

## **Title: Automated Feature Extraction for High-Throughput Synchrotron Characterization**

*Authors: Robert Tang-Kong\*, Suchismita Sarker, Jeff Scragle, Apurva Mehta*

As synchrotron light sources continuously push to higher and higher intensities, experiments can be performed faster than ever. While data collection speeds continue to increase, new discoveries and knowledge lag. The primary cause of the bottleneck is that because scientific information extraction from measurements is continues to be performed manually by humans. To alleviate this bottleneck raw data must be transformed into meaningful information in nearly real time, and that implies without human intervention. The first critical step in this process is to reduce a complex and noisy spectrum to a sparse set of characteristic features that carries the information contained in the spectrum. For wide angle scattering patterns (WAX), these are position, intensity and widths of the peaks.

We have developed automated background removal and peak identification algorithms capable of extracting peak information for use in machine learning from raw 2D WAXS images. Automated peak extraction is performed by isolating peaks from the slowly vary background either by Bayesian Block Analysis and an iterative superposition of Voigt lineshapes to reconstruct the line profile. This algorithm is fast and judged after application to over a thousand patterns, matching expert human curation. This development is crucial in automatic transformation of measurements to knowledge.