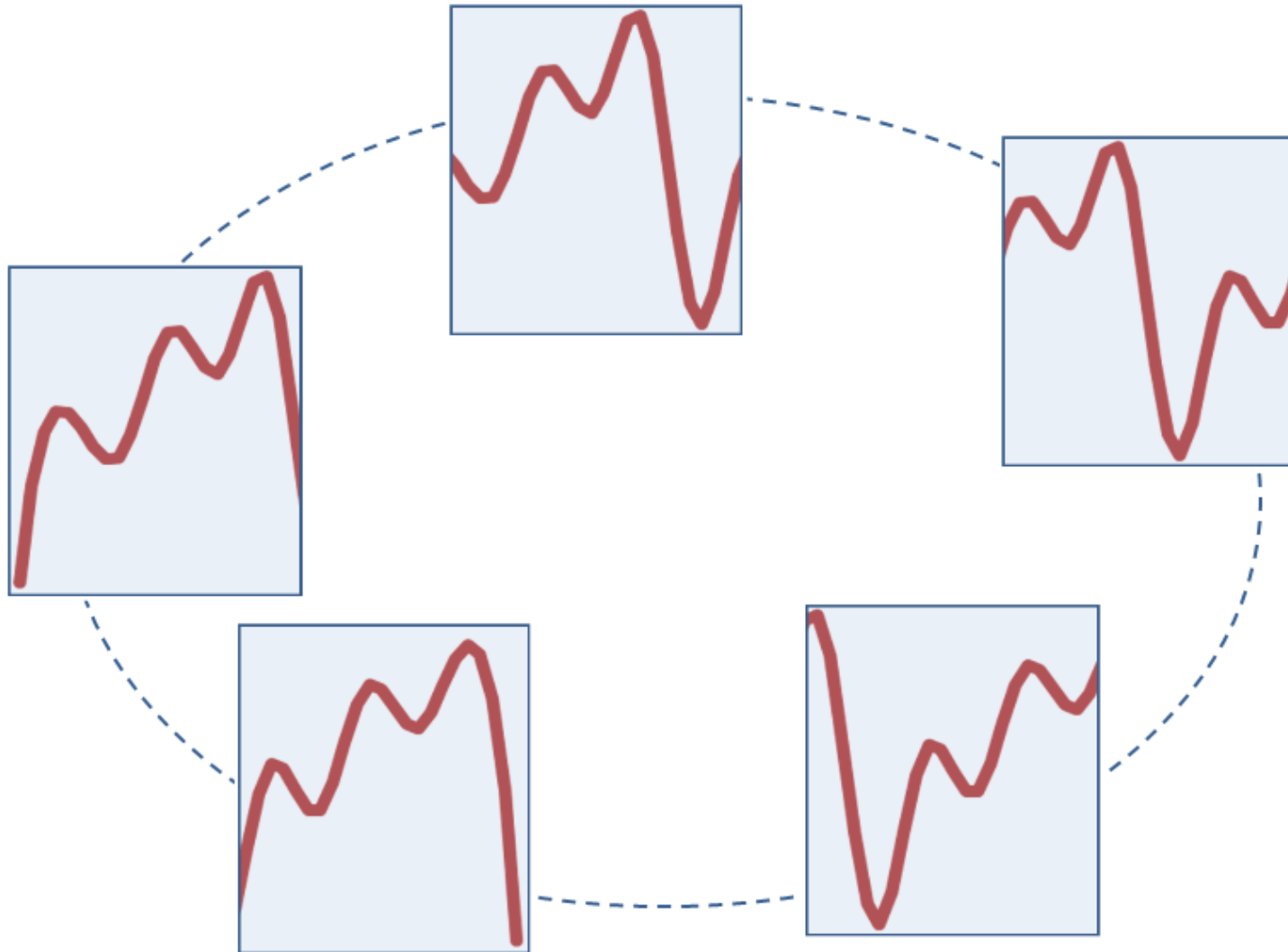
What is recurrence, and
how do we quantify it?



Sliding Windows

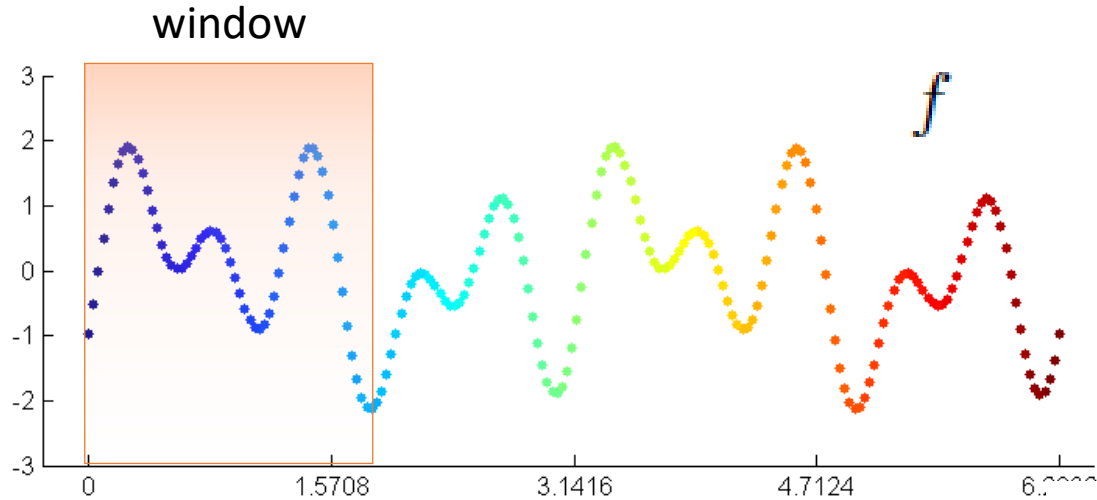


Sliding Windows



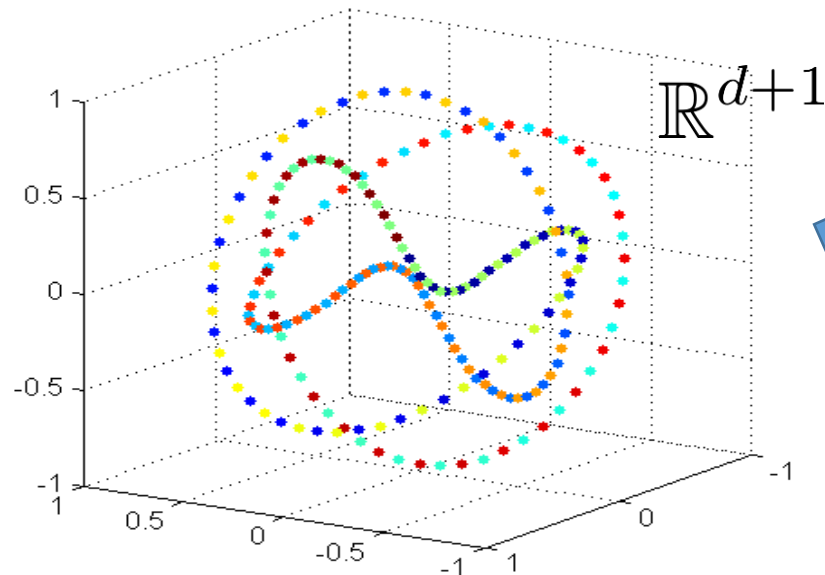
SW1PerS: Sliding Windows and 1-Persistence Scoring

Sliding window embedding



τ \rightarrow Step/delay
 $d\tau$ \rightarrow Window size
 $d + 1$ \rightarrow Dimension

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

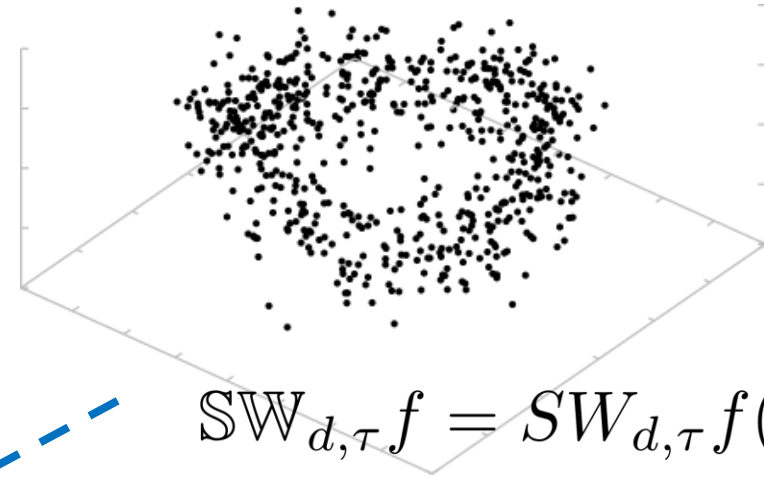
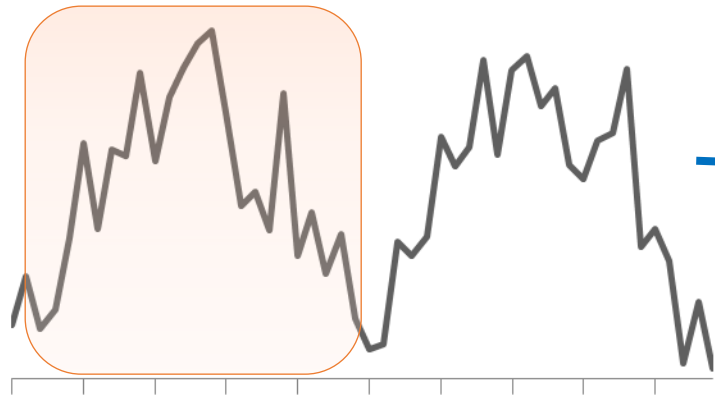


