Mobius Beats: The Twisted Spaces of Sliding Window Audio Novelty Functions with Rhythmic Subdivisions

Christopher J. Tralie. ctralie@alumni.princeton.edu
Postdoctoral Associate, Department of Mathematics, Duke University

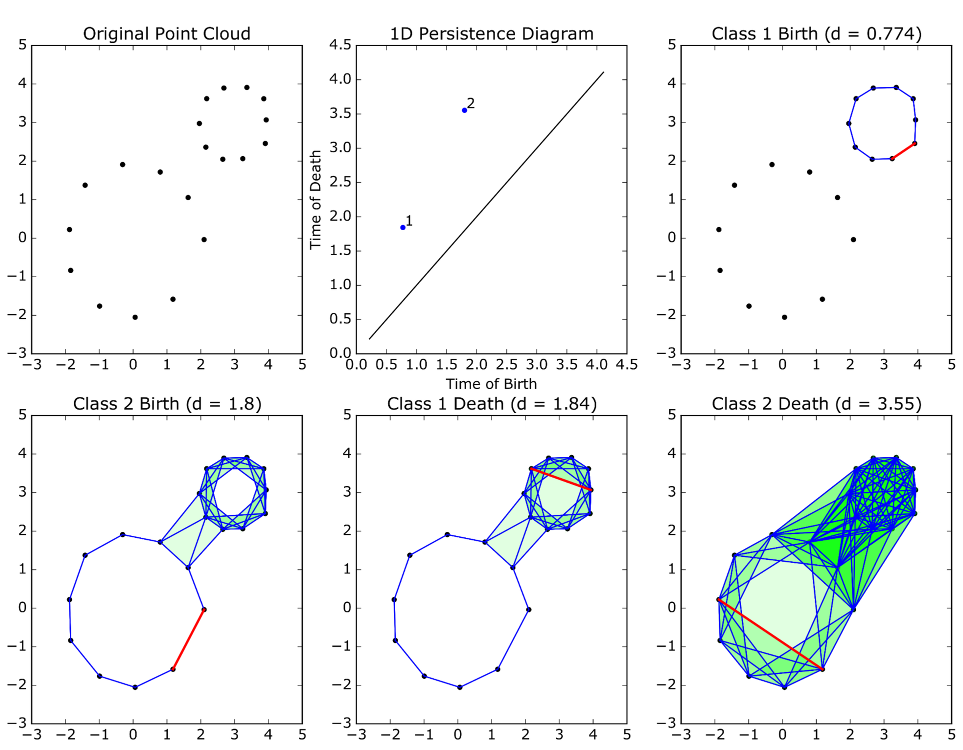
John Harer, harer@math.duke.edu
Professor, Department of Mathematics, Duke University
Chief Scientist / CEO, Geometric Data Analytics

Abstract

In this work, we show that the sliding window embeddings of certain audio novelty functions (ANFs) representing songs with rhythmic subdivisions concentrate on the boundary of non-orientable surfaces such as the Mobius strip. This insight provides a radically different *topological* approach to classifying types of rhythm hierarchies. In particular, we use tools from topological data analysis (TDA) to detect subdivisions, and we use thresholds derived from TDA to build graphs at different scales. The Laplacian eigenvectors of these graphs contain information which can be used to estimate tempos of the subdivisions. We show a proof of concept example on two audio snippets from the MIREX tempo training dataset, and we hope in future work to find a place for this in other MIR pipelines.

