

Fast step transition and state identification (STASI) for discrete single-molecule data analysis

Bo Shuang[†], David Coopert[†], J. Nick Taylor[‡], Lydia Kisley[†], Jixin Chen[†], Wenxiao Wang[§], Chun Biu Li[‡], Tamiki Komatsuzaki[‡], and Christy F. Landes^{*†§}

[†] Department of Chemistry, Rice University, MS 60, Houston, Texas 77251-1892, United States

[‡] Molecule & Life Nonlinear Sciences Laboratory, Research Institute for Electronic Science, Hokkaido University, Sapporo 001-0020, Japan

[§] Department of Electrical and Computer Engineering, Rice Quantum Institute, Rice University, MS 60, Houston, Texas 77251-1892, United States

We introduce a step transition and state identification (STaSI) method for piecewise constant single-molecule data with a newly derived minimum description length equation as the objective function. We detect the step transitions using the Student's t test and group the segments into states by hierarchical clustering. The optimum number of states is determined based on the minimum description length equation. This method provides comprehensive, objective analysis of multiple traces requiring few user inputs about the underlying physical models and is faster and more precise in determining the number of states than established and cutting-edge methods for single-molecule data analysis. Perhaps most importantly, the method does not require either time-tagged photon counting or photon counting in general and thus can be applied to a broad range of experimental setups and analyses.

