Syntropy: a Python package for discrete and continuous information theory.

Thomas F. Varley [1]

1 Vermont Complex Systems Institute, University of Vermont, Burlington, VT, USA.

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Chris Vernon

Reviewers:

- @manolomartinez
- @cbueth

Submitted: 05 September 2025
Published: unpublished

Summary:

Information theory (a branch of mathematics concerned with the logic of inference and uncertainty) has emerged as a kind of *lingua franca* for the study of complex systems (Varley, 2024b), finding applications in neuroscience [Timme et al. (2014)](Varley et al., 2024), climatology [Goodwell (2020)](Goodwell et al., 2020), developmental biology (Blackiston et al., 2025), economics [Kim et al. (2020)](Rajpal & Guerrero, 2025), sociology (Varley & Kaminski, 2022) and more. The goal of the Syntropy package is to provide an accessible, easy-to-use API for scientists interested in applying information-theoretic analyses to arbitrary data sets. To this end, Syntropy implements a variety of different measures, as well as a variety of estimators for different types of data (discrete, continuous, time series, etc). Being primarily concerned with multivariate information theory and the application to complex systems, a core part of the Syntropy package is implementing the various information decompositions and higher-order mutual information measures that have been developed in recent years for the study of redundant and synergistic interactions in complex systems.

The Syntropy package takes considerable inspiration from the Networkx package (Hagberg et al., 2008). By providing an accessible, high-level interface for complex analyses, the developers of Networkx have made network science and graph-theoretic tools available to a much more diverse range of scientists than would ever have been possible if every researchers was required to develop their own functions. The goal of Syntropy is to be the Networkx of information theory"; lowering the barrier to entry fro information theory so that powerful tools are accessible to scientists from many different fields. Accordingly, the Syntropy package is written in pure Python, with minimal external dependencies, all of which are standard libraries for scientific computing in Python, including `numpy`, `scipy`, and `networkx` packages.

While the core information-theoretic functions have been implemented and tested for correctness, the goal is for Syntropy to be a "living package": updating the library of functions and estimators as the field of multivariate information theory continues to develop. Future work may also include porting the library to other scientific programming languages such as Julia, R, or MATLAB.

Description of package

The Syntropy package is broken up into two main arms: discrete estimators (which operate on discrete, multivariate probability distribution represented by Python dictionaries), and continuous estimators (currently focused on parametric Gaussian estimators based on time series and covariance matrices). Within each branch, we have several sub-modules that describe different families of analysis. There is a `shannon` module which contains basic Shannon information theory functions (entropy, conditional entropy, mutual information, Kullback-Leibler divergence, etc). There is a `multivariate mi` module that includes higher-order generalizations of mutual information and their extensions (total correlation, dual total

The Journal of Open Source Software

correlation, O-information, etc (Rosas et al., 2019)). The `decomposition` module contains the partial information (Williams & Beer, 2010), partial entropy (Ince, 2017), and generalized information decompositions (Varley, 2024a). The `temporal` module contains a small number of functions designed for time series analysis specifically [Blanc et al. (2008)](Sparacino et al., 2025). Finally, both branches have a `utils` module which contains basic operations on discrete and Gaussian probability distributions, as well as some example distributions of theoretical note (such as the James Triadic and Dyadic distribution (R. G. James & Crutchfield, 2017). These primarily support the higher-level functions, but may be of interest to those working on more basic probability theory.

Wherever possible, we have tried to ensure overlap between the two branches, ensuring that each information theoretic function has both discrete and continuous implementations (e.g. there is a discrete and continuous entropy estimator, a discrete and continuous mutual information estimator, etc). This ensures that researchers can select the most appropriate tool for the data they have, rather than being forced to transform their data to make it fit (e.g. discretizing continuous data).

Statement of need

While many different computing packages for information-theoretic analysis exist, they are typically restricted to only a single subset of analyses or data type. For example, the DIT package (R. James et al., 2018) is specific to discrete information theory and uses a customized distribution object. The IDTxl (Wollstadt et al., 2019) package implements several different classes of estimators, but is focused specifically on information dynamics and time series analysis, as is the JIDT package (Lizier, 2014). Information decomposition analysis in particular is an outstanding issue: DIT and IDTxl have limited support for the partial information decomposition (PID) of discrete random variables, but no other package also implements the PID for continuous Gaussian random variables, or supports the partial entropy decomposition (PED), generalized information decomposition (GID), or alpha-synergy decomposition. Several other classes of analyses are also included in Syntropy that are not present in other primary packages, including spectral estimators for Gaussian autoregressive processes, and Lempel-Ziv estimators for discrete dynamic processes.

Acknowledgments

This work was supported by the Cold Regions Research and Engineering Laboratory, under contract number W913E524C0012. TFV would also like to thank Dr. Olaf Sporns and Dr. Maria Pope for extensive support and discussion about all things information theory over many years, and Dr. Josh Bongard for support while finalizing this package.

