

Article

# Evaluating Approximations and Heuristic Measures of Integrated Information

André Sevenius Nilsen <sup>1,\*</sup> , Bjørn Erik Juel <sup>1</sup>  and William Marshall <sup>2,3</sup>

<sup>1</sup> Brain Signalling Group, Department of Physiology, Institute of Basic Medicine, University of Oslo, Sognsvannsveien 9, 0315 Oslo, Norway; b.e.juel@medisin.uio.no

<sup>2</sup> Department of Psychiatry, University of Wisconsin, Madison, WI 53719, USA; wmarshall3@wisc.edu

<sup>3</sup> Department of Mathematics and Statistics, Brock University, St. Catharines, ON L2S 3A1, Canada

\* Correspondence: sevenius.nilsen@gmail.com; Tel.: +47-908-07-044

Received: 8 March 2019; Accepted: 22 May 2019; Published: 24 May 2019



**Abstract:** Integrated information theory (IIT) proposes a measure of integrated information, termed Phi ( $\Phi$ ), to capture the level of consciousness of a physical system in a given state. Unfortunately, calculating  $\Phi$  itself is currently possible only for very small model systems and far from computable for the kinds of system typically associated with consciousness (brains). Here, we considered several proposed heuristic measures and computational approximations, some of which can be applied to larger systems, and tested if they correlate well with  $\Phi$ . While these measures and approximations capture intuitions underlying IIT and some have had success in practical applications, it has not been shown that they actually quantify the type of integrated information specified by the latest version of IIT and, thus, whether they can be used to test the theory. In this study, we evaluated these approximations and heuristic measures considering how well they estimated the  $\Phi$  values of model systems and not on the basis of practical or clinical considerations. To do this, we simulated networks consisting of 3–6 binary linear threshold nodes randomly connected with excitatory and inhibitory connections. For each system, we then constructed the system's state transition probability matrix (TPM) and generated observed data over time from all possible initial conditions. We then calculated  $\Phi$ , approximations to  $\Phi$ , and measures based on state differentiation, coalition entropy, state uniqueness, and integrated information. Our findings suggest that  $\Phi$  can be approximated closely in small binary systems by using one or more of the readily available approximations ( $r > 0.95$ ) but without major reductions in computational demands. Furthermore, the maximum value of  $\Phi$  across states (a state-independent quantity) correlated strongly with measures of signal complexity (LZ,  $r_s = 0.722$ ), decoder-based integrated information ( $\Phi^*$ ,  $r_s = 0.816$ ), and state differentiation (D1,  $r_s = 0.827$ ). These measures could allow for the efficient estimation of a system's capacity for high  $\Phi$  or function as accurate predictors of low- (but not high-)  $\Phi$  systems. While it is uncertain whether the results extend to larger systems or systems with other dynamics, we stress the importance that measures aimed at being practical alternatives to  $\Phi$  be, at a minimum, rigorously tested in an environment where the ground truth can be established.

**Keywords:** integrated information theory; differentiation; integration; complexity; consciousness; computational; IIT; Phi

## 1. Introduction

The nature of consciousness, defined as a subjective experience, has been a philosophical topic for centuries but has only recently become incorporated into mainstream neuroscience [1]. However, as consciousness is a subjective phenomenon, and thus not directly measurable, it must be operationalized to allow for empirical investigation of its nature and underlying mechanisms [2]. In other words,

the scientific study of consciousness requires an objective measure. One such measure has been developed within the framework of the integrated information theory (IIT), introduced and elaborated by Giulio Tononi and colleagues [3–5]. The theory has attracted much interest because of its axiomatic quantitative approach towards illuminating fundamental aspects of consciousness. The theory proposes that consciousness is identical to a particular type of integrated information ( $\Phi$ ;  $\Phi$ ) which is defined and quantified within the theory as a measure of a system's informational irreducibility, or how much information a system in a definite state specifies about its own past and future above and beyond how much such information is specified by its parts.

