Multilevel irreversibility reveals higher-order organisation of non-equilibrium interactions in human brain dynamics

Ramón Nartallo-Kaluarachchi[1,2,*], Leonardo Bonetti[3,2], Gemma Fernández-Rubio[3], Peter Vuust[3], Gustavo Deco[4,5,6], Morten L. Kringelbach[2,3,7], Renaud Lambiotte[1,8], and Alain Goriely[1,†]

1. Mathematical Institute, University of Oxford, Oxford, OX2 6GG, United Kingdom

2. Centre for Eudaimonia and Human Flourishing, University of Oxford, Oxford, OX3 9BX, United Kingdom

3. Center for Music in the Brain, Aarhus University, & The Royal Academy of Music, Aarhus/Aalborg, Denmark

4. Centre for Brain and Cognition, Computational Neuroscience Group, Universitat Pompeu Fabra, 08018 Barcelona, Spain

5. Department of Information and Communication Technologies, Universitat Pompeu Fabra, 08018 Barcelona, Spain

6. Institució Catalana de la Recerca i Estudis Avancats (ICREA), 08010 Barcelona, Spain

7. Department of Psychiatry, University of Oxford, Oxford, OX3 7JX United Kingdom

8. Turing Institute, London, NW1 2DB United Kingdom

**nartallokalu@maths.ox.ac.uk, †goriely@maths.ox.ac.uk*

(Dated: May 2, 2024)

Information processing in the human brain can be modelled as a complex dynamical system operating out of equilibrium with multiple regions interacting nonlinearly. Yet, despite extensive study of non-equilibrium at the global level of the brain, quantifying the irreversibility of interactions among brain regions at multiple levels remains an unresolved challenge. Here, we present the Directed Multiplex Visibility Graph Irreversibility framework, a method for analysing neural recordings using network analysis of timeseries. Our approach constructs directed multi-layer graphs from multivariate time-series where information about irreversibility can be decoded from the marginal degree distributions across the layers, which each represents a variable. This framework is able to quantify the irreversibility of every interaction in the complex system. Applying the method to magnetoencephalography recordings during a long-term memory recognition task, we quantify the multivariate irreversibility of interactions between brain regions and identify the combinations of regions which showed higher levels of non-equilibrium in their interactions. For individual regions, we find higher irreversibility in cognitive versus sensorial brain regions whilst for pairs, strong relationships are uncovered between cognitive and sensorial pairs in the same hemisphere. For triplets and quadruplets, the most non-equilibrium interactions are between cognitive-sensorial pairs alongside medial regions. Finally, for quintuplets, our analysis finds higher irreversibility when the prefrontal cortex is included in the interaction. Combining these results, we show that multilevel irreversibility offers unique insights into the higher-order organisation of neural dynamics and presents a new perspective on the analysis of brain network dynamics.

I. INTRODUCTION

The human brain produces complex spatiotemporal neural dynamics across multiple time and length scales. Abstracting the brain as a large-scale network of discrete interacting regions has proved fruitful in the analysis and modelling of neural dynamics [1]. Moreover, this abstraction lends neuroscientists the language and tools of statistical physics in the hope of uncovering the central mechanisms driving brain function and their links to observed neural dynamics [2, 3]. For instance, recent data captured by functional imaging showed large scale violations of detailed balance in human brain dynamics, suggesting that the brain is operating far from equilibrium [4]. This fundamental observation has prompted the development of a range of techniques to provide a measure for the degree of non-equilibrium in neuroimaging time-series recorded in different conditions [5–10]. These measures have shown that the degree of non-equilibrium is elevated during cognitive tasks [4–7] whilst reduced in both impairments of consciousness [11], sleep [10] and Alzheimer's disease [12] indicating that non-equilibrium may be a key signature of healthy consciousness and cognition in the brain [13]. Despite this, current methods are restricted to aggregate measures of non-equilibrium. We present a novel approach to non-equilibrium brain dynamics that is able to measure the irreversibility of individual, higher-order interactions to gain valuable insight into the organisation of regional interactions in neural dynamics.

The second law of thermodynamics asserts that, in the absence of entropy sinks, the average entropy of a system increases as time flows forwards [14, 15]. More specifically, a system at a steady-state dissipating heat to

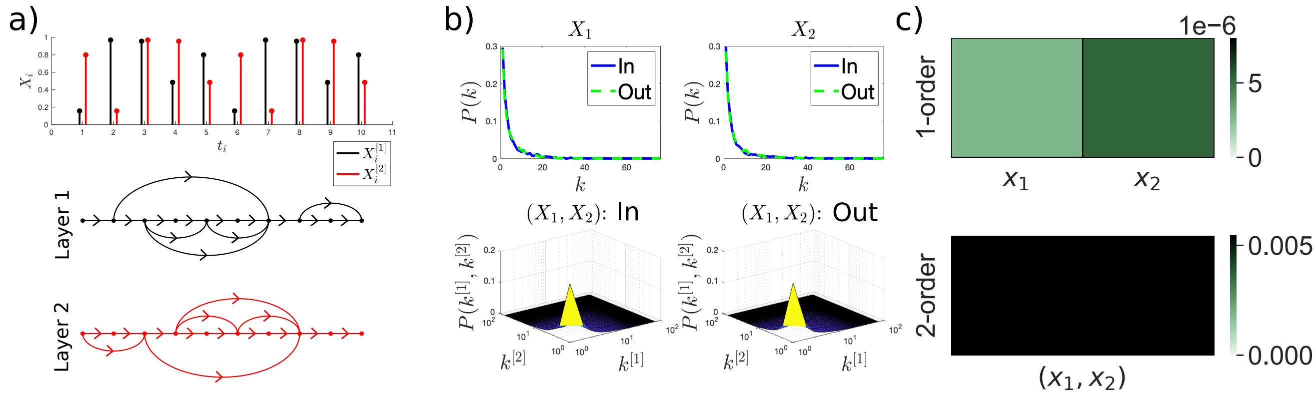

FIG. 1. The DiMViGI workflow. (a) A 2-layer directed multiplex visibility graph from a random time-series. (b) In- and out- degree distributions for each tuple at each level i.e. (x_1), (x_2) and (x_1, x_2). (c) Jensen-Shannon divergence of in- and out- degree distributions at each level.

its environment causes an increase in entropy [16, 17]. This results in the system breaking the detailed balance condition and an asymmetry in the probability of transitioning between system states [18]. This, in turn, yields macroscopically irreversible trajectories from reversible microscopic forces inducing, what Eddington denoted, 'the arrow of time' (AoT) [19]. Results in modern non-equilibrium thermodynamics have shown that the entropy production rate of a system that is out of equilibrium is equal to the information-theoretic evidence for the AoT quantified by the divergence

$$\sigma = \sum_{\Gamma} P(\Gamma) \log \frac{P(\Gamma)}{P(\Gamma')}, \quad (1)$$

where Γ is a trajectory, Γ' is its time-reversal and $P(\Gamma)$ is the 'path-probability', the probability of observing that specific trajectory [20, 21]. This divergence provides a distance from equilibrium [22–24]. Two complimentary interpretations of the AoT in the brain have been given. First, the hierarchical organisation of positions in state-space, that results from asymmetrical transition probabilities, has been linked to the dynamic hierarchical organisation of brain regions [7, 25, 26]. Second, the AoT has been interpreted as inducing a 'causal flow' in the system where some regions emerge as information 'sources' and others as 'sinks' with these relationships identifiable from irreversibility analysis [7, 8]. These studies for quantifying non-equilibrium in the brain approximate the global evidence for the AoT in time-series using techniques such as estimating transitions between coarse-grained states [4], with time-shifted correlations [5], machine learning [6] or with model-based approaches [7–10]. However, the AoT and the corresponding production of entropy is a macroscopic property of the system, emerging from interactions between the microscopic variables at multiple scales. Recent theoretical research has shown that the AoT can be decomposed into unique contributions arising at each scale within the system [23, 27]. Motivated by this decomposition, we present the Directed Multiplex Visibility Graph Irreversibility (DiMViGI) framework, as shown in Fig. 1, for analysing the irreversibility of multivariate signals at multiple levels using network analysis of time-series, in particular the visibility graph [28, 29]. Using the DiMViGI framework, we investigate the irreversibility of human brain signals, captured by magnetoencephalography (MEG), during a long-term recognition task of musical sequences that utilised long-term memory [30–35]. Our analysis covers all possible levels in the system and is able to capture the higher-order organisation of brain regional interactions yielding interpretable and novel insights into the neural dynamics underpinning long-term memory.

II. DECOMPOSING THE ARROW OF TIME IN MULTIVARIATE SYSTEMS

We first motivate our framework by decomposing the information theoretic evidence for the AoT into contributions from different levels. Consider a multivariate system of N interacting units assumed to be in a thermodynamic steady-state. Then, we define the global rate of entropy production, σ, for the system. Depending on whether the system is discrete or continuous and Markovian or non-Markovian, the path probability, and hence the entropy production rate, has a different formulation (see SI for explicit formulation). Following [23], we decompose the global rate of entropy production as

$$\sigma = \sum_{k=1}^{N} \sigma^{(k)}. \quad (2)$$

where $\sigma^{(1)}$ is produced by the units evolving individually and $\sigma^{(k)}$ is the contribution of interactions between tuples of size k. We further decompose $\sigma^{(k)}$ into the sum of unique contributions to the entropy production rate $\sigma^{(x_1,\ldots,x_k)}$ produced by the interactions between the particular k-tuple of variables $(x_1, \ldots, x_k)$:

$$\sigma^{(k)} = \sum_{(x_1,\ldots,x_k)\in\Lambda_k} \sigma^{(x_1,\ldots,x_k)}, \tag{3}$$

where Λ_k denotes the set of all tuples of variables of size k. Hence, we have

$$\sigma = \sum_{k=1}^{N} \sum_{(x_1,\ldots,x_k)\in\Lambda_k} \sigma^{(x_1,\ldots,x_k)}. \tag{4}$$

Motivated by this decomposition, we note that individual interactions can have differential contributions to the overall irreversibility of the system. Our framework aims to compute the irreversibility of individual k-tuples of variables in a multivariate time-series in order to compare interactions at each level, defined by k. We identify tuples of variables whose multivariate trajectory is highly irreversible indicating a strongly non-equilibrium interaction between the variables in this tuple. We note that the decompositions in equations (2-3) are only well-defined for multi-partite dynamics, meaning only one variable changes in a single time-step, an assumption that does not hold in continuous data [27]. However, we do not directly measure the unique contribution $\sigma^{(x_1,\ldots,x_k)}$ to the global entropy production rate, but instead approximate the irreversibility of a given tuple,

$$\varsigma^{(x_1,\ldots,x_k)} = \sum_{\Gamma^{(x_1,\ldots,x_k)}} P(\Gamma^{(x_1,\ldots,x_k)}) \log \frac{P(\Gamma^{(x_1,\ldots,x_k)})}{P(\Gamma'^{(x_1,\ldots,x_k)})}, \tag{5}$$

where $\Gamma^{(x_1,\ldots,x_k)}$ is the projection of a trajectory into the portion of state-space defined by the variables $(x_1, \ldots, x_k)$ (see SI for details).

III. MEASURING IRREVERSIBILITY WITH THE MULTIPLEX VISIBILITY GRAPH

We build on the growing paradigm of network analysis of time-series that has gained traction in the analysis of neural signals [36]. These methods are characterised by mapping a time-series into a corresponding network. For instance, the visibility algorithm maps a univariate time-series into a so-called 'visibility graph' (VG) [28]. Explicitly, given a time-series $\{X_i\}_{i\in I}$ with time indices $\{t_i\}_{i\in I}$, where $X_i \in \mathbb{R}$ and I is the index set, the VG has one node for each $i \in I$. Nodes $i, j \in I$ are connected by an edge if the corresponding data-points (t_i, X_i) and (t_j, X_j) are 'mutually visible' i.e. that they satisfy that, for any intermediate data-point (t_k, X_k) with $t_i < t_k < t_j$,

$$X_k < X_j + (X_i - X_j)\frac{t_j - t_k}{t_j - t_i}. \tag{6}$$

In geometric terms, this condition is met if (t_i, X_i) is visible from (t_j, X_j). That is, the line connecting (t_i, X_i) and (t_j, X_j) does not cross any intermediate data-points as shown in Panel b) of Fig. 2. Trivially, each node is connected to its neighbours whilst large positive fluctuations become hubs with many connections due to their greater visibility. This construction can be naturally extended

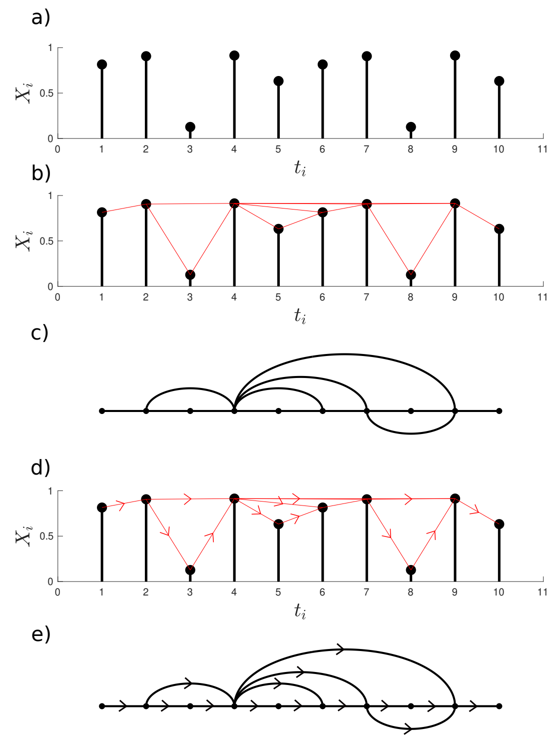

FIG. 2. An example of a visibility and directed visibility graph on a random time-series. (a) A random equi-spaced time-series. (b) The red lines connected data points that mutually visible. (c) The visibility graph associated with the random series. (d) A time-series showing visibility directed forward in time. (e) The directed visibility graph corresponding to the above series.

to multivariate time-series by considering the 'multiplex visibility graph' (MVG) [37]. Given a multivariate time-series with N variables, the MVG is a multi-layer graph, a so-called 'multiplex', with N independent layers with the same node base. Applying the visibility algorithm to each variable in turn yields a series of VGs which each define one layer of the MVG.

We can further generalise the VG to measure irreversibility in univariate time-series by extending the undirected VG to a time-directed counterpart (DVG) [29, 38]. To do so, we

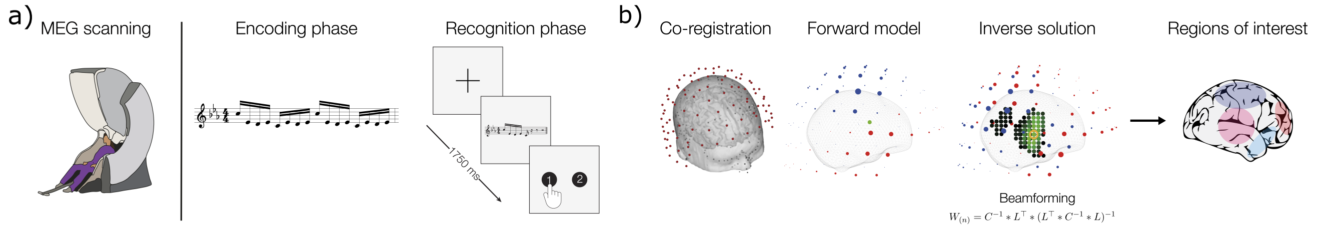

FIG. 3. The experimental paradigm for collection and processing of MEG data. (a) The brain activity in 51 participants was collected using magnetoencephalography (MEG) while they performed a long-term auditory recognition task. Participants memorised a 5 tone musical sequence. They were then played 5 further sequences of tones that were either the original sequence or a modified version. They then were requested to state whether the sequence belonged to the original music or was a varied version of the original sequences. In this analysis we only consider the experimental condition where participants were played the original memorised sequence. (b) The MEG data was co-registered with the individual anatomical MRI data, and source reconstructed using a beamforming algorithm. This procedure returned one time-series for each of the 3559 reconstructed brain sources. Six main functional brain regions (ROIs) were derived. The neural activity for each ROI was extracted yielding a multivariate time-series. For further details on the experimental set-up see Materials and Methods and SI. For a comparison between experimental conditions see Bonetti et al [30].

simply direct the edges 'forward in time'. For example, an edge connecting time-points $t_i < t_j$ is now directed $i \to j$ (see Panels d-e) of Fig. 2) and decompose the degree d of a node into the sum of the in-going and out-going degree,

$$d = d_{\text{in}} + d_{\text{out}}. \tag{7}$$

A univariate stationary process, $X(t)$, is time-reversible if the trajectory $\{X(t_1), ..., X(t_T)\}$ is as probable as $\{X(t_T), ..., X(t_1)\}$ [39]. Therefore, in the case of a reversible process, the in- and out-going degree distributions of the associated DVG should converge [29, 38]. It follows that the level of irreversibility can be captured by measuring the divergence between the in- and out-going degree distributions. We extend this method to the case of multivariate time-series. We direct the edges of the MVG such that they go forward in time yielding a directed MVG (DMVG). Since this is a multiplex graph, we can calculate the multivariate joint, over all layers, in- and out-going degree distributions, and all associated marginals.

Explicitly, we consider a multivariate time-series with N variables and T time points, given by $\{\mathbf{X}(t_1), ..., \mathbf{X}(t_T)\}$, where $\mathbf{X}(t_i) = (x_1(t_i), ..., x_N(t_i)) \in \mathbb{R}^N$ and build its associated DMVG. For a given k-tuple of variables, $(n_1, ..., n_k)$, we calculate the multivariate marginal in-going and out-going degree distributions:

$$P_{\text{in}}^{(n_1,...,n_k)}(d_1, ..., d_k), \quad P_{\text{out}}^{(n_1,...,n_k)}(d_1, ..., d_k), \tag{8}$$

where $P^{(n_1,...,n_k)}(d_1, ..., d_k)$ is the probability of a node having degree d_i in layer n_i for all i simultaneously. We then compute the divergence between these particular in- and out-going marginal distributions using Jensen-Shannon divergence (JSD) (see Materials and Methods). Due to the multidimensional distributions, we are quantifying irreversibility in the multivariate state-space. Repeating this for all possible k-tuples in the system, we quantify the relative irreversibility that is being produced by each interaction at a given level. We can repeat this process for all values of k, thus measuring irreversibility at all levels.

In summary, the DiMViGI framework, shown in Fig. 1, begins with a multivariate time-series of neural activity. The series is mapped into the associated DMVG using the visibility algorithm. We calculate the joint in and out-degree distributions and all the possible marginal in- and out- degree distributions. We measure the JSD between the pairs of in- and out-marginals for each tuple in the system to quantify the irreversibility of that interaction. At each level k, we can then compare the relative irreversibility of each k-order interaction to identify the dominant irreversible interactions.