Sliding window point-cloud

$$SW_{d,\tau} f \parallel SW_{d,\tau} f(I)$$

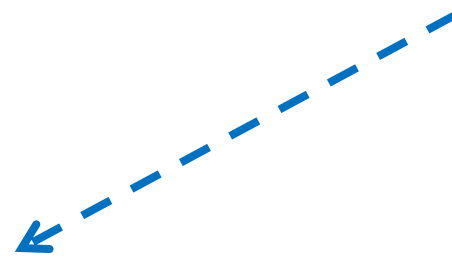
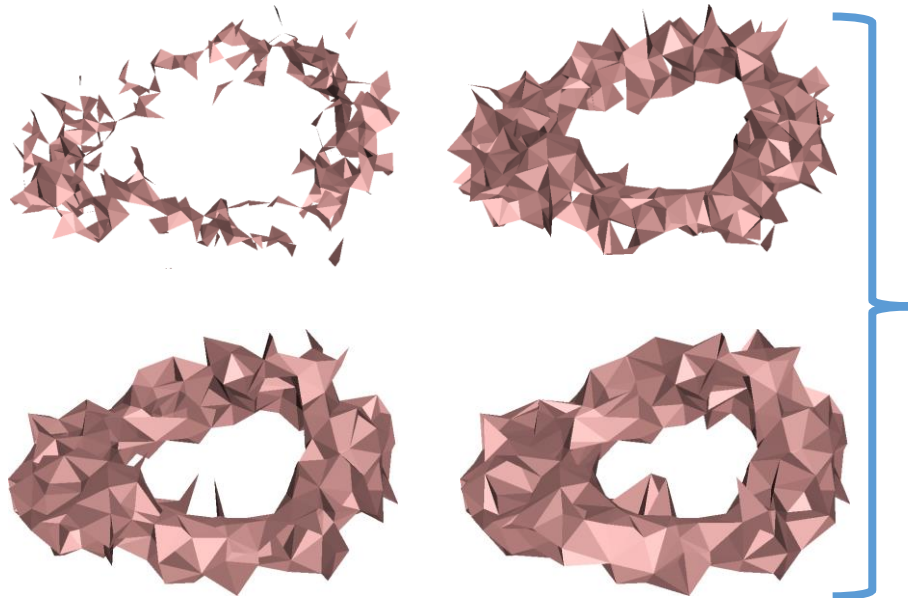
$$I \subset \mathbb{R}$$

Sliding window point-cloud



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

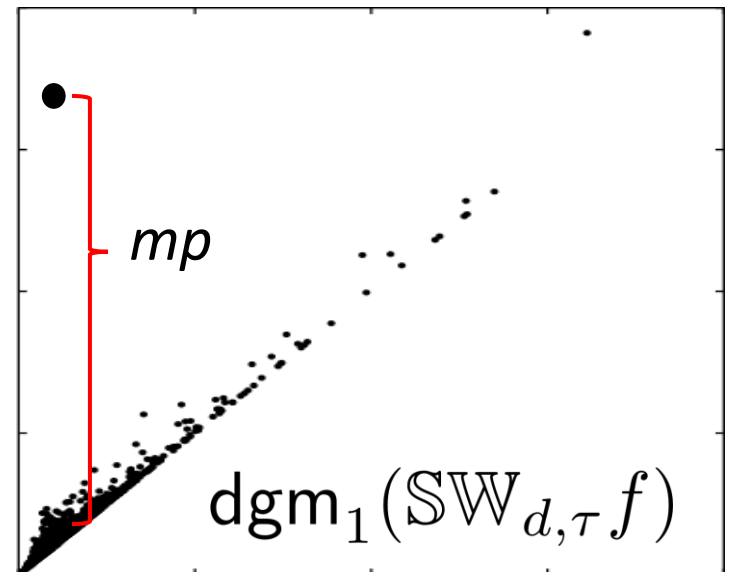
$\{R_\epsilon(X)\}_{\epsilon \geq 0}$ Rips complexes



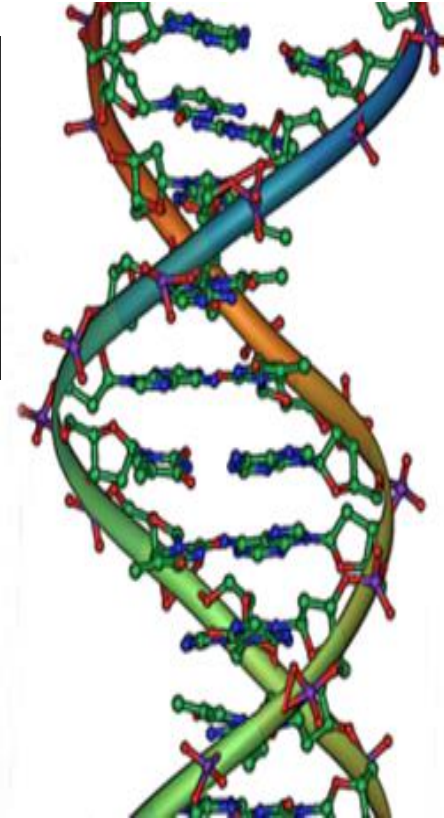
dgm



1D- Persistence Diagram



Biological Clocks



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 Deckard³, Steve B. Haase^{4,5} and John Harer^{1,4,6}

FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW DESIGN

Clipboard: Paste, Cut, Copy, Format Painter

Font: Avenir Next Re, 11, Bold, Italic, Underline, Text Color, Background Color

Alignment: Wrap Text, Merge & Center

Number: General, \$, %, , , .00, .00












Styles: Conditional Formatting, Format as Table, Cell Styles

Cells: Insert, Delete, Format

Editing: AutoSum, Fill, Clear, Sort & Find & Filter, Select

M54

fx

	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG
1	Norm PI	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
2		11895.2	12898	11911.3	10194.2	11634.3	9322.2	11362.2	9880.3	11402.3	9936.2	8346.8	9777.8	10107.9	6893.2	7986.2	8590.2	8712	9665.2	9078.2	8100.0
3		19494.2	17352.1	18554.1	18380.1	19133.7	17574.4	16778.7	14786.6	17291.2	17252.2	13845.6	16773.6	18236.5	17100.1	17017.1	18457.5	16770.5	18643.4	21593.4	18700.0
4		25261.2	24843.7	23924.7	27841	24070.2	26161.9	23777.7	27511.9	21394.7	22450.8	24288.4	26018.1	26731.8	22372.5	22325.2	22496	23327.2	24087.7	24628	24800.0
5		3117.3	3732.7	3964.2	3111.2	3379	2959.5	2460.3	3250.3	2853	2272.1	3588.5	2691.6	2488.9	1873.1	1733.9	1710.5	2064.8	1919.4	2168.1	17500.0
6		5019.7	4949.6	4661.2	4852.6	4287.6	3668.6	4460.2	4930.9	3365.7	3093.2	3951.8	5614.6	6065.7	3268.9	4349.1	3656.8	5440.7	5067.7	5025.1	5500.0
7		2291.9	1924.7	1974.4	2114.4	2601.3	2184	2781.6	2603.2	2565.3	2421.2	2609.9	2189	2186.3	2167	2106	1908.3	2348.1	2218.4	2171.6	2200.0
8		31688.7	32265.8	34335.7	32932.3	34843.3	28885.1	26908.4	28641.3	26995.8	29937.8	32894.9	32400.3	33468.2	26061.7	25781	24484.8	29700.4	31629	32161	25600.0
9		32270.2	32399.8	30527.7	33447	28200.7	26104.6	24098.6	27367.5	17303.9	17192.4	18305.2	29987.3	28857.8	18378.8	22644.6	22858.7	26571.7	27955.2	30859	23900.0
10		12132.3	12016.1	12976.3	11842.6	11746.3	12261.7	11203.4	9859.5	11791.3	13934.8	13435.6	12962.6	14130.5	15687.7	15616.2	14913.9	16606.3	16124.5	16905	16000.0
11		27906.8	24578	28547.9	25954.4	27405	28009.8	26748.6	26927.5	23337.6	25312.5	27580.3	26473.5	28164	25583.9	27012.6	25628.8	28332.6	29554	27519.3	27600.0
12		4213.7	3803.1	4142	3671.7	4110.5	3930.1	4336.3	3128.7	4498.3	4249.8	4238.2	3936.1	3804.3	4242.1	4079.7	3252.3	4084.6	4041	4102.6	3700.0

hughes-liver-v1_swft3_rescomb_d

FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW

Paste Cut Copy Format Painter Clipboard

Avenir Next Re 11 Font

Wrap Text Alignment

General Number

Conditional Formatting Styles

Cell Styles

Insert Delete Format Cells

AutoSum Fill Clear Editing

Sort & Find & Filter Select

A2

	A	C	E	G	I	K	L	M	N	O	P	Q	R
	Probe	Symbol	SW_rank	DL_rank	LS_rank	JTK_rank	Max-Min	Norm Plot	18	19	20	21	22
1													
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

FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW

Clipboard: Paste, Cut, Copy, Format Painter

Font: Avenir Next Re, 11, Bold, Italic, Underline, Color

Alignment: Wrap Text, Merge & Center

Number: General, \$, %, , €, .00, .0

Styles: Conditional Formatting, Format as Table, Cell Styles

Cells: Insert, Delete, Format


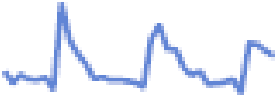


Editing: AutoSum, Fill, Clear, Sort & Find & Filter, Select

A2: 1450869_at

	A	C	E	G	I	K	L	M	N	O	P	Q	R
	Probe	Symbol	SW_rank	DL_rank	LS_rank	JTK_rank	Max-Min	Norm Plot	18	19	20	21	22
45096	1459877_x_at		44920	39869.5	40494	37728.5	2233.2		4848.7	4365.1	4505.4	4685.3	4858
45097	1459917_at	Ggnbp2	44920	24121	30730.5	37728.5	1659.9		1132.3	1077.1	1122.8	1250	1164
45098	1459948_at		44920	6142	21561	37728.5	7536.4		3512.8	3705.9	4896	2755.5	2715
45099	1459957_at		44920	27801.5	26601.5	21692.5	172.4		144.6	149.9	111.6	127.9	122.7
45100	1460126_at		44920	27184	40494	37728.5	1024		1065.1	734.9	986.6	810	883.5
45101	1460610_at	Ost4	44920	39204.5	31251	37728.5	55.7		95	94	95	95.3	90.4
45102	FX-MURINE_b1	NA	44920	29424	21164.5	29740	12663.2		12215.2	10674.9	10611.6	11054.7	13470.4

hughes-liver-v1_swft3_rescomb_d

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

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



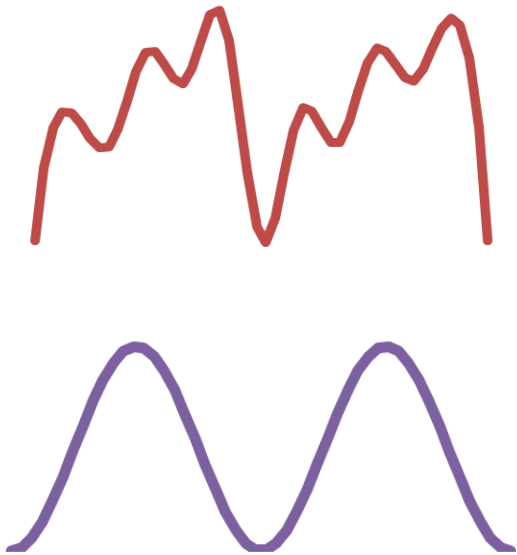
Table 3 Number and percentage of probes in the top 10% of rankings from each algorithm that are in a consensus set. That is, those which appear in the top 10% of rankings for at least three algorithms.

SW1PerS:

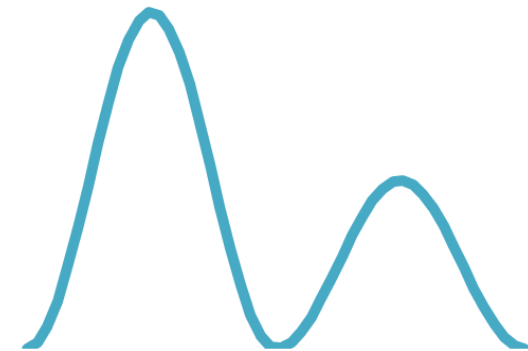
Sliding Windows and 1-Persistence Scoring



Shape agnostic



Resistant to dampening



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

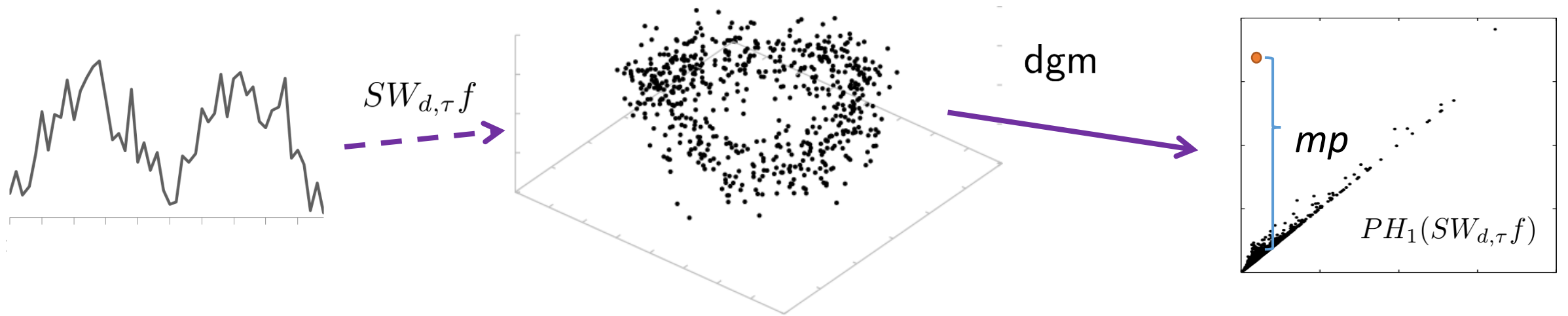


CrossMark

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)



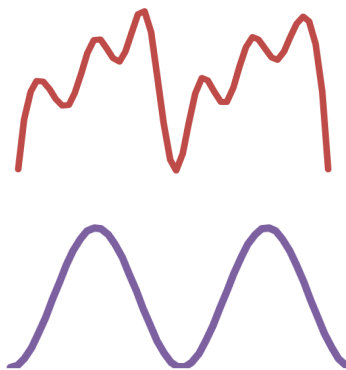
$N \sim$ Sample size

$$d = 2N$$

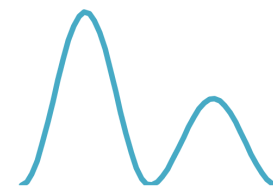
$L =$ number of periods

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

Shape agnostic



Resistant to dampening



mp : maximum persistence
(recurrence score)

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

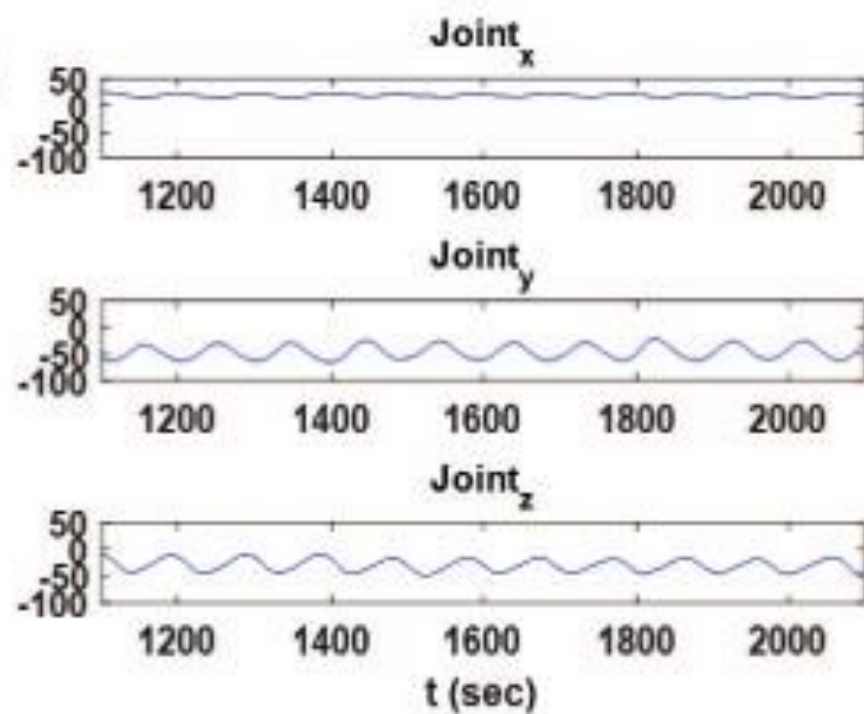
(a)



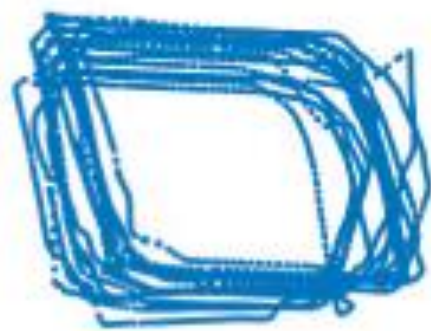
(b)



(c)



Delay Embedding



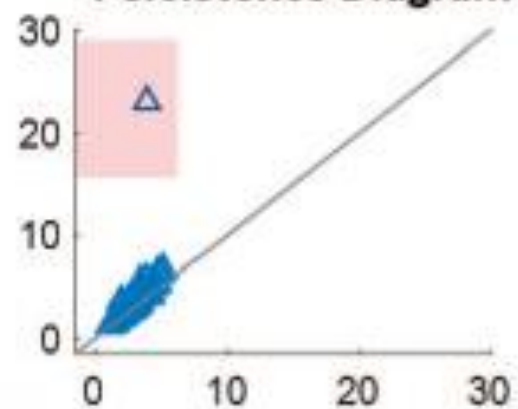
(d)

Subsampled PC



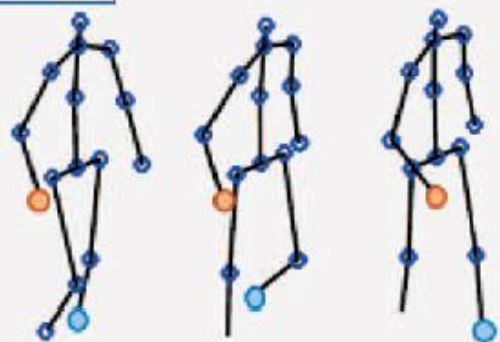
(e)