A major practical limitation of IIT is the computational cost of calculating  $\Phi$ , which, according to the current formulation (version 3.0 [5]; here referred to as  $\Phi_{3.0}$ , implemented through PyPhi [6]), grows as  $O(n53^n)$  [6] for binary systems where  $n$  is the number of elements in the system. In addition, computing  $\Phi_{3.0}$  requires full knowledge of a system's transition probabilities (the probability of the system transitioning from any state to any other state). Taken together, these knowledge and computational requirements place strong constraints on both the system size and the level of possible precision for which  $\Phi_{3.0}$  can be calculated. Therefore, the exact value of  $\Phi_{3.0}$  is intractable for most biological or artificial systems of interest. Currently, the largest systems being investigated are in the order of 20–30 binary elements [7,8], with a practical limit of ~10–12 elements, unless special assumptions are made about the system under investigation (e.g., see [9]).

As  $\Phi_{3.0}$  quickly becomes computationally intractable as a function of network size, one approach is to implement approximations (computational shortcuts) within the framework of IIT<sub>3.0</sub> that reduce the computational cost [6]. Another approach is to use heuristic measures that capture central intuitions of IIT such as information differentiation and integration via more tractable methods [10–15]. While many heuristics have been applied to electrophysiological data (e.g., [10,13,14,16–18]), simulated time series of continuous variables (e.g., [11,19]), and discrete variables (e.g., [15,20]), only [15] have tested a few approximations and heuristics with respect to  $\Phi_{3.0}$  in evolved logic-gate-based animats. Notably, a study [19] compared the behavior of several heuristic measures developed for time-series data; however, the authors were interested in the consistency among the methods, rather than in a comparison with  $\Phi_{3.0}$ .

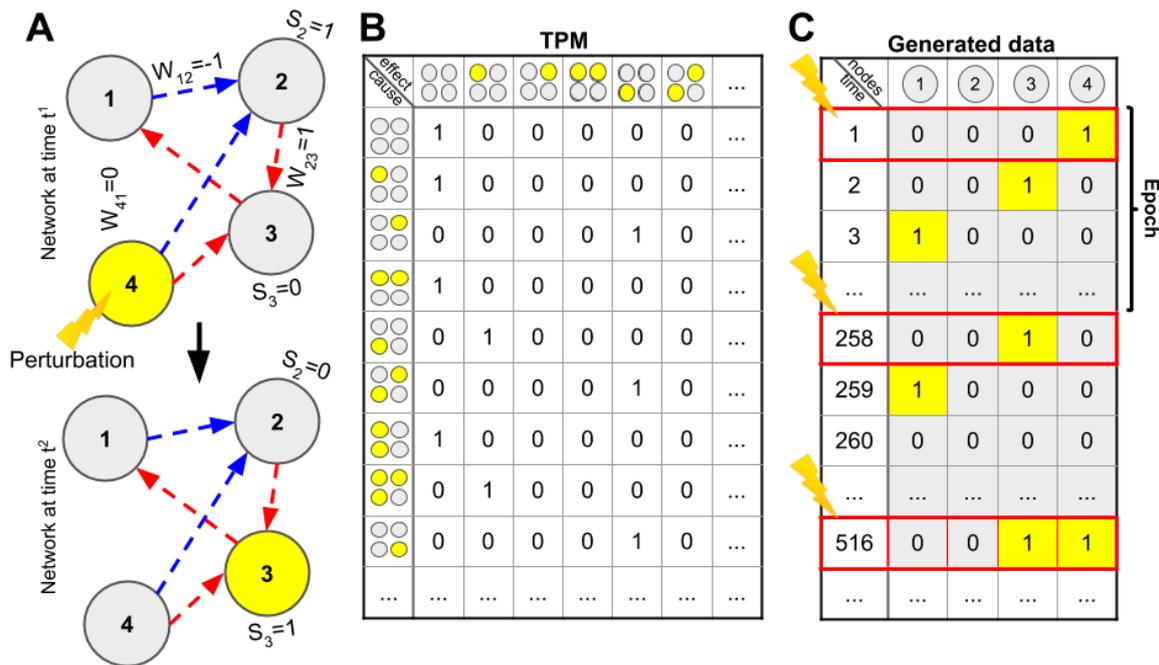
The lack of direct comparisons with  $\Phi_{3.0}$  is a gap in the current literature of integrated information methods. If an approximation or heuristic is to be used in an attempt to falsify IIT, then the results are only valid to the extent that the measure accurately estimates  $\Phi_{3.0}$  (similarly, for evidence in favor of IIT). It is not possible to validate the proposed measures in the networks of interest (due to the computational considerations outlined above); however, we can validate the measures in smaller systems where  $\Phi_{3.0}$  can be calculated directly. We claim that correspondence in smaller systems is a necessary condition for any measure used to evaluate IIT. Therefore, by using deterministic, isolated, discrete networks of binary logic gates of similar type as those employed in IIT<sub>3.0</sub> [5], this paper aims to evaluate the accuracy relative to  $\Phi_{3.0}$  of (1) approximations that speed up parts of  $\Phi_{3.0}$  calculations and (2) heuristic measures of integrated information.

