

## Machine Learning for Atomic Scale Chemical and Morphological Assessment

Scott Broderick<sup>1</sup>, Tianmu Zhang<sup>1</sup>, Bhargava Urala Kota<sup>2</sup>, Ramachandran Subramanian<sup>2</sup>, Srirangaraj Setlur<sup>2</sup>, Venu Govindaraju<sup>2</sup> and Krishna Rajan<sup>1</sup>

<sup>1</sup> Department of Materials Design and Innovation, University at Buffalo, Buffalo, NY.

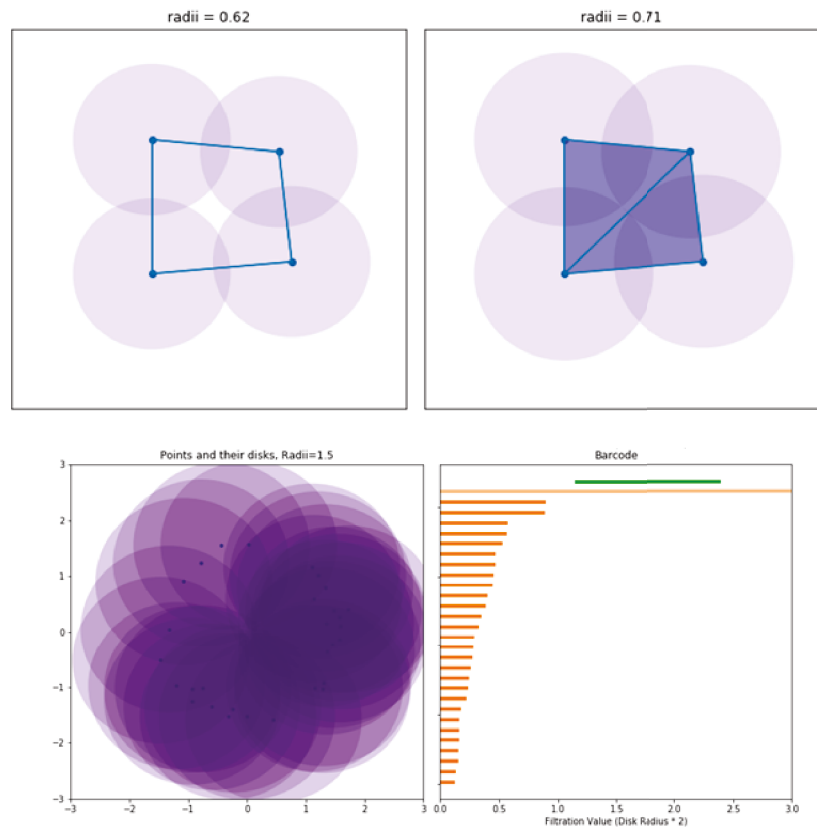
<sup>2</sup> Department of Computer Science and Engineering, University at Buffalo, Buffalo, NY.

We address the challenge in assessing the crystal structure and reconstruction inputs in atom probe tomography (APT) through the application of a machine learning analysis. Particularly, the use of topological data analysis will be applied to characterize the atomic scale morphology and corresponding structural effects in APT, which provides a powerful approach for characterizing nanoparticle surface chemistry and surface structure at the atomic/ nanostructural scale. Quantitative analysis of atom probe data aided by machine learning leads to a new understanding and utilization of the information measured. For example, we have applied persistence diagrams to characterize the structure based on packing density and relative atomic positions (Figure 1). By tracking changes in neighborhoods, where the neighborhoods are characterized by the topological measures, where the bar sizes define groupings of atoms and define the structural regions which otherwise are hidden in the massive data space [1]. This further allows us to assess the data systematics and connect the data across a larger length scale.

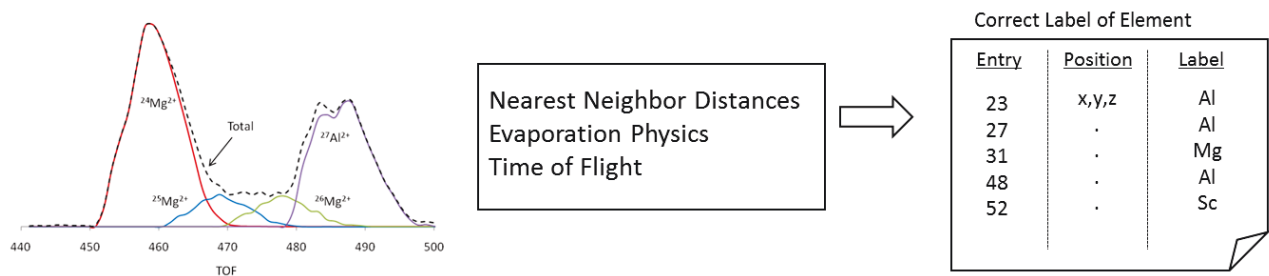
To further understand the role of noise and uncertainty in data, particularly as related to chemical identification, we employ a new approach to discriminate isotopes which have overlapping mass-to-charge ratios ( $m/n$ ), resulting due to the complex evaporation physics which we have previously explored [2], as well as the uncertainty. We previously developed an approach which captured the kinetic energy behavior of the atoms through a multi-tier data dimensionality reduction analysis, allowing us to discriminate the isotopes [3]. However, further assessment and understanding of the data systematics is required. To address this challenge, we have tested multiple unsupervised clustering algorithms and Gaussian mixture models (GMM) to cluster each row of the APT data into groups corresponding to a single isotope. Since the ground truth (true isotope identity for each row) is unknown, we evaluate our algorithms by measuring the distribution of isotopes obtained and comparing them to natural isotope distributions (Figure 2). The presented work highlights the utility of machine learning and data driven analyses for extracting and understanding information from the atomic to microstructural scale, while accounting for complexity in data uncertainty and noise [4].

### References:

- [1] T. Zhang, S.R. Broderick, K. Rajan in “Nanoinformatics”, ed. I. Tanaka, (Springer) p. 133.
- [2] C. Loyola *et al*, Journal of Vacuum Science and Technology A **34** (2016), p. 061404.
- [3] S.R. Broderick *et al*, Ultramicroscopy **132** (2013), p. 121.
- [4] We acknowledge support from NSF DIBBs project award number ARI 1640867.



**Figure 1.** Demonstration of Persistent Homology. Each entry corresponds to a unique value of disk radius. The 0-th homology classes are represented by the orange bars and the 1st homology class is represented by the green bar. All the bars for the 0-th homology class are sorted in the order of their death time. This provides a unique approach to define topological features and allow us to characterize atomic scale morphology and capture microstructural features that are otherwise difficult to extract.



**Figure 2.** The logic developed for defining atom types through machine learning analysis. As an example, the isotopes of Mg and Al overlap when plotted versus TOF. By incorporating the APT data with clustering and GMM algorithms, we are able to correctly assign the atom label.