Persistence Diagram

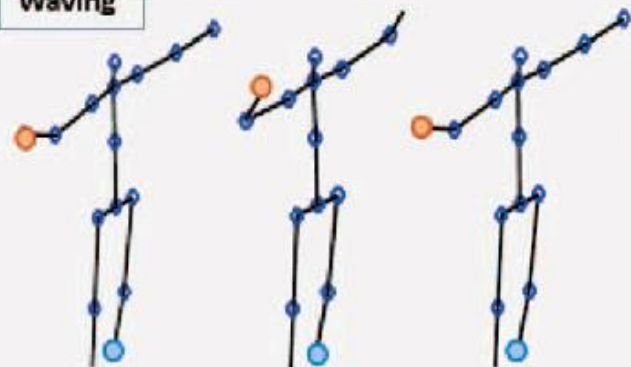


(f)

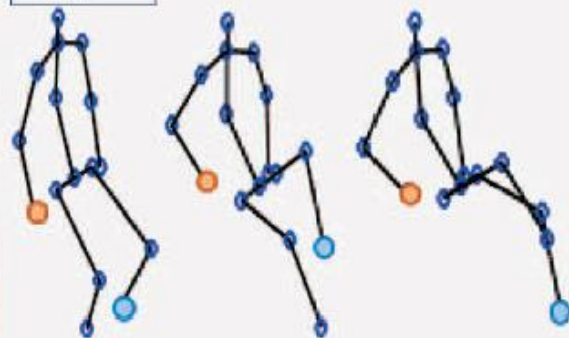
Walking



Waving



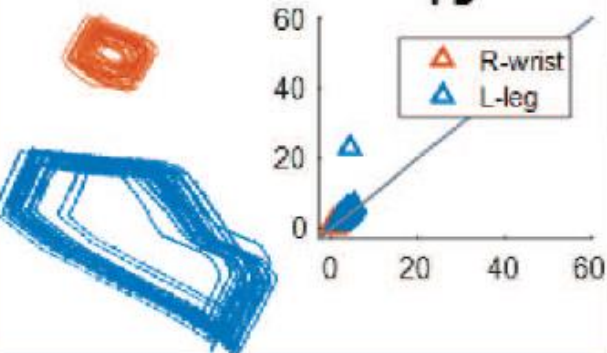
Bicycle



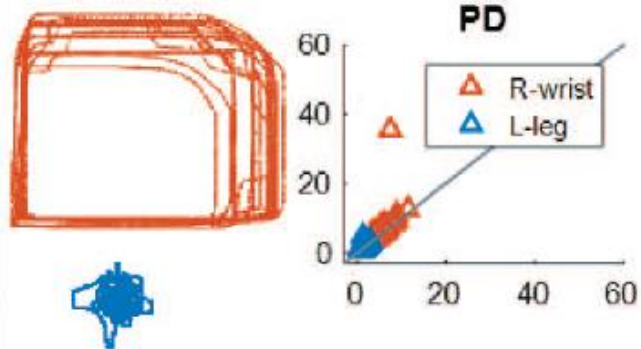
Golfing



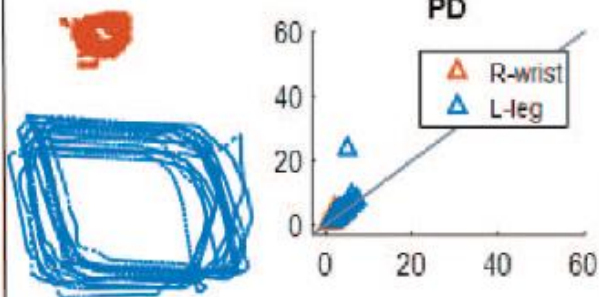
PD



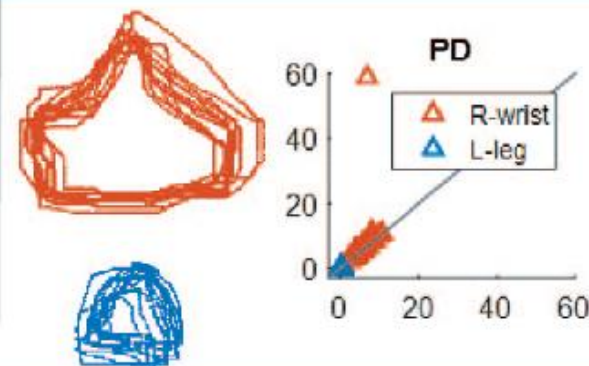
PD



PD



PD



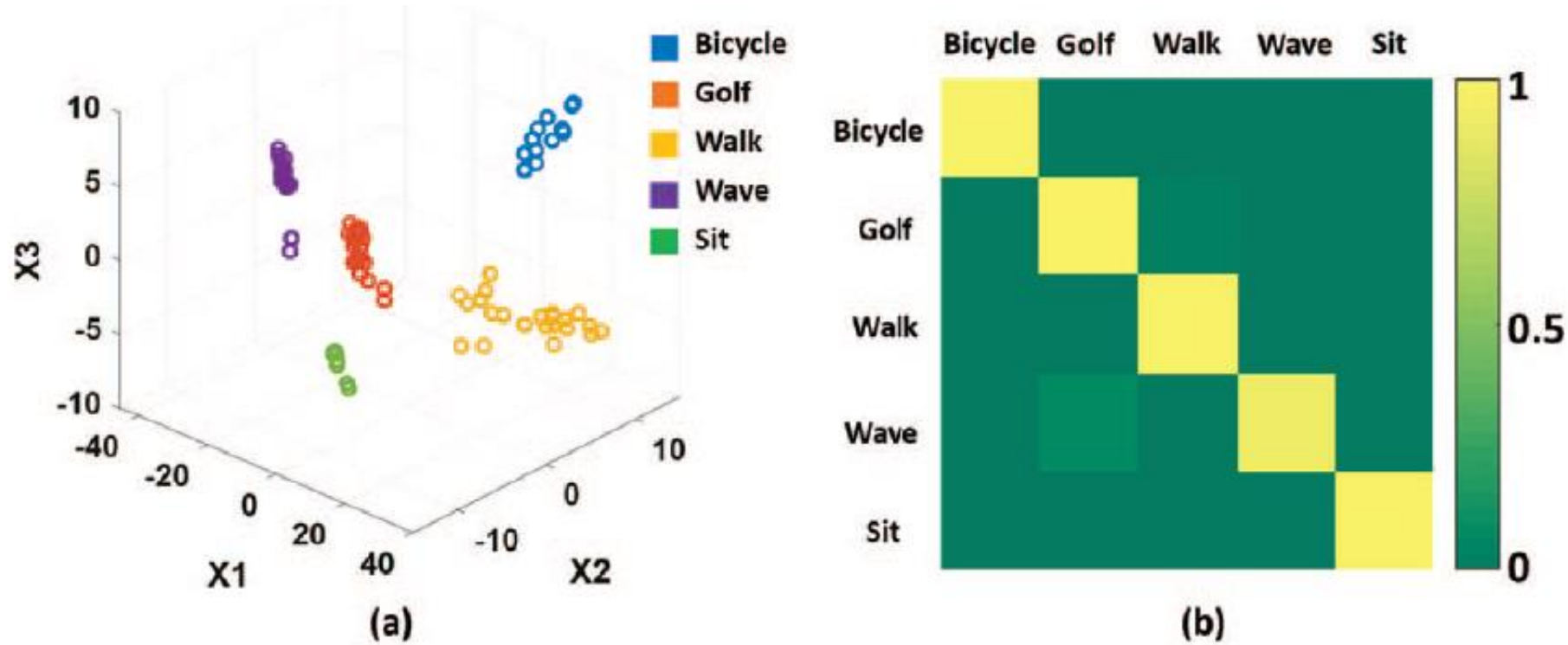


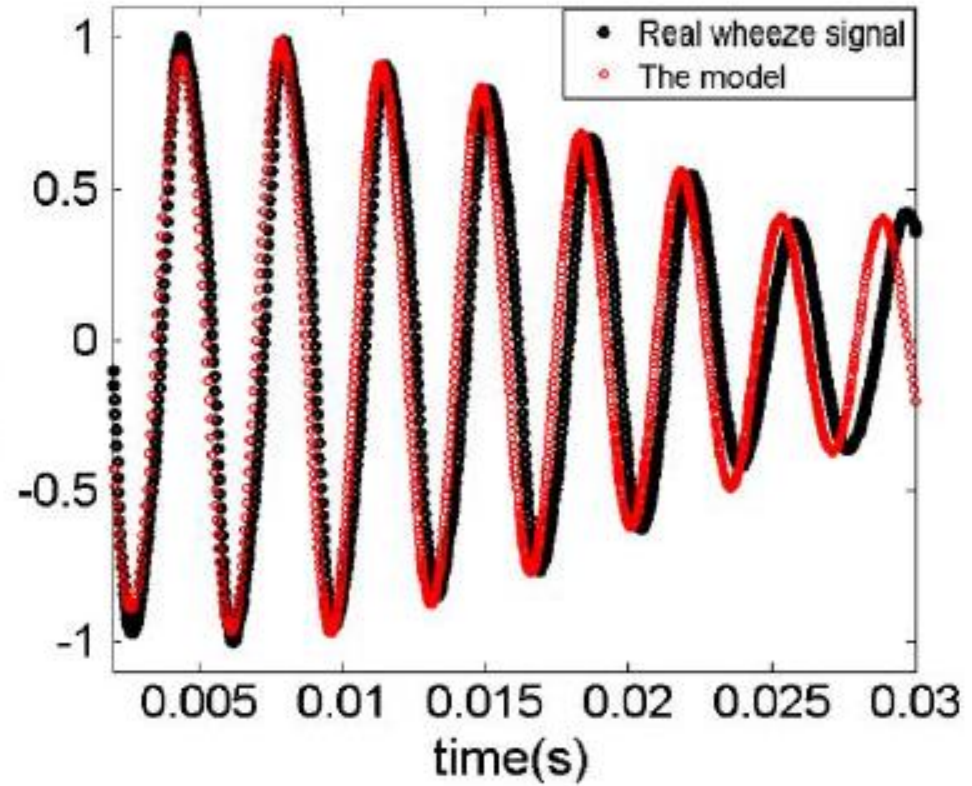
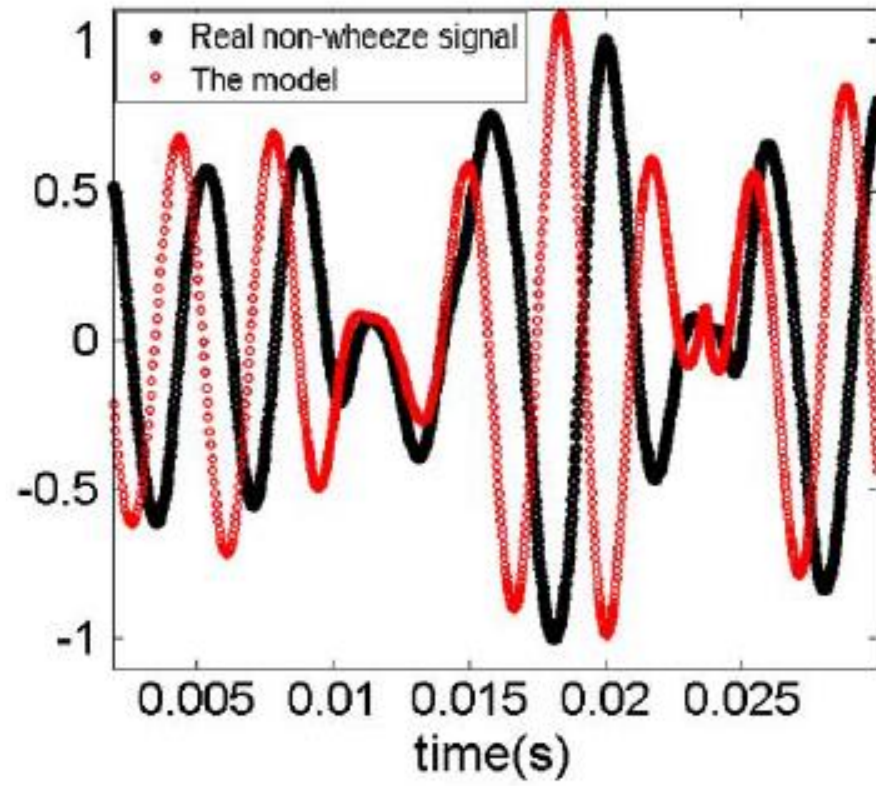
Fig. 4. (a) Separation of classes from the training set, (b) Confusion matrix over the predicted and true classes

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

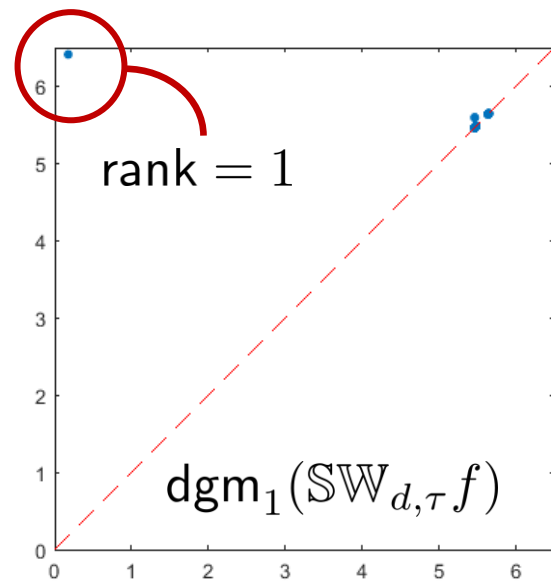
