MDEntropy: Information-Theoretic Analyses for Molecular Dynamics

Carlos X. Hernández[1] and Vijay S. Pande[1]

[1]Stanford University

1 October 2017

Paper DOI: http://dx.doi.org/
Software Repository: https://github.com/msmbuilder/mdentropy
Software Archive: https://doi.org/10.5281/zenodo.1000997

Summary

MDEntropy is a Python package for information-theoretic (IT) analysis of data generated from molecular dynamics simulations. While correlation studies have long been of interest to the molecular dynamics (MD) community (McCammon, Gelin, and Karplus 1977, McClendon et al. (2009)), IT tools to analyze MD trajectories have been much less developed. *MDEntropy* seeks to fill this niche by providing an easy-to-use Python API that works seamlessly with other Python packages, such as `mdtraj`, `msmbuilder`, and `numpy` (R. T. McGibbon et al. 2015, Walt, Colbert, and Varoquaux (2011), M. P. Harrigan et al. (2017)).

Functionality in *MDEntropy* is centered around `mdtraj` trajectories and the statistical tools available in `msmbuilder`. Leveraging these tools allows for statistically robust analyses of many IT estimators across a variety of biomolecular feature-spaces (Schreiber 2000, Kraskov, Stögbauer, and Grassberger (2004)).

MDEntropy is actively developed and maintained by researchers at Stanford University. Source code for *MDEntropy* is hosted on GitHub and is continuously archived to Zenodo (Hernández and Pande 2017). Full documentation, including Jupyter Notebook tutorials, can be found at http://msmbuilder.org/mdentropy.

References

Harrigan, Matthew P., Mohammad M. Sultan, Carlos X. Hernández, Brooke E. Husic, Peter Eastman, Christian R. Schwantes, Kyle A. Beauchamp, Robert T. McGibbon, and Vijay S. Pande. 2017. "MSMBuilder: Statistical Models for Biomolecular Dynamics." *Biophysical Journal* 112 (1): 10–15. doi:10.1016/j.bpj.2016.10.042.

Hernández, Carlos X., and Vijay S. Pande. 2017. "Msmbuilder/mdentropy: V0.3." doi:10.5281/zenodo.1000997.

Kraskov, Alexander, Harald Stögbauer, and Peter Grassberger. 2004. "Estimating mutual information." *Physical Review E* 69 (6). American Physical Society: 066138. doi:10.1103/PhysRevE.69.066138.

McCammon, J. Andrew, Bruce R. Gelin, and Martin Karplus. 1977. "Dynamics of folded proteins." *Nature* 267 (5612). Nature Publishing Group: 585–90. doi:10.1038/267585a0.

McClendon, Christopher L, Gregory Friedland, David L Mobley, Homeira Amirkhani, and Matthew P Jacobson. 2009. "Quantifying Correlations Between Allosteric Sites in Thermodynamic Ensembles." *Journal*

of Chemical Theory and Computation 5 (9): 2486–2502. doi:10.1021/ct9001812.

McGibbon, Robert T, Kyle A Beauchamp, Matthew P Harrigan, Christoph Klein, Jason M Swails, Carlos X Hernández, Christian R Schwantes, Lee-Ping Wang, Thomas J Lane, and Vijay S Pande. 2015. "MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories." *Biophysical Journal* 109 (8). Elsevier: 1528–32. doi:10.1016/j.bpj.2015.08.015.

Schreiber, Thomas. 2000. "Measuring Information Transfer." *Physical Review Letters* 85 (2): 461–64. doi:10.1103/PhysRevLett.85.461.

Walt, Stéfan van der, S Chris Colbert, and Gael Varoquaux. 2011. "The NumPy Array: A Structure for Efficient Numerical Computation." *Computing in Science & Engineering* 13 (2): 22–30. doi:10.1109/MCSE.2011.37.