

SOFTWARE

MINT: A toolbox for the analysis of multivariate neural information coding and transmission

Gabriel Matías Lorenz^{1,2,3☯}, Nicola Marie Engel^{1☯}, Marco Celotto^{1,4}, Loren Koçillari^{1,5}, Sebastiano Curreli², Tommaso Fellin², Stefano Panzeri^{1*}

1 Institute for Neural Information Processing, Center for Molecular Neurobiology (ZMNH), University Medical Center Hamburg-Eppendorf (UKE), Hamburg, Germany, **2** Optical Approaches to Brain Function Laboratory, Istituto Italiano di Tecnologia, Genova, Italy, **3** Department of Pharmacy and Biotechnology, University of Bologna, Bologna, Italy, **4** Department of Brain and Cognitive Sciences, Picower Institute for Learning and Memory, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America, **5** Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf (UKE), Hamburg, Germany

☯ These authors contributed equally

* s.panzeri@uke.de



OPEN ACCESS

Citation: Lorenz GM, Engel NM, Celotto M, Koçillari L, Curreli S, Fellin T, et al. (2025) MINT: a toolbox for the analysis of multivariate neural information coding and transmission. *PLoS Comput Biol* 21(4): e1012934. <https://doi.org/10.1371/journal.pcbi.1012934>

Editor: Hugues Berry, Inria, FRANCE

Received: November 5, 2024

Accepted: March 6, 2025

Published: April 15, 2025

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pcbi.1012934>

Copyright: © 2025 Lorenz et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Information theory has deeply influenced the conceptualization of brain information processing and is a mainstream framework for analyzing how neural networks in the brain process information to generate behavior. Information theory tools have been initially conceived and used to study how information about sensory variables is encoded by the activity of small neural populations. However, recent multivariate information theoretic advances have enabled addressing how information is exchanged across areas and used to inform behavior. Moreover, its integration with dimensionality-reduction techniques has enabled addressing information encoding and communication by the activity of large neural populations or many brain areas, as recorded by multichannel activity measurements in functional imaging and electrophysiology. Here, we provide a Multivariate Information in Neuroscience Toolbox (MINT) that combines these new methods with statistical tools for robust estimation from limited-size empirical datasets. We demonstrate the capabilities of MINT by applying it to both simulated and real neural data recorded with electrophysiology or calcium imaging, but all MINT functions are equally applicable to other brain-activity measurement modalities. We highlight the synergistic opportunities that combining its methods afford for reverse engineering of specific information processing and flow between neural populations or areas, and for discovering how information processing functions emerge from interactions between neurons or areas. MINT works on Linux, Windows and macOS operating systems, is written in MATLAB (requires MATLAB version 2018b or newer) and depends on 4 native MATLAB toolboxes. The calculation of one possible way to compute information redundancy requires the installation and compilation of C files (made available by us also as pre-compiled files). MINT

Data availability statement: All data and code used to produce this paper is made publicly available. All code is downloadable in source code (<https://github.com/panzerilab/MINT> with DOI doi.org/10.5281/zenodo.13998526) and is licensed under GNU GPLv3. The neural CA1 data are attached to this submission as Supplemental Material.

Funding: This work was supported by the Simons Foundation for Autism Research Initiative (SFARI) grant number 982347 (to SP) and the NextGenerationEU FAIR grant number PE0000013 (to TF). The Funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: I have read the journal's policy and the authors of this manuscript have the following competing interests: Stefano Panzeri is a member of the editorial board of PLoS Computational Biology. No other conflict of interest is declared by any other author.

is freely available at <https://github.com/panzerilab/MINT> with DOI doi.org/10.5281/zenodo.13998526 and operates under a GNU GPLv3 license.

Introduction

Brain functions are based on the ability of groups of neurons or brain areas to encode, process and transmit information [1,2]. Consequently, information theory [3], the mathematical theory of communication, has deeply influenced the conceptualization of brain operations. It has become a method of choice to analyze neural activity because of its many advantages [4–8]. It provides single-trial measures of how neural activity encodes variables important for cognitive functions such as sensory stimuli, and it is thus more relevant for single-trial behavior than trial-averaged measures. It captures contributions of both linear and non-linear interactions between variables at all orders, and thus allows hypotheses-free measures of information encoding that place upper bounds to the performance of any decoder. Because of its generality, it can be applied to any type of brain activity recordings. Also, it facilitates direct comparisons between the predictions of normative neural theories and real neural data [6,9].

Earlier work using information theory to analyze empirical neural data has focused on low-dimensional measures of neural activity such as single neurons, small neural populations or aggregate measures (LFPs, M/EEG, fMRI). These studies have considered only how information is encoded in neural activity, regardless of how it may be used downstream. Such seminal studies have demonstrated, e.g., how the temporal structure of neural activity (from single-neuron spike timing to network oscillations [10–16]) contributes to sensory encoding, or how neural mechanisms such as adaptation contribute to brain information processing [8,17].

Over the last decade, neuroscience has seen major progress in the ability to record simultaneously the activity of many neurons and/or brain areas. These advances have driven the development of novel information theoretic analytical tools to investigate how information processing emerges from the interaction and communication among neurons or areas. Studies have provided multivariate information tools to individuate when synergy and redundancy arise in small populations, or to understand the mechanisms for generating redundancy and synergy, for example to characterize how correlations between the activity of different neurons shape information processing [1,18–22]. Recent work has also coupled information theory with dimensionality-reduction techniques to study how information is encoded in populations of tens to hundreds of cells [23–30]. Other studies have developed multivariate information theory to quantify transmission, rather than encoding, of information across neurons or areas [31–41]. These methods measure the overall or stimulus-specific information exchanged between simultaneously recorded neurons and areas and determine whether transmission relies on synergistic integration of information across nodes. Another major direction of progress has been in recording neural activity during behavior [42]. To support the growing interest on how neural computations shape behavior, information theory has produced tools to characterize

the multivariate simultaneous relationship between sensory stimuli, neural activity and behavioral output to enable quantifying the impact on behavior of the information encoded in a certain area or population [24,43,44].

While the use and dissemination of information theoretic algorithms has been aided by software toolboxes [45–61], no toolbox yet provides a comprehensive implementation of tools to compute both information encoding and transmission, to break down information into components reflecting the effect of interactions and to quantify behavioral or downstream relevance of the encoded information (see Table A in S1 Appendix). To fill these gaps and address the need to collect organically these tools in a format that allows immediately multiple analyses, here we introduce a new Multivariate Information in Neuroscience Toolbox (MINT). MINT provides a comprehensive set of information theoretic functions (including Shannon Entropy and Mutual Information, directed information transmission measures, information decompositions) and estimators (binned probability estimators, limited-sampling bias corrections). The implemented information-theoretic functions are detailed in S1 Appendix. What they compute, and how they can be used in neuroscience is summarized in Table 1. The accuracy and applicability of these algorithms has been validated and demonstrated extensively with both discrete neural variables, such as spikes in electrophysiological recordings [8,13,14,62–64], and continuous neural variables, e.g., LFP, M/EEG, fMRI and calcium traces [24,30,65–67].

Importantly, as we demonstrate with examples, combining these multivariate tools enables addressing questions that cannot be addressed with a single tool. For example, combining tools to identify the specific contribution of correlations to population encoding or the amount of encoded information that informs behavior with dimensionality-reduction techniques allows understanding how large neural populations influence behaviors. Combining information encoding tools with content-specific information transmission tools can reverse engineer information flow in neural networks with unprecedented understanding. We thus anticipate that MINT will lead to uncover numerous new insights into neural information processing.

Design and implementation

MINT is written in MATLAB (version 2018b or newer) and depends on the Statistics and Machine Learning, Optimization, Parallel Computing and Signal Processing Toolboxes. MINT takes as input neural data (array of neural activity recorded in each trial) and task variables (sensory stimuli or behavioral responses presented or produced in each trial). It outputs information values and their null-hypothesis values for computing statistical significance. Fig A in S1 Appendix illustrates MINT functions, options, and core routines.

MINT computes Entropy (H.m), which measures neural variability; and Mutual Information (MI.m), which measures information encoding (Fig 1B and 1C). It computes the Information Breakdown of Shannon Information into contributions due to correlations between neurons [18,20,21,68,69]. It also computes Partial Information Decompositions (PID) [70, 71] of the information about a target variable carried by two or more source variables into unique, synergistic and redundant information (Fig 1C). Computation of PID requires specifying a redundancy measure, which can be selected by the user among options [52,70,72,73] with complementary advantages. (Redundancy of [52] requires either a MATLAB-compatible C compiler or pre-compiled files made available by us for Windows 11, macOS, and Linux Debian). MINT computes additional functions of neuroscientific value: Intersection Information (II, function II.m; Fig 1B, see [44]), the amount of stimulus information in neural activity that is used to inform behavior; Transfer Entropy (TE, see [74]) and Feature-Specific Information Transfer (FIT, see [36]), which measure overall and stimulus-feature-specific information transmission between nodes of neural networks (TE.m, FIT.m and cFIT.m; Fig 1D).

The information quantities depend on the probabilities of task variables (e.g., presented sensory stimuli) and neural responses. MINT implements the *direct method* [14,75] estimator based on discretizing neural responses and task variable values and computing the empirical occurrences across experimental trials of the discrete or binned responses. These estimators have been widely used in neuroscience information theoretic studies, because neural spiking activity is intrinsically discrete and is usually quantified as the number of spikes emitted in one- or multiple-time windows of

Table 1. Glossary of main information theoretic functions.