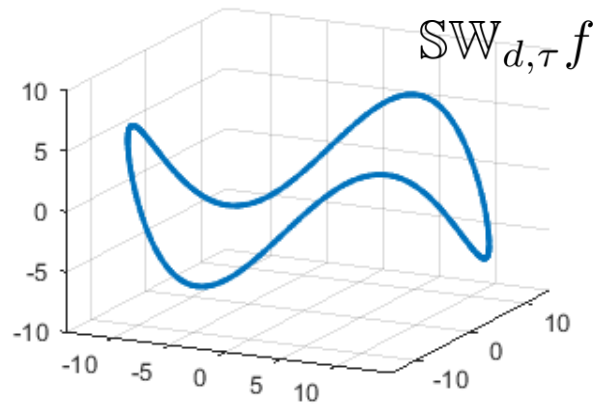
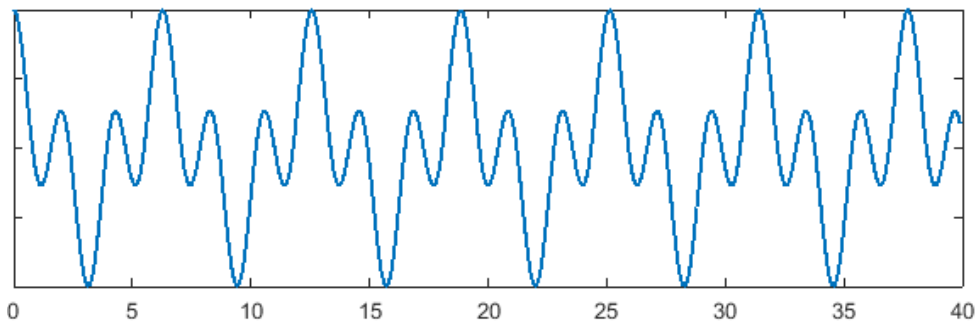
Saba Emrani, Thanos Gentimis, and Hamid Krim, *Fellow, IEEE*



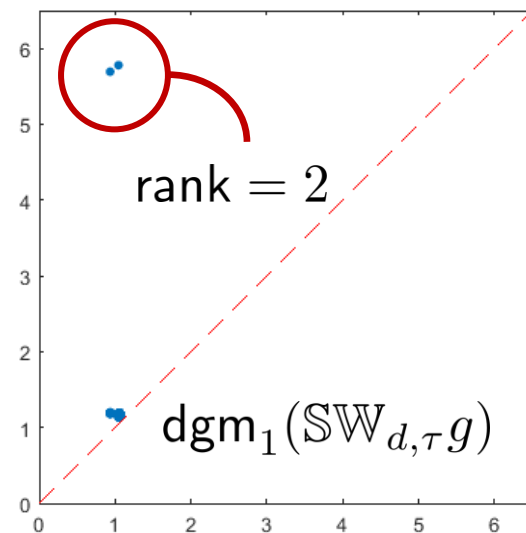
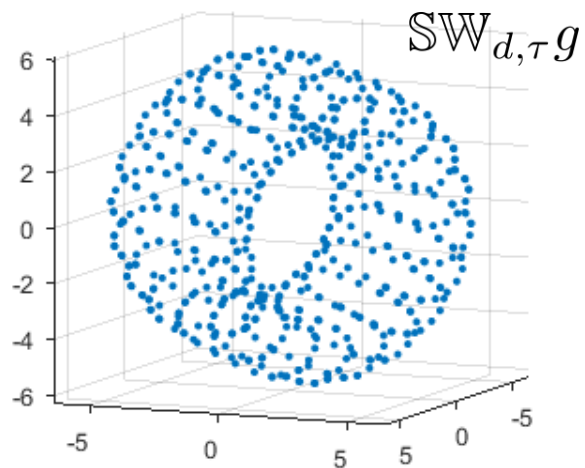
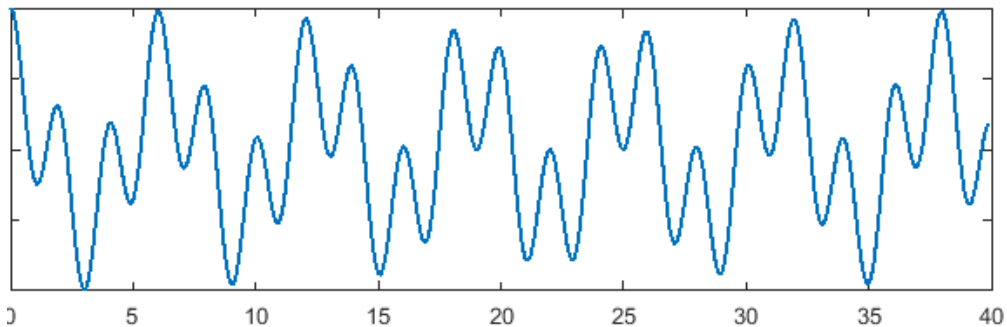
The accuracy of our proposed technique is 98.39% while the accuracy of the techniques proposed in [8] and [9] are 86.2% and 95.5%, respectively.

Geometric QuasiPeriodicity

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



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



Time Series

Sliding Window Point Cloud

Persistent Homology



Audio for this story is unavailable.

