

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

Edoardo Chidichimo <sup>1</sup>

<sup>1</sup> Department of Psychology, University of Cambridge, UK

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

## Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: [Open Journals](#) 

## Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

## License

Authors of papers retain copyright  
and release the work under a  
Creative Commons Attribution 4.0  
International License ([CC BY 4.0](#)).

## 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 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 scales of organisation. The toolbox allows 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](#)).

## 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.

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-generalty 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 earned a notable place in neuroscience ([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. The information-theoretic measures available in HyperIT 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.

$$I(X; Y) = \sum_x \sum_y p_{XY}(x, y) \log_2 \frac{p_{XY}(x, y)}{p_X(x)p_Y(y)} \quad (1)$$

Transfer entropy ( $TE_{Y \rightarrow X}$ ) 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).

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

with parameters including embedding history length for source ( $l$ ) and target ( $k$ ), embedding delay for source ( $\tau_Y$ ) and target ( $\tau_X$ ), 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 non-overlapping and form an additive set, exhaustively describing the informational composition of a multivariate system (Equation 3).

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

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 between two multivariate processes (Equation 4).

$$I(\mathbf{X}_t; \mathbf{X}_{t'}) = \text{Syn}(X_t^1, X_t^2; \mathbf{X}_{t'}) + \text{Unq}(X_t^1; \mathbf{X}_{t'} | X_t^2) + \text{Unq}(X_t^2; \mathbf{X}_{t'} | X_t^1) + \text{Rdn}(X_t^1, X_t^2; \mathbf{X}_{t'}) \quad (4)$$

by some time delay  $t' > t$ .

## 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). Users can choose to specify the scale of organisation, too, 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 and transfer entropy 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 heavily 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. Users will need to store the `infodynamics.jar` file locally in their working directory as well as ensuring the `jpye` dependency is installed. Finally, HyperIT employs the `phyid` package to compute integrated information atoms (Luppi et al., 2022; Mediano et al., 2021). Notably, we encourage neuroscientists engaged with the hyperscanning paradigm to complement this toolbox with the well-established Hyperscanning Python Pipeline (HyPyP) of which HyperIT is compatible with (Ayrolles et al., 2021).

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

## Acknowledgements

Thanks to Pedro A. M. Mediano, Eric C. Dominguez, Andrea I. Luppi, Richard Bethlehem, and Guillaume Dumas during the genesis 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>

- 110 Barnett, L., Barrett, A. B., & Seth, A. K. (2009). Granger causality and transfer entropy  
111 are equivalent for Gaussian variables. *Physical Review Letters*, 103(23), 238701. <https://doi.org/10.1103/PhysRevLett.103.238701>  
112
- 113 Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory* (2nd ed). Wiley-  
114 Interscience. ISBN: 978-0-471-24195-9
- 115 Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., & Garnero, L. (2010). Inter-brain  
116 synchronization during social interaction. *PLoS ONE*, 5(8), e12166. <https://doi.org/10.1371/journal.pone.0012166>  
117
- 118 Lizier, J. T. (2014). JIDT: An information-theoretic toolkit for studying the dynamics of complex  
119 systems. *Frontiers in Robotics and AI*, 1. <https://doi.org/10.3389/frobt.2014.00011>
- 120 Luppi, A. I., Mediano, P. A. M., Rosas, F. E., Holland, N., Fryer, T. D., O'Brien, J. T.,  
121 Rowe, J. B., Menon, D. K., Bor, D., & Stamatakis, E. A. (2022). A synergistic core  
122 for human brain evolution and cognition. *Nature Neuroscience*, 25(6), 771–782. <https://doi.org/10.1038/s41593-022-01070-0>  
123
- 124 Mediano, P. A. M., Rosas, F. E., Luppi, A. I., Carhart-Harris, R. L., Bor, D., Seth, A. K.,  
125 & Barrett, A. B. (2021). *Towards an extended taxonomy of information dynamics via*  
126 *Integrated Information Decomposition*. <https://doi.org/10.48550/ARXIV.2109.13186>
- 127 Mediano, P. A. M., Rosas, F., Carhart-Harris, R. L., Seth, A. K., & Barrett, A. B. (2019).  
128 *Beyond integrated information: A taxonomy of information dynamics phenomena*. <https://doi.org/10.48550/ARXIV.1909.02297>  
129
- 130 Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical*  
131 *Journal*, 27(3), 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- 132 Timme, N. M., & Lapish, C. (2018). A tutorial for information theory in neuroscience. *Eneuro*,  
133 5(3), ENEURO.0052–18.2018. <https://doi.org/10.1523/ENEURO.0052-18.2018>
- 134 Williams, P. L., & Beer, R. D. (2010). *Nonnegative decomposition of multivariate information*.  
135 <https://doi.org/10.48550/ARXIV.1004.2515>