Detection and hypothesis testing for extremely noisy videos using TDA

Cornell University[1], ASU[2]

Andrew M. Thomas[1], Peter A. Crozier[2], Yuchen Xu[1], David S. Matteson[1]

Idea: investigate the dynamics of nanoparticle structure via cubical persistent homology (**cPH**).

- Sophisticated methods are needed to capture dynamics because of the presence of extreme amounts of Poisson shot noise at a high temporal resolution.
- We can use topological methods such as **cPH** to capture the amount of structure in an image, but also to capture *where* this structure is present.

Detection: detect structure via the pixels that create components in **cPH**. These pixels *are the local minima in the smoothed image.* Note that we smooth (or filter) our image via a Gaussian filter.

- Exclude identified pixels that are present outside of a pre-chosen polygonal region.
- For location of structure (see figure to the right), we may choose pixels whose lifetimes in **cPH** are only above a certain threshold *η.*
- If our image is *I*, we denote the output of the algorithm to the right as *$A(I)$*—the locations of local minima & their lifetimes.

Testing procedure

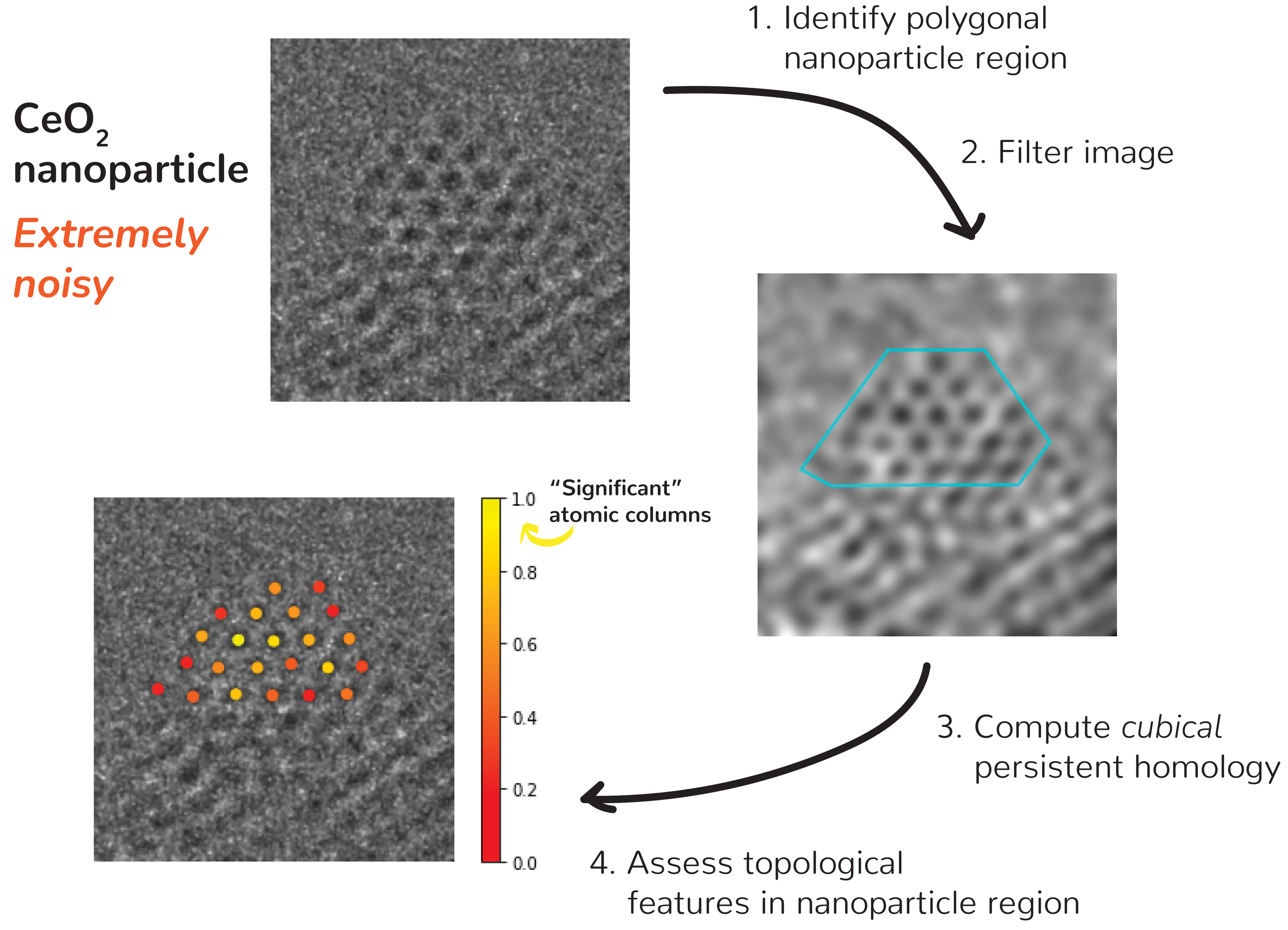
If we assume that all the pixels in our image are generated by noise, we can simulate images according to this noise distribution and compare summaries of the persistence diagrams, such as ***persistent entropy*** (***H***) and the (new!) *ALPS* statistic (***Δ***), which is more conservative than persistent entropy.

