

Abstract Submitted  
for the 4CS20 Meeting of  
The American Physical Society

**MPI-parallel Molecular Dynamics Trajectory Analysis with the H5MD Format in MDAnalysis** EDIS JAKUPOVIC, OLIVER BECKSTEIN, Arizona State University Department of Physics — With the growing size of molecular dynamics trajectory files, file I/O has become a bottleneck in the analysis of MD trajectories. If one could open and analyze a trajectory file in parallel, analysis speeds could go from hours to minutes. Previous work [1] found that parallel I/O via MPI-IO and HDF5 led to near ideal strong scaling. However, the previous feasibility study did not provide a usable implementation of a true MD trajectory format. The goal of this work was to add a parallel HDF5 file format coordinate reader to MDAnalysis, a widely used Python library for the analysis of MD simulation data. We added a trajectory reader and performed benchmarks on two typical workloads with different performance characteristics: An I/O bound task and a compute-bound task. These benchmarks were performed on a typical desktop resource and on ASU's Agave supercomputer with the BeeGFS parallel file system, and both showed substantial speedups with our parallel reader. The addition of the HDF5 reader provides a foundation for the development of parallel trajectory analysis with MPI and the MDAnalysis package. [1] M. Khoshlessan, I. Paraskevagos, G. C. Fox, S. Jha, and O. Beckstein. In *Conc. & Comp.: Prac. & Exp.*, 2020. doi: 10.1002/cpe.5789

Edis Jakupovic  
Arizona State University

Date submitted: 29 Sep 2020

Electronic form version 1.4