IV. ANALYSIS OF MEG DURING LONG-TERM RECOGNITION

We consider MEG recordings from 51 participants with 15 trials per participant source-localised into six regions of interest (ROIs) collected according to the experimental paradigm presented in Fig. 3, described in Materials and Methods, SI and in Ref. [30]. The ROIs include the auditory cortices in the left and right hemispheres (ACL, ACR); the hippocampal and inferior temporal cortices in the the left and right hemispheres (HITL, HITR) and two medial regions, the bilateral medial cingulate gyrus (MC) and the bilateral ventro-medial prefrontal cortex (VMPFC). Panel a) of Fig. 4 shows a schematic representation of the regions. The participants performed an auditory recognition task during the MEG recordings (Panel a), Fig 3). First, they memorised a short musical piece. Next, they were presented musical sequences and were requested

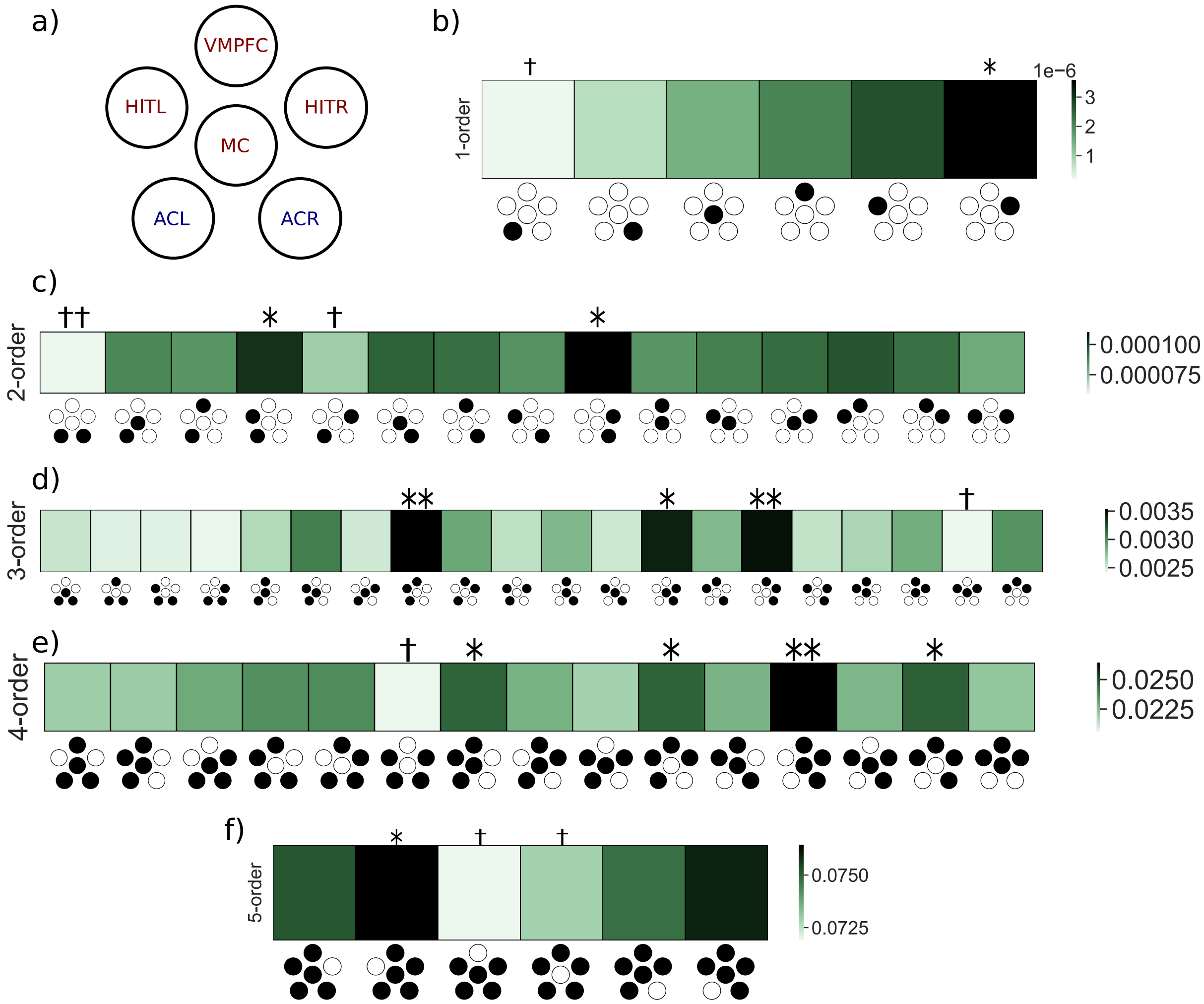

FIG. 4. DiMViGI analysis of 6-ROI MEG recordings during a long-term memory task. The number of (*)/(†) represents the number of standard deviations above/below the mean for a particular tuple at that level. (a) Schematic diagram showing the organisation of the ROIs in the MEG recordings. The ROIs are ACL/R: auditory cortex left/right; MC: medial cingulate gyrus; VMPFC: ventro-medial prefrontal cortex; HITL/R: hippocampal inferior temporal cortex left/right. Cognitive regions are in red and sensory regions in blue. (b) 1-order irreversibility at cohort-level. At this level, we consider irreversibility of each signal in isolation. The hippocampal regions are the most irreversible whilst the sensory regions are the most reversible. (c) 2-order irreversibility at cohort-level. The pairs that show the most irreversibility are those that include a sensory and hippocampal pair in the same hemisphere (ACL/R, HITL/R). The most reversible pair is (ACL, ACR) which is made up of two sensory regions. (d) 3-order irreversibility at cohort-level. The triplets that are most irreversible are those that include an intra-hemispheric sensory and hippocampal pair as well as the prefrontal cortex (ACL/R, HITL/R, VMPFC). The most reversible contains both hippocampal regions and the medial cingulate gyrus, (HITL, HITR, MC). (e) 4-order irreversibility at cohort-level. The quadruplets that are most irreversible are those that include a hippocampal and sensory pair and both medial regions (ACL/R, HITL/R, MC, VMPFC) and those that include both hippocampal regions, a sensory region and the VMPFC. The most reversible is the quadruplet that contains no medial regions. (f) 5-order irreversibility at cohort-level. The most reversible quintuplets are those that omit a medial region, in particular the quintuplet that omits the VMPFC.

to state whether the sequence belonged to the original music or was a varied version of the original sequences. Since differences between experimental conditions have been described in detail by Bonetti et al [30] and are beyond the scope of this work, here, we consider only one experimental condition, where participants recognised the original, previously memorised sequences.

For each participant and trial, we construct the DMVG. Next we estimate every marginal in- and out- degree distribution using each DMVG as a sample and calculate the JSD. We denote the JSD between k-dimensional degree distributions as the k-order irreversibility. Alternatively, for each participant in isolation, the degree distributions can be calculated using only their associated trials to get an estimate of the $k-$order irreversibility for each participant and each tuple (see SI). However, due to the higher number of samples, the cohort-level analysis is more robust and hence is our focus in this report. The results of the DiMViGI analysis are presented in Figure 4. We note that the darker colours represent tuples with greater irreversibility whilst the lighter colours reflect more reversible interactions. The icon along the x-axis indicates which tuple is being considered, with reference to the schematic in Panel a) of Fig. 4, with the included regions coloured in black. Furthermore, we highlight statistically significant tuples at each level. The number of $(*)/(\dagger)$ indicates the number of standard deviations above/below the $k-$level mean.

We begin our analysis at 1-order. Whilst individual (microscopic) variables are often reversible in a non-equilibrium complex system, the ROIs considered here reflect a very coarse parcellation of the brain. At this level, we are considering each ROI, which is composed of many truly microscopic variables, in isolation and note that each one shows significant irreversibility. It is clear from Panel b) of Fig. 4, that the ROIs have a clear disparity in their levels of irreversibility. The sensory ROIs are more reversible than the medial and hippocampal ROIs. Furthermore, there is a skew towards the right hemisphere being more irreversible than the left. This result emerges consistently across all levels. Next, we consider the irreversibility of pairwise interactions ($k = 2$). Panel c) of Fig. 4 shows the 2-order irreversibility for all pairs. We are able to identify strongly irreversible pairs such as the intra-hemispheric pairs (ACL, HITL) and (ACR, HITR). On the other hand, cross-hemispheric pairs, e.g. (ACL, ACR), are the most reversible, indicating a lack of interaction between them. The strong hemispheric symmetry in the results validates the findings, as it is an expected and intuitive observation. Panel d) of Fig. 4 shows the irreversibility for each triplet interaction in the system. The highly irreversible triplets are those that include a hemispheric pair alongside a medial region, with those containing the VMPFC, a region known to drive brain dynamics during task [40], being particularly irreversible. Panel e) of Fig. 4 shows that the most irreversible quadruplet interactions are composed of a hemispheric pair alongside both medial regions as well as those that contain (VMPFC, HITL, HITR) alongside a sensory region. Conversely, the quadruplet containing no medial regions, is the most reversible, and therefore has the least interaction. This is particularly interesting as this quadruplet is made up of the two most irreversible pairs yet they do not appear to interact as a foursome. Therefore, this framework is truly capturing higher-order interactions that cannot simply be decomposed into a sum of independent interactions of lower order. Finally, Panel f) of Fig. 4 shows that quintuplets that contain both medial ROIs are the most irreversible. Furthermore, the quintuplet that does not contain the VMPFC has the most reversible interaction.

We can interpret this result in the context of predictive coding and its links to sensory tasks [41–43]. The participants are exposed to a memorised tonal sequence that does not deviate from their expectation of what they were about to hear. Under the theory of predictive coding, this would result in an adjustment of a participant's prior expectations, facilitated by asymmetric, hierarchical interactions between brain regions at multiple levels, in order to reinforce the prior expectations in light of the new sensory information [44]. This in turn would lead to a cascade of non-equilibrium, asymmetric interactions between key ensembles of regions whose function is optimised for the process of auditory recognition [7]. With the DiMViGI framework, we are able to identify these ensembles in empirical data and reach conclusions about the interactions between ROIs during long-term recognition.

V. DISCUSSION

In this study, we describe a novel framework for measuring the emergence of non-equilibrium dynamics, through multivariate irreversibility, at multiple system levels. We are able to capture the irreversibility of each possible interaction in a multivariate time-series of signals. Applying the DiMViGI framework to neural recordings obtained during a long-term memory recognition task, we investigate the higher-order organisation, and the associated non-equilibrium interactions, of brain regions and how they break time-reversal symmetry during an auditory recognition task. The results clearly show a broad distribution of irreversibility at each system level; hence we are able identify which interactions are driving the global entropy production rate. Furthermore, we link irreversibility to hierarchical predictive coding and theorise that non-equilibrium interactions could emerge as a consequence of the modulation of prior expectations in light of new sensory information [44]. According to the theory of predictive coding, this might be realised through hierarchically asymmetric interactions that, in turn, induce the emergence of irreversibility at multiple system levels [7, 45, 46]. Whilst a recent analysis of these neural recordings with standard methods was able to identify a hierarchy of information processing in the brain during long-term recognition [30], the introduction of the DiMViGI framework appears crucial to uncovering the higher-order and non-equilibrium nature of the interactions. Such insights are opaque to traditional analyses but emerge from the unique lens of non-equilibrium statistical physics.

The implications of the framework and the associated results are multi-fold. Firstly, we go beyond aggregate

[4–7, 9, 10] or univariate [29, 38] measures of irreversibility, expanding the exisiting quiver of techniques for studying non-equilibrium in the brain to include a multilevel approach. Our technique is able to capture multilevel contributions to irreversibility in continuous time-series, so far only considered for binary variables [27], that is nonspecific and can be applied to multivariate time-series from any domain to identify particular highly non-equilibrium interactions. Secondly, we build on the sustained interest in identifying higher-order organisation in neural recordings and other multivariate time-series [47–51], particularly in information theoretic decompositions of brain data that reveal how higher-order functional interactions shape neural dynamics [52–54]. Notably, many higher-order frameworks are either computationally, or by formulation, restricted to studying either triplet [48, 49, 51, 52] or system-wide interactions [47], whilst our results extend easily to all possible levels in the system. Our framework attempts to bridge the broader discussion on higher-order mechanisms and behaviours in complex systems [55–57] with non-equilibrium thermodynamics [58] through the quantification and interpretation of multilevel irreversibility. Finally, our work further solidifies the visibility algorithm, and network analysis of time-series, as an empirically useful tool in the analysis of neural data [36, 59].

Despite these promising results, we note some nuanced limitations in our framework. Whilst the visibility algorithm and the degree distribution approach reduces the dimension of the data, we are still computing an entropy between high-dimensional distributions which is computationally restrictive. This can be circumvented limiting the support of the degree-distribution to exponentially improve computational efficiency whilst minimally affecting numerical accuracy (see SI). Nevertheless, analysing all possible interactions yields a combinatorial explosion, hence we opt for a coarse, low-dimensional, parcellation of the brain that allows us to analyse the system at all possible levels. However, the highlighting of individual tuples is most meaningful when there is a strong intuition about the nature of the interaction, which can be only be expected in low-dimensional parcellations where ROIs are clear, functionally segregated brain areas. Finally, we note that our measure is undirected within the tuple, meaning we cannot identify the direction of information flow as one can with classical measures of causality [60, 61] or some approaches to the AoT [7, 8]. However, we note that the AoT represents directed flow between states and not variables, meaning it is not a direct measure of causality, but instead capturing a distinct, but related, phenomena in interacting dynamics.

A key advantage of the DiMViGI framework is the ability to scale between levels with a consistent approach. Strictly local measures such as auto- and cross-correlations are limited to individual and pairwise interactions [62, 63]. On the other hand, simply applying global measures to each subset of variables in the time-series, such as coarse-graining or using a model-based measure, yields an inconsistent approach where different tuples cannot be compared. Our framework extends consistently to all levels thus yielding directly comparable quantities.

VI. CONCLUSIONS

In this work, we have introduced the Directed Multiplex Visibility Graph Irreversibility framework for measuring the irreversibility of multivariate interactions at all levels within a system. We applied this method to neural recordings during a long-term auditory recognition task to study the relative irreversibility of different interactions between brain regions. Doing so, we were able to demonstrate the hierarchical higher-order organisation of brain dynamics during tasks. This analysis suggests that reinforcement of prior expectations during an auditory recognition task is facilitated through a hierarchy of irreversible higher-order interactions in the brain, an observation that we link to the mechanisms of predictive coding. Furthermore, we highlighted the particular combinations of cognitive and sensorial regions that are preferentially recruited during audition and long-term recognition. This framework is nonspecific and provides a general tool for investigating higher-order interactions and non-equilibrium dynamics in multivariate time-series emerging from other complex systems.

VII. MATERIALS AND METHODS

A. Estimating degree distributions from finite samples

For each sample, a multivariate time-series, we construct the DMVG, defined by the multiplex adjacency matrix, A,

$$A_{ij}^{[l]} = \begin{cases} 1 & \text{if } i \to j \text{ in layer } l \\ 0 & \text{else} \end{cases} . \tag{9}$$

Then we calculate the in- and out-degree of each node in each layer

$$\tilde{d}_i^{[l],\mathrm{in}} = \sum_j A_{ji}^{[l]}, \tag{10}$$

$$\tilde{d}_i^{[l],\mathrm{out}} = \sum_j A_{ij}^{[l]}, \tag{11}$$

where $d_i^{[l],\mathrm{in}}, d_i^{[l],\mathrm{out}}$ are the in-and out-degree of node i in layer l respectively.

For a $k-$tuple $(n_1, \ldots, n_k)$, we calculate $P_{\text{in}}^{(n_1,\ldots,n_k)}(d_1, \ldots, d_k)$ by counting the number of nodes i, across all samples, where

$$\tilde{d}_i^{[l],\text{in}} = d_l, \tag{12}$$

for each $l \in \{1, \ldots, k\}$ simultaneously and for $d_l \in \{1, \ldots, d_{\text{max}}\}$ where d_{max} is the maximum degree of a node in the network, and then dividing through by the total number of nodes in all samples. We calculate the same for $P_{\text{out}}^{(n_1,\ldots,n_k)}(d_1, \ldots, d_k)$.

As we are using a finite number of samples, we then perform Laplace smoothing to eliminate singularities of the form $P(x) = 0 < Q(x)$ for which divergence is ill-defined. Instead of using,

$$P^{(n_1,\ldots,n_k)}(d_1, \ldots, d_k) = \frac{N}{M}, \tag{13}$$

where N is the number of nodes satisfying condition 12 and M is the total number of nodes across samples, we perform the following replacement,

$$P^{(n_1,\ldots,n_k)}(d_1, \ldots, d_k) = \frac{N+1}{M + d_{\text{max}}^k}. \tag{14}$$

Such an approach is equivalent to assuming a uniform Bayesian prior for the degree distributions [64].

B. Computing Jensen-Shannon divergence

We quantify the divergence between the in- and out-degree distributions using Jensen-Shannon divergence (JSD) which is a symmetrised version of Kullback-Leibler divergence (KLD) that does not suppose a model-data relationship [65]. This is defined between two probability distributions P, Q as

$$J(P|Q) = \frac{1}{2}D(P|M) + \frac{1}{2}D(Q|M), \tag{15}$$

where $M = \frac{1}{2}(P + Q)$ is an averaged distribution and $D(\cdot)$ represents the KLD, given by,

$$D(P|Q) = \sum_{x \in \mathcal{X}} P(x) \log \frac{P(x)}{Q(x)}. \tag{16}$$

As $\mathcal{X}$ represents the support of the distribution, it takes the form $\{1, \ldots, d_{\text{max}}\}^k$ where k is the dimension of the probability distributions and d_{max} is the maximum degree of a node in the multiplex. For computational feasibility, d_{max} can be limited during the calculation of JSD, truncating the sum. For 5-order analysis, we limit d_{max} to 75. For a systematic analysis of the effect of degree limiting see SI.

C. Magnetoencephalography (MEG) data

1. *Participants*

The participant cohort consisted of 83 healthy volunteers [33 males and 50 females] with ages in the range 19 to 63 (mean age 28.76 ± 8.06). Participants were recruited in Denmark, came from Western countries, reported normal hearing and gave informed consent before the experiment. The project was approved by the Institutional Review Board (IRB) of Aarhus University (case number: DNC-IRB-2020-006) and experimental procedures complied with the Declaration of Helsinki – Ethical Principles for Medical Research. After pre-processing, the 51 participants with at least 15 non-discarded trials in the first experimental condition were included in the analysis. For those participants with more than 15 trials, 15 trials were randomly sampled.