## 2. Materials and Methods

### 2.1. Networks

We randomly generated networks consisting of  $n \in \{3, \dots, 6\}$  binary linear threshold nodes (state  $S \in \{0,1\}$ ), with fixed threshold ( $\theta = 1$ ) and weighted connections between nodes ( $W_{ij} \in \{1,0,-1\}$ , for  $i, j = 1, \dots, n$ ). There were no self-connections ( $W_{ii} = 0$ ). Connections were generated as follows: First, for all  $i \neq j$ , we set  $W_{ij} = 1$  with a probability  $p \in \{0.2, 0.3, \dots, 1.0\}$ , a parameter that was fixed for each network. Second, we changed the sign of non-zero connections to  $W_{ij} = -1$  with probability  $q \in \{0.0, 0.1, \dots, 0.8\}$ ; this parameter was also fixed for each network. The remaining weights were kept at  $W_{ij} = 0$ , i.e., no connection. Altogether, the connections were independent, with  $\Pr(W_{ij} = 1) = p(1 - q)$  and  $\Pr(W_{ij} = -1) = pq$ , and  $\Pr(W_{ij} = 0) = 1 - p$ . To avoid duplicate network architectures,

all networks were checked for uniqueness up to an isomorphism of nodes, i.e., two networks were considered equal if they could be mapped to each other by a relabeling of nodes (using a brute force algorithm). The networks were isolated (no external inputs or modulators). In sum, we generated networks with nodes that could take one of two states ( $S_t = 0, 1$ ) and would be activated ( $S_{t+1} = 1$ ) if the weighted sum of the inputs to the node was equal to or larger than its threshold ( $\theta = 1$ ). If a node was activated, it would then output to other nodes according to its outgoing connection weights. Importantly, this allowed for networks with excitatory ( $W_{ij} = 1$ ), inhibitory ( $W_{ij} = -1$ ), and no ( $W_{ij} = 0$ ) connection between any given pair of nodes (see Figure 1a).



**Figure 1.** (A) Networks were randomly generated with  $n$  binary linear threshold nodes ( $S_i \in \{0, 1\}$ ,  $\theta \geq 1.0$ ) and connections ( $W_{ij} \in \{-1, 0, 1\}$ ). Each network was perturbed into each possible initial state, and the following state transitions were recorded. (B) The networks’ node mechanism and connection weights were used to generate a transition probability matrix (TPM), containing the probability of one state leading to any other state. (C) From the TPM, we generated an “observed” time series using frequent perturbations of the initial states. The sequence of state transitions following an initial state perturbation is termed an epoch.

To investigate various measures and approximations, we needed functional information about the networks in the form of a probabilistic description of the transitions from any given state to any other state, i.e., a transition probability matrix (TPM). For each network, a TPM was constructed based on the node mechanism (linear threshold with  $\theta = 1$ ) and the connection weights  $W_{ij}$ . As the generated networks were deterministic, the TPM contained only a single ‘1’ in each row representing the next state of the network.

From the TPM, given an initial condition, we were able to generate “observed” time-series data for each network. From a given initial condition, a network may only explore part of its state space before reaching an attracting fixed point or periodic sequence. While generating the observed data, we periodically perturbed the network into a new state, ensuring that our data fully explored the state space of the network and that the results were not dependent on our choice of initial condition. This procedure resembles the perturbations applied by transcranial magnetic stimulation (TMS) during empirical studies of consciousness [14]. The generated time-series data consisted of  $2^n$  epochs, where one epoch was generated by initializing/perturbing a network to an initial state and then was simulated for a total of  $\alpha(n)(2^n + 1)$  timesteps. The function  $\alpha(n)$  ensured parity of bits between the generated

time series for networks of different sizes (see Appendix A.1). This perturbation and simulation process was repeated for all possible network states ( $2^n$ ) sequentially, with each epoch appended to the last preceding epoch. The resulting simulated time series (sequence of epochs) produced an  $\alpha(n)(2^n + 1)2^n$ -by- $n$  matrix where each of the  $n$  columns reflected the state of a single node over time, and each row reflected the current state of each network node (0/1) at a given time. In sum, we derived a TPM from the mechanism and connectivity profile of individual nodes and then, using the TPM and perturbations, generated a time series of observed data that explored the entire state space of the network (see Figure 1b,c).

