The Journal of Open Source Software

HyperIT: A Python toolbox for an information-theoretic social neuroscience

Edoardo Chidichimo [1]

1 Department of Psychology, University of Cambridge, UK

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Open Journals

Reviewers:

- @openjournals

Submitted: 01 January 1970
Published: unpublished

Summary

`HyperIT` is an open-source, class-based Python toolbox designed for an information-theoretic social neuroscience. Specifically designed for hyperscanning paradigms (simultaneous neuroimaging and neurophysiological recordings of social interactions), `HyperIT` is equipped to compute various Shannon information theory measures including mutual information, transfer entropy, and integrated information decomposition for continuous time-series signals (fMRI, EEG, MEG, fNIRS, etc.) at different frequency bands and spatial scales of organisation. The toolbox allows flexible customisation of parameters for up to five different empirical estimators (including histogram, box kernel, kNN, Gaussian, and symbolic methods), statistical significance testing, and matrix visualisations. `HyperIT` integrates, depends upon, and provides an interface for the Java Information Dynamics Toolkit (Lizier, 2014) and `phyid` library (Luppi et al., 2022; Mediano et al., 2021). To the best of the author's knowledge, there are no Python toolboxes or libraries specifically designed for hyperscanning analyses that are equipped with information-theoretic analyses whilst dealing with epochality, frequency resolution, and spatial scales of organisation.

Statement of need

For the past two decades, social neuroscience has addressed the interpersonal neural dynamics of interacting persons through the hyperscanning paradigm (Dumas et al., 2010). Synchrony and other statistical properties have been prioritised in these analyses, although most measures remain of the parametric type despite neural activity and coordination dynamics often being distinctly nonlinear. `HyperIT` presents the first principled and unified analysis framework for hyperscanning data using nonparametric measures.

Between interacting persons, we observe coordination and communication on various scales of organisation whose unit of transfer we can loosely describe as *information*. Whilst the idea of interpersonal information transfer may appear intuitive *prima facie*, it is not exactly clear how to define information, let alone measure its content, distribution, transfer, storage, modification, and other informational dynamics. Thus, an elegant definition and quantification of information that is *mathematically precise and consistent* is essential, particularly one that is *content-invariant* and *model-free*; i.e., making no assumptions on the information content itself nor the statistical distributions or relationship model between random variables.

Information theory, lauded as *the* mathematical theory of communication (Shannon, 1948), lends itself well to this cause, proffering domain-generality and information as its common currency. Originally birthed for the development of communications engineering, various information-theoretic measures have found vast applicability in computer science (e.g., Kolmogorov complexity), economics (e.g., portfolio theory and Kelly gambling), statistics (e.g., Fisher information and hypothesis testing), and probability theory (e.g., limit theorems and large derivations; Cover & Thomas (2006)). Information theory has recently earned a notable place in neuroscience, too (Timme & Lapish, 2018).

Of relevance for neuroscientific endeavours, information theory provides measures that can detect linear *and* nonlinear dependencies between continuous time-series signals, allowing researchers to analyse both the correlation and causation between two or more random variables. In all, whilst standard measures of correlation and prediction are only sensitive to linear dependencies and only describe the variable's overall relationship, information-theoretic analyses and measures can quantify and more comprehensively describe the dynamics of complex systems that may demonstrate nonlinear dependencies whilst maintaining a particular robustness to noise. `HyperIT` computes well-recognised and validated information-theoretic measures that, for the first time, can be simultaneously applied to both hyperscanning and intra-brain analyses for various neural recordings. Usefully, and unlike other libraries, `HyperIT` is equipped to handle epoched and event-based data as well as specifying both frequency bands and the level of organisation by comparing channels pairwise or by regions of interest. These measures include *(I) mutual information*, *(II) transfer entropy*, and *(III) integrated information decomposition*, described below.