To increase signal we may consider the sequence of summed images

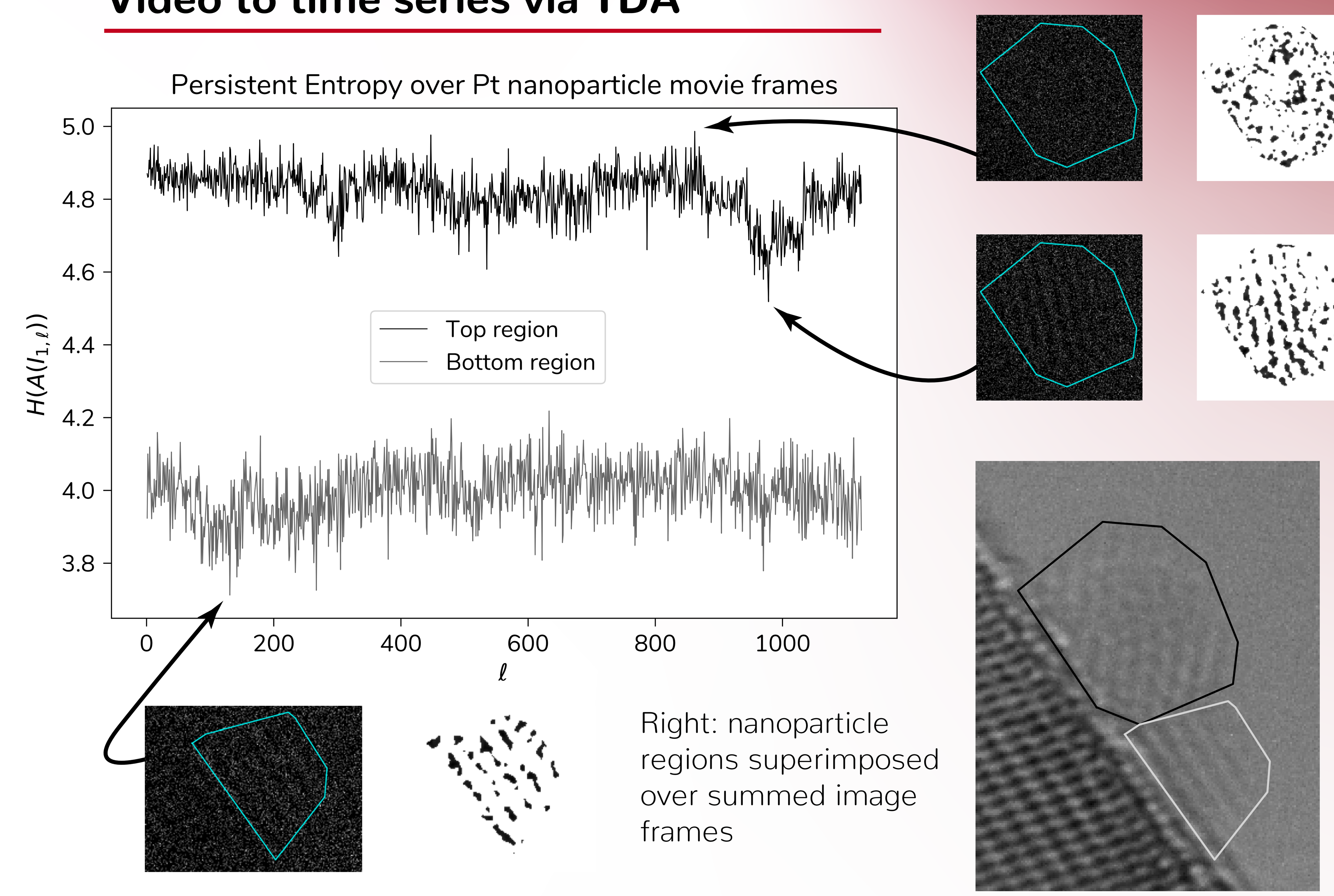
$$I_{m,\ell} = \sum_{k=0}^{m-1} I_{\ell+k}$$

Once we have simulated images (using the values of vacuum regions of the summed nanoparticle images to simulate them), we can compare the simulated values to the observed values to get ***p-values***. *Our procedure leverages the fact we assume each frame of the video has the same noise distribution to save on the number of images we need to generate and the number of times we need to compute* **cPH**.