## 2.2. Integrated Information

For the networks defined above, we calculated  $\Phi_{3,0}$  as implemented through PyPhi v1.0 [6]. Here, we just give a brief summary of how  $\Phi_{3,0}$  was defined and calculated, but see reference [5] for a more detailed account. Generally, IIT proposes that a physical system's degree of consciousness is identical to its level of state-dependent causal irreducibility ( $\Phi^{\max}$ ), i.e., the amount of information of a system in a specific state above and beyond the information of the system's parts.

The calculation of  $\Phi_{3,0}$  began with "mechanism-level" computations. For a given *candidate system* (subset of a network) in a state, we identified all possible *mechanisms* (subsets of system nodes in a state that irreducibly constrained the past and future state of the system). For each mechanism, we considered all possible *purviews* (subsets of nodes) that the mechanism constrained. For a given mechanism–purview combination, we found its *cause–effect repertoire* (CER; a probability distribution specifying how the mechanism causally constrained the past and future states of the purview). To find the irreducibility of the CER, the connections between all permissible bipartitions of elements in the purview and the mechanism were *cut* (see [6]); the bipartition producing the least difference is called the *minimum information partition* (MIP). Irreducibility, or integrated information,  $\varphi$ , is quantified by the earth mover's distance (EMD) between the CER of the uncut mechanism and the CER of the mechanism partitioned by the MIP. A mechanism, together with the purview over which its CER is maximally irreducible and the associated  $\varphi$  value, specifies a *concept*, which expresses the causal role played by the mechanism within the system. The set of all concepts is called the *cause–effect structure* of the candidate system.

Once all irreducible mechanisms of a candidate system were found, a similar set of operations was done at the "system level" to understand whether the set of mechanisms specified by the system were reducible to the mechanisms specified by its parts. The irreducibility of the candidate system was quantified by its conceptual integrated information,  $\Phi$ . This process was repeated for all candidate systems, and the candidate system that was maximally irreducible among all candidate systems was termed a *major complex* (MC). According to IIT then, the MC was the substrate that specified a particular conscious experience for the (physical) system in a state, and  $\Phi_{3,0}$  quantified the irreducibility of the cause–effect structure it specified in that state. As such,  $\Phi_{3,0}$  was calculated for every reachable state of the system, i.e., state-dependently.

As many of the heuristics and approximations outlined below are state-independent, there is no direct comparison to the state-dependent  $\Phi_{3,0}$ . To facilitate comparisons with these measures, we further computed a state-independent quantity,  $\Phi_{3,0}^{\text{peak}}$ , as the maximum value of  $\Phi_{3,0}$  across all states of the network. The quantity  $\Phi_{3,0}^{\text{peak}}$  can be thought of as a measure of a network capacity for consciousness, rather than its currently realized level of consciousness. Alternatively, we could also compute the mean value of  $\Phi_{3,0}$ , which has some relation to the state-dependent value of  $\Phi_{3,0}$  under certain regularity conditions [15], but the results were similar (see Figure 5d).

## 2.3. Approximations and Heuristics

To speed up the calculation of  $\Phi_{3,0}$ , one can implement several shortcuts or approximations based on assumptions about the system under consideration. Here, we aimed to test six specific approximations; three approximations that are already implemented in the toolbox for calculating  $\Phi_{3,0}$

(PyPhi; [6]) that reduce the complexity of evaluating information lost during partitioning of a network; two shortcuts based on estimating the elements included in the MC rather than explicitly testing every candidate subsystem; and one estimation of a system's  $\Phi_{3.0}^{peak}$  from the  $\Phi$  of a few states, rather than taking the maximum over all possible states. All approximations were likely to compare well against  $\Phi_{3.0}$ , but were unlikely to yield significant savings in computational demand.