Measures

Mutual information ($I(X;Y)$) is a positively-biased, symmetric measure indicating how much (Shannon) information is shared between two random variables X and Y (Equation 1). One may also interpret MI as quantifying the "distance" between the joint probability distribution (p_{XY}) and the product of marginal probability distributions ($p_X \otimes p_Y$), such that $I(X;Y) = 0$ *iff* $p_{XY} = p_X \otimes p_Y$. Thus, it is understood as a special instance of the Kullback-Leibler divergence measure.

$$I(X;Y) = \sum_{x} \sum_{y} p_{XY}(x,y) \log_2 \frac{p_{XY}(x,y)}{p_X(x) p_Y(y)}. \tag{1}$$

Transfer entropy ($TE_{Y \to X}$) is a measure that non-parametrically measures the statistical connectivity and time-directed transfer of information between random variables or processes. It is often taken as the non-linear equivalent of Granger causality and, indeed, equivalence has been demonstrated under Gaussianity (Barnett et al., 2009). Specifically, this measure uses conditional mutual information to quantify the reduction in uncertainty about the future of one process given knowledge about another variable and its own history (Equation 2).

$$\begin{aligned} TE^{(k,l,u,\tau)}_{Y \to X} &= I(\mathbf{Y}^{(l,\tau_Y)}_{t-u}; X_t | \mathbf{X}^{(k,\tau_X)}_{t-1}) \\ &= \sum_{\substack{x_t, \mathbf{x}^{(k,\tau_X)}_{t-1} \in \mathcal{X}, \\ \mathbf{y}^{(l,\tau_Y)}_{t-u} \in \mathcal{Y}}} p\left(x_t, \mathbf{x}^{(k,\tau_X)}_{t-1}, \mathbf{y}^{(l,\tau_Y)}_{t-u}\right) \log_2 \left(\frac{p\left(\mathbf{y}^{(l,\tau_Y)}_{t-u}, x_t | \mathbf{x}^{(k,\tau_X)}_{t-1}\right)}{p\left(x_t | \mathbf{x}^{(k,\tau_X)}_{t-1}\right)} \right), \end{aligned} \tag{2}$$

with parameters including embedding history length for source (l) and target (k), embedding delay for source (τ_Y) and target (τ_X), and some causal delay or interaction lag u.

More recently, approaches to exhaustively decompose a multivariate system's informational structure has described three modes; namely, information about a target variable may be redundant (Rdn): information that is shared between variables; unique (Unq): information that is specific to a single variable; or synergistic (Syn): information that is only learnt from the conjunction of multiple sources and not individually, with the exclusive-OR function being a canonical example (Williams & Beer, 2010). These so-called "partial information atoms" are

The Journal of Open Source Software

non-overlapping and form an additive set, exhaustively describing the informational composition of a multivariate system (Equation 3).

$$\begin{aligned} I(Y;X_1,X_2) = &\,\mathsf{Syn}(Y;X_1,X_2)+ \\ &\,\mathsf{Unq}(Y;X_1)+\mathsf{Unq}(Y;X_2)+ \\ &\,\mathsf{Rdn}(Y;X_1,X_2). \end{aligned} \tag{3}$$

A recent development, termed *integrated information decomposition*, extends this decomposition to multi-source and multi-target continuous time-series random variables to decompose information dynamics into various qualitative modes including information storage, copy, transfer, erasure, downward causation, causal decoupling, and upward causation (Mediano et al., 2019, 2021). This measure specifically decomposes the time-delayed mutual information ($t' > t$) between two multivariate processes (Equation 4).

$$\begin{aligned} I(\mathbf{X}_t;\mathbf{X}_{t'}) = &\,\mathsf{Syn}(X^1_t,X^2_t;\mathbf{X}_{t'})+ \\ &\,\mathsf{Unq}(X^1_t;\mathbf{X}_{t'}|X^2_t)+\mathsf{Unq}(X^2_t;\mathbf{X}_{t'}|X^1_t)+ \\ &\,\mathsf{Rdn}(X^1_t,X^2_t;\mathbf{X}_{t'}). \end{aligned} \tag{4}$$