2. *Experimental stimuli and design*

We employed an old/new paradigm auditory recognition task [30, 32, 33, 35]. Participants listened to a short musical piece twice and asked to memorise it to the best of their ability. The piece was the first four bars of the right-hand part of Johann Sebastian Bach's Prelude No. 2 in C Minor, BWV 847. Next, participants listened to 135 five-tone musical sequences, corresponding to 27 trials in 5 experimental conditions, of 1750 ms each and were requested to indicate if the sequence belonged to the original music or was a variation. Differences between experimental conditions have been described in detail by Bonetti et al [30]. We consider one experimental condition, where participants recognised the original, previously memorised sequences.

3. *Data acquisition*

MEG recordings were taken in a magnetically shielded room at Aarhus University Hospital, Aarhus, Denmark using an Elekta Neuromag TRIUX MEG scanner with 306 channels (Elekta Neuromag, Helsinki, Finland). The sampling rate was 1000 Hz with analogue filtering of 0.1-330 Hz. For further details on the data acquisition see SI.

4. *MEG pre-processing*

First, raw MEG sensor data was processed by MaxFilter [66] to attenuate external interferences. We then applied signal space separation (for parameters see SI). Then the data was converted into Statistical Parametric Mapping (SPM) format, preprocessed and analyzed in MATLAB (MathWorks, Natick, MA, USA) using in-house codes and the Oxford

Centre for Human Brain Activity (OHBA) Software Library (OSL) [67]. The continuous MEG data was visually inspected and large artefacts were removed using OSL. Less than 0.1% of the collected data was removed. Next, independent component analysis (ICA) was implemented to discard artefacts in the brain data from heart-beats and eye-blinks (for details see SI) [68]. Lastly, the signal was epoched in 135 trials, 27 trials for each of 5 experimental conditions and the mean signal recorded in the baseline (the post-stimulus brain signal) was removed. Each resulting trial lasted 4400 ms plus 100 ms of baseline time.

5. *Source reconstruction*

We employed the beamforming method to spatially localise the MEG signal [69]. For details on the beamforming algorithm and the implementation see SI.

D. Code and data availability

The code used to implement the DiMViGI framework will be made available at `https://github.com/rnartallo/multilevelirreversibility` following peer-review and publication.
The in-house code used for MEG pre-processing is available at `https://github.com/leonardob92/LBPD-1.0`.
The multimodal neuroimaging data analysed here is available upon reasonable request.

VIII. AUTHOR CONTRIBUTIONS

R.N.K designed and performed the analysis and wrote the manuscript. L.B. designed the analysis and the experiment and collected and pre-processed data and edited the manuscript. G.F.R. designed the experiment, collected and pre-processed the data. P.V. designed and provided funding for the experiment. G.D. edited the manuscript. M.L.K, R.L. and A.G. supervised the research, analysed the data and edited the manuscript. The authors declare no competing interests.

ACKNOWLEDGMENTS

R.N.K was supported by an EPSRC doctoral scholarship from grants EP/T517811/1 and EP/R513295/1. L.B, M.L.K and P.V were supported by the Center for Music in the Brain (MIB), funded by the Danish National Research Foundation (project number DNRF117). L.B. was also supported by Carlsberg Foundation (CF20-0239), Lundbeck Foundation (Talent Prize 2022), Linacre College of the University of Oxford (Lucy Halsall fund) and Nordic Mensa Fund. G.F.R. was supported by Fundación Mutua Madrileña. G.D. was supported by the AGAUR research support grant (ref. 2021 SGR 00917) funded by the Department of Research and Universities of the Generalitat of Catalunya, by the project NEurological MEchanismS of Injury, and the project Sleep-like cellular dynamics (NEMESIS) (ref. 101071900) funded by the EU ERC Synergy Horizon Europe and by the project PID2022-136216NB-I00 financed by the MCIN /AEI /10.13039/501100011033 / FEDER, UE., the Ministry of Science and Innovation, the State Research Agency and the European Regional Development Fund. M.L.K was further supported by Centre for Eudaimonia and Human Flourishing funded by the Pettit and Carlsberg Foundations. R.L. was supported by the EPSRC grants EP/V013068/1 and EP/V03474X/1.

[1] D. S. Bassett and O. Sporns, Network neuroscience, Nature Neuroscience **20**, 353–364 (2017).
[2] C. W. Lynn and D. S. Bassett, The physics of brain network structure, function and control, Nature Reviews Physics **1**, 318–332 (2019).
[3] D. R. Chialvo, Emergent complex neural dynamics, Nature Physics **6**, 744 (2010).
[4] C. W. Lynn, E. J. Cornblath, L. Papadopoulos, and D. S. Bassett, Broken detailed balance and entropy production in the human brain, Proceedings of the National Academy of Sciences of the United States of America **118** (2021).
[5] G. Deco, Y. Sanz-Perl, H. Bocaccio, E. Tagliazucchi, and M. L. Kringelbach, The INSIDEOUT framework provides precise signatures of the balance of intrinsic and extrinsic dynamics in brain states, Communications Biology **5** (2022).
[6] G. Deco, Y. Sanz-Perl, L. de la Fuente, J. D. Sitt, B. T. T. Yeo, E. Tagliazucchi, and M. L. Kringelbach, The arrow of time of brain signals in cognition: Potential intriguing role of parts of the default mode network, Network Neuroscience **7**, 966–998 (2023).
[7] R. Nartallo-Kaluarachchi, M. Asllani, G. Deco, M. L. Kringelbach, A. Goriely, and R. Lambiotte, Broken detailed balance and entropy production in directed networks, arXiv **2402.19157** (2024).
[8] T. Bolton, D. Van De Ville, E. Amico, M. Preti, and R. Liégeois, The arrow-of-time in neuroimaging time series identifies causal triggers of brain function, Human Brain Mapping **44**, 4077 (2023).
[9] G. Deco, C. Lynn, Y. Sanz-Perl, and M. L. Kringelbach, Violations of the fluctuation-dissipation theorem reveal distinct non-equilibrium dynamics of brain states, Physical Review

E **108** (2023).

[10] M. Gilson, E. Tagliazucchi, and R. Cofré, Entropy production of multivariate Ornstein-Uhlenbeck processes correlates with consciousness levels in the human brain, Physical Review E **107** (2023).

[11] E. Guzmán, Y. Sanz Perl, J. Vohryzek, A. Escrichs, D. Manasova, B. Türker, E. Tagliazucchi, M. Kringelbach, J. Sitt, and G. Deco, The lack of temporal brain dynamics asymmetry as a signature of impaired consciousness states, Interface Focus **13** (2023).

[12] J. Cruzat, R. Herzog, P. Prado, Y. Sanz-Perl, R. Gonzalez-Gomez, S. Moguilner, M. Kringelbach, G. Deco, E. Tagliazucchi, and A. Ibañez, Temporal irreversibility of large-scale brain dynamics in Alzheimer's disease, Journal of Neuroscience **43**, 1643–1656 (2023).

[13] L. de la Fuente, F. Zamberlan, H. Bocaccio, M. Kringelbach, G. Deco, Y. Sanz Perl, C. Pallavicini, and E. Tagliazucchi, Temporal irreversibility of neural dynamics as a signature of consciousness, Cerebral Cortex **33**, 1856 (2023).

[14] S. Carnot, *Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance* (Bachelier, Paris, 1824).

[15] R. Clausius, Ueber die bewegende kraft der wärme und die gesetze, welche sich daraus für die wärmelehre selbst ableiten lassen, Annalen der Physik **79**, 368–397 (1850).

[16] C. Jarzynski, Nonequilibrium equality for free energy differences, Physical Review Letters **78** (1997).

[17] G. E. Crooks, Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, Physical Review E **60** (1999).

[18] R. K. P. Zia and B. Schmittmann, Probability currents as principal characteristics in the statistical mechanics of non-equilibrium steady states, Journal of Statistical Mechanics: Theory and Experiment (2007).

[19] A. S. Eddington, *The Nature of the Physical World* (Cambridge University Press, 1928).

[20] U. Seifert, Entropy production along a stochastic trajectory and an integral fluctuation theorem, Physical Review Letters **95** (2005).

[21] R. Kawai, J. M. R. Parrondo, and C. V. den Broeck, Dissipation: The phase-space perspective, Physical Review Letters **98** (2007).

[22] E. H. Feng and G. E. Crooks, Length of time's arrow, Physical Review Letters **101** (2008).

[23] C. W. Lynn, C. M. Holmes, W. Bialek, and D. J. Schwab, Decomposing the local arrow of time in interacting systems, Physical Review Letters **129** (2022).

[24] A. Seif, M. Hafezi, and C. Jarzynski, Machine learning the thermodynamic arrow of time, Nature Physics **17**, 105 (2021).

[25] M. L. Kringelbach, Y. Sanz-Perl, E. Tagliazucchi, and G. Deco, Toward naturalistic neuroscience: Mechanisms underlying the flattening of brain hierarchy in movie-watching compared to rest and task, Science Advances **9** (2023).

[26] G. Deco, D. Vidaurre, and M. L. Kringelbach, Revisiting the global workspace orchestrating the hierarchical organization of the human brain, Nature Human Behaviour **5**, 497 (2021).

[27] C. W. Lynn, C. M. Holmes, W. Bialek, and D. J. Schwab, Emergence of local irreversibility in complex interacting systems, Physical Review E **106** (2022).

[28] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuño, From time series to complex networks: The visibility graph, Proceedings of the National Academy of Sciences of the United States of America **105**, 4972 (2008).

[29] L. Lacasa, A. Nuñez, E. Roldán, J. Parrondo, and B. Luque, Time series irreversibility: a visibility graph approach, The European Physical Journal B **85** (2012).

[30] L. Bonetti, G. F. Rubio, F. Carlomagno, D. Pantazis, P. Vuust, and M. Kringelbach, Revealing the spacetime hierarchical whole-brain dynamics of auditory predictive coding, Nature Communications **In press** (2024).

[31] L. Bonetti, E. Brattico, S. E. P. Bruzzone, G. Donati, G. Deco, D. Pantazis, P. Vuust, and M. L. Kringelbach, Brain recognition of previously learned versus novel temporal sequences: a differential simultaneous processing, Cerebral Cortex **33**, 5524 (2023).

[32] L. Bonetti, G. F. Rubio, M. Lumaca, F. Carlomagno, E. R. Olsen, A. Criscuolo, S. Kotz, P. Vuust, E. Brattico, and M. Kringelbach, Challenging age-related decline in brain function: Evidence from fast neuroimaging of musical sequence recognition, bioRxiv 10.1101/2023.07.13.548815 (2022).

[33] G. Fernández-Rubio, E. Brattico, S. A. Kotz, M. L. Kringelbach, P. Vuust, and L. Bonetti, Magnetoencephalography recordings reveal the spatiotemporal dynamics of recognition memory for complex versus simple auditory sequences, Communications Biology **5** (2022).

[34] G. Fernández-Rubio, E. R. Olsen, M. Klarlund, O. Mallon, F. Carlomagno, P. Vuust, M. L. Kringelbach, E. Brattico, and L. Bonetti, Investigating the impact of age on auditory short-term, long-term, and working memory, Psychology of Music **0** (2023).

[35] G. Fernández-Rubio, F. Carlomagno, P. Vuust, M. L. Kringelbach, and L. Bonetti, Associations between abstract working memory abilities and brain activity underlying long-term recognition of auditory sequences, PNAS Nexus **1**, 1 (2022).

[36] T. F. Varley and O. Sporns, Network analysis of time series: Novel approaches to network neuroscience, Frontiers in Neuroscience **15** (2022).

[37] L. Lacasa, V. Nicosia, and V. Latora, Network structure of multivariate time series, Scientific Reports **5** (2015).

[38] J. F. Donges, R. V. Donner, and J. Kurths, Testing time series reversibility using complex network methods, Europhysics Letters **102** (2012).

[39] G. Weiss, Time-reversibility of linear stochastic processes, Journal of Applied Probability **12**, 831 (1975).

[40] G. Deco, Y. S. Perl, A. Ponce-Alvarez, E. Tagliazucchi, P. C. Whybrow, J. Fuster, and M. L. Kringelbach, One ring to rule them all: The unifying role of prefrontal cortex in steering task-related brain dynamics, Progress in Neurobiology **227** (2023).

[41] K. Friston, Predictive coding, precision and synchrony, Cognitive Neuroscience **3**, 238 (2012).

[42] S. Koelsch, P. Vuust, and K. Friston, Predictive processes and the peculiar case of music, Trends in Cognitive Sciences

23, 63 (2019).

[43] P. Vuust, O. A. Heggli, K. J. Friston, and M. L. Kringelbach, Music in the brain, Nature Reviews Neuroscience **23**, 287 (2022).

[44] K. Friston, A theory of cortical responses, Philosophical Transactions of the Royal Society B **360**, 815 (2005).

[45] K. Friston, Hierarchical models in the brain, PLOS Computational Biology **4** (2008).

[46] K. Friston and S. Kiebel, Predictive coding under the free-energy principle, Philosophical Transactions of the Royal Society B **364**, 1211 (2009).

[47] G. Petri, P. Expert, F. Turkheimer, R. Carhart-Harris, D. Nutt, P. J. Hellyer, and F. Vaccarino, Homological scaffolds of brain functional networks, Journal of the Royal Society Interface **11** (2014).

[48] A. Santoro, F. Battiston, G. Petri, and E. Amico, Higher-order organization of multivariate time series, Nature Physics **19**, 221–229 (2023).

[49] M. R. R. Tabar, F. Nikakhtar, L. Parkavousi, A. Akhshi, U. Feudel, and K. Lehnertz, Revealing higher-order interactions in high-dimensional complex systems: A data-driven approach, Physical Review X **14** (2024).

[50] R. F. Betzel, J. Faskowitz, and O. Sporns, Living on the edge: network neuroscience beyond nodes, Trends in Cognitive Sciences **27**, 1068 (2023).

[51] C. Giusti, R. Ghrist, and D. S. Bassett, Two's company, three (or more) is a simplex, Journal of Computational Neuroscience **41**, 1 (2016).

[52] T. F. Varley, M. Pope, M. G. Puxeddu, J. Faskowitz, and O. Sporns, Partial entropy decomposition reveals higher-order information structures in human brain activity, Proceedings of the National Academy of Sciences of the United States of America **120** (2023).

[53] A. I. Luppi, P. A. M. Mediano, F. E. Rosas, N. Holland, T. D. Fryer, J. T. O'Brien, J. B. Rowe, D. K. Menon, D. Bor, and E. A. Stamatakis, A synergistic core for human brain evolution and cognition, Nature Neuroscience **25**, 771–782 (2022).

[54] A. I. Luppi, F. E. Rosas, P. A. Mediano, D. K. Menon, and E. A. Stamatakis, Information decomposition and the informational architecture of the brain, Trends in Cognitive Sciences **28**, 352 (2024).

[55] R. Lambiotte, M. Rosvall, and I. Scholtes, From networks to optimal higher-order models of complex systems, Nature Physics **15**, 313–320 (2019).

[56] F. Battiston, E. Amico, A. Barrat, G. Bianconi, G. F. de Arruda, B. Franceschiello, I. Iacopini, S. Kéfi, V. Latora, Y. Moreno, M. M. Murray, T. P. Peixoto, F. Vaccarino, and G. Petri, The physics of higher-order interactions in complex systems, Nature Physics **17**, 1093–1098 (2021).

[57] F. E. Rosas, P. A. M. Mediano, A. I. Luppi, T. F. Varley, J. T. Lizier, S. Stramaglia, H. J. Jensen, and D. Marinazzo, Disentangling high-order mechanisms and high-order behaviours in complex systems, Nature Physics **18**, 476 (2022).

[58] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, Reports of Progress in Physics **475** (2012).

[59] S. Sannino, S. Stramaglia, L. Lacasa, and D. Marinazzo, Visibility graphs for fMRI data: Multiplex temporal graphs and their modulations across resting-state networks, Network Neuroscience **1**, 208–221 (2017).

[60] C. W. J. Granger, Investigating causal relations by econometric models and cross-spectral methods, Econometrica **37**, 424 (1969).

[61] T. Schreiber, Measuring information transfer, Physical Review Letters **85** (2000).

[62] I. Z. Steinberg, On the time reversal of noise signals, Biophysical Journal **50**, 171 (1986).

[63] A. Crisanti, A. Puglisi, and D. Villamaina, Nonequilibrium and information: The role of cross correlations, Physical Review E **85** (2012).

[64] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval* (Cambridge University Press, 2008) p. 260.

[65] F. Nielsen, On the Jensen–Shannon symmetrization of distances relying on abstract means, Entropy **21** (2019).

[66] S. Taulu and J. Simola, Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements, Physics in Medicine and Biology **51**, 1 (2010).

[67] M. Woolrich, L. Hunt, A. Groves, and G. Barnes, MEG beamforming using Bayesian PCA for adaptive data covariance matrix regularization, NeuroImage **57**, 1466 (2011).

[68] D. Mantini, S. D. Penna, L. Marzetti, F. de Pasquale, V. Pizzella, M. Corbetta, and G. L. Romani, A signal-processing pipeline for magnetoencephalography resting-state networks, Brain Connectivity **1**, 49 (2011).

[69] A. Hillebrand and G. R. Barnes, Beamformer analysis of MEG data, International Review of Neurobiology **68**, 149 (2005).

Supplementary Information:
'Multilevel irreversibility reveals higher-order organisation of non-equilibrium interactions in human brain dynamics'

Ramón Nartallo-Kaluarachchi[1,2], Leonardo Bonetti[3,2], Gemma Fernández-Rubio[3], Peter Vuust[3], Gustavo Deco[4,5,6], Morten L. Kringelbach[2,3,7], Renaud Lambiotte[1,8], Alain Goriely[1,*]

[1]Mathematical Institute, University of Oxford, Oxford, OX2 6GG, United Kingdom
[2]Centre for Eudaimonia and Human Flourishing, University of Oxford, Oxford, OX3 9BX, United Kingdom
[3]Center for Music in the Brain, Aarhus University, & The Royal Academy of Music, Aarhus/Aalborg, Denmark
[4]Centre for Brain and Cognition, Computational Neuroscience Group, Universitat Pompeu Fabra, 08018 Barcelona, Spain
[5]Department of Information and Communication Technologies, Universitat Pompeu Fabra, 08018 Barcelona, Spain
[6]Institució Catalana de la Recerca i Estudis Avancats (ICREA), 08010 Barcelona, Spain
[7]Department of Psychiatry, University of Oxford, Oxford, OX3 7JX United Kingdom
[8]Turing Institute, London, NW1 2DB United Kingdom
[*]goriely@maths.ox.ac.uk

May 2, 2024

1 Introduction

In this supporting information, we provide additional results and analysis not included in the main manuscript. This supporting information is organised as follows. In Section 2, we show the results of the directed multiplex visibility graph irreversibility (DiMViGI) framework applied to data from individual participants to produce distributions of participant spread for the irreversibility of each tuple, rather than the cohort-level analysis presented in the main manuscript. We assess the significance of the differences between tuples using pairwise $t-$tests and one-way ANOVAs. Moreover, we calculate the correlations between the cohort and participant level analysis. In Section 3, we further validate the significance of the results obtained at the participant level by shuffling the time-series to produce surrogate data. We show that shuffling restores detailed balance indicating that the results obtained are strictly due to the specific trajectories present in the time-series. Next, in Section 4, we show that the DiMViGI framework factorises for independent variables, theoretically validating the argument that it captures 'true' higher-order interactions. Using the factorisation, in Section 5, we are able to define the unique irreversibility generated by a higher order interaction by removing the lower-level interactions. We compare this to the 'combined' results presented in the manuscript. In Section 6, we validate that our method captures a correlate of the entropy production rate by using simulated data from four specific examples of the multivariate Ornstein-Uhlenbeck process. In Section 7, we investigate the effect on the results of limiting the maximum degree in the support of the distributions, an approach that improves computational efficiency whilst only minimally reducing accuracy. In Section 8, we discuss the definitions of entropy production rate for Markovian and non-Markovian dynamics. Finally in Section 9 we present a comprehensive description of the experimental paradigm and the techniques used to record and pre-process the magnetoencephalography (MEG) data.