Another approach is to use heuristics that capture aspects of  $\Phi_{3.0}$ . These heuristics can be separated into two classes: those that require the full TPM and discrete dynamics (heuristics on discrete networks requiring perturbational data) and those that require time-series data (heuristics from observed data). While these measures may reduce the computational demands, the heuristics based on discrete dynamics still require full structural and functional knowledge of the system, which reduces their applicability. On the other hand, measures based on observed data significantly broaden the potential applicability at the cost of estimating the underlying causal structure by using the observed time series.

All approximations and heuristics that were tested are listed in Table 1, together with an identifier (from “A” to “N”) that will be used in the text for ease of reading, as well as a reference and brief description.

**Table 1.** Overview of measures.

#	S.D. Measure	S.I. Measure	Description	Ref.
	$\Phi_{3.0}$	$\Phi_{3.0}^{peak}$	Integrated information according to IIT 3.0	[5]
A	CO $\Phi_{3.0}$	CO $\Phi_{3.0}^{peak}$	Cut one connection when making partitions	[6]
B	NN $\Phi_{3.0}$	NN $\Phi_{3.0}^{peak}$	No new concepts after partitioning	[6]
C	WS $\Phi_{3.0}$	WS $\Phi_{3.0}^{peak}$	Whole system as MC	
D	IC $\Phi_{3.0}$	IC $\Phi_{3.0}^{peak}$	Elements with recurrent connections as MC	
E		Est. $n\Phi_{3.0}^{peak}$	Estimate $\Phi_{3.0}^{peak}$ from $n$ states ( $n=1,2,\dots,15$ )	
F	$\Phi_{2.0}$	$\Phi_{2.0}^{peak}$	Integrated information according to IIT 2.0	[3]
G	$\Phi_{2.5}$	$\Phi_{2.5}^{peak}$	$\Phi_{2.0}/\Phi_{3.0}$ hybrid	[12]
H		D1	Reachable states	[15]
I		D2	Cumulative variance of elements	[15]
J		S	Coalition sample entropy	[13]
K		LZ	Functional complexity	[13]
L		$\Phi^*$	Decoder based integrated information	[10]
M		SI	Integrated stochastic interaction	[11]
N		MI	Mutual information	[21]

**Abbreviations:** S.D.: state-dependent; S.I.: state-independent; Ref: reference; IIT: integrated information theory;  $\Phi$ : integrated information;  $\Phi^{peak}$ : maximum  $\Phi$  over system states; CO: cut-one approximation; NN: no-new-concepts approximation; WS: whole-system approximation; MC: major complex; IC: iterative-cut approximation; Est.  $n$ :  $\Phi_{3.0}^{peak}$  estimated from  $n$  sample states; D1/2: state differentiation; S: coalition entropy; LZ: Lempel–Ziv complexity;  $\Phi^*$ : decoder-based  $\Phi$ ; SI: stochastic interaction; MI: mutual information.

### 2.3.1. Approximations to $\Phi_{3.0}$

We calculated several approximations to  $\Phi_{3.0}$ . (A) The cut-one approximation (CO) reduced the number of partitions considered when searching for the MIP. The approximation assumes that the MIP is achieved by cutting only a single node out of the candidate system; (B) the no-new-concepts approximation (NN) eliminates the need to rebuild the entire cause–effect structure for every partition under the assumption that when a partition is made it does not give rise to new concepts. Thus, one only needs to check for changes to existing mechanisms, rather than reevaluating the entire powerset of potential mechanisms.

We also tested two approximations based on estimates of which nodes are included in the MC. These approximations assumed the MC consisted of either (C) all the nodes in the system taken as a whole (whole system; WS), or (D) the subsystem of the network where all nodes with no recursive

connectivity (no input and/or output connections) or an unreachable state (nodes that were always “on” or always “off”, such as a node with only inhibitory inputs) had been removed, iteratively (iterative cut; IC). Note that by unreachable, we mean there was no state of the network that would lead to a particular node being “on” (or “off”) in the next time step. This does not mean that we could not use an external perturbation to set the node into any state (which we did when generating the observed data). In IIT<sub>3,0</sub>, such a node (either with no inputs, no outputs, or an unreachable state) can be partitioned without loss, leading to  $\Phi_{3,0} = 0$ . Simply excluding these nodes from the MC is not an approximation but a computational shortcut, as they will necessarily be outside the MC. However, the approximation consisted in assuming that the remaining set of recursively connected nodes was the MC.