Functionality

`HyperIT`, then, addresses these approaches and offers a user-friendly, class-based toolbox where users can create a `HyperIT` object passing two multivariate sets of continuous time-series signals (dimensions including epochality and multiple channels). As mentioned, users can choose to bandpass filter their signals at specified frequency bands as well as specify the scale of organisation by setting the "regions of interest" property; namely, specifying whether (and which) channels are computed pairwise or as clusters with one another for all measures, making micro-, meso-, and macro-scale analysis readily available. From here, users can call mutual information, transfer entropy, and integrated information decomposition functions specifying estimation type and estimation-specific parameters (outlined in documentation). Estimators include (a) histogram/binning, (b) box kernel, (c) Gaussian, (d) k-nearest neighbour, and (e) symbolic approaches. Users may also choose to conduct statistical significance testing via permutation testing for each calculation and optionally visualise the measure matrices. Importantly, the toolbox relies upon the Java Information Dynamics Toolkit for computing mutual information and transfer entropy with various estimators (Lizier, 2014), although `HyperIT` makes this Java-oriented toolbox accessible and compatible for Python coders who may in turn enjoy a diverse ecosystem of other scientific libraries, most notably MNE and HyPyP (Ayrolles et al., 2021). Users will need to store the `infodynamics.jar` file locally in their working directory as well as ensuring the `jpype` dependency is installed. Finally, `HyperIT` employs the `phyid` package to compute integrated information atoms (Luppi et al., 2022; Mediano et al., 2021). Thus, `HyperIT` offers a distinct interoperability and compatibility with other Python libraries whilst providing unique features of handling hyperscanning data, and versatility with epochality, frequency resolutions, and spatial scales of organisation.

In all, social neuroscience researchers working with continuous time-series signals (including fMRI, EEG, MEG, and fNIRS), particularly in the context of hyperscanning, will be, for the first time, able to comfortably compute common and powerful information-theoretic measures using our `HyperIT` toolbox.

Acknowledgements

Thanks to Richard Bethlehem, Guillaume Dumas, Pedro A. M. Mediano, Eric C. Dominguez, and Andrea I. Luppi for their support and contributions to the fruition of this project.

References

Ayrolles, A., Brun, F., Chen, P., Djalovski, A., Beauxis, Y., Delorme, R., Bourgeron, T., Dikker, S., & Dumas, G. (2021). HyPyP: A Hyperscanning Python Pipeline for inter-brain connectivity analysis. *Social Cognitive and Affective Neuroscience*, *16*(1-2), 72–83. https://doi.org/10.1093/scan/nsaa141

Barnett, L., Barrett, A. B., & Seth, A. K. (2009). Granger causality and transfer entropy are equivalent for Gaussian variables. *Physical Review Letters*, *103*(23), 238701. https://doi.org/10.1103/PhysRevLett.103.238701

Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory* (2nd ed). Wiley-Interscience. ISBN: 978-0-471-24195-9

Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., & Garnero, L. (2010). Inter-brain synchronization during social interaction. *PLoS ONE*, *5*(8), e12166. https://doi.org/10.1371/journal.pone.0012166

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

Luppi, A. I., Mediano, P. A. M., Rosas, F. E., Holland, N., Fryer, T. D., O'Brien, J. T., Rowe, J. B., Menon, D. K., Bor, D., & Stamatakis, E. A. (2022). A synergistic core for human brain evolution and cognition. *Nature Neuroscience*, *25*(6), 771–782. https://doi.org/10.1038/s41593-022-01070-0

Mediano, P. A. M., Rosas, F. E., Luppi, A. I., Carhart-Harris, R. L., Bor, D., Seth, A. K., & Barrett, A. B. (2021). *Towards an extended taxonomy of information dynamics via Integrated Information Decomposition*. https://doi.org/10.48550/ARXIV.2109.13186

Mediano, P. A. M., Rosas, F., Carhart-Harris, R. L., Seth, A. K., & Barrett, A. B. (2019). *Beyond integrated information: A taxonomy of information dynamics phenomena*. https://doi.org/10.48550/ARXIV.1909.02297

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Timme, N. M., & Lapish, C. (2018). A tutorial for information theory in neuroscience. *Eneuro*, *5*(3), ENEURO.0052–18.2018. https://doi.org/10.1523/ENEURO.0052-18.2018

Williams, P. L., & Beer, R. D. (2010). *Nonnegative decomposition of multivariate information*. https://doi.org/10.48550/ARXIV.1004.2515