The algorithm



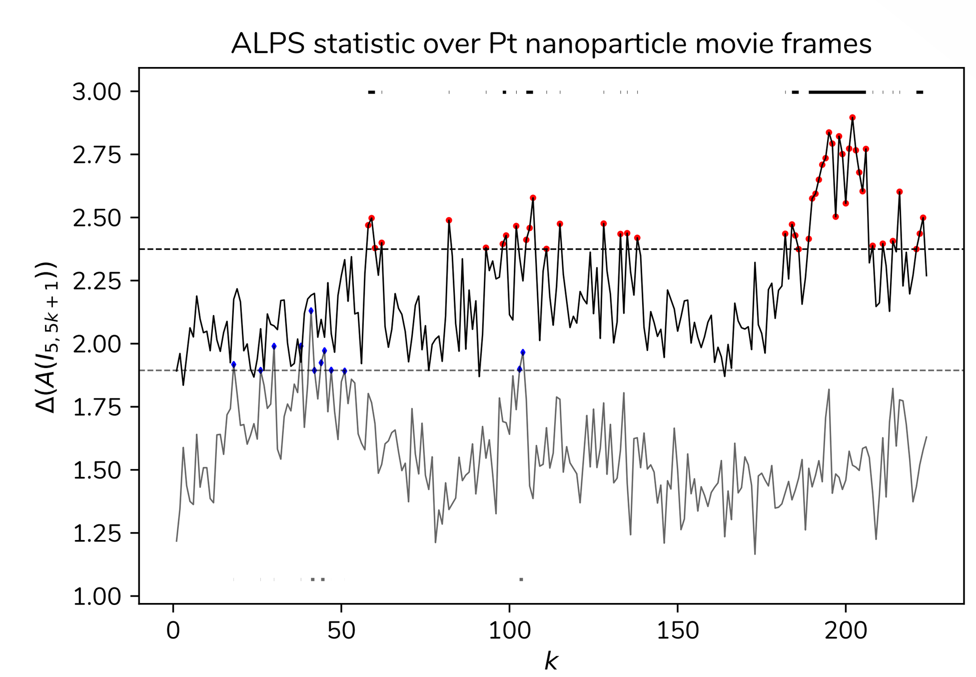
Video to time series via TDA

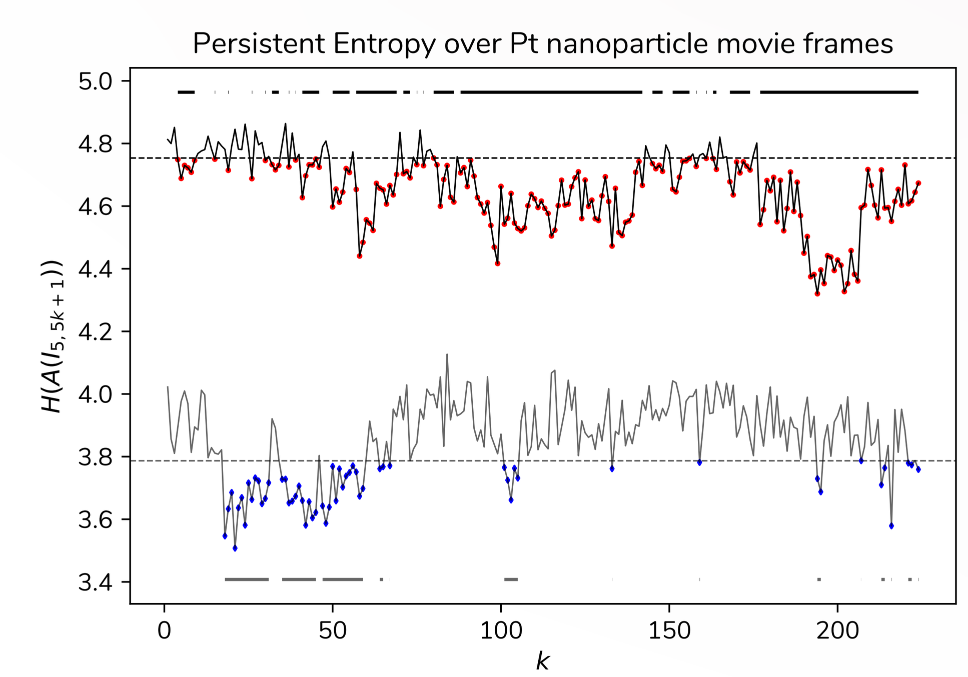


Right: nanoparticle regions superimposed over summed image frames

Hypothesis testing*

*Bars represent significant frames; i.e. frames not generated by pure noise





Selected references

Y. Benjamini and D. Yekutieli. *The control of the false discovery rate in multiple testing under dependency*. The Annals of Statistics, 29(4):1165 – 1188, 2001.

Y.-M. Chung and S. Day. *Topological fidelity and image thresholding: A persistent homology approach*. Journal of Mathematical Imaging and Vision, 60(7):1167–1179, 2018.

P. Dłotko. Cubical complex. In GUDHI User and Reference Manual. GUDHI Editorial Board, 2015.

T. Kaczynski, K. Mischaikow, and M. Mrozek. Computational homology, volume 157. Springer Science & Business Media, 2006.

R. Manzorro, Y. Xu, J. L. Vincent, R. Rivera, D. S. Matteson, and P. A. Crozier. *Exploring blob detection to determine atomic column positions and intensities in time-resolved tem images with ultra- low signal-to-noise*. Microscopy and Microanalysis, pages 1–14, 2022.

B. Rieck, T. Yates, C. Bock, K. Borgwardt, G. Wolf, N. Turk-Browne, and S. Krishnaswamy. *Uncovering the topology of time-varying fMRI data using cubical persistence*. Advances in Neural Information Processing Systems, 2020-December(NeurIPS), 2020.