Function	What is Computed	What it is Used for
Entropy (H)	Variability of random variable X	Assess variability of neural activity
Mutual Information (MI)	How well an ideal observer can predict X from single-trial observations of Y	How well activity of population of neurons encodes info about task variables (e.g., stimuli)
Information Breakdown	MI about X carried by a multivariate Y, broken down into contributions arising from correlations between different dimensions of Y	How correlations between multivariate neural activity (e.g., activity of populations of neurons) contribute to population-level encoding of task variables (e.g., sensory stimuli)
Partial Information Decomposition (PID)	Decomposes MI carried by multidimensional Y about X into unique info about X carried by each element of Y, synergistic info found only in the interactions among elements of Y, and redundant info shared among elements of Y	Whether groups of neurons or brain areas carry synergistic or redundant info about task variables (e.g., stimuli) or about activity of other neurons or brain areas
Redundancy-Synergy Index (RSI)	Quantifies whether the effect on encoding info about X of interactions between different variables within a multi-dimensional Y is predominantly redundant or synergistic	Whether groups of neurons or brain areas carry predominantly synergistic or redundant info
Intersection Information (II)	Info about X carried by Y which is used to inform Z about X	How much of the info about a task variable (e.g., sensory stimulus) encoded in neural activity is used to inform behavioral reports (e.g., choices)
Transfer Entropy (TE)	MI about past activity of sender X found in future activity of receiver Y, conditioned on past activity of Y	Measures overall transmission of info between nodes of a neural network
Feature Specific Information Transfer (FIT)	MI about feature S presently encoded by Y redundant with MI about S previously encoded by X and unique w.r.t. MI about S previously encoded by Y	Measures info transmission between nodes of a neural network of info about a specific feature of task variables (e.g., sensory stimuli)
Conditional TE (cTE) and Conditional FIT (cFIT)	Versions of TE and FIT with info flow conditioned or unique w.r.t. another variable	Measure overall or feature specific transmission of info between network nodes discounting info possibly passing through other nodes
Supervised dimensionality reduction (decoding)	Project data onto lower-dimensional space using labeled data to optimize decoding	Intermediate step for information calculations with large neural populations
Unsupervised dimensionality reduction	Project data onto lower-dimensional space using unlabeled data to optimize data explainability	Intermediate step for info calculations with large neural populations
Limited-sampling bias correction	Produces unbiased info estimates unaffected by the limited-sampling bias	Needed to obtain more accurate info estimates in all practical situations
Hierarchical Data shuffling	Random permutations of data used to create null distribution for statistical testing for assessing the role of certain data features in info encoding	Needed for statistical testing in all applications. Useful to assess role of spike timing or correlation between neurons by comparing info values obtained with these features preserved or shuffled

This table reports a short explanation of what the implemented information theoretic quantities compute and for what type of applications they may be used. X, Y, Z denote random variables.

<https://doi.org/10.1371/journal.pcbi.1012934.t001>

interest [14,75]. The direct method captures the information carried by spike counts very precisely (Fig C in S1 Appendix). Because they are simple and do not make assumptions about the probability distributions, discretized estimators have been used to compute information also from continuous-valued aggregate measures of neural activity such as LFP, M/EEG, fMRI [11,65,66,76] or continuous-valued behavioral variables [77]. If the scientific question at hand needs PID in addition to Shannon information and the data are not Gaussian, then discrete or discretized approaches are advised (as non-discrete non-parametric estimators are available only for Shannon information and entropy). MINT provides binning functions to discretize analogue data (equi-spaced or equi-populated binning, binning with user-defined bin edges, and possibly automated determination of bin numbers [78,79]).

Any real experiment only yields a finite number of trials from which probabilities must be estimated. Finite sampling when using direct methods leads to a systematic error (bias) in information estimates (Fig 1E), which can be as big as the

true information values. Thus, plugin estimators (based on just plugging in probabilities into the information equations) are biased, and bias corrections methods need to be added to it and are essential for practical neuroscience applications. Six such well-established methods are included in MINT [80–85]. These methods, along with binning, parallelization options and other features are user-specified in an input structure (opts). Information (function MI.m) is computed by default with the plugin method, as it preserves all information available in the discretized neural activity. We recommend its use for small-dimensional (up to $N=2$ or 3) neural response (e.g., responses of populations of up to 2–3 neurons) as its estimates from datasets of realistic sizes can be still effectively corrected for the limited-sampling bias (Fig C in S1 Appendix).

Alternatively, probability estimators suited for real-valued data [86–88], such as nearest-neighbors or kernel methods, can be used to estimate information and are available in MINT by specification in the input structure (opts). These methods also work well for low-dimensional data.

Neither these estimators nor the direct method, however, work on their own when considering high-dimensional neuronal responses (such as the activity of populations of many neurons), as the curse of dimensionality prevents the direct sampling of the joint response probabilities from high dimensional data (Fig C in S1 Appendix). We thus provide additional pipelines, recommended for high-dimensional neural responses such as the activity of large neural populations, that compute information from the empirical neural response probabilities but after reducing the dimensionality of neural population activity [24]. These *dimensionality-reduction pipelines* include supervised methods (Support Vector Machines, SVMs and Generalized Linear Models, GLMs) which reduce the dimensionality by providing decoding or posterior probabilities of the task variables given the single-trial neural population activity (Fig 1A) and allow reliable estimations with small datasets (Fig C in S1 Appendix). We also provide unsupervised methods (Non-negative Matrix Factorization, NMF [89]; Principal Component Analysis, PCA) which reduce dimensionality individuating small numbers of dimensions with the highest explanation power of neural activity. Supervised dimensionality-reduction algorithms that individuate the directions in neural activity space with most discriminability of the task variable (e.g., SVM) may be in general better suited than unsupervised algorithms individuating dimensions that target best reconstruction of the spike trains (e.g., PCA, NMF) when the most information is not encoded in the direction with most variations in neural activity space (Fig C in S1 Appendix).

MINT provides all these dimensionality-reduction techniques with native MATLAB functions, but it also allows easy interfacing with external libraries (e.g., libsvm [90] and glmnet [91]) (Fig B in S1 Appendix). Importantly, these dimensionality-reduction tools can be coupled (Fig 1F) with MINT's *Hierarchical Shuffling* tool (hShuffle.m) which can disrupt, by trial shuffling, specific features of population activity (such as response timing or correlations between neurons) to probe their contribution to information processing [24,92].

When deciding which estimator to apply to a given dataset, we recommend users to test different algorithms on synthetic data that match essential features of the experiments (e.g., discrete spike counts or continuous signals, number of trials and data dimensionality, information levels) and choose what suits best. MINT provides a simulator of correlated or uncorrelated neural population spike train activity that can be used for this purpose.

Results

We illustrate how to use MINT to address highly topical neuroscientific questions, emphasizing the utility of using synergistically multiple algorithms, allowed by MINT. In all examples, we use the limited-sampling bias corrections and hierarchical data shuffles of MINT, as they are essential for empirical data analyses.

Computing the role of interactions between neurons in information encoding

An important question in neuroscience is whether and how the functional interactions (measured as activity correlations) between neurons enhance or limit information encoding in neural populations [1,93]. Several information theoretic methods have been developed to address complementary aspects of this question [18–22,68,70,92,94]. Here we illustrate what we gain from their combined use enabled by MINT.

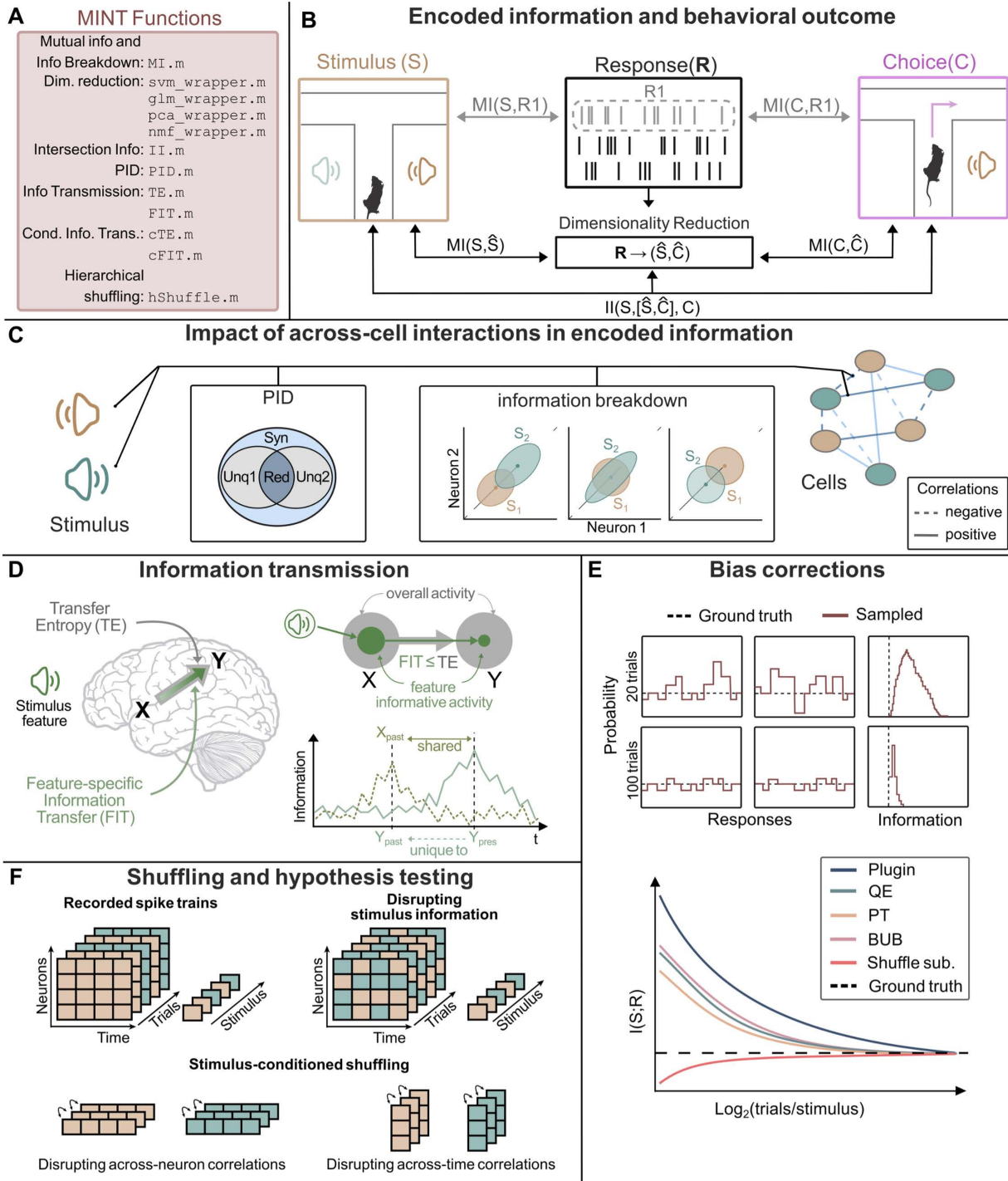


Fig 1. Overview of MINT. A: List of main MINT functions. B: MINT provides multivariate information theoretic functions to quantify the amount of information that single neurons or neural populations carry about task-relevant variables (e.g., sensory stimuli or behavioral choices). These methods are based on either direct-method calculation of information from the discretized probabilities (ideal for small neural populations but not scalable with population size), estimation through other techniques such as Kernel-based methods that can operate on real-valued data, or by using supervised or unsupervised dimensionality-reductions techniques to approximate high-dimensional neural population response probabilities with probabilities in lower-dimensional spaces (scalable with population size). It also provides tools to quantify how much of the information encoded by neural activity is used to inform behavior. C: MINT has multiple functions to compute, in small or large populations, how interactions between the activity of different neurons

shape information encoding and create synergy or redundancy. D: MINT has tools to compute transfer of information from one neural population or brain region to another. It can compute both the total or stimulus-specific information transmitted between two nodes, with the option of conditioning over the activity of other nodes. E: MINT has tools to correct plugin estimators for the limited-sampling bias, an essential tool for analysis of empirical neuroscience data. F: MINT has a set of hierarchical permutation algorithms that provide null hypothesis testing for significance of information encoding and information transmission and for the impact of correlations across neurons or time. Mouse sketch is modified from doi.org/10.5281/zenodo.3925917, brain sketch is modified from doi.org/10.5281/zenodo.3925989, and speaker sketch is modified from svgrepo.com/svg/506329/speaker-2. All resources are licensed under CC BY 4.0 (creativecommons.org/licenses/by/4.0).