2 The DiMViGI framework applied to participant-level data

In this section we show the results of applying the DiMViGI framework to data at the participant-level and obtain distributions for the irreversibility of each tuple. As mentioned in the main manuscript, we analysed MEG recordings from 51 participants with 15 trials per participant. In the cohort-level analysis presented in the main manuscript, we constructed the in- and out-degree distributions using $51 \times 15 = 765$ samples of the multiplex network. In order to examine the spread between participants, we repeat the same analysis for each participant in isolation, using only the 15 associated trials. As a result the degree distributions are much more poorly estimated and produce much higher divergences. Nevertheless, we are able to quantify the irreversibility of each tuple of brain regions for each participant and examine the distribution.

Figure 1 shows the results of the DiMViGI analysis for the participant-level data distributions of the irreversibility for each tuple in each level. Panel a) shows the schematic representation of the 6 regions of interest (ROIs) that correspond to variables in the multivariate time-series. The icons on the x-axis of the subsequent panels indicate which ROIs are included in each tuple. Panels b-f) show the participant-level distributions for 1-5 order respectively. We run one-way ANOVAs and find that, at each level 1-5, the tuple is a significant predictor of irreversibility ($p < 0.00001$). In addition, we run paired $t-$tests to see which tuples at each level are significantly different in a pairwise comparison. Figures 2-6 display the participant-level distributions with the significance results of the pairwise $t-$tests displayed. Due to the explosion in possible pairs of tuples, for levels 2-5 we restrict to pairwise comparisons between tuples that contain at least one ROI in common (note: due to the clear hierarchy at level 1, we compare adjacent ROIs only for clarity). The significance of each comparison is denoted as follows: (ns) if $p > 0.05$; (*) if $p < 0.05$; (**) if $p < 0.01$; (***) if $p < 0.001$ and (****) if $p < 0.0001$. Figure 2 shows that at level 1, the difference between each tuple in pairwise comparison is significant ($p < 0.0001$). Figures 3-6 show that at levels 2-5, there is a mixture of significant and not significant differences depending on the number of ROIs in common between the compared tuples.

Finally, we compare the participant-level analysis to the cohort-level analysis by calculating the ranking of tuples at each level for each participant and comparing it to the cohort-level ranking, using Spearman's ρ. In addition, for each participant at each level, we calculate Pearson's r (correlation coefficient) between the participant level and the cohort level. Panel a) of Fig 7 shows the ρ for each participant

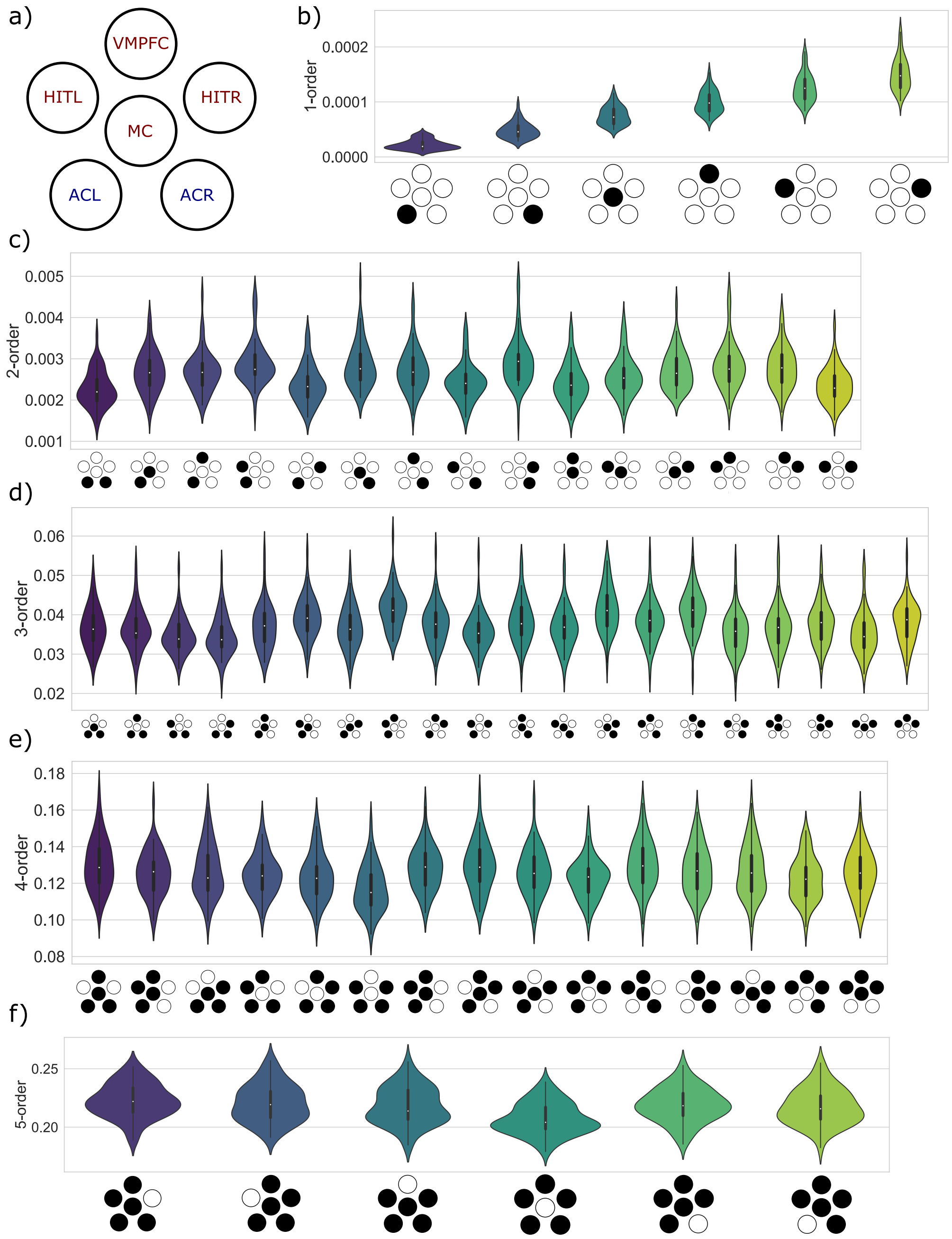

Figure 1: Participant distribution of irreversibility for each tuple. a) Schematic representation of the 6 brain regions of interest (ROIs) in the MEG recordings. The icon in the following panels indicates which regions are included in each tuple. b) 1-order irreversibility distribution for each ROI in isolation. The results follow the same hierarchy as the cohort-level analysis in the main manuscript. c) 2-order irreversibility distribution for each pairs of ROIs. d) 3-order irreversibility distribution for each triplet of ROIs. e) 4-order irreversibility distribution for each quadruplet of ROIs. f) 5-order irreversibility distribution for each quintuplet of ROIs.

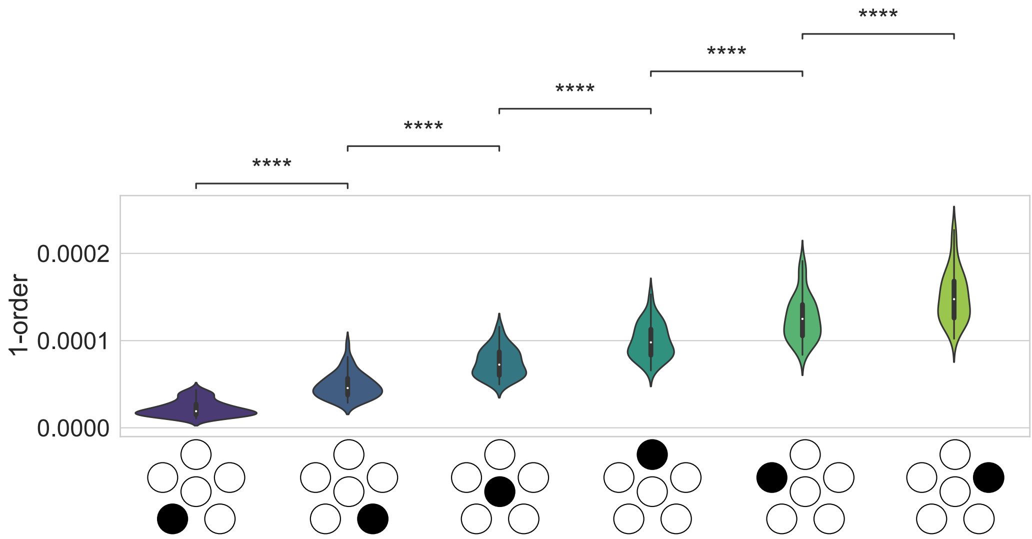

Figure 2: 1-irreversibility of the participant-level data with the results of paired $t-$tests between adjacent ROIs displayed. Each pairwise comparison is significant. (****) implies significance at level $p < 0.0001$.

at each level when compared to the cohort-level ranking. Panel b) of Fig 7 shows the r for each participant at each level when compared to the cohort-level measurements. Both show that at lower orders (1-3), the measurements, and rankings, obtained from the participants in isolation agree closely with the cohort-level results. However, at higher order (4-5), the low number of samples, 15, in the participant-level analysis is not enough to accurately estimate the high-dimensional degree distributions leading to a lack of agreement between the well-estimated cohort analysis and the poorly-estimated participant analysis.

3 Validation of results against surrogate data from shuffling time-series

In order to validate that the irreversibilities calculated from the MEG data are significant quantities reflecting the temporal structure of the neural dynamics, we must compare them to surrogate data. When generating surrogate data, we aim to break the temporal correlations and restore detailed balance. In order to do this, we randomly shuffled the time-series in time. This means that the number of occurrences of each state remains the same as the original data, but the sequence of states is now randomised thereby restoring detailed balance, as shown by Lynn et al [24].

Other common approaches for generating surrogate time-series such as phase randomisation or Fourier transform surrogates preserve the temporal structure of the time-series and are therefore unsuitable for this application [22, 30].

Figure 8 shows the comparison of the measurements in the original MEG time-series and its randomly shuffled surrogate. For each tuplet the difference between the shuffled and original time-series

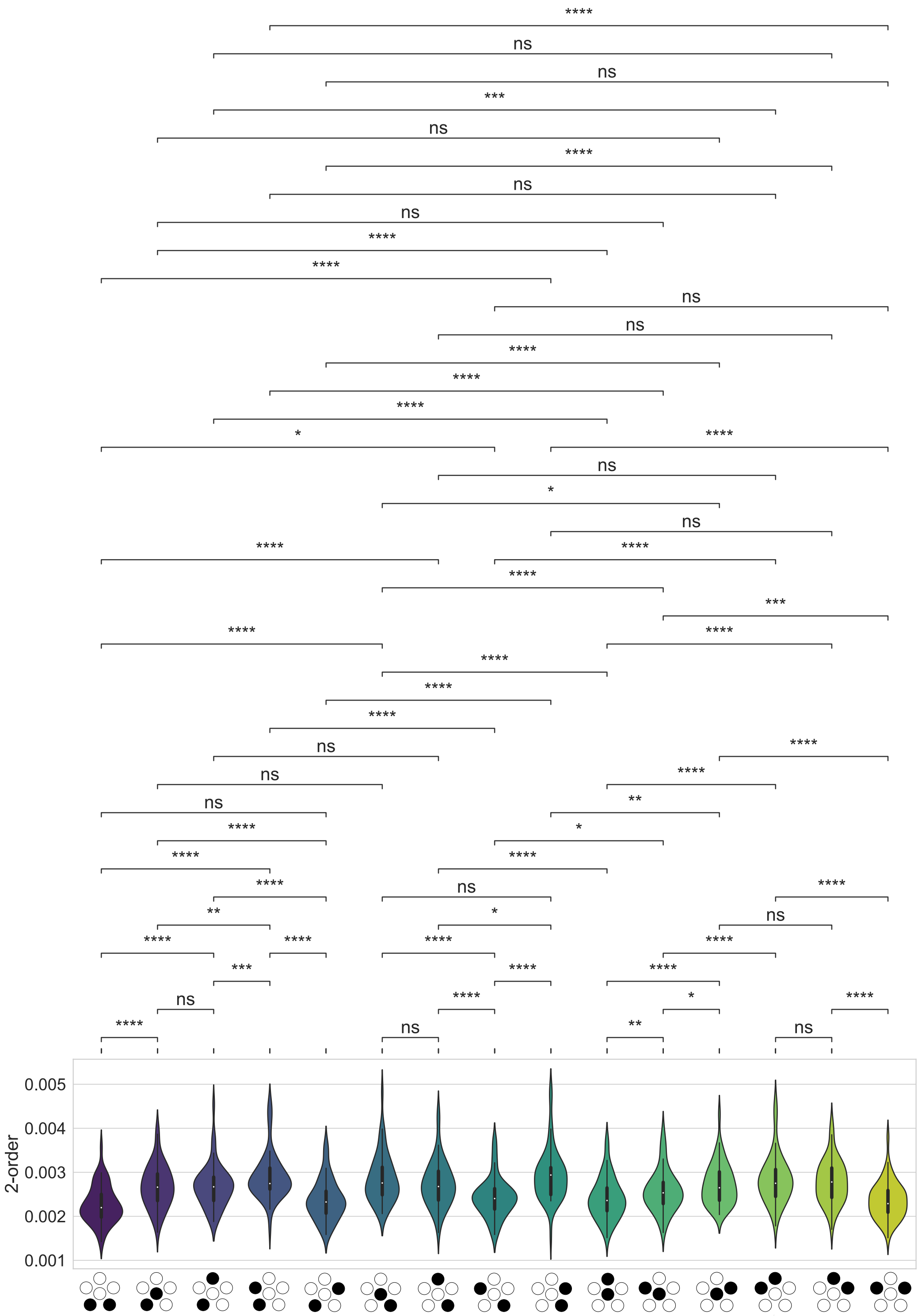

Figure 3: 2-irreversibility of the participant-level data with the results of paired $t-$tests between all pairs of pairs of ROIs that contain at least one ROI in common displayed. The significance of each comparison is denoted as follows: (ns) if $p > 0.05$; (*) if $p < 0.05$; (**) if $p < 0.01$; (***) if $p < 0.001$ and (****) if $p < 0.0001$. There is a range of not significant and significant comparisons between pairs of ROIs.

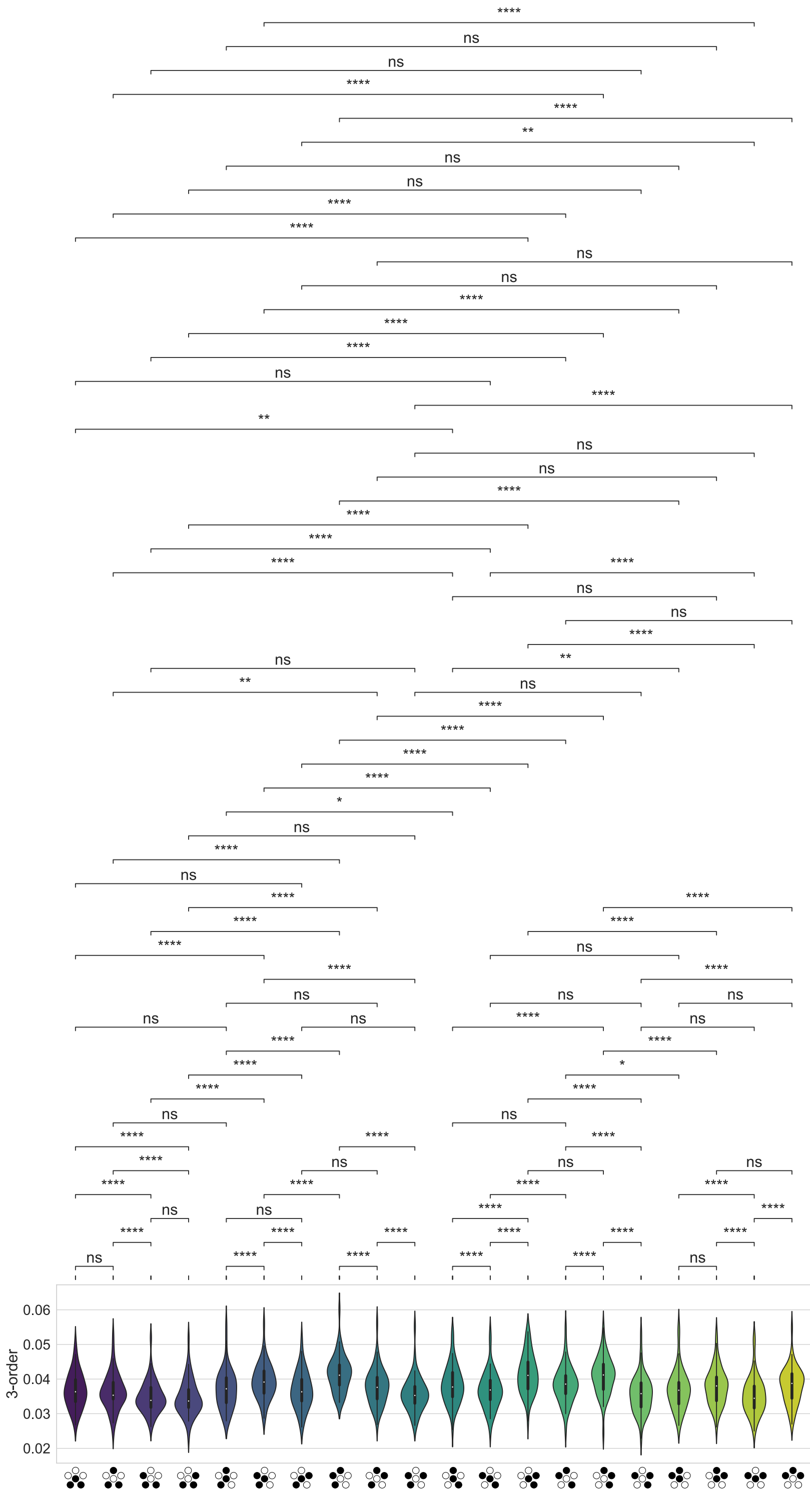

Figure 4: 3-irreversibility of the participant-level data with the results of paired t−tests between all pairs of triplet of ROIs that contain at least one ROI in common displayed. The significance of each comparison is denoted as follows: (ns) if $p > 0.05$; (*) if $p < 0.05$; (**) if $p < 0.01$; (***) if $p < 0.001$ and (****) if $p < 0.0001$. There is a range of not significant and significant comparisons between triplets of ROIs.