As with  $\Phi_{3,0}$ , these measures were calculated in a state-dependent and state-independent manner. Finally, we tested (E) if the state-independent  $\Phi_{3,0}^{peak}$  could be estimated by randomly sampling the state-dependent  $\Phi_{3,0}$ , termed here “Est. $n\Phi_{3,0}^{peak}$ ”, where  $n$  refers to the number of samples ( $n = 1, 2, \dots, 15$ ).

### 2.3.2. Heuristics on Discrete Networks

To estimate  $\Phi_{3,0}$ , we investigated several heuristic measures defined for discrete networks. While the latest iteration of IIT takes steps to make the mathematical formalism more in tune with the intended interpretation of its axioms and postulates, IIT<sub>3,0</sub> is more computationally intractable than previous versions (see S1 of [5]). To compare the results of the two newest versions of the theory, we tested (F)  $\Phi$  based on IIT<sub>2,0</sub>,  $\Phi_{2,0}$  [3], and (G)  $\Phi_{2,0}$  incorporating minimization over both cause–effect and not only cause,  $\Phi_{2,5}$  [12]. These measures are, however, still limited by the exponential growth in computational time and are included here because IIT<sub>2,0</sub> was used as inspiration for other measures, and their validity depends on the correspondence between IIT<sub>2,0</sub> and IIT<sub>3,0</sub>.

As  $\Phi_{3,0}$  is sensitive to a large state repertoire, i.e., divergent and convergent behavior-weakening cause/effect constraints (assuming irreducibility), we also included two measures that captured the dynamical differentiation of states in the system; (H) The number of reachable states, D1, quantifying the system’s available repertoire of states, and (I) cumulative variance of system elements, D2, indicating the degree of difference between system states [15]. For D1, we calculated the number of states that were reachable, i.e., states that had a valid precursor state. Accordingly, D1 was inversely related to a system’s degeneracy of state transitions. D2 calculated the cumulative variance of activity in each system node given the maximum entropy distribution of initial conditions. As such, D2 reflected how different the system’s reachable states were from each other. See [15] for a more thorough account.

Both  $\Phi_{2,0}$  and  $\Phi_{2,5}$  were calculated in a state-dependent and in a state-independent manner ( $\Phi_{2,0}^{peak}/\Phi_{2,5}^{peak}$ ), while both D1 and D2 were only defined state-independently. All the heuristics on discrete systems were calculated using the system TPM. As such, while these measures were faster to calculate and flexible in terms of network size, they still required full knowledge of the functional dynamics of the system (i.e., the full TPM).

### 2.3.3. Heuristics from Observed Data

To alleviate the full knowledge requirement, we considered heuristic measures that are defined for observed (time-series) data. Given their relative success in distinguishing conscious from unconscious states in experiments and clinical populations [13,22,23] and their apparent similarity to central IIT intuitions, we focused on measures of signal diversity. There are many candidates to choose from, but here, we included (J) coalition entropy ( $S$ ), measured by the entropy of the observed state distribution indicating a system’s average diversity of visited states [22], and (K) signal complexity measured by algorithmic compressibility through Lempel-Ziv compression (LZ), indicating the degree of order or patterns in the observed state sequences of a system [22]. Both entropy and complexity measures have been used in EEG to distinguish between states of consciousness [13,24].

In addition, several measures have been developed that share many of IITs underlying intuitions, such as capturing integrated information of a system above and beyond its parts while staying computationally tractable [10,11,19,21,25]. Although these measures can be applied to continuous data in the time domain such as EEG, here, we focused on a selection of these measures that can be applied to discrete, binary data. Specifically, we tested: (L) decoder-based integrated information ( $\Phi^*$ ) based on IIT<sub>2.0</sub> [21], (M) integrated stochastic interaction (SI) based on IIT<sub>2.0</sub> [11], and (N) mutual information (MI) based on IIT<sub>1.0</sub> [21]. The integrated information measures were implemented using the “Practical PHI toolbox for integrated information analysis” [26] with the discrete forms of the formulae, employing a MIP exhaustive search with a bipartition scheme (powerset;  $2^{n-1}-1$ ) and a normalization factor according to IIT<sub>2.0</sub> [3]. All heuristics were calculated in a state-independent manner, using the time-series data generated for the whole network (no searching through subsystems).