<https://doi.org/10.1371/journal.pcbi.1012934.g001>

We consider how a population of N neurons encodes information about a stimulus variable S . For neuron pairs ($N=2$), we computed the population information (Mutual Information between stimulus and the joint neural population response) with the direct method that estimates information directly from the empirical discretized response probabilities (see Design and implementation). The overall effect of interactions between neurons is expressed by the Redundancy-Synergy Index (RSI), the difference between the population information and the sum of single-neuron stimulus information [19,95]. Positive (negative) RSI indicates predominantly synergistic (redundant) interactions. Contributions of synergy and redundancy can be separated using PID [71,96]. The Information Breakdown [20–22] shows how RSI arises from interactions between neurons, by breaking down RSI into components $I_{sig-sim}$ (contribution of the similarity across neurons of trial-averaged responses to different stimuli, see also [19]), $I_{cor-ind}$ (contribution of the interplay between the signs of signal similarity and of noise correlations, defined as correlations between neurons in trials to the same stimulus), and of $I_{cor-dep}$ (quantifying information added by the stimulus-modulation of noise correlations, or, equivalently, bounding the information lost when using decoders trained without considering correlations [18,69]).

These small-population direct information calculations have the advantage of not making assumptions about decoding mechanisms, but do not scale up to large populations because of the curse of dimensionality [82]. Population information can be obtained by estimating probabilities in the reduced space of the stimuli decoded from single-trial neural activity. These estimates scale well with population size and can be computed robustly with small datasets (Fig C in [S1 Appendix](#)). However, specific decoders may severely underestimate total information in neural activity (see Fig C in [S1 Appendix](#)), especially when the decoder does not operate on the features of neural activity that carry most information. We illustrate below how MINT allows determining the role of correlations in population coding by comparing decoders that do or do not use information in correlated activity and by leaving intact or removing information in correlated activity using hierarchical shuffling tools [1,24,92].

We illustrate these methods first by simulating the activity of $N=20$ neurons responding to two stimuli. In the first simulated scenario (Fig 2A), only correlations between activity of different neurons, but not the single-neuron activities, are stimulus-modulated and thus encode stimulus information. The single cell information is zero, but the pairwise population information is not. Positive RSI arises because of large synergy with negligible redundancy. The Information Breakdown reveals that all the synergistic information is due to stimulus-dependent correlations. Population decoding with SVM of the $N=20$ neurons reveals that large-population information can be accessed exclusively with a non-linear decoder, and that shuffling correlations destroys all information, confirming it is exclusively encoded by correlations.

In the second simulated scenario (Fig 2B), information is encoded by single cells, correlations are only weakly stimulus-modulated, all neurons have equal stimulus tuning (responding more strongly to stimulus 2), and noise correlations are positive. In this configuration, redundancy is created (all neurons have the same trial-averaged response profiles to the stimuli) and correlations reduce information (they are elongated along the axis separating the mean firing rates of individual neurons and thus increase the overlap between the stimulus-specific distributions of neural activity) [20]. Negative RSI arises because of larger redundancy (created by so called signal similarity expressing the similarity of tuning to stimuli of individual neurons) than synergy (created by the small but present stimulus-modulation of correlations). Information Breakdown analysis reveals that indeed information is more redundant because the signal-noise similarity (captured by $I_{sig-sim}$ and $I_{cor-ind}$) is larger than the small stimulus-dependent correlations $I_{cor-dep}$. In the large ($N=20$) population most

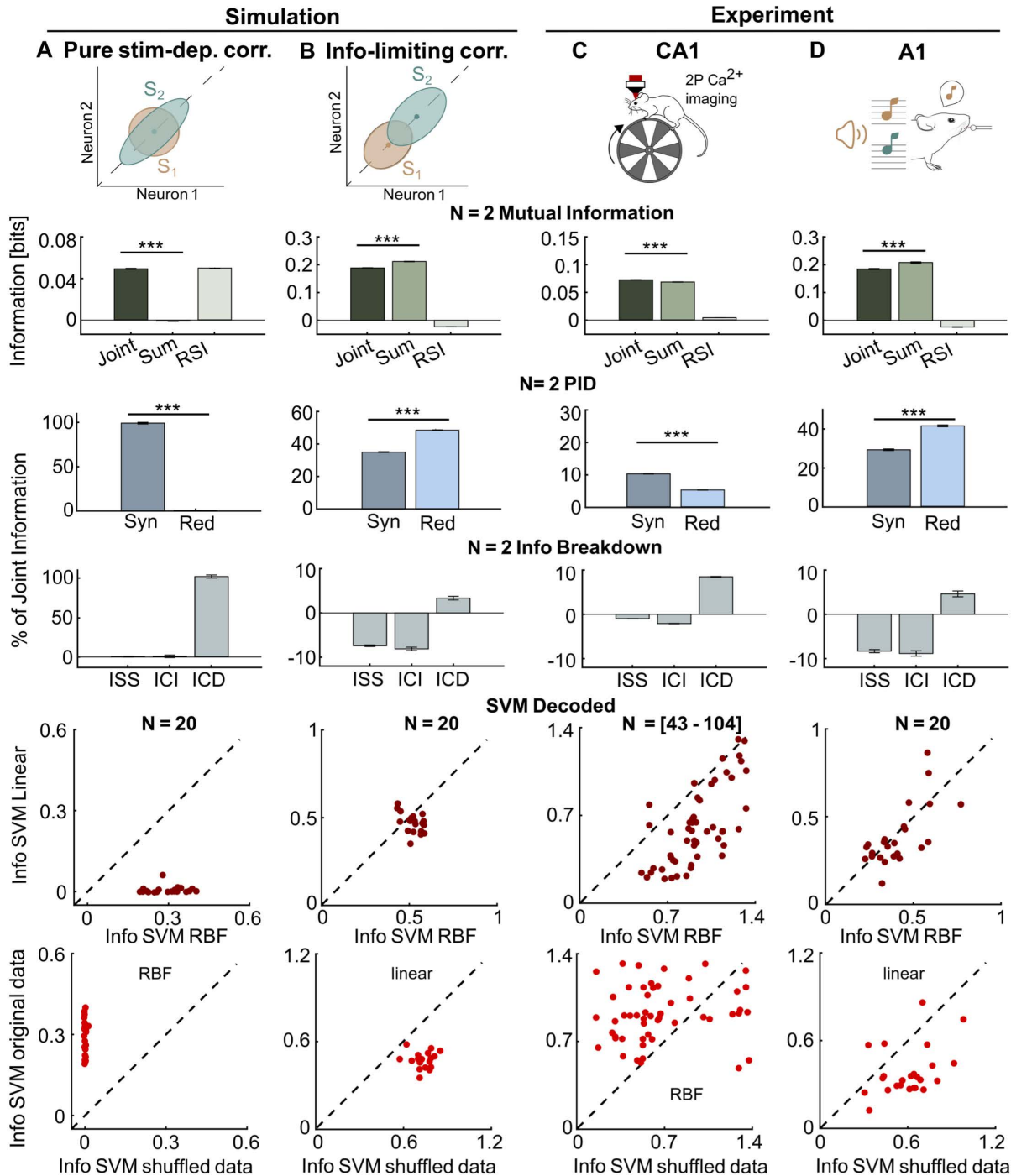


Fig 2. Assessing the role of correlations among neurons in neural population encoding. In each column we consider analysis of a different dataset. **A:** simulated population of $N=20$ neurons which carry information only by stimulus-dependent correlations, with no stimulus information provided by single-neuron firing rate modulation. **B:** simulated population of $N=20$ neurons which carry information by single neuron firing modulations and which have information-reducing correlations. **C:** CA1 recordings of $N=[43-104]$ neurons over $n=11$ sessions during spatial navigation of a linear track in virtual reality. **D:** A1 recordings of $N=20$ neurons over $n=12$ sessions during tone presentation. In each row, we plot from top to bottom: direct calculation of information for neuron pairs and of sum of single neuron information; direct calculation of redundancy-synergy index (RSI), of synergy and redundancy separately and of the Information Breakdown components for neuron pairs; calculation of encoded information of the whole population using the information in the confusion matrix of an SVM decoder (linear or RBF), computed either on the real population responses (which contain correlations

between neurons) or pseudo-population “shuffled” response obtained collecting randomly permuted trials to the same stimulus (shuffling removes correlations at fixed stimulus). In columns A-B we compute Shannon Information between neural activity and the identity of the two simulated stimuli. In column C-D we compute Shannon Information between neural activity and the identity of the presented tone ($S=2$ different tones) or the spatial location of the mouse (binning locations into $S=12$ equi-distant spatial bins), respectively. In column C, direct measures of pairwise information were obtained with $R=2$ equi-populated bins (appropriate for this dataset consisting of non-deconvolved calcium fluorescent traces). In column D, direct measures of pairwise information were obtained with $R=3$ bins, done by capping to 2 spike counts (appropriate for this dataset consisting of calcium signals deconvolved to estimate firing rates and activity counted in short windows). In each panel we plot mean \pm SEM (for simulated data in panel A-B: over $n=190$ neural pairs and $n=10$ simulation repetitions for the direct information calculations; over $n=5$ different data folds and $n=10$ simulation repeats for the decoding information values; for CA1 data in Panel C: over $n=36158$ simultaneously recorded neuron pairs for the direct information calculations, and over $n=11$ recording sessions and $n=5$ trial folds for the decoding information values; for A1 data in Panel D: over $n=2280$ simultaneously recorded neuron pairs for the direct information calculations, and over $n=12$ recording sessions and $n=2$ trial folds for the decoding information values). Symbols *, **, *** denote two-tailed $p < 0.05$, $p < 0.01$, $p < 0.001$ respectively, computed with paired t-tests. See SM6.1, SM7.2 and SM7.3 in [S1 Appendix](#) for details of simulations and real data analysis. Mouse sketch in Panel D is modified from doi.org/10.5281/zenodo.3925985, speaker sketch is modified from svgrepo.com/svg/506329/speaker-2 and music note sketch is modified from svgrepo.com/svg/458908/sound. All resources are licensed under CC BY 4.0 (creativecommons.org/licenses/by/4.0).