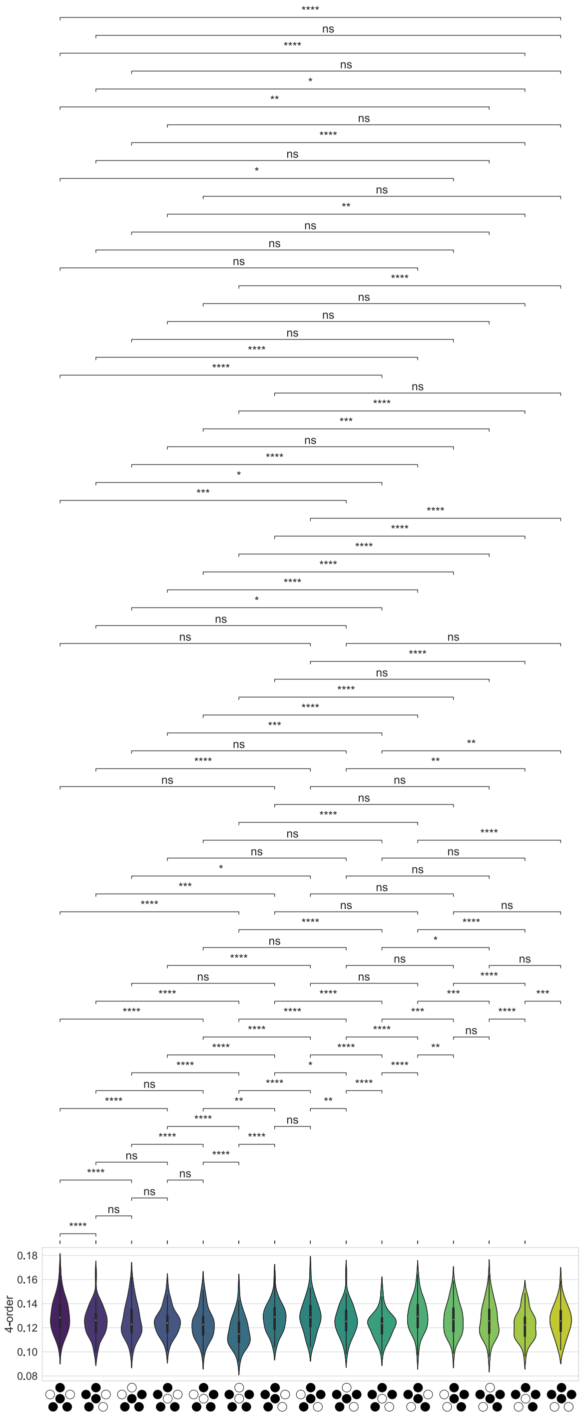

Figure 5: 4-irreversibility of the participant-level data with the results of paired $t-$tests between all pairs of quadruplets of ROIs that contain at least one ROI in common displayed. The significance of each comparison is denoted as follows: (ns) if $p > 0.05$; (*) if $p < 0.05$; (**) if $p < 0.01$; (***) if $p < 0.001$ and (****) if $p < 0.0001$. There is a range of not significant and significant comparisons between quadruplets of ROIs.

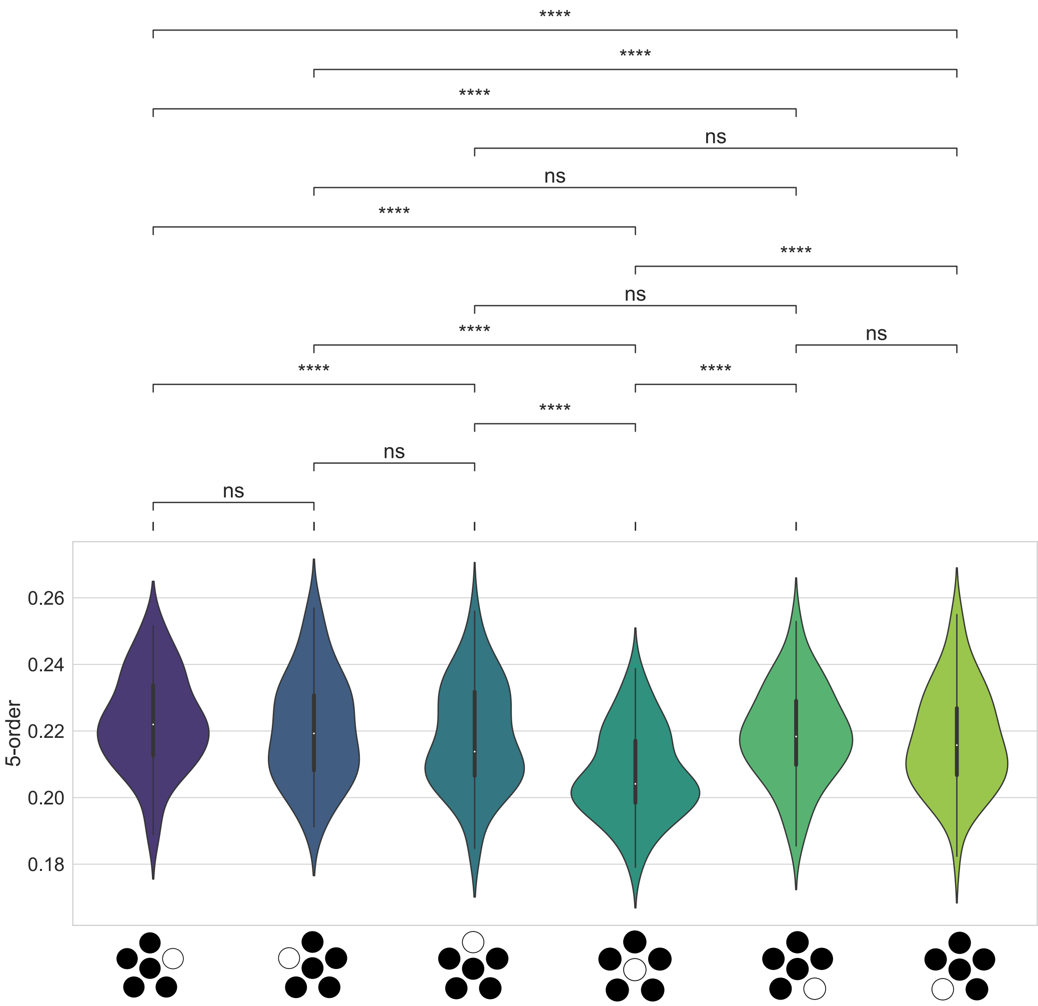

Figure 6: 5-irreversibility of the participant-level data with the results of paired t−tests between all pairs of quintuplets of ROIs that contain at least one ROI in common displayed. The significance of each comparison is denoted as follows: (ns) if $p > 0.05$; (*) if $p < 0.05$; (**) if $p < 0.01$; (***) if $p < 0.001$ and (****) if $p < 0.0001$. There is a range of not significant and significant comparisons between quintuplets of ROIs.

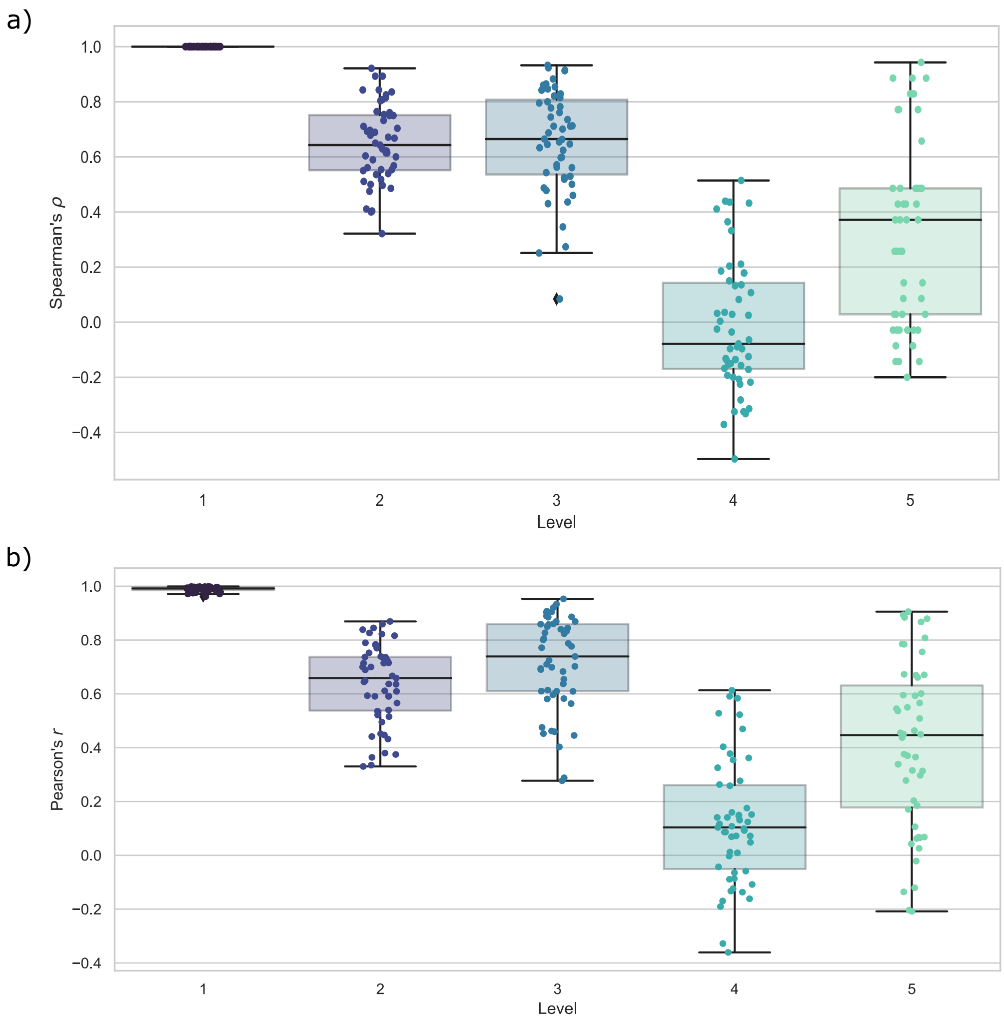

Figure 7: Correlations between participant and cohort level results. a) Spearman's ρ coefficient for the ranking of tuples at each level for each participant when compared to the ranking obtained from the cohort-level analysis. b) Pearson's r correlation coefficient for the measurements of tuples at each level for each participant when compared to the measurements obtained from the cohort-level analysis. The figure shows that at lower orders, the participant distributions agree more closely with the cohort-level analysis. However, at higher orders, the degree distributions are more poorly estimated leading to low agreement between the cohort and participant-level analysis.

is significant ($p < 0.0001$). This shows that the irreversibility measured using the DiMViGI framework is a significant statistical feature of the multivariate time-series as the shuffled data is measured to be far more reversible using the DiMViGI framework.

4 The DiMViGI framework factorises for independent variables

To illustrate that the DiMViGI framework indeed can differentiate higher order interactions from the composition of lower order ones, we consider the framework applied to a k-tuple of variables, $(x_1, ...x_k)$. First we assume that x_1 is independent of $(x_2, ..., x_k)$ and show that we can write the irreversibility of the $k-$tuple as the sum of the irreversibility of x_1 plus the irreversibility of the $(k-1)-$tuple $(x_2, ..., x_k)$. Inductively, we can show that the DiMViGI framework factorises for independent variables meaning that the irreversibility of their interaction is merely the sum of the irreversibility of each variable in isolation. This validates that the framework is truly capturing multilevel irreversibility.

Consider $(x_1, ..., x_k)$ such that x_1 is independent of the other variables. As x_1 is independent, the edges in the associated layer of the multiplex network are also independent. As a result, the joint (in- and out-) degree distribution of the multiplex factorises as follows,

$$P(d_1, ..., d_k) = P(d_1)P(d_2, ..., d_k). \tag{1}$$

Under the DiMViGI framework, we quantify the irreversibility of the triplet as

$$\begin{align} \varsigma^{(x_1,...,x_k)} &= \text{JSD}(P_{\text{in}}(d_1, ..., d_k), P_{\text{out}}(d_1, ..., d_k)), \tag{2} \\ &= \sum_{d_1,...,d_k} P_{\text{in}}(d_1, ..., d_k) \log \frac{P_{\text{in}}(d_1, ..., d_k)}{P^*(d_1, ..., d_k)} + \sum_{d_1,...,d_k} P_{\text{out}}(d_1, ..., d_k) \log \frac{P_{\text{out}}(d_1, ..., d_k)}{P^*(d_1, ..., d_k)}, \tag{3} \end{align}$$

where $P^* := \frac{1}{2}(P_{\text{in}} + P_{\text{out}})$. We focus first on the term concerning the in-degree distribution and use the independence of x_1 and the properties of logarithms to factorise and simplify this expression,

$$\begin{align} &\sum_{d_1,...,d_k} P_{\text{in}}(d_1, ..., d_k) \log \frac{P_{\text{in}}(d_1, ..., d_k)}{P^*(d_1, ..., d_k)} \nonumber \\ &\quad = \sum_{d_1,...,d_k} P_{\text{in}}(d_1)P_{\text{in}}(d_2, ..., d_k) \log \frac{P_{\text{in}}(d_1)P_{\text{in}}(d_2, ..., d_k)}{P^*(d_1)P^*(d_2, ..., d_k)} \tag{4} \\ &\quad = \sum_{d_1,...,d_k} P_{\text{in}}(d_1)P_{\text{in}}(d_2, ..., d_k) \left(\log \frac{P_{\text{in}}(d_2, ..., d_k)}{P^*(d_2, ..., d_k)} + \log \frac{P_{\text{in}}(d_1)}{P^*(d_1)} \right) \tag{5} \\ &\quad = \sum_{d_2,...,d_k} P_{\text{in}}(d_2, ..., d_k) \log \frac{P_{\text{in}}(d_2, ..., d_k)}{P^*(d_2, ..., d_k)} + \sum_{d_1} P_{\text{in}}(d_1) \log \frac{P_{\text{in}}(d_1)}{P^*(d_1)}. \tag{6} \end{align}$$

By symmetry the same is true for the term concerning the out-degree distribution. Substituting in the simplified expression, we get

$$\begin{align} \varsigma^{(x_1,...,x_k)} &= \text{JSD}(P_{\text{in}}(d_2, ..., d_k), P_{\text{out}}(d_2, ..., d_k)) + \text{JSD}(P_{\text{in}}(d_1), P_{\text{out}}(d_1)), \tag{7} \\ &= \varsigma^{(x_2,...,x_k)} + \varsigma^{(x_1)}. \tag{8} \end{align}$$

This indicates that in this k-dimensional system that does not contain a genuine k-order interaction, the irreversibility of the $k-$tuple simply decomposes into the sum of non-independent tuples. By induction,

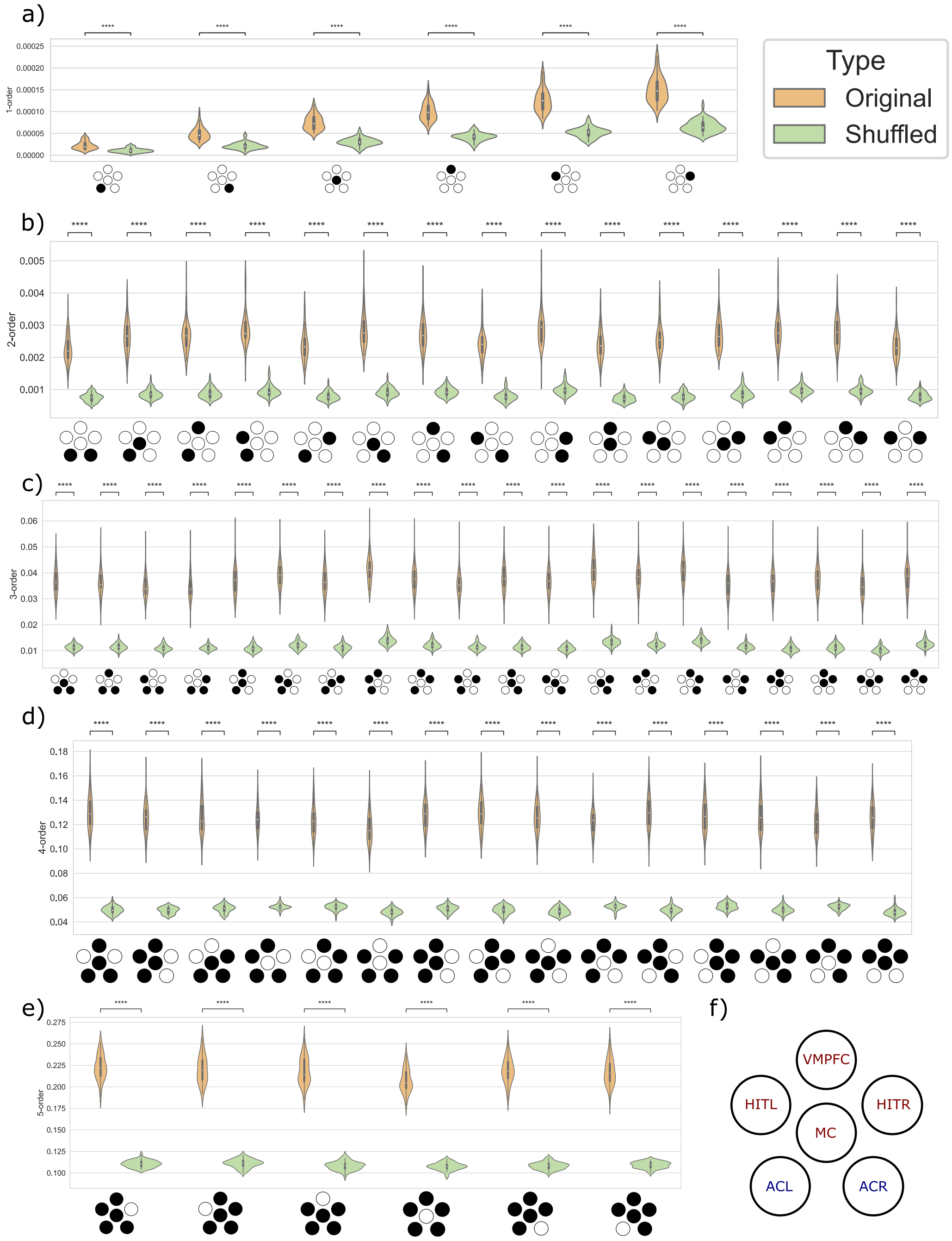

Figure 8: Comparison of original neural recording against surrogate data obtained via shuffling the time-series in time. The difference between the original and shuffled data for each tuple at each level is significant. We label (****) if $p < 0.0001$. a) 1-order irreversibility is significant ($p < 0.0001$) for each ROI when compared to shuffled data. b) 2-order irreversibility is significant ($p < 0.0001$) for each pair of ROIs when compared to shuffled data. c) 3-order irreversibility is significant ($p < 0.0001$) for each triplet of ROIs when compared to shuffled data. d) 4-order irreversibility is significant ($p < 0.0001$) for each quadruplet of ROIs when compared to shuffled data. e) 5-order irreversibility is significant ($p < 0.0001$) for each quintuplet of ROIs when compared to shuffled data. f) Schematic representation of the 6 brain regions of interest (ROIs) in the MEG recordings. The icon in the preceding panels indicates which regions are included in each tuple.

for a $k-$tuple where all variables are independent, the irreversibility fully decomposes into the sum of the 1-order irreversibilities,

$$\varsigma^{(x_1,\ldots,x_k)} = \sum_{i=1}^{k} \varsigma^{(x_i)}. \tag{9}$$

5 Unique contributions from higher order interactions

In our analysis, we have considered the irreversibility of multilevel interactions. However, for a given $k-$order interaction, we have measured the irreversibility of the combined $k-$tuple. This is in contrast with the 'unique' irreversibility that is contributed purely by the k-body interaction, discounting the $j-$body interactions for $j < k$ that are included within this $k-$tuple.