MUSIC LISTS

The Unsettling Sound Of Tritones, The Devil's Interval

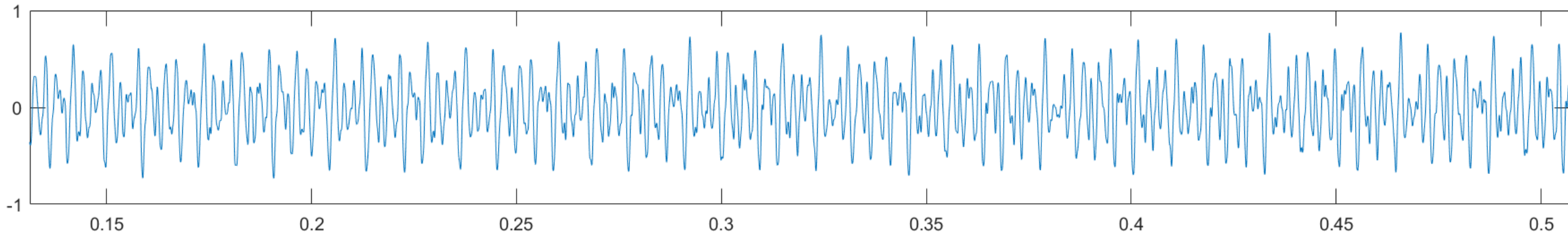
(joint with Adam Huston)

October 31, 2017 · 10:52 AM ET

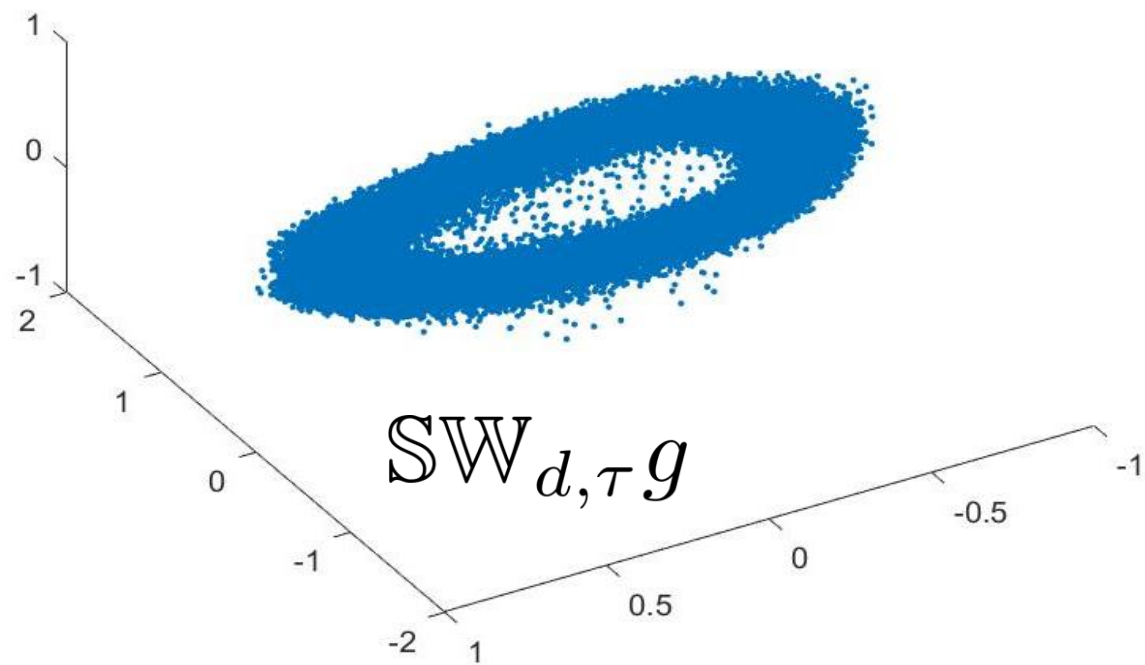
JUDITH KOGAN



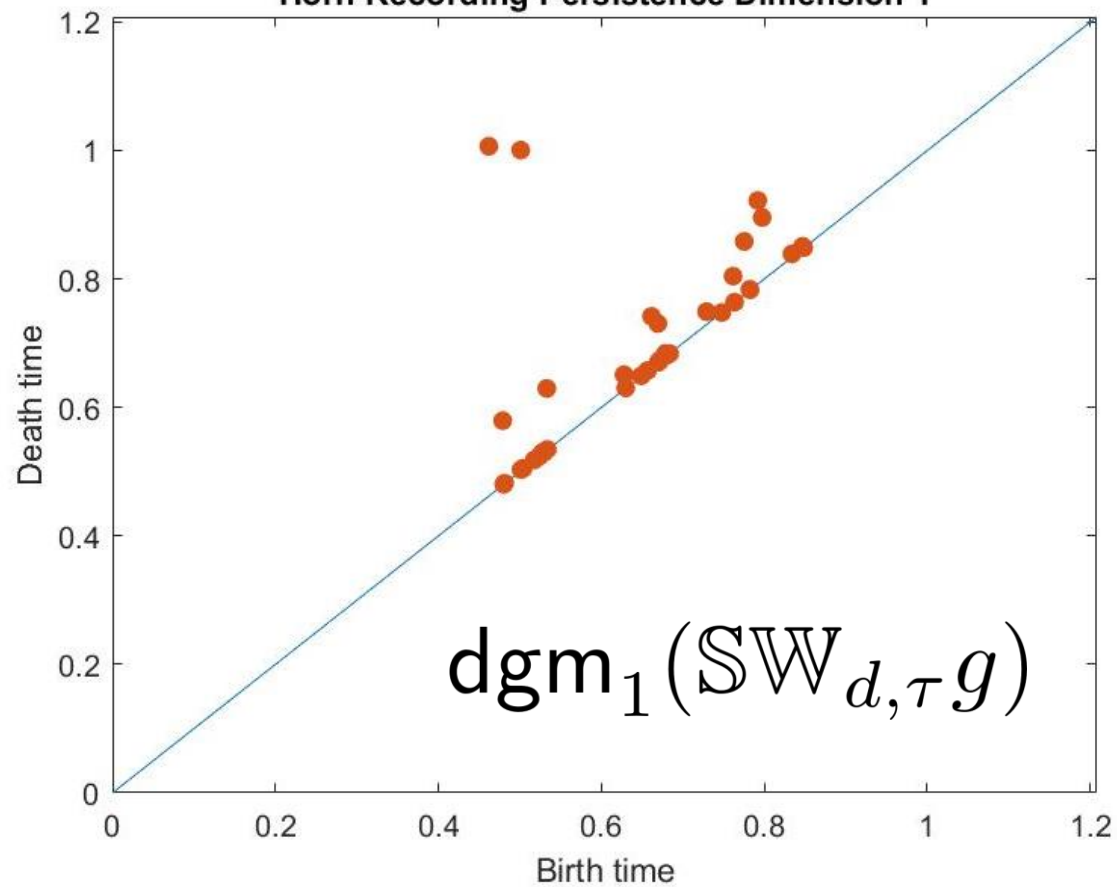
Frans Francken's *Death Playing the Violin*. In music theory, the tritone came to be known as the devil's interval.



PCA Representation of Horn Recording



Horn Recording Persistence Dimension 1



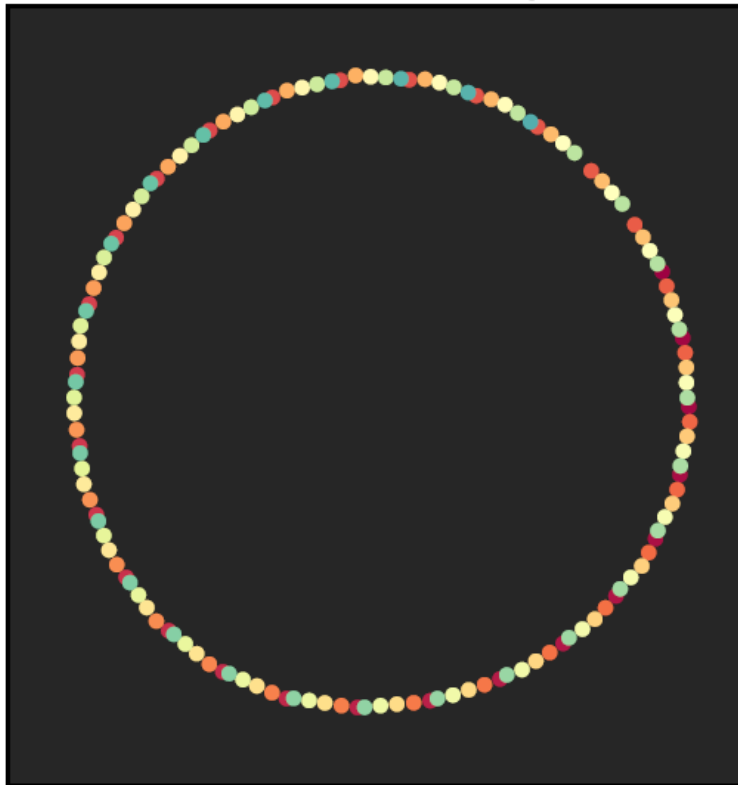
(Quasi)Periodicity Quantification in Video Data, Using Topology*

Christopher J. Tralie[†] and Jose A. Perea[‡]

Recurrence in video data (SW1PerS-video)