Topological Data Analysis

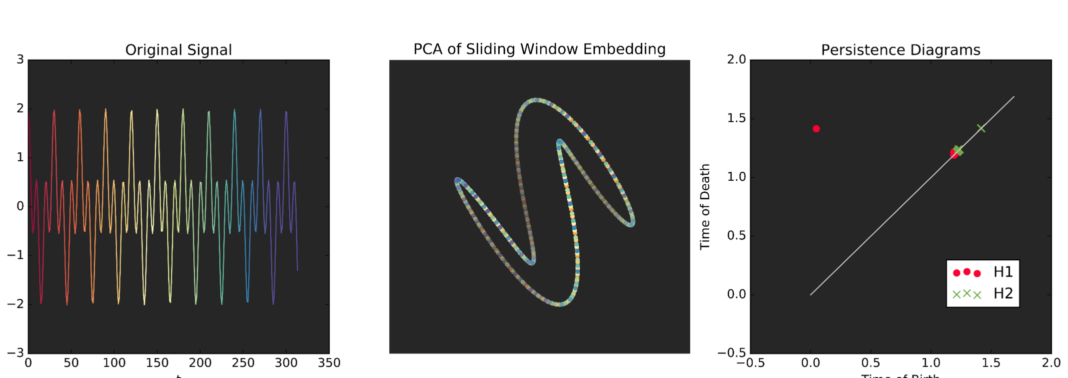
- Persistence diagrams capture multiscale topological structure in point clouds (see [2])
- Add edges with length less than or equal to a threshold in a Rips Filtration.
  - For 1D homology, loops classes are "born" at distance equal to first edge that completes them, and loop classes "die" when their boundaries are expressible as formal sums of triangle boundaries (a linear algebra problem)
  - Usually use binary coefficients in computations, but we use other fields in this work to pick up on "twists"



Sliding Window Embeddings

- Way of turning 1D time series into a point cloud
- Periodic signals (f(t) = f(t + kT)) turn into topological loops, but geometry may be complicated
- Can use persistence diagrams to quantify periodicity after embedding

$$S_{M,\tau}[f](t) = \begin{bmatrix} f(t) \\ f(t+\tau) \\ \vdots \\ f(t+M\tau) \end{bmatrix} \in \mathbb{R}^{M+1}$$



References/Code

[1] Hadar Averbuch-Elor and Daniel Cohen-Or. Ringit: Ring-ordering casual photos of a temporal event. ACM Trans. Graph., 34(3):33–1, 2015.
[2] Herbert Edelsbrunner and John Harer. Computational topology: an introduction. American Mathematical Soc., 2010.
[3] Daniel PW Ellis. Beat tracking by dynamic programming. Journal of New Music Research, 36(1):51–60, 2007.
[4] Chris Godsil and Gordon F Royle. Algebraic graph theory, volume 207. Springer Science & Business Media, 2013.
[5] Jose A Perea and John Harer. Sliding windows and persistence: An application of topological methods to signal analysis. Foundations of Computational Mathematics, 15(3):799–838, 2015.
[6] Sorkine, Olga. "Laplacian mesh processing." *Eurographics (STARs)*. 2005.
Please see our paper for a more complete list of references

Code
https://github.com/ctralie/GeometricBeatTracking

Acknowledgements

Chris Tralie and John Harer were supported under an NSF big data grant DKA-1447491. Chris Tralie was also supported under an NSF Graduate Fellowship NSF under grant DGF-1106401

Geometry of Sliding Windows of Subdivided Pulse Trains

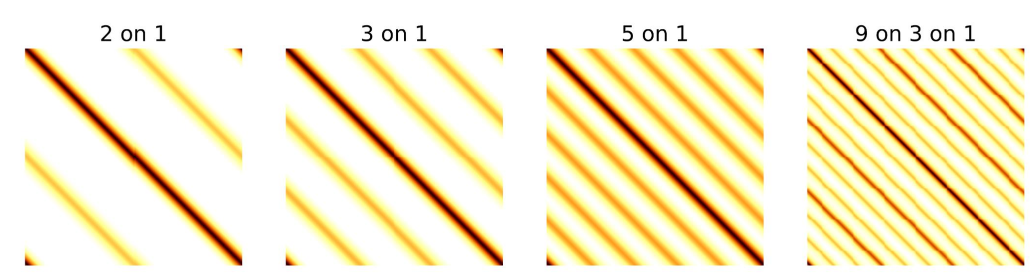
- Study synthetic model for audio novelty functions with subdivision
- k on 1 harmonic pulse trains are defined as:

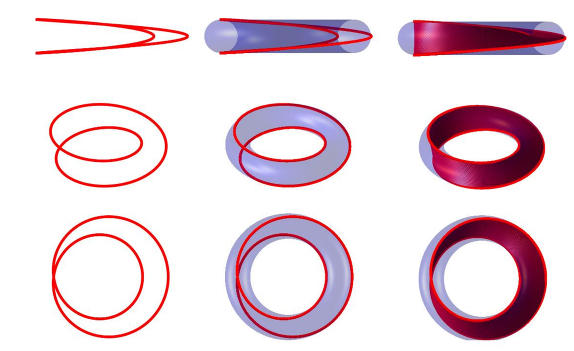
$$f(t) = a\delta(t(\text{mod}T)) + b\delta(kt(\text{mod}T))$$