Within the theory of higher order interactions, this distinction represents the difference between a hyper-graphical structure and a simplicial complex [1]. In the former, a k-body interaction does not comprise of lower-order components, whereas in the latter, every lower order relationship must exist to define a higher order one i.e. a $3-$order triangle relationship requires all the edges of the triangle to be included.

Within the lens of irreversibility, we note that the decomposition proposed by Lynn et al [25, 26] specifically considers the unique contributions to the global irreversibility. Alternatively, in the manuscript, we present a method that captures the irreversibility of path projected into the portion of state-space defined by a tuple,

$$\varsigma^{(x_1,\ldots,x_k)} = \sum_{\Gamma^{(x_1,\ldots,x_k)}} P(\Gamma^{(x_1,\ldots,x_k)}) \log \frac{P(\Gamma^{(x_1,\ldots,x_k)})}{P(\Gamma'^{(x_1,\ldots,x_k)})}, \tag{10}$$

which is not equivalent to the unique contribution. In this section, we relate our framework more closely, but still not equivalently, to the decomposition in Ref. [25] by measuring the unique contribution of the k-body interaction to $\varsigma^{(x_1,\ldots,x_k)}$. We do this by recursively subtracting the irreversibility of sub-tuples $\Omega \subset \{x_1, \ldots, x_k\}$, from the quantity $\varsigma^{(x_1,\ldots,x_k)}$. In such a way we define the unique contribution to the $\varsigma^{(x_1,\ldots,x_k)}$ of the $k-$body interaction $(x_1, \ldots, x_k)$ as,

$$\eta^{(x_1,\ldots,x_k)} = \varsigma^{(x_1,\ldots,x_k)} - \sum_{\Omega \subset \{x_1,\ldots,x_k\}} \eta^{\Omega}, \tag{11}$$

which is calculable by noting that $\eta^{(x_i)} = \varsigma^{(x_i)}$. We are able to show that, using this framework, the results are highly correlated, indicating that higher order interactions dominate the irreversibility in these large-scale neural recordings. This stands in contrast with results obtained in spike-train data that indicate that, at the neuronal level, pairwise interactions dominate [25, 26].

We note that this approach captures the unique contributions of k-body interactions by considering the following

$$\varsigma^{(x_i)} = \eta^{(x_i)} \quad \forall i \in \{1, \ldots, N\}, \tag{12}$$

i.e. the combined and unique irreversibility at 1-order is equivalent. Next we note that for two independent variables, the irreversibility factorises, under the DiMViGI framework,

$$\varsigma^{(x_i,x_j)} = \eta^{(x_i)} + \eta^{(x_j)}. \tag{13}$$

For independent variables x_i, x_j, we expect $\eta^{(x_i,x_j)} = 0$. Therefore, it is natural to define,

$$\eta^{(x_i,x_j)} = \varsigma^{(x_i,x_j)} - \eta^{(x_i)} - \eta^{(x_j)}, \tag{14}$$

which is positive for correlated variables and vanishes for independent variables. By definition, it captures the irreversibility of the pairwise interaction, discounting the singleton dynamics. In this fashion, we can recursively calculate the unique contributions at k-order using the unique contributions at j-order for $1 \leq j < k$. Concretely, we have,

$$\eta^{(x_1,\ldots,x_k)} = \varsigma^{(x_1,\ldots,x_k)} - \sum_{\Omega \subset \{x_1,\ldots,x_k\}} \eta^{\Omega}. \tag{15}$$

Figure 9 shows the contrast between the unique and combined irreversibilities for tuples at levels $k = 2, 3, 4$ at the cohort-level. We do not consider 1-order as the unique and combined values are equivalent. Furthermore, we cannot consider $k = 5$ as we employ degree-limiting (see Section 7) for computational efficiency at this level. As a result, we consciously underestimate the irreversibility at 5-order which leads to negative values when inputting these measurements into equation 15. Panel a) of Fig 9 shows a small level of contrast between the unique and combined pairwise dynamics. This indicates that the irreversibility of pairwise interactions dominates the irreversibility of singleton dynamics. Furthermore the general hierarchy is preserved. Panels b-c) show similar results with increasing levels of contrast. However, this increasing contrast is due to the combinatorics of higher order interactions. In particular, as we increase the level k, we are subtracting more terms when isolating the unique contribution. However, this difference is overstated, as panel d) shows that the correlation between unique and combined measurements is almost perfect. This indicates that at a given level k, the k-body interaction dominates the lower level interactions and contributes the most to the irreversibility. This result is both a consequence of the method, and the spatially-coarse, low-dimensional data under consideration. It further suggests that, whilst the DiMViGI framework can be used to compare irreversibility between levels, it is most useful for comparing tuples within a given level.

6 Validation using simulated data from the multivariate Ornstein-Uhlenbeck process

Next, we aim to validate our technique against simulated time-series. We choose the multivariate Ornstein-Uhlenbeck as this is one of the few models that has a known rate of entropy production [14]. Furthermore, this model has been fit to neural recordings in the past in order to estimate the entropy production rate [13].

6.1 Multivariate Ornstein-Uhlenbeck process

The Orstein-Uhlenbeck process models the velocity of a particle in Brownian motion [36]. In its generalised multivariate form, we consider N particles with coupled stochastic dynamics given by the

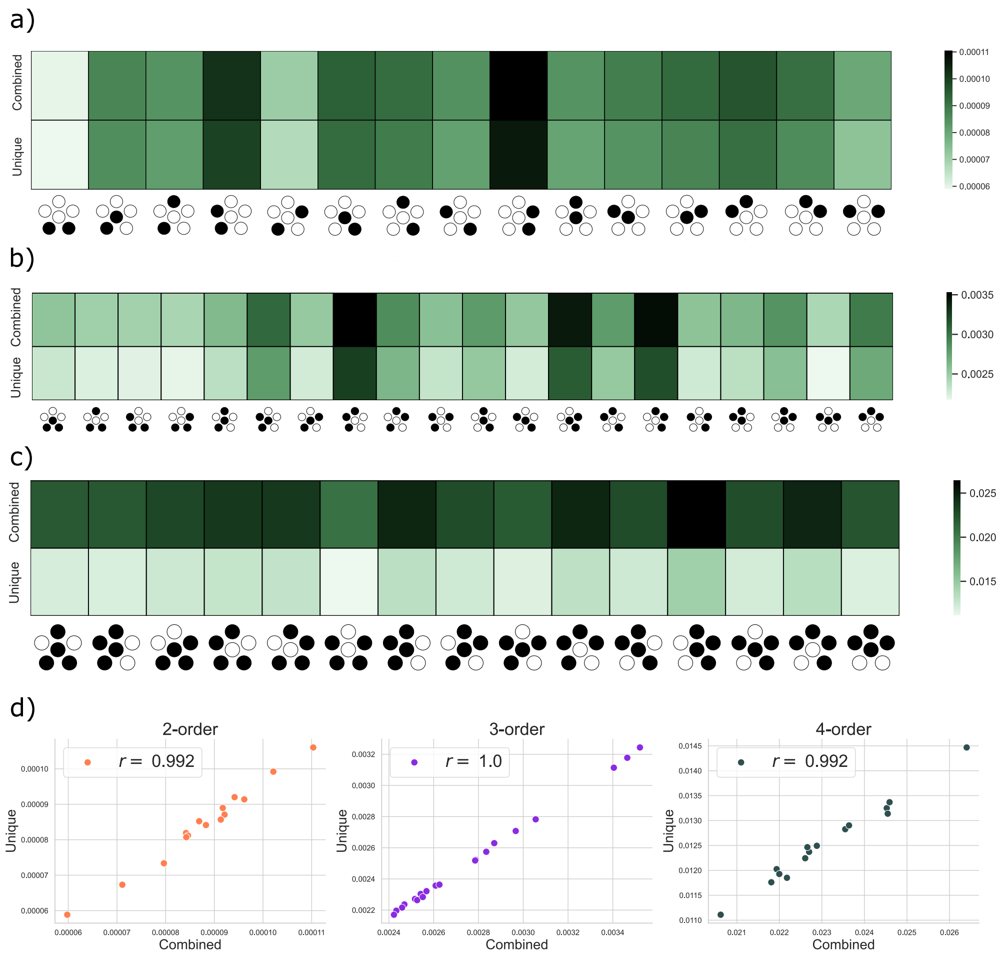

Figure 9: Comparison of the combined k-order contributions to irreversibility against the unique contributions from each k-tuple for $k = 2, 3, 4$. a) $k = 2$. As the irreversibility of the pairwise interactions is much larger than the individual trajectories, the unique and combined irreversibilities are very similar. b) $k = 3$. Whilst there is some contrast between the unique and combined irreversibilities, the general hierarchy is preserved. c) $k = 4$. Again there is some contrast between the unique and combined irreversibilities with the general hierarchy being preserved. d) We show the almost perfect correlation between the unique and combined irreversibilities at each level. This indicates that the $k-$order interactions dominate the irreversibility at level k and suggest little difference when considering unique or combined irreversibilities.

equation,

$$dX(t) = -\boldsymbol{B}X(t)\;\, dt + \boldsymbol{\eta}(t), \tag{16}$$

where $X(t) \in \mathbb{R}^N$. The friction $-\boldsymbol{B} \in \mathbb{R}^{N\times N}$ is a stable matrix, meaning that every eigenvalue has strictly negative real part. The additive noise, $\boldsymbol{\eta}(t)$, is Gaussian and has covariance $\boldsymbol{D} \in \mathbb{R}^{N\times N}$,

$$\langle \boldsymbol{\eta}(t)\boldsymbol{\eta}^\top(t')\rangle = 2\boldsymbol{D}\delta(t-t'). \tag{17}$$

$\boldsymbol{D}$ is a symmetric, positive definite matrix, and so we can calculate $\boldsymbol{L}$, its Cholesky decomposition, where $\boldsymbol{L}$ satisfies $\boldsymbol{D} = \boldsymbol{L}\boldsymbol{L}^\top$ and $\boldsymbol{L}$ is lower triangular [14]. As a result, we can write the system as a Langevin equation,

$$dX(t) = -\boldsymbol{B}X(t)\;\, dt + \boldsymbol{L}\, dW(t), \tag{18}$$

where $W(t)$ represents a N-dimensional Wiener process with independent components. The individual trajectories of a mOU are always reversible, yet at the macroscopic level, irreversibility can emerge. The macroscopic process is known to be reversible if $\boldsymbol{BD}$ is symmetric i.e.,

$$\boldsymbol{BD} = \boldsymbol{DB}^\top. \tag{19}$$

Note that $\boldsymbol{D}$ is always symmetric, whereas, in general, $\boldsymbol{B}$ is not. Furthermore, the covariance, $\boldsymbol{S}$, of the stationary state can be defined implicitly in terms of $\boldsymbol{B}$ and $\boldsymbol{D}$ by the Lyapunov equation,

$$\boldsymbol{BS} + \boldsymbol{SB}^\top = 2\boldsymbol{D}. \tag{20}$$

In the case that the process is reversible, we can use the criterion (19) to write $\boldsymbol{S}$ explicitly,

$$\boldsymbol{S} = \boldsymbol{B}^{-1}\boldsymbol{D}. \tag{21}$$

In the case that the process is irreversible, obtaining an explicit form for $\boldsymbol{S}$ is not as simple. Instead, we parameterise the level of asymmetry using the Onsager matrix of kinetic coefficients and a matrix $\boldsymbol{Q}$, that represents the asymmetry,

$$\boldsymbol{L} = \boldsymbol{BS} = \boldsymbol{D} + \boldsymbol{Q}, \tag{22}$$

$$\boldsymbol{L}^\top = \boldsymbol{SB}^\top = \boldsymbol{D} - \boldsymbol{Q}. \tag{23}$$

As shown in [14], the entropy production rate for the multivariate Ornstein-Uhlenbeck process can be written in terms of the matrices $\boldsymbol{B}$, $\boldsymbol{D}$ and $\boldsymbol{Q}$. The rate of entropy production is given by,

$$\Phi = (\boldsymbol{B}^\top\boldsymbol{D}^{-1}\boldsymbol{B}) = -(\boldsymbol{D}^{-1}\boldsymbol{BQ}). \tag{24}$$

Clearly, when the process is reversible, $\boldsymbol{Q} = 0$ and thus $\Phi = 0$. In general, the matrices $\boldsymbol{S}$ and $\boldsymbol{Q}$ cannot be determined in closed form and so Φ does not have a closed form expression. However, in the case $N = 2$ or in the presence of appropriate symmetries in the matrices $\boldsymbol{B}$ and $\boldsymbol{D}$, a closed form expression can be derived for Φ [14].

6.1.1 The case $N = 2$

In the case $N = 2$, the Lyapunov equation (20) has a closed form solution, and therefore the entropy production rate can be explicitly expressed as a function of the entries of $\boldsymbol{B}$ and $\boldsymbol{D}$ [14].

Consider the matrices,

$$\boldsymbol{B} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}; \qquad \boldsymbol{D} = \begin{pmatrix} u & w \\ w & v \end{pmatrix}. \tag{25}$$

In this case, the rate of entropy production is given explicitly by the formula,

$$\Phi = \frac{(cu - bv + (d-a)w)^2}{(a+d)(uv - w^2)}. \tag{26}$$

Clearly we have $\Phi = 0$ if and only if the reversibility criterion,

$$cu - bv + (d-a)w = 0, \tag{27}$$

is satisfied [14].

6.1.2 Cyclic symmetry

Consider the situation where the variables live on a ring with N sites where the dynamics are invariant to translations of the ring. This results in the matrices $\boldsymbol{B}$ and $\boldsymbol{D}$ being circulant, i.e.

$$\boldsymbol{B} = \begin{pmatrix} b_0 & b_1 & \dots & b_{N-1} \\ b_{N-1} & b_0 & \dots & b_{N-2} \\ \vdots & \vdots & \dots & \vdots \\ b_1 & b_2 & \dots & b_0 \end{pmatrix}; \qquad \boldsymbol{D} = \begin{pmatrix} d_0 & d_1 & \dots & d_{N-1} \\ d_{N-1} & d_0 & \dots & d_{N-2} \\ \vdots & \vdots & \dots & \vdots \\ d_1 & d_2 & \dots & d_0 \end{pmatrix}. \tag{28}$$

As $\boldsymbol{D}$ is assumed to be symmetric, this imposes the additional restriction that $d_{N-i} = d_i$. In this case, the rate of entropy production has a closed form expression,

$$\Phi = \sum_{k=0}^{N-1} \frac{(\Im(\tilde{b}_k))^2}{\Re(\tilde{b}_k)}, \tag{29}$$

where $(\tilde{b}_0, ..., \tilde{b}_{N-1})$ is the discrete Fourier transform of the vector $(b_0, ..., b_{N-1})$, $\Im(\cdot)$ represents the imaginary part of a number and $\Re(\cdot)$ represents the real part [14]. Recall that for a circulant matrix, the Fourier modes of $(b_0, ..., b_{N-1})$ coincide with the eigenvectors of $\boldsymbol{B}$.

6.2 Example processes validating the DiMViGI framework

Using the cases where we can calculate the explicit rate of entropy production, such as those detailed above, we can construct example processes and compare the measurements from our technique to the global rate of irreversibility. Figure 10 shows the results of these numerical experiments.

6.2.1 Example 1

We first consider Example 1, a 2-dimensional process with friction and noise given by,

$$\boldsymbol{B} = \begin{pmatrix} 4 & 1 \\ 2 & 1 \end{pmatrix}; \qquad \boldsymbol{D} = \begin{pmatrix} 1 + \frac{2}{2x+1} & 1 \\ 1 & 1 + \frac{2}{2x+1} \end{pmatrix}. \tag{30}$$

This gives rate of entropy production,

$$\Phi = \frac{4x^2}{5(2x+2)}, \tag{31}$$

which vanishes for $x = 0$, corresponding to a reversible process. Furthermore, as x increases from 0, the rate of entropy production grows linearly with x.

We numerically sample paths from this process for values $x = 0, 0.5, 1, ..., 10$ using an Euler-Maruyama scheme. We sample paths of length $T = 500$ with a time-step of $\Delta t = 0.01$ and keep only the last 2000 time-steps of the process, to avoid boundary effects.

As shown in Panel a) Fig. 10, we can see that the first order irreversibility captured by the DiMViGI techniques shows no correlation with the global rate of entropy production. This is because individual trajectories of the mOU are reversible. As a result, what is plotted is numerical error associated with finite trajectories which has no correlation with the parameters or Φ. On the other hand, the second order irreversibility of the pair (x_1, x_2) is capturing the global rate of entropy production as this interaction produces all the entropy in the system. As a result we can see the strong linear correlation between the 2-order irreversibility and Φ.

6.2.2 Example 2

We consider Example 2, a circulant 3-dimensional process with a strong triplet interaction,

$$\boldsymbol{B} = \begin{pmatrix} 1 & a & -a \\ -a & 1 & a \\ a & -a & 1 \end{pmatrix}; \qquad \boldsymbol{D} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}. \tag{32}$$

This gives rate of entropy production,

$$\Phi = 6a^2. \tag{33}$$

For $a = 0.5, 1, ..., 8.5$, we sample paths with the same methods as before and estimate the irreversibility of the interactions in the system.

