## Cluster-mining: An approach for determining core structures of metallic nanoparticles from atomic pair distribution function (PDF) data

Soham Banerjee<sup>a</sup>, Chia-Hao Liu<sup>a</sup>, Kirsten M. O. Jensen<sup>b</sup>, Pavol Juhas<sup>c</sup>, Jennifer D. Lee<sup>d</sup>, Marcus Tofanelli<sup>e</sup>, Christopher J. Ackerson<sup>e</sup>, Christopher B. Murray<sup>d,f</sup>, Simon J. L. Billinge<sup>a,g</sup>

Email Contact: sb3519@columbia.edu

<sup>a</sup>Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027, USA

<sup>b</sup>Department of Chemistry, University of Copenhagen, Copenhagen, DK-2100, Denmark <sup>c</sup>Computational Science Initiative, Brookhaven National Laboratory, Upton, NY 11973, USA <sup>d</sup>Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104, USA <sup>e</sup>Department of Chemistry, Colorado State University, Fort Collins, CO, 80523, USA <sup>f</sup>Department of Materials Science and Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA

<sup>g</sup>Condensed Matter Physics and Materials Science Department, Brookhaven National Laboratory, Upton, New York, 11973, USA

Accurate determination of the structure of metallic nanomaterials is a key step towards understanding and controlling their properties. This is especially challenging for ultrasmall particles where x-ray diffraction data is often not amenable to quantitative crystallographic analysis. In these cases, a better approach is to use PDF analysis. We present a database-driven method for finding and evaluating structural models for these technologically important materials. Rather than fitting a single model with many degrees of freedom, we algorithmically built libraries of nanoparticle clusters from multiple structural motifs. The approach, called cluster-mining, then returns all candidate structure models that are consistent with the data as measured by a goodness of fit. It is highly automated, easy to use, and yields models that are more physically realistic and result in better agreement to the data than models based on cubic close-packed crystallographic cores, often reported in the literature for metallic nanoparticles.