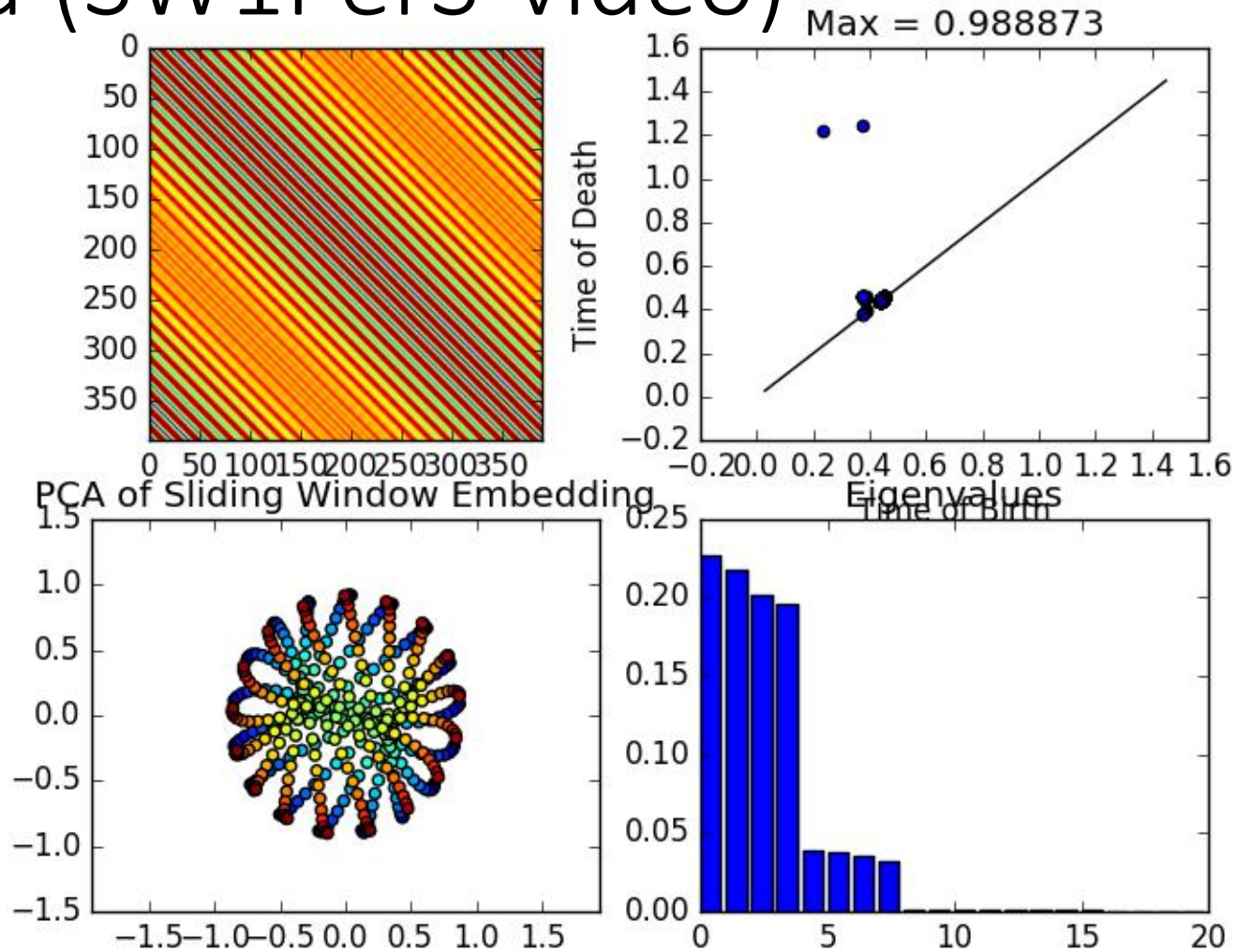
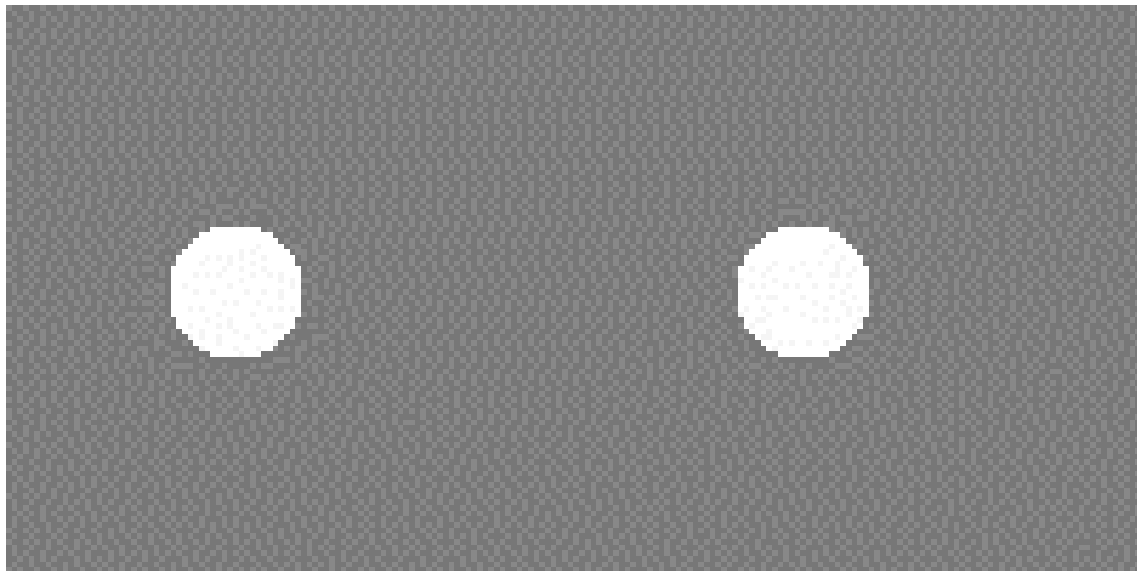


2D PCA, $\tau = 1$, $d = 25$
61.5 % Variance Explained

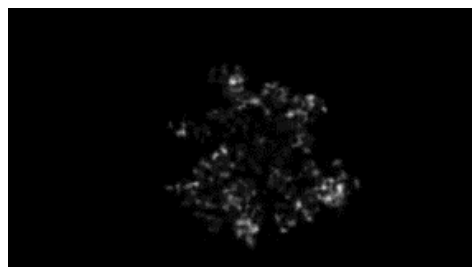
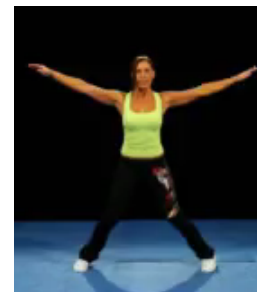
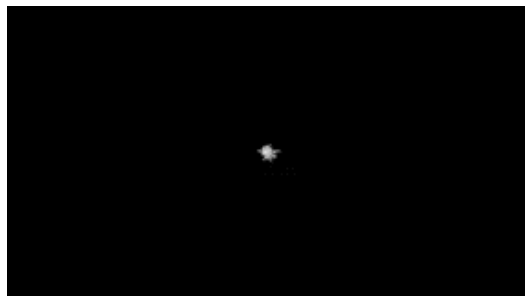
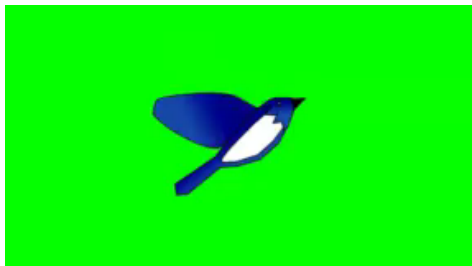


Sliding window
embedding

Recurrence in video data (SW1PerS-video)



Experiments

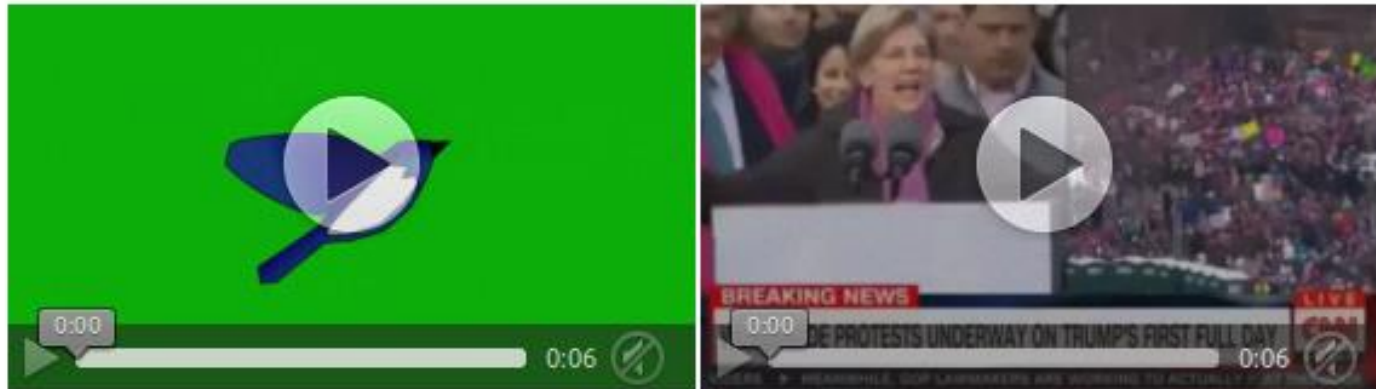


C. Tralie and J. Perea, *(Quasi)Periodicity Quantification in Video Data, Using Topology*, 2018.

Experiment: **amazon** Mechanical Turk

Instructions

There are two 5 second videos below. Enter the 3 digit number at the end of the video which has more perfect repetitions of motion both in time and location within the video frame.