As shown in Panel b) of Fig. 10, the individual trajectories are again uncorrelated with the global rate as they are reversible. Both the pair and triplet dynamics are strongly correlated with the global rate of entropy production. Whilst we do not know how much each pair contributes to Φ, the circular symmetry of the process suggests the dynamics of pairs should be identical, which we see here.

6.2.3 Example 3

We consider Example 3, a circular 3-dimensional process with only pairwise drift interactions, given by,

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 & a \\ a & 1 & 0 \\ 0 & a & 1 \end{pmatrix}; \qquad \boldsymbol{D} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \tag{34}$$

which, again, gives rate of entropy production,

$$\Phi = 6a^2. \tag{35}$$

Paths are sampled for value $a = 0.1, 0.2, ..., 1.8$. Whilst each component is only coupled to itself and one other component directly in the drift matrix, it is coupled to the entire system via the noise matrix and indirectly via the dynamics of the other components. For example, even though x_2 does not appear in the drift term for x_1, they are correlated through shared noise and via x_3. For this reason, the difference between Example 2 and Example 3 is not extreme. As can be seen in Panel c) of Fig. 10, we get almost identical dynamics of the measure. Whilst, we aim to distinguish between Example 2 and Example 3, by restricting to pairwise or triplet dynamics, we note that the mOU is a linear system that can be decomposed into its pairwise interactions, meaning it cannot produce genuine higher order effects [1]. However, we are restricted in this analysis to this model as the explicit entropy production rate is known.

6.2.4 Example 4

We consider Example 4 which is 4-dimensional and non-circulant. As a result, we no longer have the exact solution for the entropy production rate and must estimate this quantity numerically. Example 4 has drift and covariance,

$$\boldsymbol{B} = \begin{pmatrix} a^2 & a & 0 & 0 \\ a & a^2 & 0 & 0 \\ 0 & 0 & a^2 & a \\ 0 & 0 & a & a^2 \end{pmatrix}; \qquad \boldsymbol{D} = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}. \tag{36}$$

This system is interacting as a 4-dimensional system as it is coupled through the noise dynamics. However, in the drift term we have two subsystems where (x_1, x_2) interact strongly as do (x_3, x_4), but these pairs are drift-wise disjoint. In order to numerically estimate the entropy production rate, we estimate the covariance matrix from the sampled paths,

$$\boldsymbol{S} = \langle X, X^\top \rangle. \tag{37}$$

Next, we can calculate the asymmetric part of the Onsager matrix,

$$\boldsymbol{Q} = \frac{1}{2}\left(\boldsymbol{BS} - \boldsymbol{SB}^\top\right), \tag{38}$$

which can be used to calculate the entropy production rate,

$$\Phi = -(\boldsymbol{D}^{-1}\boldsymbol{BQ}). \tag{39}$$

We sample paths for values $a = 2.5, 2.7, ..., 4.9$, but we do not know how Φ scales with a. Panel d) of Fig. 10 is harder to interpret than for the previous examples, as the numerical approach produces greater variance in the plot. However, we plot least-square regression lines for each tuple. At 1-order, the lines are almost flat, as expected as there should be no correlation between the reversible individual trajectories and the underlying rate of entropy production. At 2-order, we have the most important result, which is that whilst the reversibility of all pairs scales with entropy production, the strongly interacting pairs (x_1, x_2) and (x_3, x_4), the upper two lines, are more irreversible. At 3-order, all the interactions produce almost identical amounts of irreversibility, which is to be expected as each triplet contains a strongly interacting pair and one component from the other pair, leading to a symmetry in the dynamics. Finally, the irreversibly of the quadruplet, the entire system, scales linearly with the underlying entropy production rate.

We note that the mOU is not a truly higher-order system as it is linear and the interactions can be seen as pairwise, but we are restricted to this model as it has a known rate of entropy production and producing continuous dynamics. Other techniques have validated their techniques on chaotic processes [7] or symbolic dynamics i.e. Ising model [24]. However, deterministic chaos and thermodynamic irreversibility are not equivalent. Furthermore, these processes do not allow one to scale the number of variables, nor the level of irreversibility, arbitrarily. On the other hand, by varying the thermodynamic temperature, one can vary the irreversibility of the Ising model, but the visibility graph is designed to capture correlations in continuous rather than binary series yielding this model unsuitable. For this reason, we opt exclusively for the mOU as studied here.

7 Varying the maximum degree in the support of the degree distributions

The DiMViGI framework projects the high-dimensional, continuous state-space of the multivariate time-series into a discrete and low-dimensional representation using the visibility graph, thus reducing the computational cost of calculating information-theoretic quantities [21, 37]. However, the combinatorial complexity of considering every possible tuple in a system can be restrictive. Furthermore, estimating high-dimensional degree distributions can also be computationally demanding in terms of computer memory. A simple method for improving the memory efficiency of the DiMViGI framework is to cap the maximum degree in the support of the degree distribution. The degree distribution of the visibility graph typically decays exponentially as the degree increases [21, 20, 19]. As a result, when limiting the degree, we are removing minimal information. Moreover, a k-dimensional distribution with maximum degree $d_{\max}$ contains $d_{\max}^k$ entries. Therefore, degree-limiting has an exponential reduction in the memory usage of the DiMViGI implementation. In our analysis presented in the main manuscript, we employed degree limiting in the case of $k = 5$, where we enforced $d_{\max} = 75$. In this section, we present a systematic analysis of the effect of degree limiting for each tuple at each level. We implement this limiting by enforcing that if a node has a degree greater than $\tilde{d}_{\max}$, we set its degree to $\tilde{d}_{\max}$ in the distribution.

First, we note that, in practice, the restriction causes us to underestimate the irreversibility of the tuple. However, this is not mathematically guaranteed. For a tuple, $(x_1, ..., x_k)$, we denote the irreversibility with full support to be $\varsigma^{(x_1,...,x_k)}$ and the irreversibility with limited support to be $\tilde{\varsigma}^{(x_1,...,x_k)}$. Therefore, the difference is,

$$\Delta = \varsigma^{(x_1,...,x_k)} - \tilde{\varsigma}^{(x_1,...,x_k)}. \tag{40}$$

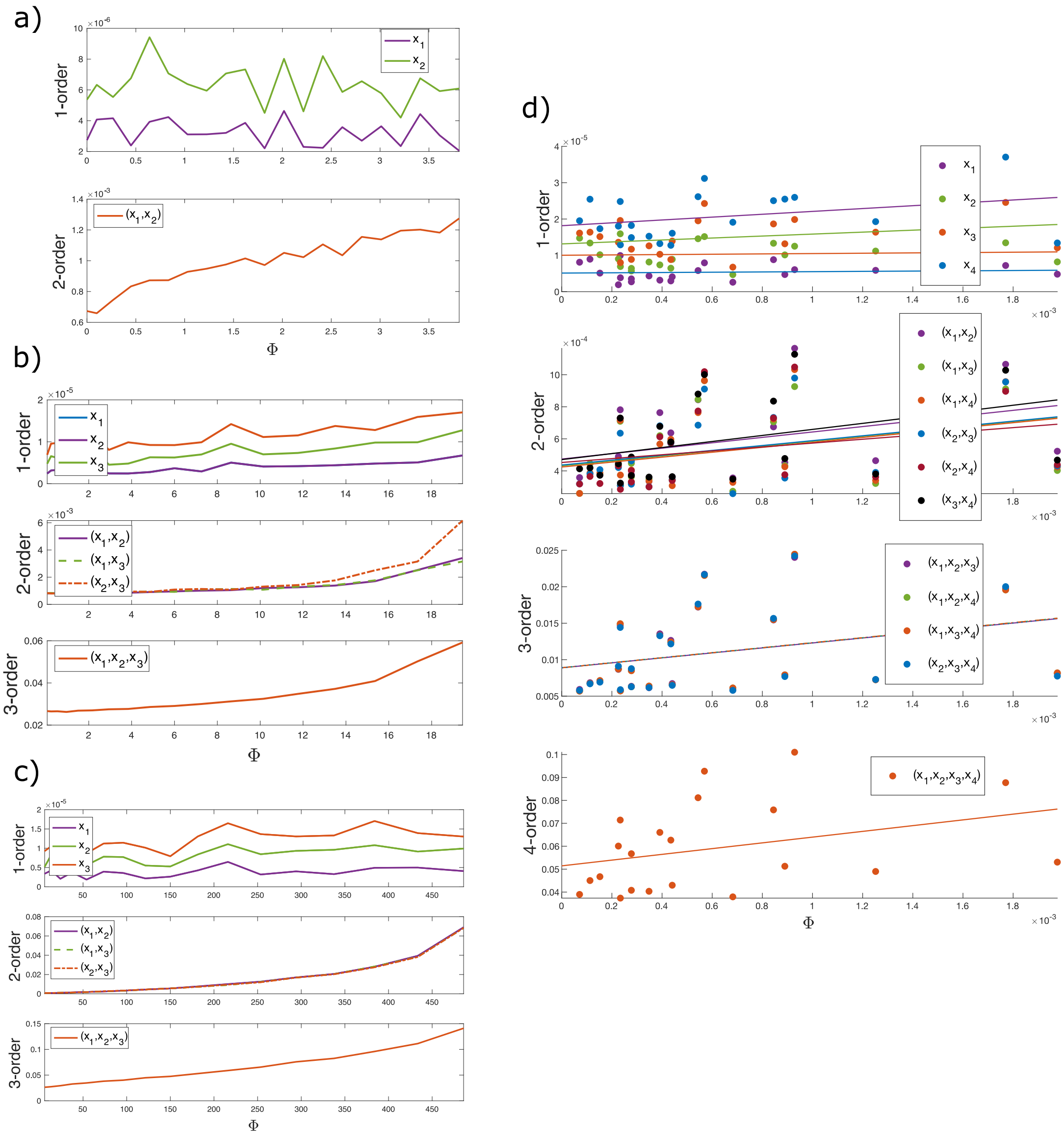

Figure 10: Validation of the DiMViGI framework using simulated data from the mOU. a) Example 1 - a 2-dimensional process. The pairwise irreversibility scales with the global rate Φ, whilst the individual variables do not. b) Example 2 - a 3 dimensional process with drift-disjoint pairs. The pairs and triplet irreversibilities scale with the global rate Φ whilst the individual trajectories do not. c) Example 3 - a 3 dimensional process with 3 way interactions in the drift and noise. The pairs and triplet irreversibilities scale with the global rate Φ whilst the individual trajectories do not. d) Example 4 - a 4 dimensional process with 2 strongly interacting pairs. The global rate Φ is estimated numerically producing variance in the plot so we plot least-square regressions. The pairs and triplets and quadruplet irreversibilities scale with the global rate Φ whilst the individual trajectories do not. Notably, the strongly interacting pairs (x_1, x_2) and (x_3, x_4) have a higher level of irreversibilities than the other pairs.

The sign of Δ reflects whether we are over or underestimating the irreversibility using the limited support. We recall the definition of JSD between distributions P and Q,

$$J(P|Q) = \frac{1}{2}D(P|M) + \frac{1}{2}D(Q|M), \tag{41}$$

where $M = \frac{1}{2}(P+Q)$ is an averaged distribution and $D(\cdot)$ represents the KLD, given by,

$$D(P|Q) = \sum_{x\in\mathcal{X}} P(x)\log\frac{P(x)}{Q(x)}. \tag{42}$$

Therefore,

$$\Delta = \left(\frac{1}{2}\sum_{d_1=\tilde{d}_{\max}+1,\dots,d_k=\tilde{d}_{\max}+1}^{d_{\max}} P_{\text{in}}\log\frac{P_{\text{in}}}{P^*} + P_{\text{out}}\log\frac{P_{\text{out}}}{P^*}\right) \tag{43}$$

$$+\frac{1}{2}\left(\sum_{d_i=\tilde{d}_{\max}\text{ for some } i} P_{\text{in}}\log\frac{P_{\text{in}}}{P^*} + P_{\text{out}}\log\frac{P_{\text{out}}}{P^*} - \tilde{P}_{\text{in}}\log\frac{\tilde{P}_{\text{in}}}{\tilde{P}^*} - \tilde{P}_{\text{out}}\log\frac{\tilde{P}_{\text{out}}}{\tilde{P}^*}\right), \tag{44}$$

where $P_{\text{in}}, P_{\text{out}}$ are the in- and out-degree distributions with the full support; $\tilde{P}_{\text{in}}, \tilde{P}_{\text{out}}$ are the in- and out-degree distributions with the limited support and $P^*, \tilde{P}^*$ are the averaged in-out distributions. In other words, the limited and full degree distributions overlap for all degrees $d < \tilde{d}_{\max}$ and therefore cancel when we take the difference. By truncating the distribution we are neglecting a number of positive terms from the full support. However, we must also consider that these edges have not been simply deleted, but are now included in overestimating the probability that a node has degree $\tilde{d}_{\max}$, hence P and $\tilde{P}$ differ when a node has degree $\tilde{d}_{\max}$ in some layer. We can rewrite Δ as,

$$\Delta = A - \frac{1}{2}\left(\sum_{d_i=\tilde{d}_{\max}\text{ for some } i} \tilde{P}_{\text{in}}\log\frac{\tilde{P}_{\text{in}}}{\tilde{P}^*} + \tilde{P}_{\text{out}}\log\frac{\tilde{P}_{\text{out}}}{\tilde{P}^*} - P_{\text{in}}\log\frac{P_{\text{in}}}{P^*} - P_{\text{out}}\log\frac{P_{\text{out}}}{P^*}\right), \tag{45}$$

where $A \geq 0$. Therefore, if the subtracted sum is less than A, we will underestimate the irreversibility, but if it greater than A we will overestimate the irreversibility. As shown in Figure 11, in practice, we consistently underestimate the irreversibility by limiting the degree, indicating that A is much larger than the subtracted term.

Figure 11 shows a systematic variation of the maximum degree. We perform the analysis with $\tilde{d}_{\max} = \alpha d_{\max}$ and $\alpha \in [0, 1]$, where $d_{\max}$ is the maximum degree of the full multiplex visibility graph (MVG) from the data for $k = 1, \dots, 4$ and $d_{\max} = 75$ for $k = 5$. We vary $\alpha = 0, 0.1, \dots, 0.9, 1$ and calculate the irreversibility of each tuple using the DiMViGI framework but with restricted degree distribution. In addition, we also show, for each value of α, the Pearson correlation coefficient, r, and Spearman's rank, ρ, between the limited and full support values at each level. Panels a-e) show the effect on the irreversibilities of k-tuples with $k = 1, \dots, 5$ respectively and panel f) recalls the schematic representation of the regions of interest. For each level, the irreversibility monotonically increases as we increase α, confirming that degree-limiting underestimates the irreversibility. For lower orders (1-2), we see that the increase is linear. In particular, for the pairwise results, to get a strong correlation with the full support irreverisibility, one needs to use a large proportion of the $d_{\max}$. On the other hand, for the higher orders (3-5), the increase is sigmoidal. Panels c-e) indicate that even limiting to half of the maximum degree is sufficient for an almost perfect correlation with the original results. For order 4-5, we see that this also captures approximately 90% of the irreversibility. With an exponentially smaller distribution, one can capture almost equivalent information. This analysis indicates that degree limiting is a very practical and useful tool to maximise the memory efficiency of the DiMViGI framework at higher orders.

8 Formulations of the entropy production rate for Markovian and non-Markovian systems

As discussed in the main manuscript, the entropy production rate of a system that is out of equilibrium is equal to the information-theoretic evidence for the AoT quantified by,

$$\sigma = \sum_{\Gamma} P(\Gamma) \log \frac{P(\Gamma)}{P(\Gamma')}, \tag{46}$$

where Γ is a trajectory, Γ' is its time-reversal and $P(\Gamma)$ is the 'path-probability', the probability of observing that specific trajectory [34, 18]. Calculating this divergence quantifies the distance from equilibrium [8, 25, 33].

The formulation of the path probability is different in discrete and continuous systems as well as for Markovian and non-Markovian dynamics.

In the case of discrete time, discrete space, Markovian dynamics, the entropy production rate simplifies to,

$$\sigma = \sum_{i,j} P_{ij} \log \frac{P_{ij}}{P_{ji}}, \tag{47}$$

where P_{ij} is the join transition probability, $P(x_{t+1} = j, x_t = i)$ [32, 31].

For $l-$order Markovian dynamics, the rate of entropy production is given by,

$$\sigma = \frac{1}{l} \sum_{x_1,\ldots x_{l+1}} P_{x_1,\ldots,x_{l+1}} \log \frac{P_{x_1,\ldots,x_{l+1}}}{P_{x_{l+1},\ldots,x_1}}, \tag{48}$$

where $P_{x_1,\ldots,x_{l+1}}$ is the probability of observing the exact sequence of states $x_1, \ldots, x_{l+1}$.

For a general system with discrete states in discrete time, the rate of entropy production of a single trajectory is given by,

$$\sigma = \lim_{t\to\infty} \frac{1}{t} \sum_{x_1,\ldots x_{t+1}} P_{x_1,\ldots,x_{t+1}} \log \frac{P_{x_1,\ldots,x_{t+1}}}{P_{x_{t+1},\ldots,x_1}}, \tag{49}$$

where $P_{x_1,\ldots,x_{t+1}}$ is the probability of observing the exact sequence of states $x_1, \ldots, x_{t+1}$.