<https://doi.org/10.1371/journal.pcbi.1012934.g002>

information can be accessed with a linear SVM, with the non-linear SVM adding relatively little, and noise correlations reduce information (shuffling them away increases information).

We then applied the same analyses to two real neural datasets. We first analyze encoding of the mouse position (within a linear track) by populations of 43–104 simultaneously recorded neurons from the CA1 region of the mouse hippocampus [27] (Fig 2C). With the pairwise analysis, PID shows that both synergy and redundancy are present, but synergy is larger and the Information Breakdown shows that this is due to modulation of the noise correlations strength with the position ($I_{cor-dep} \sim 10\%$ of the pairwise information). Using a nonlinear decoder of the whole population increases information by $\sim 80\%$ over what could be achieved with linear decoders, and shuffling data to destroy correlations decreases the nonlinearly decoded information by $\sim 80\%$, revealing a large effect of hippocampal noise correlations in position encoding by large neural populations, whose size could not be inferred by neuron pairs analysis.

We then analyzed encoding of sound intensity by populations of 20 neurons simultaneously recorded from the mouse auditory cortex (A1) during pure-tone sound presentation (Fig 2D). These networks were selected, among all recorded neurons, based on their encoding of task-relevant information in [97]. With the pairwise analysis, PID shows that both synergy and redundancy are present, but redundancy is larger. Information Breakdown analysis shows that this is due to negative $I_{sig-sim}$ (neuron pairs have similar tuning to the stimuli) and $I_{cor-ind}$ (most neural pairs have also positive correlations), with $I_{cor-dep}$ contributing much less. Decoding whole-population activity with a nonlinear SVM did not increase the information decoded with a linear SVM (stimulus-dependent correlations were weak), and shuffling away noise correlations increases information substantially (thus correlations strongly reduced information).

Together, these results illustrate the power of combining MINT tools to understand deeply how interaction between neurons shape neural population coding.

Computing how stimulus information encoded in neural population activity informs behavioral discriminations

Traditional approaches to neural information encoding of sensory stimuli have focused solely, as in the above examples, on how neurons or populations encode information about these stimuli. However, it could be that little or none of the information they encoded is actually utilized to inform behavior. It is thus important to have instruments to understand how much information in neural activity contributes to behavior.

Intersection Information (II) measures how much of the sensory information encoded in neural population activity is read out to inform behavior (Fig 3A; see sketch in Fig 1B), and is computed with PID (using the tri-variate probabilities of stimuli, neural activity and behavioral choices) as the component of neural information that is both about stimulus and choice [24,43,44]. To demonstrate its use, we applied it to analyze the activity of populations of neurons recorded with 2-photon calcium imaging in mice in auditory cortex (A1) during pure-tone perceptual discrimination [97] (Fig 3B).

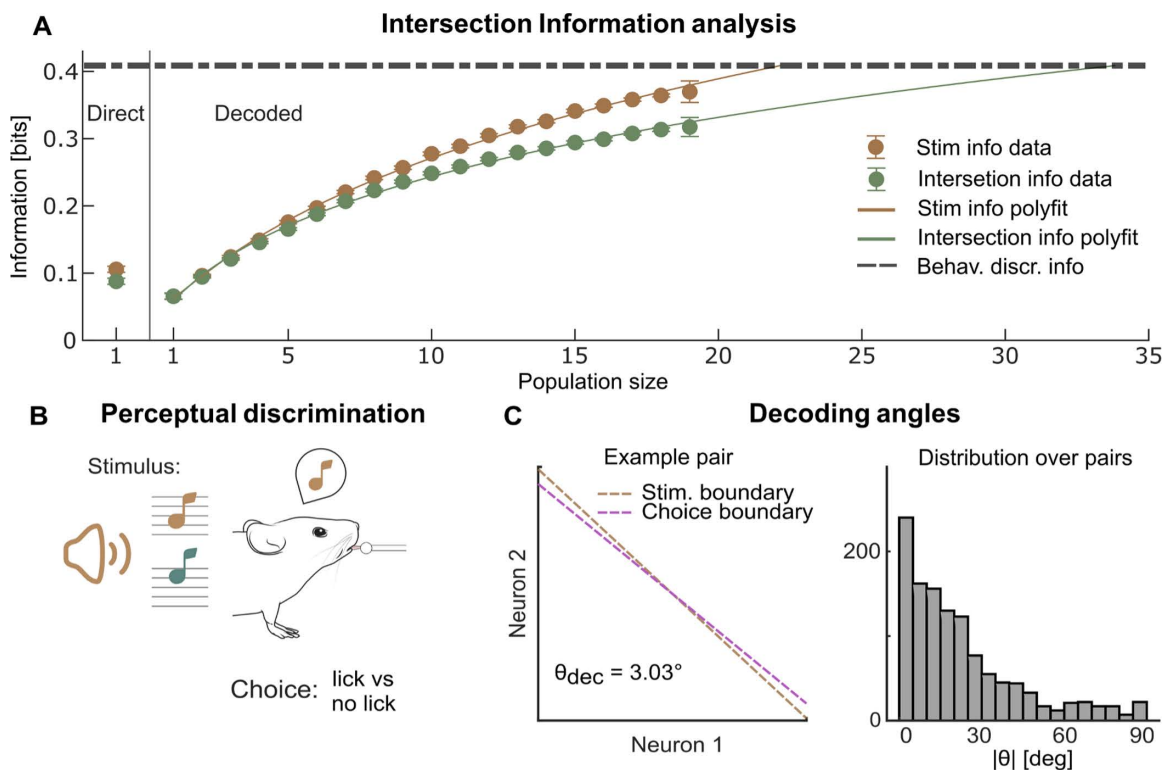


Fig 3. Stimulus and Intersection information coding in populations of cortical auditory neurons during a tone discrimination task. A: Stimulus information and Intersection Information encoded in neural activity recorded during a sound tone discrimination task. Left: single cell estimates using the direct method. Right: estimates of the information quantities using a RBF SVM (2-fold cross validation) as function of the population size. We plot the mean \pm SEM over all $n = 12$ Field of Views and over all folds and over all subpopulations used. For population sizes $N = 2-18$, for >100 independent subpopulations can be obtained, we shortened computation time using only $n = 100$ randomly sampled subpopulations. For population size $N = 1$ and 19, we used all the $n = 20$ different subpopulations available. For the direct information calculation, we used 3 bins for 0 spikes, 1 spike and any value above 1. For all information analyses, we used the shuffle subtraction to correct for the limited-sampling bias. The dashed horizontal line plots the averaged information needed to explain behavioral discrimination accuracy (computed as the information between stimulus and choice). Full lines show log-polynomial fits to the dependence of stimulus and intersection information on population size. The population size with information sufficient to explain behavioral discrimination accuracy is the x-axis intercept of the point at which the fit lines cross. B: Schematic of the behavioral task in mice used when recording the data analyzed in this figure. C: Stimulus and choice boundary computed with MINT in the space of paired neural activity for one example neural pair in the dataset. The value of the angle between the two axes is reported in the inset. Right: distribution of the absolute value of the angle between the stimulus and choice boundaries for the $n = 2280$ neural pairs in this dataset. See SM6.2 and SM7.3 in [S1 Appendix](#) for details of simulations and real data analysis. Mouse sketch is modified from doi.org/10.5281/zenodo.3925985, speaker sketch is modified from svgrepo.com/svg/506329/speaker-2 and music note sketch is modified from svgrepo.com/svg/458908/sound. All resources are licensed under CC BY 4.0 (creativecommons.org/licenses/by/4.0).

<https://doi.org/10.1371/journal.pcbi.1012934.g003>

We first considered information encoded in single neurons, computed with the direct method. If the readout of the stimulus information in neural activity was optimal (respectively, completely suboptimal), I_l would equal the stimulus information, (respectively, be zero). We found that for single neurons, I_l was $\sim 90\%$ of the total single-neuron stimulus information, showing that information encoded by these neurons is not read out optimally but still efficiently.

For sampling reasons explained above, the direct calculation of I_l can be done for small ($N = 1-3$), but not for large populations. How can we use I_l to address how information relevant to behavior scales with population size? Specifically, how large must a population be to account for perceptual discrimination ability? To answer this, in MINT we combined I_l with dimensionality-reduction techniques. In this application, we used an SVM to compress neural activity (using `svm_wrapper.m` before `I_l.m`). This compression loses some information (the information values obtained with the plugin method are

~20% higher than the single cell values obtained with SVM decoders; Fig 3A). However, II population information computed with SVM decoders are scalable and data-robust (Fig C in S1 Appendix). Computing how information scales with population size (Fig 3A) shows that as population size increases, the gap between stimulus information and II widened. This means that behaviorally-relevant information is more redundant across neurons compared to information that is not used to inform behavior, confirming the usefulness of redundancy for behavioral readout [98]. Had we considered only stimulus information, we would have incorrectly concluded that ~23 such neurons are sufficient to account for the mouse discrimination performance (Fig 3A). However, taking intersection information into account reveals that ~34 such neurons are instead needed to fully account for the perceptual discrimination ability, as not all stimulus information encoded in neural populations is read out (Fig 3A).

We endowed MINT with instruments to characterize neural mechanisms of readout. Suboptimality may arise because of a misalignment between how information is encoded and how the brain reads it out to inform choices [43]. MINT returns the axes in neural activity space trained to discriminate between stimuli and the axes trained to discriminate different choices (using `svm_wrapper.m`, see Fig D in S1 Appendix for examples on simulated data). Computing decoding angles of pairs of A1 neurons (Fig 3C) shows that most pairs had a small but non-zero misalignment between stimulus and choice decoders, which explains the efficient but sub-optimal readout.

In sum, combining Intersection Information and dimensionality reduction can give precise insights about the behavioral relevance of information encoded by neural populations.

Mapping content-specific encoding and transmission of information within a network

MINT provides both algorithms to study information encoding in individual network nodes and information transmission across nodes. We here illustrate how to combine them for reverse-engineering the information flow within neural networks.