#### 2.4. Analysis

Comparisons between  $\Phi_{3.0}$  and approximate measures (CO, NN, WS, IC) were analyzed using Pearson correlations ( $r$ ) and separate ordinary least-squares linear regression models as the approximations were expected to be closely related to  $\Phi_{3.0}$ . Statistics of linear fits are reported. For comparisons between  $\Phi_{3.0}$  and all other measures we used Spearman’s correlation ( $r_s$ ) to investigate the monotonicity of the relationship, as a linear relationship was not necessarily expected. All state-dependent measures were compared to  $\Phi_{3.0}$ , while all state-independent measures were compared to  $\Phi_{3.0}^{peak}$ . Metrics of significance ( $p$  values) are not reported because of our large sample size; for our sample ( $n > 1981$ ), correlations as small as  $|r| = 0.044$  were statistically significant at the 0.05 level, but such small correlations were not meaningful in the context of the study. As we focused on high correspondence, we instead report correlations as weak,  $0.5 < r < 0.7$ , medium  $0.7 < r < 0.8$ , strong  $0.8 < r < 0.9$ , and very strong,  $r > 0.9$  (for both  $r$  and  $r_s$ ).

#### 2.5. Setup

Calculation of measurements was performed in Python (v3.6) with PyPhi (v1.0) [6] for  $\Phi_{3.0}$ , CO, NN, WS, and IC; Matlab (v2016b) with “Practical PHI toolbox for integrated information analysis” (v1.0) [26] for  $\Phi^*$ , SI, MI; custom code in Python (v3.6) for  $\Phi_{2.0}$ ,  $\Phi_{2.5}$ , D1, D2; and Python (v3.6) with scripts from [13] for LZ, and S. Statistics were done with custom code in Python (v3.6) and Statsmodels (v0.8.0). Everything else was done with custom code in Python (v3.6), Numpy (v1.13.1), SciPy (v0.19.1), and Pandas (v0.20.3).

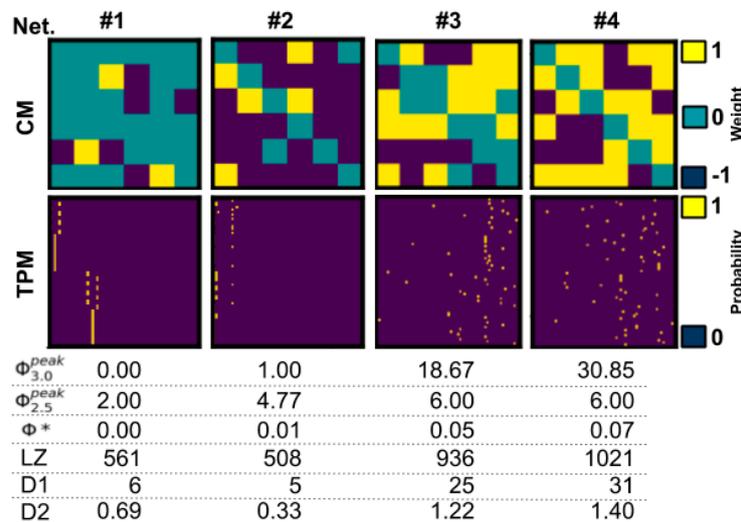
### 3. Results

We analyzed 2032 randomly generated networks, with 131 three-node, 675 four-node, 866 five-node, and 360 six-node networks. In total, 61,224 states were analyzed. Note that the heuristic measures were only analyzed in 309 of the six-node networks due to time constraints. See Table 2 for an overview of the main results and Figure 2 for four example networks.

Table 2. Overview of results.

#	S.D. Measure	<i>r</i>	S.I. Measure	<i>r</i>
	$\Phi_{3.0}$		$\Phi_{3.0}^{peak}$	
A	CO $\Phi_{3.0}$	0.999	CO $\Phi_{3.0}^{peak}$	0.999
B	NN $\Phi_{3.0}$	0.999	NN $\Phi_{3.0}^{peak}$	0.999
C	WS $\Phi_{3.0}$	0.936	WS $\Phi_{3.0}^{peak}$	0.977
D	IC $\Phi_{3.0}$	0.955	IC $\Phi_{3.0}^{peak}$	0.987
E			Est <sub>5</sub> $\Phi_{3.0}$	0.859
F	$\Phi_{2.0}$	0.622	$\Phi_{2.0}^{peak}$	0.838
G	$\Phi_{2.5}$	0.473	$\Phi_{2.5}^{peak}$	0.832
H			D1	0.827
I			D2	0.718
J			S	0.711
K			LZ	0.722
L			$\Phi^*$	0.816
M			SI	0.537
N			MI	0.306