References

Blackiston, D., Dromiack, H., Grasso, C., Varley, T. F., Moore, D. G., Srinivasan, K. K., Sporns, O., Bongard, J., Levin, M., & Walker, S. I. (2025). Revealing non-trivial information structures in aneural biological tissues via functional connectivity. *PLOS Computational Biology*, *21*(4), e1012149. https://doi.org/10.1371/journal.pcbi.1012149

Blanc, J.-L., Schmidt, N., Bonnier, L., Pezard, L., & Lesne, A. (2008, October). Quantifying Neural Correlations Using Lempel-Ziv Complexity. *Neurocomp08*.

Goodwell, A. E. (2020). "It's Raining Bits": Patterns in Directional Precipitation Persistence across the United States. *Journal of Hydrometeorology*, *21*(12), 2907–2921. https://doi.org/10.1175/JHM-D-20-0134.1

Goodwell, A. E., Jiang, P., Ruddell, B. L., & Kumar, P. (2020). Debates—Does Information Theory Provide a New Paradigm for Earth Science? Causality, Interaction, and Feedback. *Water Resources Research*, *56*(2), e2019WR024940. https://doi.org/10.1029/2019WR024940

Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring Network Structure, Dynamics, and Function using NetworkX. In G. Varoquaux, T. Vaught, & J. Millman (Eds.), *Proceedings of the 7th Python in Science Conference* (pp. 11–15).

Ince, R. A. A. (2017). The Partial Entropy Decomposition: Decomposing multivariate entropy and mutual information via pointwise common surprisal. *arXiv:1702.01591 [Cs, Math, q-Bio, Stat]*. http://arxiv.org/abs/1702.01591

James, R. G., & Crutchfield, J. P. (2017). Multivariate Dependence beyond Shannon Information. *Entropy*, *19*(10), 531. https://doi.org/10.3390/e19100531

James, R., Ellison, C., & Crutchfield, J. (2018). Dit: A Python package for discrete information theory. *Journal of Open Source Software*, *3*(25), 738. https://doi.org/10.21105/joss.00738

Kim, S., Ku, S., Chang, W., & Song, J. W. (2020). Predicting the Direction of US Stock Prices Using Effective Transfer Entropy and Machine Learning Techniques. *IEEE Access*, *8*, 111660–111682. https://doi.org/10.1109/ACCESS.2020.3002174

Lizier, J. T. (2014). JIDT: An Information-Theoretic Toolkit for Studying the Dynamics of Complex Systems. *Frontiers in Robotics and AI*, *1*. https://doi.org/10.3389/frobt.2014.00011

Rajpal, H., & Guerrero, O. (2025). Synergistic small worlds that drive technological sophistication. *PNAS Nexus*, *4*(4), pgaf102. https://doi.org/10.1093/pnasnexus/pgaf102

Rosas, F., Mediano, P. A. M., Gastpar, M., & Jensen, H. J. (2019). Quantifying High-order Interdependencies via Multivariate Extensions of the Mutual Information. *Physical Review E*, *100*(3), 032305. https://doi.org/10.1103/PhysRevE.100.032305

Sparacino, L., Mijatovic, G., Antonacci, Y., Ricci, L., Marinazzo, D., Stramaglia, S., & Faes, L. (2025). *Decomposing Multivariate Information Rates in Networks of Random Processes*. arXiv. https://doi.org/10.48550/arXiv.2502.04555

Timme, N., Alford, W., Flecker, B., & Beggs, J. M. (2014). Synergy, redundancy, and multivariate information measures: An experimentalist's perspective. *Journal of Computational Neuroscience*, *36*(2), 119–140. https://doi.org/10.1007/s10827-013-0458-4

Varley, T. F. (2024a). Generalized decomposition of multivariate information. *PLOS ONE*, *19*(2), e0297128. https://doi.org/10.1371/journal.pone.0297128

Varley, T. F. (2024b). *Information Theory for Complex Systems Scientists*. arXiv. https://doi.org/10.48550/arXiv.2304.12482

Varley, T. F., Havert, D., Fosque, L., Alipour, A., Weerawongphrom, N., Naganobori, H., O'Shea, L., Pope, M., & Beggs, J. (2024). The serotonergic psychedelic N,N-dipropyltryptamine alters information-processing dynamics in in vitro cortical neural circuits. *Network Neuroscience*, 1–35. https://doi.org/10.1162/netn_a_00408

Varley, T. F., & Kaminski, P. (2022). Untangling Synergistic Effects of Intersecting Social Identities with Partial Information Decomposition. *Entropy*, *24*(10), 1387. https://doi.org/10.3390/e24101387

Williams, P. L., & Beer, R. D. (2010). Nonnegative Decomposition of Multivariate Information. *arXiv:1004.2515 [Math-Ph, Physics:physics, q-Bio]*. http://arxiv.org/abs/1004.2515

Wollstadt, P., Lizier, J. T., Vicente, R., Finn, C., Martinez-Zarzuela, M., Mediano, P., Novelli, L., & Wibral, M. (2019). IDTxl: The Information Dynamics Toolkit xl: A Python package for the efficient analysis of multivariate information dynamics in networks. *Journal of Open Source Software*, *4*(34), 1081. https://doi.org/10.21105/joss.01081