Submit

Results: Humans (Amazon turk) vs Computers

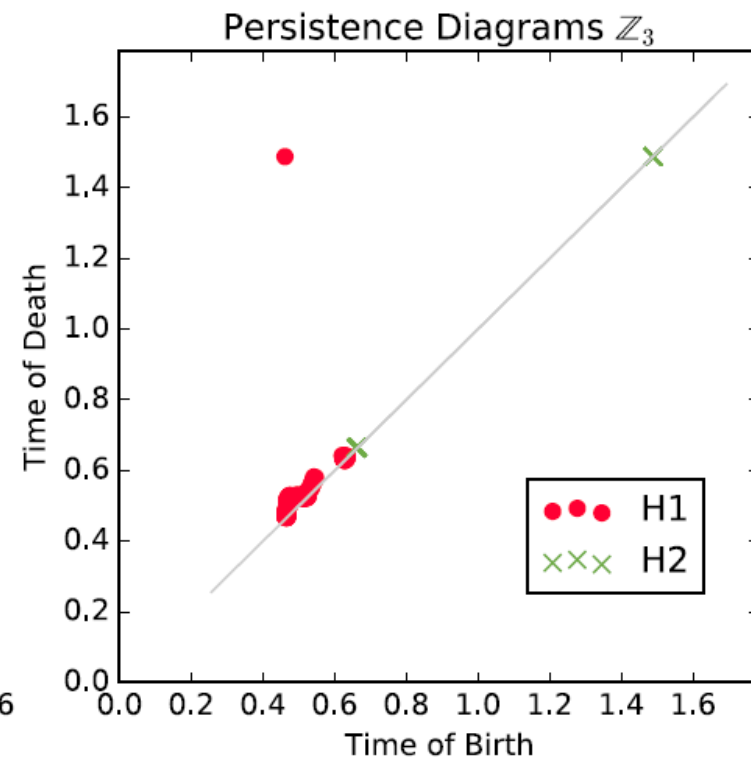
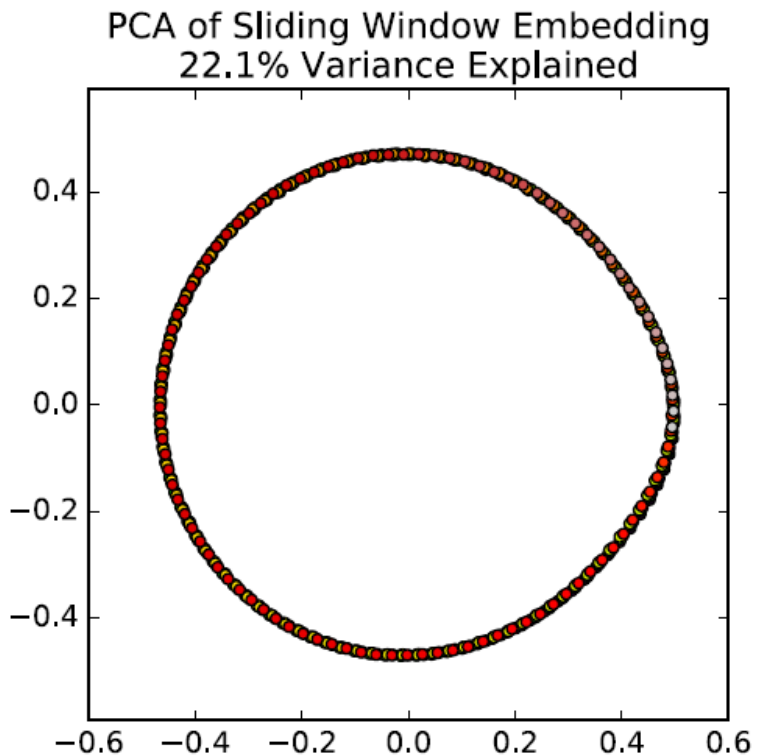
Correlation of rankings
(from most periodic to least periodic)
across 20 videos

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

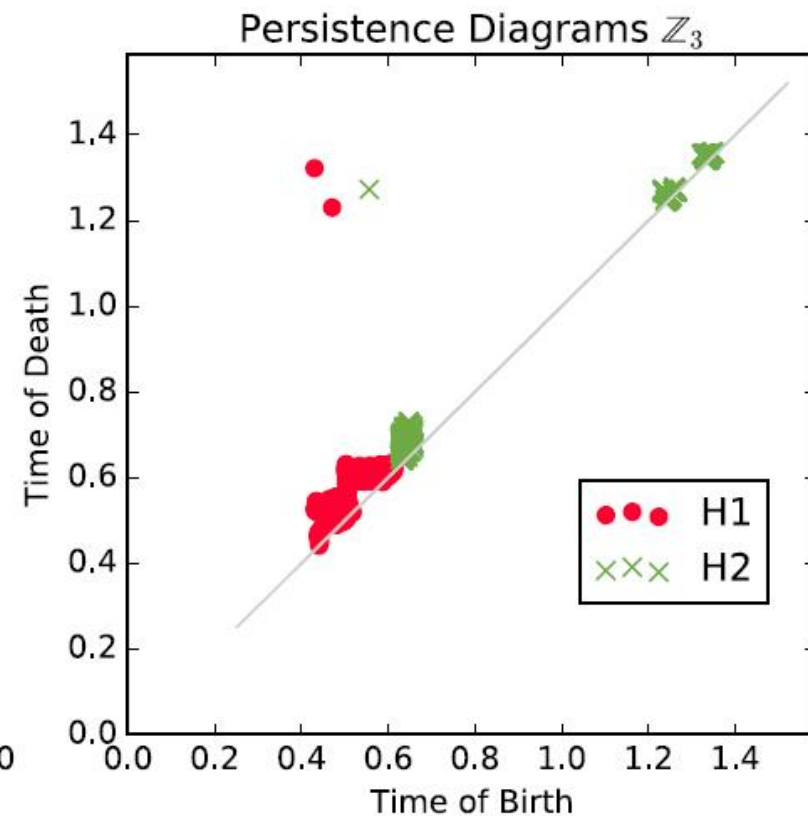
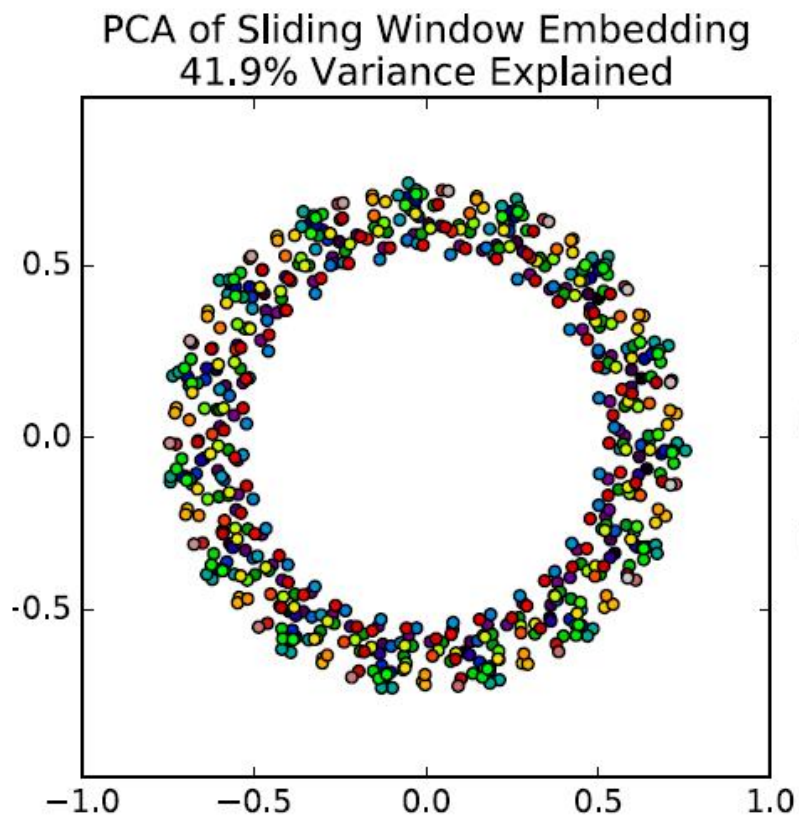
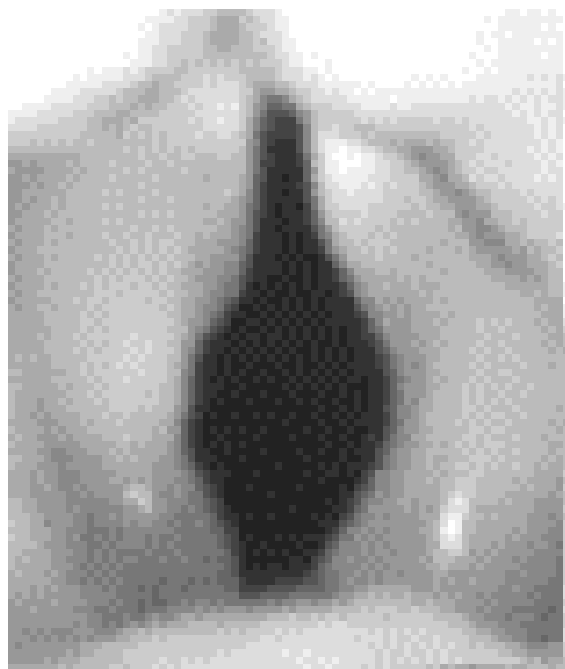
Recurrence in video data (SW1PerS-video)



normal



Recurrence in video data (SW1PerS-video)

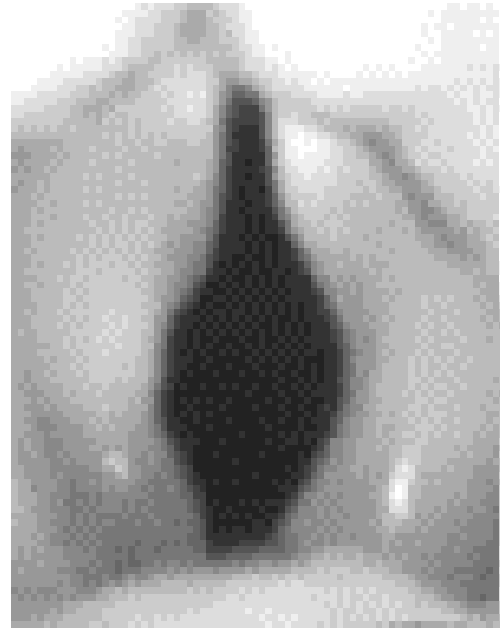


Clinical asymmetry

Laryngeal video-endoscopy



normal



Clinical
asymmetry



Mucus irregular



AP-Biphonation

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