We first simulated a network with four nodes X_1, \dots, X_4 each modeling the aggregate activity of a brain area (as, e.g., measured by aggregate neural signals such as LFPs, M/EEG or fMRI, see SM 6.3 in S1 Appendix). This network has a well-defined ground-truth flow of information about two independent stimulus features S_1, S_2 (Fig 4A). Information about S_1 is received from the outside by nodes X_1 and X_4 in a short time window (3–12 ms from simulation start for X_1 and 15–24 ms for X_4), and is then sent from X_1 to X_2 and X_3 with a 5 ms delay. Information about S_2 is received (in the 3–12 ms window) from the outside by X_2 which then sends it to X_1 with a 5 ms delay. Nodes X_3 and X_4 exchange information (also with a 5 ms delay) which is not about S_1 or S_2 . To disentangle the information flow, we computed (using the direct method) information encoded or transmitted at each time (in Fig 4 we plot for each node and link the maximal information values over time, but we show in Fig E in S1 Appendix that time-resolved analysis reconstructs correctly the ground-truth information encoding windows and communication delays), and we used MINT's non-parametric permutations tests to identify significant encoding or transmission. Using Mutual Information between individual stimulus features and individual node activity reveals correctly that all nodes have information about S_1 and that only X_1 and X_2 have information about S_2 (Fig 4B). To study how this information is exchanged within the network, we first computed overall information transfer with Transfer Entropy, finding correctly significant transfer from X_1 to X_2 and X_3 , from X_2 to X_1 , and from X_3 to X_4 (Fig 4C). To reveal the information content of this exchange we computed Feature Specific Information Transfer (FIT), revealing correctly that the information transferred from X_1 is about S_1 but not about S_2 , and that the information transferred from X_2 is about S_2 (Fig 4D). FIT finds no information transfer from X_3 to X_4 about S_1 or S_2 , thus determining correctly that the overall information transfer from X_3 to X_4 detected with TE is not about any of the two stimulus features. Finally, the finding that X_1 and X_4 encode information about S_1 while they do not receive it from other network nodes implies that X_1 and X_4 receive external S_1 information. Similarly, because X_2 encodes information about S_2 while not receiving within-network S_2 information demonstrates that X_2 receive external S_2 information. Thus, combining encoding with transmission analyses could correctly reverse engineer the within-network specific information flow.

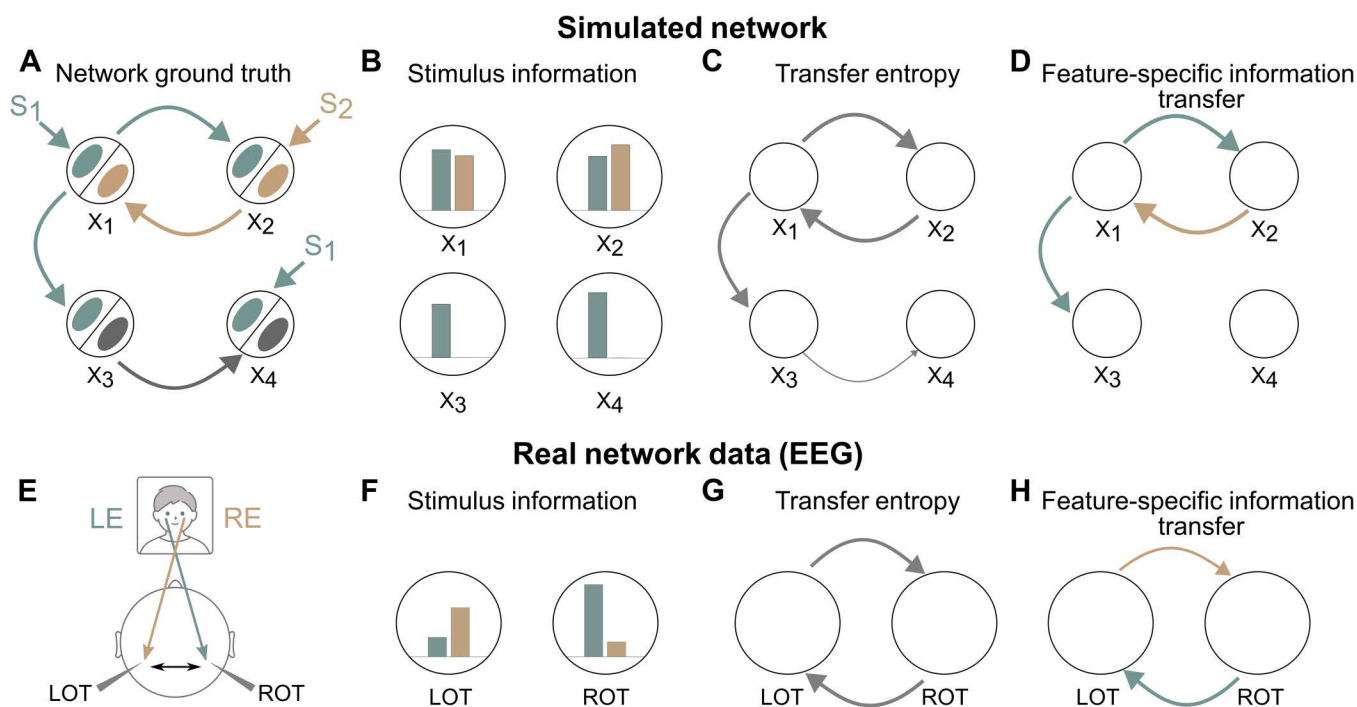


Fig 4. Reverse engineering information flow using stimulus-encoding and stimulus-transfer estimation algorithms. Panels A-D test MINT on simulated network data. A: Schematic of the simulation. The network comprises four neural nodes (black circles) X_1, \dots, X_4 , each containing two subpopulations (ellipses within the circles) encoding two independent binary stimulus features X_1, S_2 . The ground-truth stimulus specific information communication is plotted in Panel A, with grey color used to indicate no stimulus selectivity, and green and brown colors used to indicate information selectivity to X_4 and S_1 respectively. B: Maximum Mutual Information across time between each neural population S_2 and the stimuli S_1 and S_2 . C: Transfer Entropy (TE) between nodes. D: FIT about X_i and S_1 between nodes. In panels C-D, only significant ($p < 0.01$, permutation test) links are plotted, with thickness proportional to the computed value. In each panel we plot the average information values across $n = 10$ simulation repeats. Panels E-F test MINT on real human EEG data. E: Schematic of the putative information flow inter-hemispheric information flow. LOT (ROT) denote Left (Right) Occipito-Temporal regions. LE (respectively RE) denote the Left (respectively Right) Eye face visibility feature. F: Maximum Mutual Information across time about the left or right eye visibility present in left of right OT region. G: Significant Transfer Entropy between LOT and ROT brain regions. H: Significant FIT between LOT and ROT brain regions. In panels G-H, only significant ($p < 0.01$, permutation test) links are plotted, with thickness proportional to the computed value. In each panel we plot the average information values across $n = 15$ experimental subjects. See SM6.3 and SM7.1 in [S1 Appendix](#) for details of simulations and real data analysis. Human face sketch is modified from [svgrepo.com/svg/493087/men-in-their-20s-and-30s-face](#), and head sketch is modified from [doi.org/10.5281/zenodo.3926093](#). All resources are under license CC BY 4.0 ([creativecommons.org/licenses/by/4.0](#)).

<https://doi.org/10.1371/journal.pcbi.1012934.g004>

We next tested how MINT reverse-engineers information flow in real brain networks by applying it to an existing EEG dataset recorded from human participants detecting the presence of either a face or a random texture from images covered by random bubble masks [99]. Prior work [36,99,100] revealed that the visibility of the eye region (proportion of visible pixels in the eye area) is critical for successful face discrimination and that the Occipito-Temporal (OT) EEG electrodes are those encoding most Mutual Information about both left and right eye visibility (Fig 4E and 4F). To understand if some of this information was exchanged across the OT regions in different hemispheres, we used TE and FIT to analyze transmission of left or right eye visibility information across OTs. TE across hemispheres was found in both directions (right-to-left and left-right), suggesting a bi-directional inter-hemispheric communication (Fig 4G). However, specific information transfer was precisely directional: FIT about the left eye was only from right-to-left and FIT about the right eye was only from left-to-right (Fig 4H). Thus, using MINT allowed establishing encoding and directional transfer of different eye features across hemispheres with high specificity. These analyses could also temporally localize both encoding and inter-hemispheric transfer (Fig F in [S1 Appendix](#)).

Together, these results illustrate the power of combining MINT tools to reverse-engineer encoding and flow of specific information across brain networks.

Availability and future directions

MINT is downloadable in source code (<https://github.com/panzerilab/MINT> with DOI doi.org/10.5281/zenodo.13998526), including a Dockerfile, and is licensed under GNU GPLv3. It contains documentation on using it and on building and installing it from source, unit tests, use examples, and replication of paper figures (https://github.com/panzerilab/MINT_figures).

The modularity of MINT allows it to be used alongside any other MATLAB function or toolbox. As exemplified above, we already provide pipelines for interfacing with decoding toolboxes. We plan to add plugins to generate neural and behavioral data from data acquisition and preprocessing toolboxes (e.g., [101]) with MINT's input-data format requirements, and to generate MINT's outputs suitable to be fed directly into toolboxes for further advanced analyses, e.g., for network analysis of information-transfer outputs [102].

We plan to further extend the range of information-theoretic methodology implemented in MINT. MINT's current version emphasizes discretized maximum likelihood estimators. However, we provide only a handful of data-discretization techniques that go with it. We plan to endow them with optimal discretization algorithms based on model selection techniques (Akaike and Bayesian information criteria). While MINT already implements a number of probability estimators for real-valued data we plan to extend them to include other binless and kernel-based estimators [103,104], and parametric probability models (Gaussian, Poisson) proposed in the neuroscience literature. Although we provide several tools for assessing the role of correlated activity, we plan to implement currently missing Maximum Entropy estimators [63]. Finally, the derivation of new neuroscience-related information quantities with PID is highly active [71,105] and the open source and modularity of MINT will allow rapid integration of new developments.

A limitation that may restrict MINT's usage is that it is developed only in MATLAB at this stage. We are thus developing a translated Python version of MINT to widen usage. However, we verified that MINT is usable from Python using the MATLAB Engine API for Python and we provide instructions in SM2 in [S1 Appendix](#).

Supporting information

S1 Appendix. Supplementary material. Theoretical definitions, descriptions of the available methods in the toolbox, detailed descriptions of the simulation specifications used in the figures as well as supplementary analysis and figures. (PDF)

S1 Data. CA1 recordings in hippocampus. Calcium imaging recordings used to produce part of the results in [Fig 2](#). The file should be first unzipped with any unzip software and then can be uploaded using the software provided in our MINT toolbox. (GZ)

Acknowledgements

We thank C. Becchio, G. Iurilli and H. Nili for feedback.

Author contributions

Conceptualization: Gabriel Matías Lorenz, Nicola Marie Engel, Stefano Panzeri.

