Progress update

Philip Hartout

March 25, 2022

Introduction

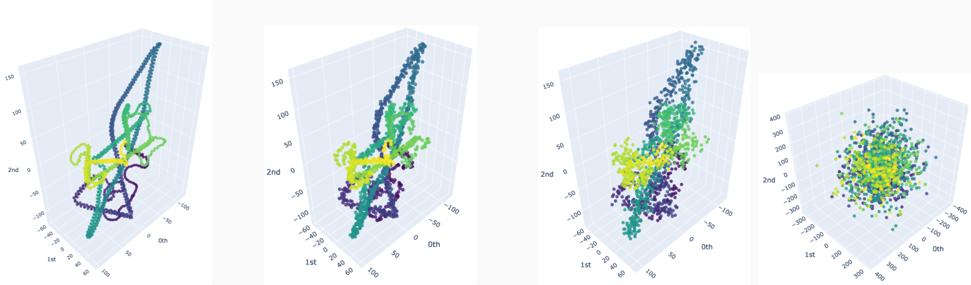
- TDA stuff
- Perturbations
- Next steps?

Kernels on persistence diagrams

Implemented properly from GUDHI [2].
A number of kernels are available:

- The sliced Wasserstein kernel (approximates Wasserstein similarity between PDs and is p.s.d.). [1]
- The persistence weighted Gaussian kernel (slower to compute + approximates). [3]
- The persistence scale space kernel [5] (approximates, is slower as well). [5] proves that the p-Wasserstein distance is not n.s.d.
- The persistence Fisher kernel [4]. Looks the fastest and does not approximate any other distance to be p.s.d.

Nice visualizations



(a) No noise **(b)** 2 **(c)** 6 **(d)** 99

Figure 1: Progressive injection of Gaussian Noise

Single experiment

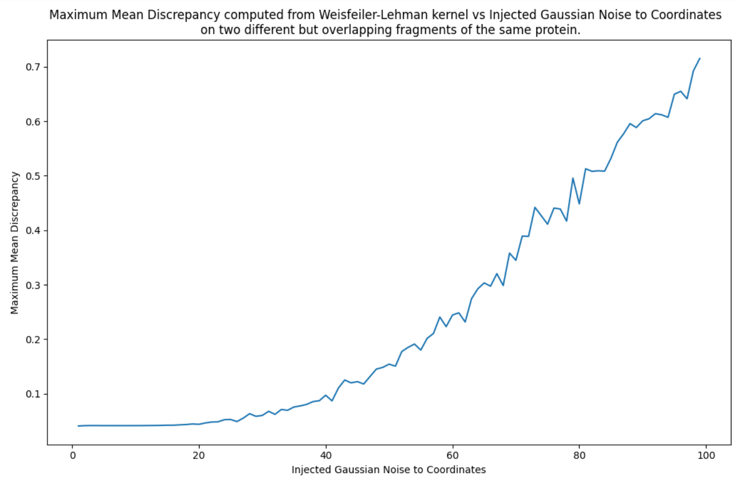


Figure 2: What happens to the MMD for $\varepsilon = 20$?

Multiple experiments varying ε for the ε-graphs

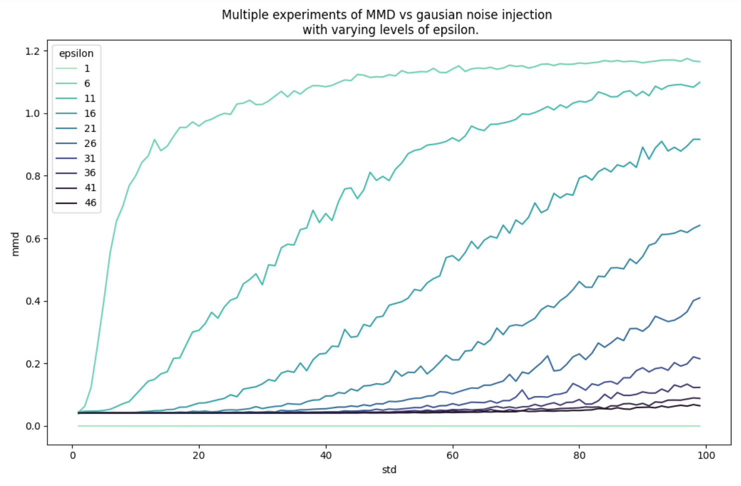


Figure 3: What happens to the MMD if ε varies?

Compressed representations

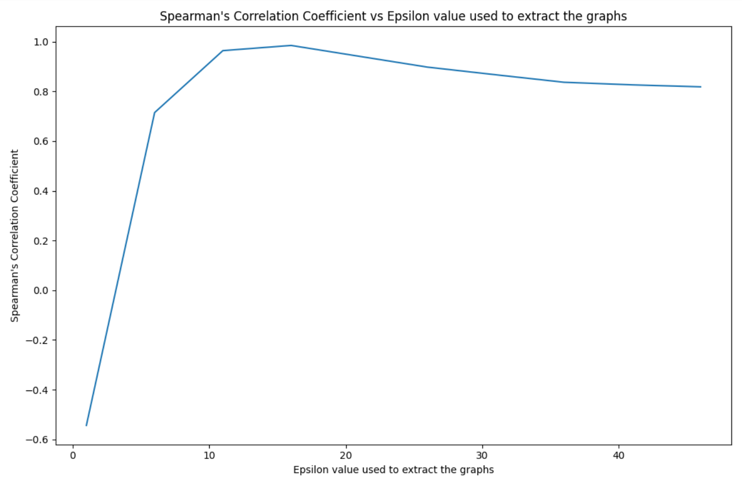


Figure 4: How can we represent the previous plot in a more compressed way?

Next steps

- Data version control and better pipelining using dvc
- Apply perturbations to subdomain (apply rotation to part of protein)
- Clashing descriptors, Ramachandran angles
- Non-MMD based meaure. [6]
- TDA experiments using aforementioned kernel
- Actually make progress on background?

References i

M. Carriere, M. Cuturi, and S. Oudot.
Sliced wasserstein kernel for persistence diagrams.
In *International conference on machine learning*, pages 664–673. PMLR, 2017.

P. Dlotko.
Persistence representations.
In *GUDHI User and Reference Manual*. GUDHI Editorial Board, 2017.

G. Kusano, Y. Hiraoka, and K. Fukumizu.
Persistence weighted gaussian kernel for topological data analysis.
In *International Conference on Machine Learning*, pages 2004–2013. PMLR, 2016.

T. Le and M. Yamada.
Persistence fisher kernel: A riemannian manifold kernel for persistence diagrams.
Advances in Neural Information Processing Systems, 31, 2018.

J. Reininghaus, S. Huber, U. Bauer, and R. Kwitt.
A stable multi-scale kernel for topological machine learning.
In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4741–4748, 2015.

R. Thompson, B. Knyazev, E. Ghalebi, J. Kim, and G. W. Taylor.
On evaluation metrics for graph generative models.
arXiv preprint arXiv:2201.09871, 2022.