- Distance functions between sliding windows of length 2T satisfy:

$$d(s,t) = \begin{Bmatrix} 0 & |s-t| = lT \\ 2|b-a| & |s-t| = l'T + T/k \\ \infty & \text{otherwise} \end{Bmatrix}$$

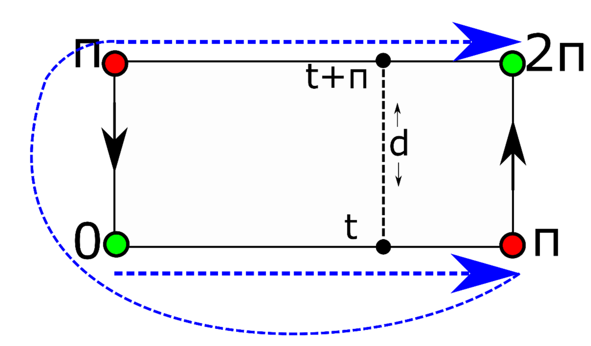
- Self-similarity matrices (SSMs) shown below for different cases



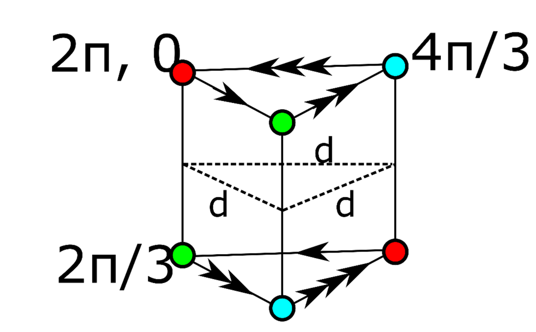


- Mobius strip is equally close to itself halfway through its boundary. The 2 on 1 geometry also has this property
- Similarly twisted geometries for other k values