For continuous time Markovian dynamics, the time-dependent rate of entropy production is given

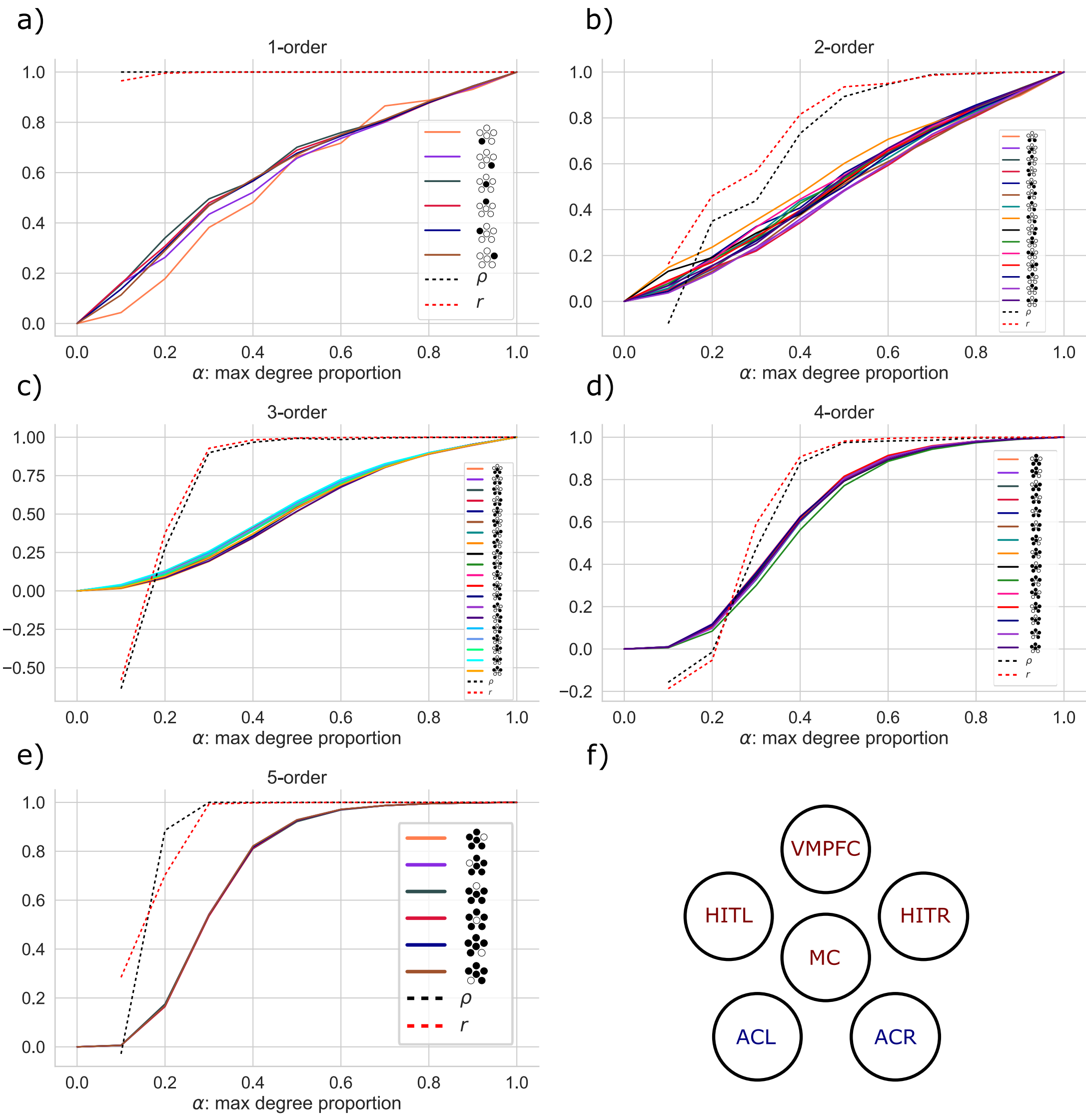

Figure 11: Systematic analysis of the effect of degree-limiting on the irreversibility of each tuple. Panels a-e) show the proportion of irreversibility captured with degree limited to $\tilde{d}_{\text{max}} = \alpha d_{\text{max}}$ for $\alpha \in [0, 1]$. Using a limited degree underestimates the irreversibility of the tuple. In addition a-e) show the correlation between the limited degree results at each level, and the full degree results. For higher orders, degree limiting is shown to lose minimal information, both in terms of the correlations between tuples and absolute values. This indicates that is is a valuable technique for maximising the efficiency of the DiMViGI framework. Panel f) recalls the schematic representation for the tuples in the legends.

by,

$$\sigma(t) = \frac{1}{2} \sum_{i,j} [p_i(t) w_{ij} - p_j(t) w_{ji}] \log \frac{p_i(t) w_{ij}}{p_j(t) w_{ji}}, \tag{50}$$

where $p_i(t)$ is the instantaneous probability distribution and w_{ij} are the transition rates [6].

For Markovian dynamics in continuous space and time, given by the Langevin equation,

$$\dot{\mathbf{x}} = A(\mathbf{x}, t) + B(\mathbf{x}, t) \cdot \eta(t), \tag{51}$$

we can write an equivalent Fokker-Planck equation,

$$\partial_t p(\mathbf{x}, t) = -\nabla \mathbf{j}(\mathbf{x}, t) \tag{52}$$

$$\mathbf{j}_i(\mathbf{x}, t) = A_i(\mathbf{x}, t) p(\mathbf{x}, t) - \sum_j \nabla_j (D_{ij}(\mathbf{x}, t) p(\mathbf{x}, t)), \tag{53}$$

where,

$$D(\mathbf{x}, t) = \frac{1}{2} B(\mathbf{x}, t) B(\mathbf{x}, t)^\top. \tag{54}$$

Following [6], the time-dependent entropy production rate is given by,

$$\sigma(t) = \int \mathbf{x} \; F(\mathbf{x}, \mathbf{t}) \mathbf{j}(\mathbf{x}, t), \tag{55}$$

where,

$$F(\mathbf{x}, t) = \frac{\mathbf{j}^\top(\mathbf{x}, t) D(\mathbf{x}, t)^{-1}}{p(\mathbf{x}, t)}. \tag{56}$$

9 Experimental paradigm and MEG recordings

In this section, we provide additional information about the experimental paradigm, acquisition and pre-processing of the MEG recordings.

9.1 Experimental paradigm

We employed an old/new paradigm auditory recognition task [4, 3, 2, 9, 11, 10]. Participants listened the first four bars of the right-hand part of Johann Sebastian Bachs Prelude No. 2 in C Minor, BWV 847, twice and were asked to memorise it to the best of their ability. Next, participants listened to 135 five-tone musical sequences, corresponding to 27 trials in 5 experimental conditions, of 1750 ms each and were requested to indicate if the sequence belonged to the original music or was a variation. The experimental conditions corresponded to systematic variations on the position of the first varied tone in the sequence. For a detailed description and analysis of the different experimental conditions, see Bonetti et al [4]. We consider one experimental condition, where participants recognised the original, previously memorised sequences.

9.2 Data acquisition

The MEG recordings were taken in a magnetically shielded room at Aarhus Univeristy Hospital (AUH), Aarhus, Denmark on an Elekta Neuromag TRIUX MEG scanner with 306 channels (Elekta Neuromag, Helsinki, Finland). The sampling rate was 1000 Hz with analogue filtering of 0.1-330 Hz. Before taking the recordings, we registered the head shape of participants and the position of four Head Position Indicator (HPI) coils with respect to three anatomical landmarks using a 3D digitiser (Polhemus Fastrak, Colchester, VT, USA). We used this recording to co-register MRI scans with the MEG recordings. During the MEG recordings, the HPI coils continuously registered the localisation of the participant's head which was then used for movement correction. Furthermore, heartbeats and eye-blinks were recorded with two sets of bipolar electrodes which were then used, further along the pre-processing pipeline, to remove artefacts from the MEG recordings. The MRI scans were taken on a CE-approved 3T Siemens MRI-scanner at AUH. The MRI data consisted of structural T1 (mprage with fat saturation) with a spatial resolution of 1.0 x 1.0 x 1.0 mm and the following sequence parameters: echo time (TE) = 2.61 ms, repetition time (TR) = 2300 ms, reconstructed matrix size = 256 x 256, echo spacing = 7.6 ms, bandwidth = 290 Hz/Px. The MRI and MEG recordings were acquired on two separate days.

9.3 MEG data pre-processing

Firstly, the raw MEG sensor data (204 planar gradiometers and 102 magnetometers) was pre-processed using MaxFilter to attenuate external interferences [35]. Next, signal space separation was applied with MaxFilter parameters: spatiotemporal signal space separation [SSS], down-sample from 1000Hz to 250Hz, movement compensation using cHPI coils [default step size: 10 ms], correlation limit between inner and outer subspaces used to reject overlapping intersecting inner/outer signals during spatiotemporal SSS: 0.98). Then the data was converted into Statistical Parametric Mapping (SPM) formatting and further pre-processed in MATLAB (MathWorks, Natick, MA, USA) using in-house-built codes (LBPD, available at `https://github.com/leonardob92/LBPD-1.0.git`) and the Oxford Centre for Human Brain Activity (OHBA) Software Library (OSL) (available at `https://ohba-analysis.github.io/osl-docs/`) [38]. OSL is freely available software that builds on the Fieldtrip [29], FSL [39], and SPM [12] toolboxes. Next the continuous MEG data was visually inspected and large artefacts were removed. This removal discarded less than 0.1% of the data. Independent component analyses (ICA) were used to removed artefacts stemming from heart-beats and eye-blinks [27]. Firstly, the original signal was decomposed into independent components. Next, we isolated and discarded the components that picked up activity from eye-blinks and heartbeats. Then the signal was rebuilt using the remaining components. Lastly, the signal was epoched into 135 trials, 27 trials in 5 experimental conditions, and the mean baseline signal, obtained from the post-stimulus brain signal, was removed. Each resulting trial lasted 4500 ms, made up of 4400 ms plus 100 ms of baseline time.

9.4 Source reconstruction

Whilst MEG recordings have excellent temporal resolution when compared to other imaging modalities, one must employ source-reconstruction to spatially locate activity in the brain. We employed the beam-forming algorithm [15, 16, 5] implemented in both in-house codes and OSL, SPM and FieldTrip.

In the following, we give a thorough description of the inverse model employed in the beam-forming algorithm. The algorithm is made up of two steps: (1) designing a forward model, (2) computing the inverse solution.

The forward model is a theoretical model that considers each brain source as a voxel/active source. The model describes how the strength of each dipole would be reflected onto each of the MEG sensors. We employed magnetometer channels and an 8-mm grid, which returned 3559 dipole locations (voxels) within the whole brain. We co-registered the individual, structural T1 data with the fiducial points and then computed a forward model by adopting the 'Single Shelf' method [28]. This outputs the so-called 'leadfield' model which is an $S \times M$ matrix, L, where S is the number of sources and M is the number of MEG channels. In three cases, the structural T1 was not available and so we performed the leadfield computation with the 'MNI152-T1 with 8-mm spatial resolution' template.

Next, we used the beam-forming algorithm to compute the inverse solution. By sequentially applying a set of weights to the source locations, the algorithm can isolate the contribution of each source to the activity recorded by each MEG channel at each time-point of the recording. We summarise the beamforming algorithm in the following steps.

Firstly, the data recorded by the MEG sensors B at time t is described by the equation,

$$B_t = LQ_t + \nu \tag{57}$$

where L is the leadfield model, Q is the 'dipole matrix' which carries the activity of each dipole over time and ν is noise [16]. The aim is to compute Q by solving the inverse problem. In the beam-forming algorithm, weights are computed and then applied to B_t i.e. for a single dipole, q, we have,

$$q_t = W^\top B_t. \tag{58}$$

Beam-forming computes weights, W_n, for each brain source n using the covariance matrix of the MEG sensors, C, calculated on the continuous signal with all trials concatenated, in the following fashion,

$$W_n = (L_n^\top C^{-1} L_n)^{-1} L_n^\top C^{-1}. \tag{59}$$

Following Nolte [28], the computation of the leadfield method was performed for three orientations of each brain source. Using singular value decomposition (SVD), the three orientations were reduced to one,

$$L = \mathrm{SVD}(\tilde{L}^\top C^{-1} \tilde{L})^{-1}, \tag{60}$$

a common technique for simplifying beam-forming output [17, 23]. Here, $\tilde{L}$ represents the leadfield model with three orientations. Lastly, the obtained weights were applied to each brain source at each time-point and normalised according to Luckhoo et al [23]. In addition to individual trials, the weights were applied to averaged neural activity over all trials. The procedure returned a time-series for each of the 3559 brain sources for each trial, referred to as the 'neural activity index'. The sign ambiguity of the evoked responses time series was adjusted for each brain source using its sign in correspondence with the N100 response to the first tone of the auditory sequences (see Refs. [4, 2, 11, 9]).

Finally, the 3559 voxels obtained through source reconstruction were reduced to six functional brain parcels (or regions of interest (ROIs)) that roughly correspond to auditory cortices in the left and right hemispheres (ACL, ACR); the hippocampal and inferior temporal cortices in the left and right hemispheres (HITL, HITR) and two medial regions, the bilateral medial cingulate gyrus (MC) and the bilateral ventro-medial prefrontal cortex (VMPFC). The data was analysed at a temporal resolution of 4 ms so the resulting multivariate time-series was of dimension 6×1026 (variables $\times$ time-points).

References

[1] Federico Battiston et al. "The physics of higher-order interactions in complex systems". In: *Nature Physics* 17 (2021), pp. 1093–1098.

[2] L Bonetti et al. "Brain recognition of previously learned versus novel temporal sequences: a differential simultaneous processing". In: *Cerebral Cortex* 33.9 (2023), pp. 5524–5537.

[3] L. Bonetti et al. "Challenging age-related decline in brain function: Evidence from fast neuroimaging of musical sequence recognition". In: *bioRxiv* (2022). DOI: 10.1101/2023.07.13.548815. eprint: https://www.biorxiv.org/content/early/2023/07/13/2023.07.13.548815.full.pdf. URL: https://www.biorxiv.org/content/early/2023/07/13/2023.07.13.548815.

[4] L. Bonetti et al. "Revealing the spacetime hierarchical whole-brain dynamics of auditory predictive coding". In: *bioRxiv* (2022). DOI: 10.1101/2022.11.19.517195. eprint: https://www.biorxiv.org/content/early/2022/12/17/2022.11.19.517195.full.pdf. URL: https://www.biorxiv.org/content/early/2022/12/17/2022.11.19.517195.

[5] Matthew J Brookes et al. "Beamformer reconstruction of correlated sources using a modified source model". In: *Neuroimage* 34.4 (2007), pp. 1454–1465.

[6] Andreas Dechant. "Minimum entropy production, detailed balance and Wasserstein distance for continuous-time Markov processes". In: *Journal of Physics A: Mathematical and Theoretical* 55.094001 (2022).

[7] Jonathan F. Donges, Reik V. Donner, and Jurgen Kurths. "Testing time series irreversibility using complex network methods". In: *Europhysics Letters* 102.10004 (2013).

[8] Edward H. Feng and Gavin E. Crooks. "Length of Times Arrow". In: *Physical Review Letters* 101.090602 (2008).

[9] Gemma Fernández-Rubio et al. "Associations between abstract working memory abilities and brain activity underlying long-term recognition of auditory sequences". In: *PNAS Nexus* 1.4 (2022), pp. 1–10.

[10] Gemma Fernández-Rubio et al. "Investigating the impact of age on auditory short-term, long-term, and working memory". In: *Psychology of Music* 0 (2023).

[11] Gemma Fernández-Rubio et al. "Magnetoencephalography recordings reveal the spatiotemporal dynamics of recognition memory for complex versus simple auditory sequences". In: *Communications Biology* 5.1272 (2022).

[12] Karl Friston et al. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press, 2006.

[13] Matthieu Gilson, Enzo Tagliazucchi, and Rodrigo Cofré. "Entropy production of multivariate Ornstein-Uhlenbeck processes correlates with consciousness levels in the human brain". In: *Physical Review E* 107.024121 (2023).

[14] Claude Godrèche and Jean-Marc Luck. "Characterising the nonequilibrium stationary states of OrnsteinUhlenbeck processes". In: *Journal of Physics A: Mathematical and Theoretical* 52.3 (2018).

[15] Arjan Hillebrand and Gareth R. Barnes. "Beamformer Analysis of MEG Data". In: *International Review of Neurobiology* 68 (2005), pp. 149–171.

[16] M X Huang, J C Mosher, and R M Leahy. "A sensor-weighted overlapping-sphere head model and exhaustive head model comparison for MEG". In: *Physics in Medicine and Biology* 44.2 (1999), pp. 423–440.

[17] M X Huang et al. "Commonalities and differences among vectorized beamformers in electromagnetic source imaging". In: *Brain Topography* 16.3 (2004), pp. 139–158.

[18] R. Kawai, J. M. R. Parrondo, and C. Van den Broeck. "Dissipation: The Phase-Space Perspective". In: *Physical Review Letters* 98.080602 (2007).

[19] L. Lacasa and R. Flanagan. "Time reversibility from visibility graphs of nonstationary processes". In: *Physical Review E* 92.022817 (2015).

[20] L. Lacasa et al. "Time series irreversibility: a visibility graph approach". In: *The European Physical Journal B* 85.217 (2012).

[21] Lucas Lacasa et al. "From time series to complex networks: The visibility graph". In: *Proceedings of the National Academy of Sciences of the United States of America* 105.13 (2008), pp. 4972–4975.

[22] Gemma Lancaster et al. "Surrogate data for hypothesis testing of physical systems". In: *Physics Reports* 748 (2018), pp. 1–60.

[23] Henry T Luckhoo, Matthew J Brookes, and Mark W Woolrich. "Multi-session statistics on beamformed MEG data". In: *Neuroimage* 95.100 (2014), pp. 330–335.

[24] Christopher W. Lynn et al. "Broken detailed balance and entropy production in the human brain". In: *Proceedings of the National Academy of Sciences of the United States of America* 118.47 (2021).

[25] Christopher W. Lynn et al. "Decomposing the Local Arrow of Time in Interacting Systems". In: *Physical Review Letters* 129.118101 (2022).

[26] Christopher W. Lynn et al. "Emergence of local irreversibility in complex interacting systems". In: *Physical Review E* 106.034102 (2022).

[27] Dante Mantini et al. "A Signal-Processing Pipeline for Magnetoencephalography Resting-State Networks". In: *Brain Connectivity* 1.1 (2011), pp. 49–59.

[28] Guido Nolte. "The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors". In: *Physics in Medicine and Biology* 48.22 (2003), pp. 3637–3652.

[29] Robert Oostenveld et al. "FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data". In: *Computational Intelligence and Neuroscience* 156869 (2011).

[30] Dean Prichard and James Theiler. "Generating surrogate data for time series with several simultaneously measured variables". In: *Physical Review Letters* 73.951 (1994).

[31] E Roldán and JMR Parrondo. "Entropy production and Kullback-Leibler divergence between stationary trajectories of discrete systems". In: *Physical Review E* 85.031129 (2012).

[32] E Roldán and JMR Parrondo. "Estimating Dissipation from Single Stationary Trajectories". In: *Physical Review Letters* 105.150607 (2010).

[33] Alireza Seif, Mohammad Hafezi, and Christopher Jarzynski. "Machine learning the thermodynamic arrow of time". In: *Nature Physics* 17 (2021), pp. 105–113.

[34] Udo Seifert. "Entropy Production along a Stochastic Trajectory and an Integral Fluctuation Theorem". In: *Physical Review Letters* 95.040602 (2005).

[35] Samu Taulu and Juha Simola. "Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements". In: *Physics in Medicine and Biology* 51 (2010), pp. 1–10.

[36] G Uhlenbeck and L Ornstein. "On the Theory of the Brownian Motion". In: *Physical Review* 36 (823 1930).

[37] Thomas F. Varley and Olaf Sporns. "Network Analysis of Time Series: Novel Approaches to Network Neuroscience". In: *Frontiers in Neuroscience* 15.787068 (2022).

[38] Mark Woolrich et al. "MEG beamforming using Bayesian PCA for adaptive data covariance matrix regularization". In: *Neuroimage* 57.4 (2011), pp. 1466–1479.

[39] Mark W Woolrich et al. "Bayesian analysis of neuroimaging data in FSL". In: *Neuroimage* 45.1 (2009), S173–S186.