Data curation: Gabriel Matías Lorenz, Nicola Marie Engel, Loren Koçillari, Sebastiano Curreli.

Formal analysis: Gabriel Matías Lorenz, Nicola Marie Engel.

Funding acquisition: Tommaso Fellin, Stefano Panzeri.

Investigation: Gabriel Matías Lorenz, Nicola Marie Engel, Marco Celotto, Loren Koçillari.

Methodology: Gabriel Matías Lorenz, Nicola Marie Engel, Marco Celotto, Loren Koçillari, Stefano Panzeri.

Project administration: Stefano Panzeri.

Resources: Stefano Panzeri.

Software: Gabriel Matías Lorenz, Nicola Marie Engel, Marco Celotto.

Supervision: Stefano Panzeri.

Validation: Gabriel Matías Lorenz, Nicola Marie Engel, Marco Celotto, Loren Koçillari.

Visualization: Gabriel Matías Lorenz, Nicola Marie Engel, Loren Koçillari, Sebastiano Curreli.

Writing – original draft: Gabriel Matías Lorenz, Nicola Marie Engel, Marco Celotto, Stefano Panzeri.

Writing – review & editing: Gabriel Matías Lorenz, Nicola Marie Engel, Marco Celotto, Loren Koçillari, Sebastiano Curreli, Tommaso Fellin, Stefano Panzeri.

References

1. Panzeri S, Moroni M, Safaai H, Harvey CD. The structures and functions of correlations in neural population codes. *Nat Rev Neurosci*. 2022;23(9):551–67. <https://doi.org/10.1038/s41583-022-00606-4> PMID: 35732917
2. Urai A, Doiron B, Leifer A, Churchland A. Large-scale neural recordings call for new insights to link brain and behavior. *Nat Neurosci*. 2022;25(1):1–9. <https://doi.org/10.1038/s41593-021-00980-9>
3. Shannon CE. A mathematical theory of communication. *Bell Syst. Tech. J.* 1948;27(3):379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
4. Borst A, Theunissen FE. Information theory and neural coding. *Nat Neurosci*. 1999;2(11):947–57. <https://doi.org/10.1038/14731> PMID: 10526332
5. Quiñero R, Panzeri S. Extracting information from neuronal populations: information theory and decoding approaches. *Nat Rev Neurosci*. 2009;10(3):173–85. <https://doi.org/10.1038/nrn2578> PMID: 19229240
6. Fairhall A, Shea-Brown E, Barreiro A. Information theoretic approaches to understanding circuit function. *Curr Opin Neurobiol*. 2012;22(4):653–9. <https://doi.org/10.1016/j.conb.2012.06.005> PMID: 22795220
7. Azeredo da Silveira R, Rieke F. The geometry of information coding in correlated neural populations. *Annu Rev Neurosci*. 2021;44:403–24. <https://doi.org/10.1146/annurev-neuro-120320-082744> PMID: 33863252
8. Fairhall AL, Lewen GD, Bialek W, de Ruyter Van Steveninck RR. Efficiency and ambiguity in an adaptive neural code. *Nature*. 2001;412(6849):787–92. <https://doi.org/10.1038/35090500> PMID: 11518957
9. Atick JJ, Redlich AN. Towards a Theory of Early Visual Processing. *Neural Comput*. 1990;2(3):308–20. <https://doi.org/10.1162/neco.1990.2.3.308>
10. Panzeri S, Brunel N, Logothetis NK, Kayser C. Sensory neural codes using multiplexed temporal scales. *Trends Neurosci*. 2010;33(3):111–20. <https://doi.org/10.1016/j.tins.2009.12.001> PMID: 20045201
11. Belitski A, Gretton A, Magri C, Murayama Y, Montemurro MA, Logothetis NK, et al. Low-frequency local field potentials and spikes in primary visual cortex convey independent visual information. *J Neurosci*. 2008;28(22):5696–709. <https://doi.org/10.1523/JNEUROSCI.0009-08.2008> PMID: 18509031
12. Douchamps V, di Volo M, Torcini A, Battaglia D, Goutagny R. Gamma oscillatory complexity conveys behavioral information in hippocampal networks. *Nat Commun*. 2024;15(1):1849. <https://doi.org/10.1038/s41467-024-46012-5> PMID: 38418832
13. Kayser C, Montemurro MA, Logothetis NK, Panzeri S. Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. *Neuron*. 2009;61(4):597–608. <https://doi.org/10.1016/j.neuron.2009.01.008> PMID: 19249279
14. de Ruyter van Steveninck RR, Lewen GD, Strong SP, Koberle R, Bialek W. Reproducibility and variability in neural spike trains. *Science*. 1997;275(5307):1805–8. <https://doi.org/10.1126/science.275.5307.1805> PMID: 9065407
15. Victor JD. Approaches to information-theoretic analysis of neural activity. *Biol Theory*. 2006;1(3):302–16. <https://doi.org/10.1162/biot.2006.1.3.302> PMID: 19606267
16. Gollisch T, Meister M. Rapid neural coding in the retina with relative spike latencies. *Science*. 2008;319(5866):1108–11. <https://doi.org/10.1126/science.1149639> PMID: 18292344
17. Młynarski WF, Hermundstad AM. Efficient and adaptive sensory codes. *Nat Neurosci*. 2021;24(7):998–1009. <https://doi.org/10.1038/s41593-021-00846-0> PMID: 34017131
18. Latham PE, Nirenberg S. Synergy, redundancy, and independence in population codes, revisited. *J Neurosci*. 2005;25(21):5195–206. <https://doi.org/10.1523/JNEUROSCI.5319-04.2005> PMID: 15917459

19. Schneidman E, Bialek W, Berry MJ 2nd. Synergy, redundancy, and independence in population codes. *J Neurosci*. 2003;23(37):11539–53. <https://doi.org/10.1523/JNEUROSCI.23-37-11539.2003> PMID: [14684857](https://pubmed.ncbi.nlm.nih.gov/14684857/)
20. Panzeri S, Schultz SR, Treves A, Rolls ET. Correlations and the encoding of information in the nervous system. *Proc Biol Sci*. 1999;266(1423):1001–12. <https://doi.org/10.1098/rspb.1999.0736> PMID: [10610508](https://pubmed.ncbi.nlm.nih.gov/10610508/)
21. Pola G, Thiele A, Hoffmann KP, Panzeri S. An exact method to quantify the information transmitted by different mechanisms of correlational coding. *Network*. 2003;14(1):35–60. <https://doi.org/10.1088/0954-898x/14/1/303> PMID: [12613551](https://pubmed.ncbi.nlm.nih.gov/12613551/)
22. Nigam S, Pojoga S, Dragoi V. Synergistic coding of visual information in columnar networks. *Neuron*. 2019;104(2):402–411.e4. <https://doi.org/10.1016/j.neuron.2019.07.006> PMID: [31399280](https://pubmed.ncbi.nlm.nih.gov/31399280/)
23. Froudarakis E, Berens P, Ecker AS, Cotton RJ, Sinz FH, Yatsenko D, et al. Population code in mouse V1 facilitates readout of natural scenes through increased sparseness. *Nat Neurosci*. 2014;17(6):851–7. <https://doi.org/10.1038/nn.3707> PMID: [24747577](https://pubmed.ncbi.nlm.nih.gov/24747577/)
24. Runyan CA, Piasini E, Panzeri S, Harvey CD. Distinct timescales of population coding across cortex. *Nature*. 2017;548(7665):92–6. <https://doi.org/10.1038/nature23020> PMID: [28723889](https://pubmed.ncbi.nlm.nih.gov/28723889/)
25. Onken A, Liu JK, Karunasekara PPCR, Delis I, Gollisch T, Panzeri S. Using matrix and tensor factorizations for the single-trial analysis of population spike trains. *PLoS Comput Biol*. 2016;12(11):e1005189. <https://doi.org/10.1371/journal.pcbi.1005189> PMID: [27814363](https://pubmed.ncbi.nlm.nih.gov/27814363/)
26. Kühn NK, Gollisch T. Activity correlations between direction-selective retinal ganglion cells synergistically enhance motion decoding from complex visual scenes. *Neuron*. 2019;101(5):963–976.e7. <https://doi.org/10.1016/j.neuron.2019.01.003> PMID: [30709656](https://pubmed.ncbi.nlm.nih.gov/30709656/)
27. Curreli S, Bonato J, Romanzi S, Panzeri S, Fellin T. Complementary encoding of spatial information in hippocampal astrocytes. *PLoS Biol*. 2022;20(3):e3001530. <https://doi.org/10.1371/journal.pbio.3001530> PMID: [35239646](https://pubmed.ncbi.nlm.nih.gov/35239646/)
28. Sharpee TO, Berkowitz JA. Linking neural responses to behavior with information-preserving population vectors. *Curr Opin Behav Sci*. 2019;29:37–44. <https://doi.org/10.1016/j.cobeha.2019.03.004> PMID: [36590862](https://pubmed.ncbi.nlm.nih.gov/36590862/)
29. Lurilli G, Datta SR. Population coding in an innately relevant olfactory area. *Neuron*. 2017;93(5):1180–1197.e7. <https://doi.org/10.1016/j.neuron.2017.02.010> PMID: [28238549](https://pubmed.ncbi.nlm.nih.gov/28238549/)
30. Kira S, Safaai H, Morcos AS, Panzeri S, Harvey CD. A distributed and efficient population code of mixed selectivity neurons for flexible navigation decisions. *Nat Commun*. 2023;14(1):2121. <https://doi.org/10.1038/s41467-023-37804-2> PMID: [37055431](https://pubmed.ncbi.nlm.nih.gov/37055431/)
31. Wibral M, Vicente R, Lizier JT. *Directed information measures in neuroscience*. Springer; 2014.
32. Vicente R, Wibral M, Lindner M, Pipa G. Transfer entropy—a model-free measure of effective connectivity for the neurosciences. *J Comput Neurosci*. 2011;30(1):45–67. <https://doi.org/10.1007/s10827-010-0262-3> PMID: [20706781](https://pubmed.ncbi.nlm.nih.gov/20706781/)
33. Colenbier N, Van de Steen F, Uddin LQ, Poldrack RA, Calhoun VD, Marinazzo D. Disambiguating the role of blood flow and global signal with partial information decomposition. *Neuroimage*. 2020;213:116699. <https://doi.org/10.1016/j.neuroimage.2020.116699> PMID: [32179104](https://pubmed.ncbi.nlm.nih.gov/32179104/)
34. Besserve M, Lowe SC, Logothetis NK, Schölkopf B, Panzeri S. Shifts of gamma phase across primary visual cortical sites reflect dynamic stimulus-modulated information transfer. *PLoS Biol*. 2015;13(9):e1002257. <https://doi.org/10.1371/journal.pbio.1002257> PMID: [26394205](https://pubmed.ncbi.nlm.nih.gov/26394205/)
35. Luppi AI, Mediano PAM, Rosas FE, Holland N, Fryer TD, O'Brien JT, et al. A synergistic core for human brain evolution and cognition. *Nat Neurosci*. 2022;25(6):771–82. <https://doi.org/10.1038/s41593-022-01070-0> PMID: [35618951](https://pubmed.ncbi.nlm.nih.gov/35618951/)
36. Celotto M, Bim J, Tlaie A, De Feo V, Toso A, Lemke SM, et al. An information-theoretic quantification of the content of communication between brain regions. *Adv Neural Inf Process Syst · NeurIPS*. 2023;36:64213–65.
37. Lemke SM, Celotto M, Maffulli R, Ganguly K, Panzeri S. Information flow between motor cortex and striatum reverses during skill learning. *Curr Biol*. 2024;34(9):1831–1843.e7. <https://doi.org/10.1016/j.cub.2024.03.023> PMID: [38604168](https://pubmed.ncbi.nlm.nih.gov/38604168/)
38. Sherrill SP, Timme NM, Beggs JM, Newman EL. Partial information decomposition reveals that synergistic neural integration is greater downstream of recurrent information flow in organotypic cortical cultures. *PLoS Comput Biol*. 2021;17(7):e1009196. <https://doi.org/10.1371/journal.pcbi.1009196> PMID: [34252081](https://pubmed.ncbi.nlm.nih.gov/34252081/)
39. Reid AT, Headley DB, Mill RD, Sanchez-Romero R, Uddin LQ, Marinazzo D, et al. Advancing functional connectivity research from association to causation. *Nat Neurosci*. 2019;22(11):1751–60. <https://doi.org/10.1038/s41593-019-0510-4> PMID: [31611705](https://pubmed.ncbi.nlm.nih.gov/31611705/)
40. Stramaglia S, Wu G-R, Pellicoro M, Marinazzo D. Expanding the transfer entropy to identify information circuits in complex systems. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2012;86(6 Pt 2):066211. <https://doi.org/10.1103/PhysRevE.86.066211> PMID: [23368028](https://pubmed.ncbi.nlm.nih.gov/23368028/)
41. Wibral M, Pampu N, Priesemann V, Siebenhühner F, Seiwert H, Lindner M, et al. Measuring information-transfer delays. *PLoS One*. 2013;8(2):e55809. <https://doi.org/10.1371/journal.pone.0055809> PMID: [23468850](https://pubmed.ncbi.nlm.nih.gov/23468850/)
42. Pereira TD, Shaevitz JW, Murthy M. Quantifying behavior to understand the brain. *Nat Neurosci*. 2020;23(12):1537–49. <https://doi.org/10.1038/s41593-020-00734-z> PMID: [33169033](https://pubmed.ncbi.nlm.nih.gov/33169033/)
43. Panzeri S, Harvey CD, Piasini E, Latham PE, Fellin T. Cracking the neural code for sensory perception by combining statistics, intervention, and behavior. *Neuron*. 2017;93(3):491–507. <https://doi.org/10.1016/j.neuron.2016.12.036> PMID: [28182905](https://pubmed.ncbi.nlm.nih.gov/28182905/)
44. Pica G, Piasini E, Safaai H, Runyan C, Harvey C, Diamond M. Quantifying how much sensory information in a neural code is relevant for behavior. *Adv Neural Inf Process Syst*. 2017;30:3686–96.
45. Ince RAA, Petersen RS, Swan DC, Panzeri S. Python for information theoretic analysis of neural data. *Front Neuroinform*. 2009;3:4. <https://doi.org/10.3389/neuro.11.004.2009> PMID: [19242557](https://pubmed.ncbi.nlm.nih.gov/19242557/)