2 on 1

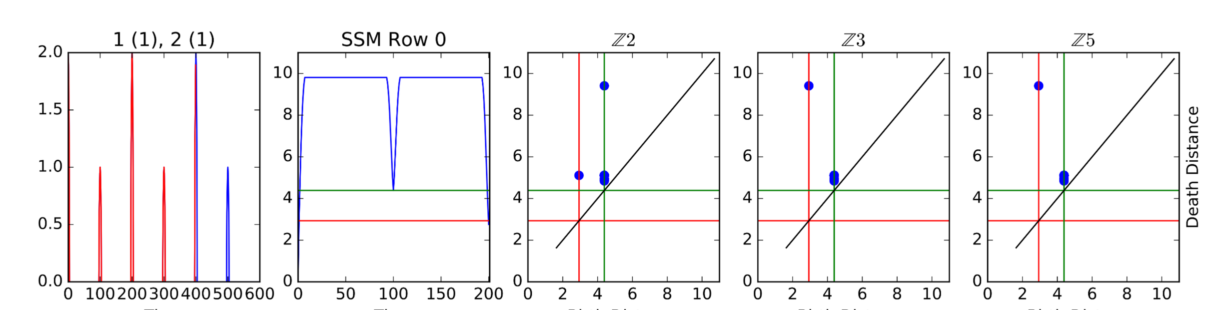


3 on 1

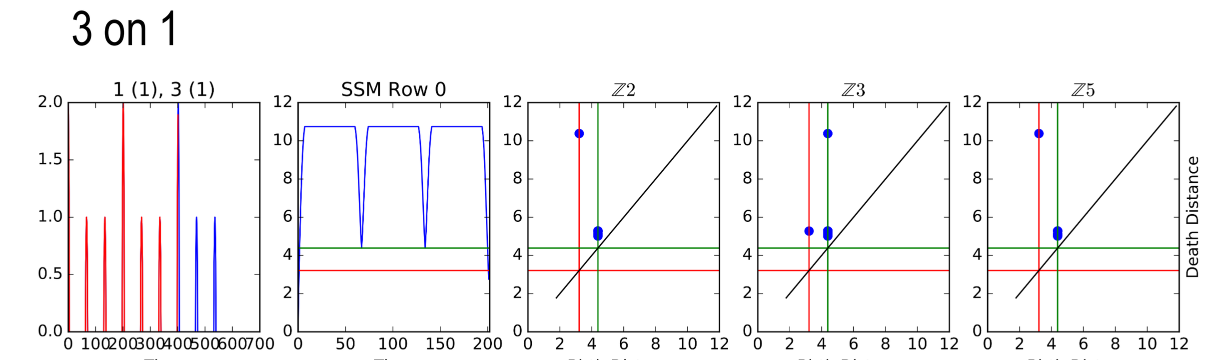


Persistence Diagrams of Synthetic Pulse Trains

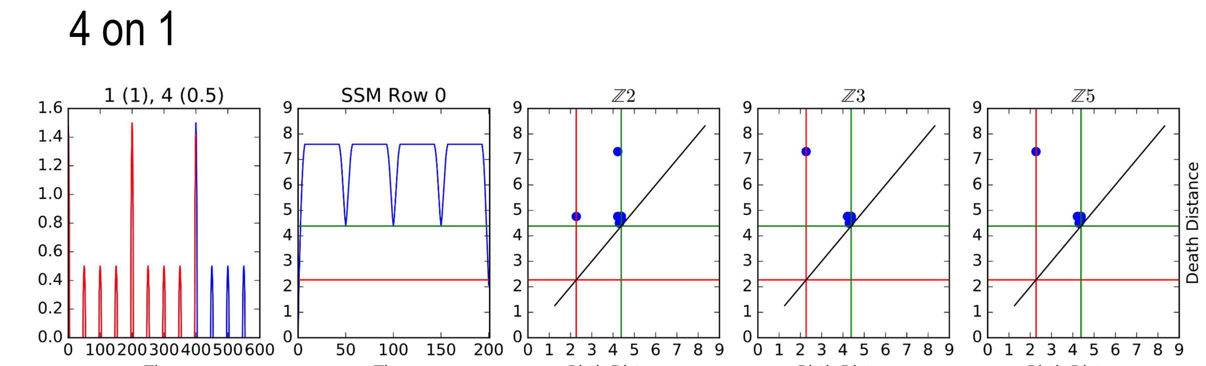
2 on 1



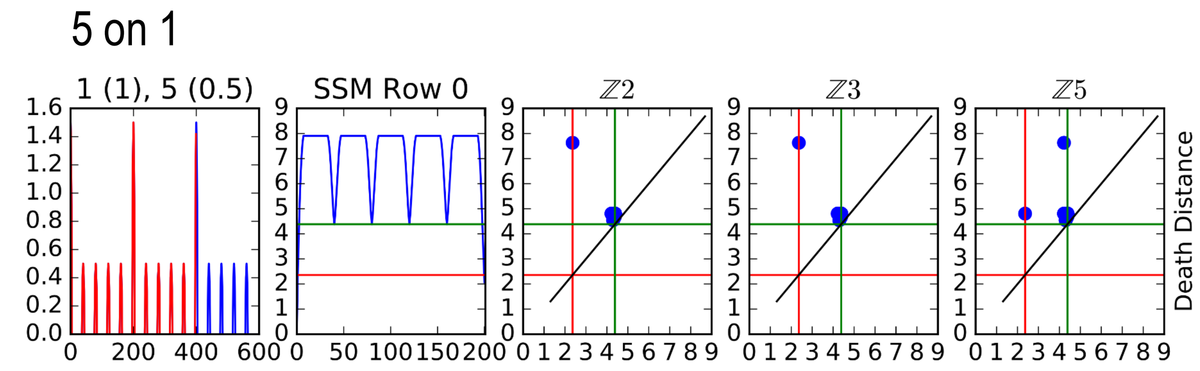
3 on 1



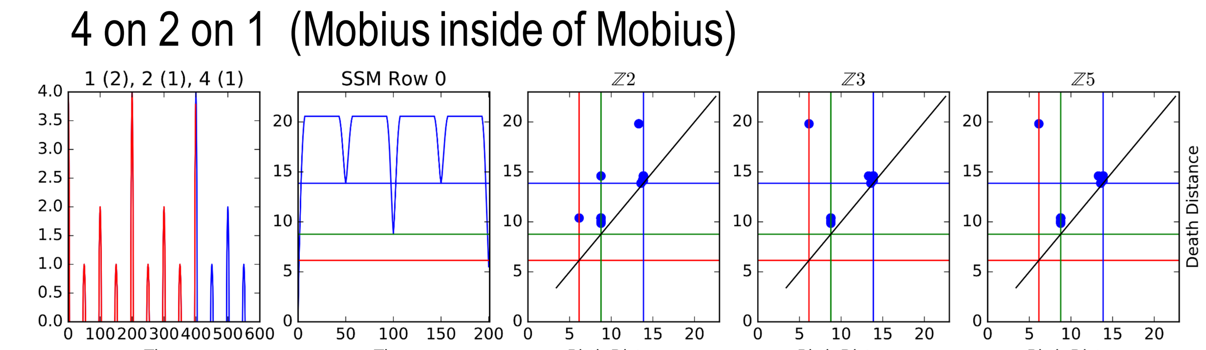
4 on 1



5 on 1

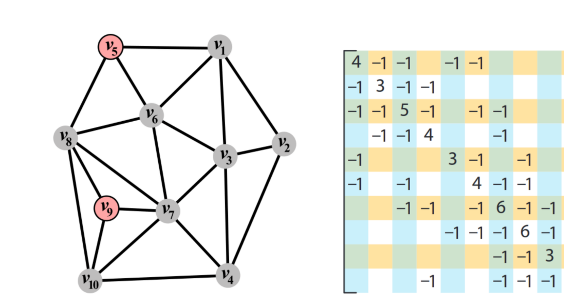


4 on 2 on 1 (Mobius inside of Mobius)



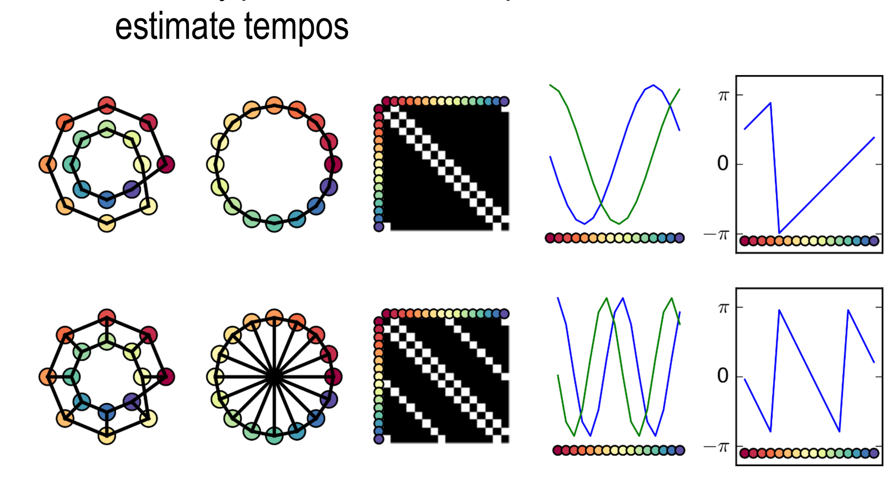
- Because of twists, field of coefficients can change persistence diagrams depending on subdivided structure.
- For k on 1 rhythms, a class from [a, b] splits into two classes [a, c] and [c, b] if the field is a prime factor of k
- We can use these observations to detect subdivision without explicitly parameterizing it