46. Magri C, Whittingstall K, Singh V, Logothetis NK, Panzeri S. A toolbox for the fast information analysis of multiple-site LFP, EEG and spike train recordings. *BMC Neurosci.* 2009;10:81. <https://doi.org/10.1186/1471-2202-10-81> PMID: [19607698](https://pubmed.ncbi.nlm.nih.gov/19607698/)
47. Lizier JT. JIDT: an information-theoretic toolkit for studying the dynamics of complex systems. *Front Robot AI.* 2014;1. <https://doi.org/10.3389/frobt.2014.00011>
48. Timme NM, Lapish C. A tutorial for information theory in neuroscience. *eNeuro.* 2018;5(3). <https://doi.org/10.1523/eneuro.0052-18.2018>
49. Combrisson E, Allegra M, Basanisi R, Ince RAA, Giordano BL, Bastin J, et al. Group-level inference of information-based measures for the analyses of cognitive brain networks from neurophysiological data. *Neuroimage.* 2022;258:119347. <https://doi.org/10.1016/j.neuroimage.2022.119347> PMID: [35660460](https://pubmed.ncbi.nlm.nih.gov/35660460/)
50. Ince RAA, Giordano BL, Kayser C, Rousselet GA, Gross J, Schyns PG. A statistical framework for neuroimaging data analysis based on mutual information estimated via a gaussian copula. *Hum Brain Mapp.* 2017;38(3):1541–73. <https://doi.org/10.1002/hbm.23471> PMID: [27860095](https://pubmed.ncbi.nlm.nih.gov/27860095/)
51. Climer JR, Dombek DA. Information theoretic approaches to deciphering the neural code with functional fluorescence imaging. *eNeuro.* 2021;8(5):ENEURO.0266-21.2021. <https://doi.org/10.1523/ENEURO.0266-21.2021> PMID: [34433574](https://pubmed.ncbi.nlm.nih.gov/34433574/)
52. Makkeh A, Theis DO, Vicente R. BROJA-2PID: a robust estimator for bivariate partial information decomposition. *Entropy (Basel).* 2018;20(4):271. <https://doi.org/10.3390/e20040271> PMID: [33265362](https://pubmed.ncbi.nlm.nih.gov/33265362/)
53. Moore DG, Valentini G, Walker SI, Levin M, editors. Inform: A toolkit for information-theoretic analysis of complex systems. 2017 IEEE Symposium Series on Computational Intelligence (SSCI); 2017.
54. Montalto A, Faes L, Marinazzo D. MuTE: a MATLAB toolbox to compare established and novel estimators of the multivariate transfer entropy. *PLoS One.* 2014;9(10):e109462. <https://doi.org/10.1371/journal.pone.0109462> PMID: [25314003](https://pubmed.ncbi.nlm.nih.gov/25314003/)
55. Szabó Z. Information theoretical estimators toolbox. *J. Mach. Learn. Res.* 2014;15(1):283–7.
56. Lindner M, Vicente R, Priesemann V, Wibral M. TRENTOOL: a Matlab open source toolbox to analyse information flow in time series data with transfer entropy. *BMC Neurosci.* 2011;12:119. <https://doi.org/10.1186/1471-2202-12-119> PMID: [22098775](https://pubmed.ncbi.nlm.nih.gov/22098775/)
57. Ito S, Hansen ME, Heiland R, Lumsdaine A, Litke AM, Beggs JM. Extending transfer entropy improves identification of effective connectivity in a spiking cortical network model. *PLoS One.* 2011;6(11):e27431. <https://doi.org/10.1371/journal.pone.0027431> PMID: [22102894](https://pubmed.ncbi.nlm.nih.gov/22102894/)
58. Goldberg DH, Victor JD, Gardner EP, Gardner D. Spike train analysis toolkit: enabling wider application of information-theoretic techniques to neurophysiology. *Neuroinformatics.* 2009;7(3):165–78. <https://doi.org/10.1007/s12021-009-9049-y> PMID: [19475519](https://pubmed.ncbi.nlm.nih.gov/19475519/)
59. Neri M, Vinchhi D, Ferreyra C, Robiglio T, Ates O, Ontivero-Ortega M, et al. HOI: A Python toolbox for high-performance estimation of Higher-Order Interactions from multivariate data. *JOSS.* 2024;9(103):7360. <https://doi.org/10.21105/joss.07360>
60. G. James R, J. Ellison C, P. Crutchfield J. dit: a Python package for discrete information theory. *JOSS.* 2018;3(25):738. <https://doi.org/10.21105/joss.00738>
61. Candadai M, Izquierdo E. infotheory: A C++/Python package for multivariate information theoretic analysis. *JOSS.* 2020;5(47):1609. <https://doi.org/10.21105/joss.01609>
62. Emanuel AJ, Lehnert BP, Panzeri S, Harvey CD, Ginty DD. Cortical responses to touch reflect subcortical integration of LTMR signals. *Nature.* 2021;600(7890):680–5. <https://doi.org/10.1038/s41586-021-04094-x> PMID: [34789880](https://pubmed.ncbi.nlm.nih.gov/34789880/)
63. Schneidman E, Berry MJ 2nd, Segev R, Bialek W. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature.* 2006;440(7087):1007–12. <https://doi.org/10.1038/nature04701> PMID: [16625187](https://pubmed.ncbi.nlm.nih.gov/16625187/)
64. Milosavljevic N, Storchi R, Eleftheriou CG, Colins A, Petersen RS, Lucas RJ. Photoreceptive retinal ganglion cells control the information rate of the optic nerve. *Proc Natl Acad Sci U S A.* 2018;115(50):E11817–26. <https://doi.org/10.1073/pnas.1810701115> PMID: [30487225](https://pubmed.ncbi.nlm.nih.gov/30487225/)
65. Ostwald D, Porcaro C, Bagshaw AP. An information theoretic approach to EEG-fMRI integration of visually evoked responses. *Neuroimage.* 2010;49(1):498–516. <https://doi.org/10.1016/j.neuroimage.2009.07.038> PMID: [19632339](https://pubmed.ncbi.nlm.nih.gov/19632339/)
66. Pessoa L, Padmala S. Decoding near-threshold perception of fear from distributed single-trial brain activation. *Cereb Cortex.* 2007;17(3):691–701. <https://doi.org/10.1093/cercor/bhk020> PMID: [16627856](https://pubmed.ncbi.nlm.nih.gov/16627856/)
67. Schultz SR, Kitamura K, Post-Uiterweer A, Krupic J, Häusser M. Spatial pattern coding of sensory information by climbing fiber-evoked calcium signals in networks of neighboring cerebellar Purkinje cells. *J Neurosci.* 2009;29(25):8005–15. <https://doi.org/10.1523/JNEUROSCI.4919-08.2009> PMID: [19553440](https://pubmed.ncbi.nlm.nih.gov/19553440/)
68. Nirenberg S, Carcieri SM, Jacobs AL, Latham PE. Retinal ganglion cells act largely as independent encoders. *Nature.* 2001;411(6838):698–701. <https://doi.org/10.1038/35079612> PMID: [11395773](https://pubmed.ncbi.nlm.nih.gov/11395773/)
69. Oizumi M, Ishii T, Ishibashi K, Hosoya T, Okada M. Mismatched decoding in the brain. *J Neurosci.* 2010;30(13):4815–26. <https://doi.org/10.1523/JNEUROSCI.4360-09.2010> PMID: [20357132](https://pubmed.ncbi.nlm.nih.gov/20357132/)
70. Williams P, Beer R. Nonnegative decomposition of multivariate information. *arXiv preprint.* 2010;arXiv:10042515.
71. Wibral M, Priesemann V, Kay JW, Lizier JT, Phillips WA. Partial information decomposition as a unified approach to the specification of neural goal functions. *Brain Cogn.* 2017;112:25–38. <https://doi.org/10.1016/j.bandc.2015.09.004> PMID: [26475739](https://pubmed.ncbi.nlm.nih.gov/26475739/)
72. Bertschinger N, Rauh J, Olbrich E, Jost J, Ay N. Quantifying unique information. *Entropy.* 2014;16(4):2161–83. <https://doi.org/10.3390/e16042161>
73. Barrett AB. Exploration of synergistic and redundant information sharing in static and dynamical Gaussian systems. *Phys Rev E Stat Nonlin Soft Matter Phys.* 2015;91(5):052802. <https://doi.org/10.1103/PhysRevE.91.052802> PMID: [26066207](https://pubmed.ncbi.nlm.nih.gov/26066207/)