Laplacian Circular Coordinates

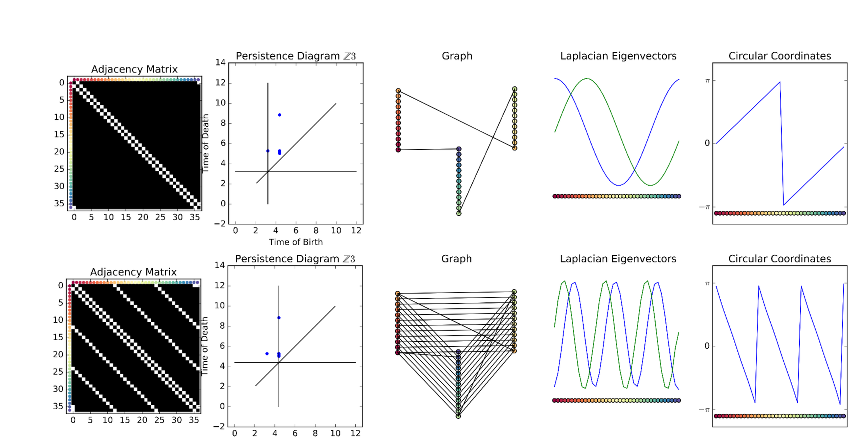


Above figure courtesy of [6]

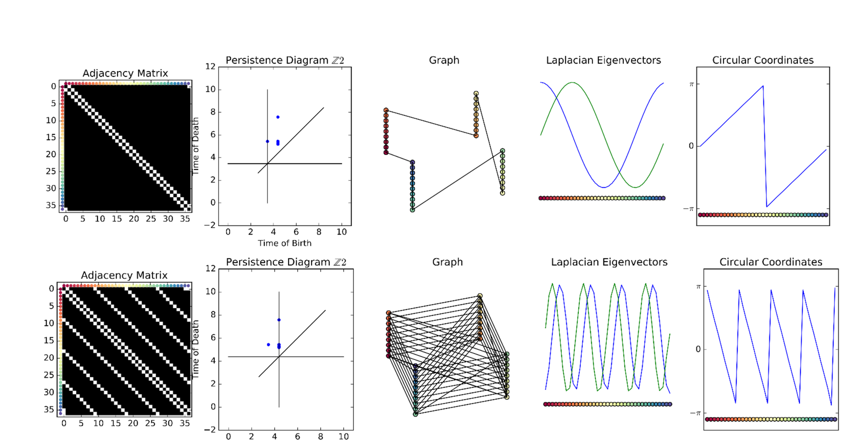
- Inspired by [1]
- Use eigenvectors of graph Laplacian **L = D – A** to parameterize sliding window embeddings at different scales
- Thresholds are birth times from split classes in persistence diagrams. Adjacency matrices are binary thresholded versions of SSMs
- If v1 and v2 are eigenvectors corresponding to smallest nonzero eigenvalues, v1 and v2 are arbitrary orthogonal linear combinations of sine and cosine, because adjacency matrix is **circulant** [4]
- Circular coordinates are atan(v2[n]/v1[n]).
- Arbitrary linear combinations due to numerical precision leads to arbitrary phase shift, but slope remains the same and can be used to estimate tempos



2 on 1
"Mobius ladder"



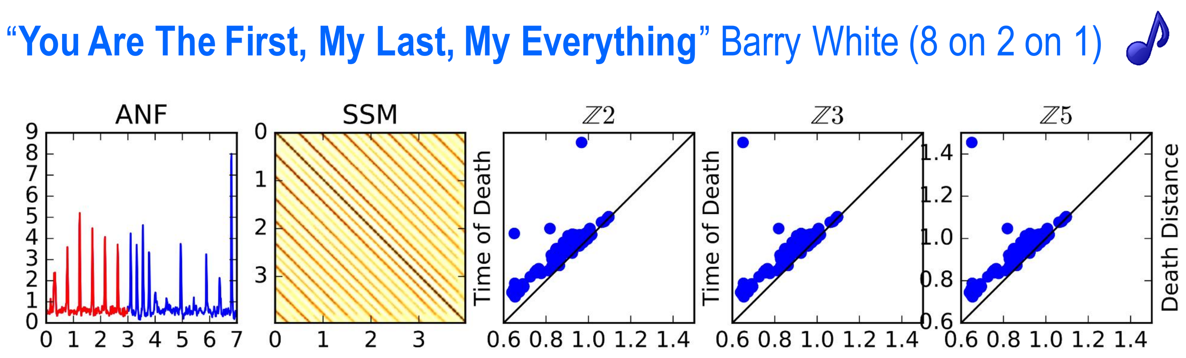
3 on 1



4 on 1

Real Audio Novelty Function Examples

"**You Are The First, My Last, My Everything**" Barry White (8 on 2 on 1)



"**Green Eyes**" Erykah Badu (9 on 3 on 1)