74. Schreiber T. Measuring information transfer. *Phys Rev Lett*. 2000;85(2):461–4. <https://doi.org/10.1103/PhysRevLett.85.461> PMID: [10991308](https://pubmed.ncbi.nlm.nih.gov/10991308/)
75. Strong SP, Koberle R, de Ruyter van Steveninck RR, Bialek W. Entropy and information in neural spike trains. *Phys Rev Lett*. 1998;80(1):197–200. <https://doi.org/10.1103/physrevlett.80.197>
76. Gross J, Hoogenboom N, Thut G, Schyns P, Panzeri S, Belin P, et al. Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biol*. 2013;11(12):e1001752. <https://doi.org/10.1371/journal.pbio.1001752> PMID: [24391472](https://pubmed.ncbi.nlm.nih.gov/24391472/)
77. Orłowska-Feuer P, Ebrahimi AS, Zippo AG, Petersen RS, Lucas RJ, Storchi R. Look-up and look-down neurons in the mouse visual thalamus during freely moving exploration. *Curr Biol*. 2022;32(18):3987–3999.e4. <https://doi.org/10.1016/j.cub.2022.07.049> PMID: [35973431](https://pubmed.ncbi.nlm.nih.gov/35973431/)
78. Scott DW. Scott's rule. *WIREs Computational Stats*. 2010;2(4):497–502. <https://doi.org/10.1002/wics.103>
79. Freedman D, Diaconis P. On the histogram as a density estimator:L 2 theory. *Z Wahrscheinlichkeitstheorie verw Gebiete*. 1981;57(4):453–76. <https://doi.org/10.1007/bf01025868>
80. Panzeri S, Treves A. Analytical estimates of limited sampling biases in different information measures. *Network*. 1996;7(1):87–107. <https://doi.org/10.1080/0954898X.1996.11978656> PMID: [29480146](https://pubmed.ncbi.nlm.nih.gov/29480146/)
81. Paninski L. Estimation of entropy and mutual information. *Neural Comput*. 2003;15(6):1191–253. <https://doi.org/10.1162/089976603321780272>
82. Panzeri S, Senatore R, Montemurro MA, Petersen RS. Correcting for the sampling bias problem in spike train information measures. *J Neurophysiol*. 2007;98(3):1064–72. <https://doi.org/10.1152/jn.00559.2007> PMID: [17615128](https://pubmed.ncbi.nlm.nih.gov/17615128/)
83. Optican LM, Gawne TJ, Richmond BJ, Joseph PJ. Unbiased measures of transmitted information and channel capacity from multivariate neuronal data. *Biol Cybern*. 1991;65(5):305–10. <https://doi.org/10.1007/BF00216963> PMID: [1742368](https://pubmed.ncbi.nlm.nih.gov/1742368/)
84. Montemurro MA, Senatore R, Panzeri S. Tight data-robust bounds to mutual information combining shuffling and model selection techniques. *Neural Comput*. 2007;19(11):2913–57. <https://doi.org/10.1162/neco.2007.19.11.2913> PMID: [17883346](https://pubmed.ncbi.nlm.nih.gov/17883346/)
85. Ince RAA, Senatore R, Arabzadeh E, Montani F, Diamond ME, Panzeri S. Information-theoretic methods for studying population codes. *Neural Netw*. 2010;23(6):713–27. <https://doi.org/10.1016/j.neunet.2010.05.008> PMID: [20542408](https://pubmed.ncbi.nlm.nih.gov/20542408/)
86. Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2004;69(6 Pt 2):066138. <https://doi.org/10.1103/PhysRevE.69.066138> PMID: [15244698](https://pubmed.ncbi.nlm.nih.gov/15244698/)
87. Holmes CM, Nemenman I. Estimation of mutual information for real-valued data with error bars and controlled bias. *Phys Rev E*. 2019;100(2–1):022404. <https://doi.org/10.1103/PhysRevE.100.022404> PMID: [31574710](https://pubmed.ncbi.nlm.nih.gov/31574710/)
88. Nemenman I, Bialek W, de Ruyter van Steveninck R. Entropy and information in neural spike trains: progress on the sampling problem. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2004;69(5 Pt 2):056111. <https://doi.org/10.1103/PhysRevE.69.056111> PMID: [15244887](https://pubmed.ncbi.nlm.nih.gov/15244887/)
89. Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature*. 1999;401(6755):788–91. <https://doi.org/10.1038/44565> PMID: [10548103](https://pubmed.ncbi.nlm.nih.gov/10548103/)
90. Chang C-C, Lin C-J. Libsvm: a library for support vector machines. *ACM Trans Intell Syst Technol*. 2011;2(3):1–27. <https://doi.org/10.1145/1961189.1961199>
91. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Soft*. 2010;33(1). <https://doi.org/10.18637/jss.v033.i01>
92. Graf ABA, Kohn A, Jazayeri M, Movshon JA. Decoding the activity of neuronal populations in macaque primary visual cortex. *Nat Neurosci*. 2011;14(2):239–45. <https://doi.org/10.1038/nn.2733> PMID: [21217762](https://pubmed.ncbi.nlm.nih.gov/21217762/)
93. Averbeck BB, Latham PE, Pouget A. Neural correlations, population coding and computation. *Nat Rev Neurosci*. 2006;7(5):358–66. <https://doi.org/10.1038/nrn1888> PMID: [16760916](https://pubmed.ncbi.nlm.nih.gov/16760916/)
94. Reich DS, Mechler F, Victor JD. Independent and redundant information in nearby cortical neurons. *Science*. 2001;294(5551):2566–8. <https://doi.org/10.1126/science.1065839> PMID: [11752580](https://pubmed.ncbi.nlm.nih.gov/11752580/)
95. Chechik G, Globerson A, Anderson M, Young E, Nelken I, Tishby N. Group redundancy measures reveal redundancy reduction in the auditory pathway. *Adv Neural Inf Process Syst*. 2001;14:1.
96. Beer RD, Williams PL. Information processing and dynamics in minimally cognitive agents. *Cogn Sci*. 2015;39(1):1–38. <https://doi.org/10.1111/cogs.12142> PMID: [25039535](https://pubmed.ncbi.nlm.nih.gov/25039535/)
97. Francis NA, Mukherjee S, Koçillari L, Panzeri S, Babadi B, Kanold PO. Sequential transmission of task-relevant information in cortical neuronal networks. *Cell Rep*. 2022;39(9):110878. <https://doi.org/10.1016/j.celrep.2022.110878> PMID: [35649366](https://pubmed.ncbi.nlm.nih.gov/35649366/)
98. Valente M, Pica G, Bondanelli G, Moroni M, Runyan CA, Morcos AS, et al. Correlations enhance the behavioral readout of neural population activity in association cortex. *Nat Neurosci*. 2021;24(7):975–86. <https://doi.org/10.1038/s41593-021-00845-1> PMID: [33986549](https://pubmed.ncbi.nlm.nih.gov/33986549/)
99. Rousselet GA, Ince RAA, van Rijsbergen NJ, Schyns PG. Eye coding mechanisms in early human face event-related potentials. *J Vis*. 2014;14(13):7. <https://doi.org/10.1167/14.13.7> PMID: [25385898](https://pubmed.ncbi.nlm.nih.gov/25385898/)
100. Ince RAA, Jaworska K, Gross J, Panzeri S, van Rijsbergen NJ, Rousselet GA, et al. The Deceptively Simple N170 reflects network information processing mechanisms involving visual feature coding and transfer across hemispheres. *Cereb Cortex*. 2016;26(11):4123–35. <https://doi.org/10.1093/cercor/bhw196> PMID: [27550865](https://pubmed.ncbi.nlm.nih.gov/27550865/)
101. Oostenveld R, Fries P, Maris E, Schoffelen J-M. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci*. 2011;2011:156869. <https://doi.org/10.1155/2011/156869> PMID: [21253357](https://pubmed.ncbi.nlm.nih.gov/21253357/)

102. Rubinov M, Sporns O. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*. 2010;52(3):1059–69. <https://doi.org/10.1016/j.neuroimage.2009.10.003> PMID: [19819337](https://pubmed.ncbi.nlm.nih.gov/19819337/)
103. Safaai H, Onken A, Harvey CD, Panzeri S. Information estimation using nonparametric copulas. *Phys Rev E*. 2018;98(5):053302. <https://doi.org/10.1103/PhysRevE.98.053302> PMID: [30984901](https://pubmed.ncbi.nlm.nih.gov/30984901/)
104. Victor JD. Binless strategies for estimation of information from neural data. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2002;66(5 Pt 1):051903. <https://doi.org/10.1103/PhysRevE.66.051903> PMID: [12513519](https://pubmed.ncbi.nlm.nih.gov/12513519/)
105. Luppi AI, Rosas FE, Mediano PAM, Menon DK, Stamatakis EA. Information decomposition and the informational architecture of the brain. *Trends Cogn Sci*. 2024;28(4):352–68. <https://doi.org/10.1016/j.tics.2023.11.005> PMID: [38199949](https://pubmed.ncbi.nlm.nih.gov/38199949/)