**Abbreviations:** *r*: correlation values, with measures A–F using Pearson’s *r*, and G–O using Spearman’s *r<sub>s</sub>*; S.D.: state-dependent; S.I.: state-independent;  $\Phi$ : integrated information;  $\Phi^{peak}$ : maximum  $\Phi$  over system states; CO: cut-one approximation; NN: no-new-concepts approximation; WS: whole-system approximation; IC: iterative-cut approximation; Est<sub>5</sub>:  $\Phi_{3.0}^{peak}$  estimated from five sample states; D1/2: state differentiation; S: coalition entropy; LZ: Lempel–Ziv complexity;  $\Phi^*$ : decoder-based  $\Phi$ ; SI: stochastic interaction; MI: mutual information.



**Figure 2.** Four example networks with connection matrices (CM) and TPMs, with  $\Phi_{3.0}^{peak}$  and corresponding values for selected state-independent heuristics. Note that network #1 does not consist of a feedforward network if you consider all connections in the CM but is a feedforward network if only excitatory (yellow) connections are considered, which is consistent with  $\Phi_{3.0}^{peak} = 0$ . Network #2 consists of a simple ring-shaped network only if excitatory connections are considered, which is consistent with  $\Phi_{3.0}^{peak} = 1$ .

### 3.1. Descriptive Statistics

Mean and variance of  $\Phi_{3.0}$  grew as a function of network elements ( $n = 3$ :  $M = 0.015 \pm 0.121SD$  to  $n = 6$ :  $M = 0.386 \pm 0.487SD$ ). As the systems increased in size, the fraction of  $\Phi_{3.0}^{peak} = 0$  networks (indicating a completely reducible system, e.g., a feedforward network) decreased. We also monitored a class of networks with  $\Phi_{3.0}^{peak} = 1$ , as this typically indicated that the MC was a stereotyped unidirectional

“loop”. The fraction of these stereotyped networks stayed relatively stable as  $n$  increased, while the fraction of networks with  $\Phi_{3.0}^{peak} > 1$  increased. See Figure 3.

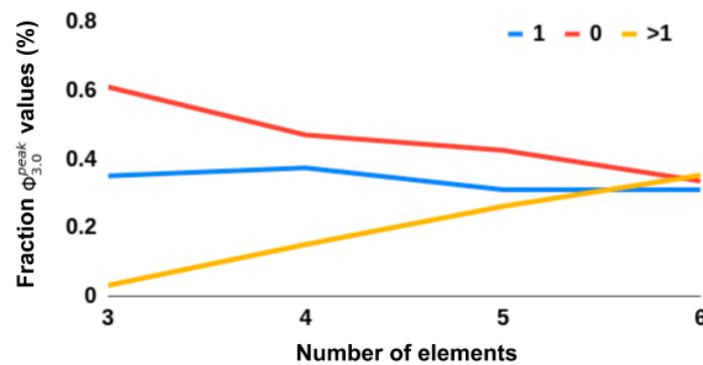


Figure 3. Overview of fraction of networks with  $\Phi_{3.0}^{peak} \in \{1, 0, >1\}$ .

### 3.2. Approximations

Both the no-new-concepts (NN) and the cut-one (CO) approximations were nearly perfectly correlated with state-dependent (S.D.)  $\Phi_{3.0}$  and state-independent (S.I.)  $\Phi_{3.0}^{peak}$  ( $r > 0.996$ ). Regression analysis showed that both no-new-concepts and cut-one approximations were strong linear predictors; S.I.:  $R^2 > 0.999$ ,  $NN\Phi_{3.0}^{peak} = 0.00 + 1.00\Phi_{3.0}^{peak}$ . S.D.:  $R^2 > 0.999$ ,  $NN\Phi_{3.0} = 1.00\Phi_{3.0}$ , and, S.I.:  $R^2 = 0.994$ ,  $CO\Phi_{3.0}^{peak} = 0.00 + 1.04\Phi_{3.0}^{peak}$ . S.D.:  $R^2 = 0.995$ ,  $CO\Phi_{3.0} = 1.02\Phi_{3.0}$ , respectively. See Figure 4a,b.

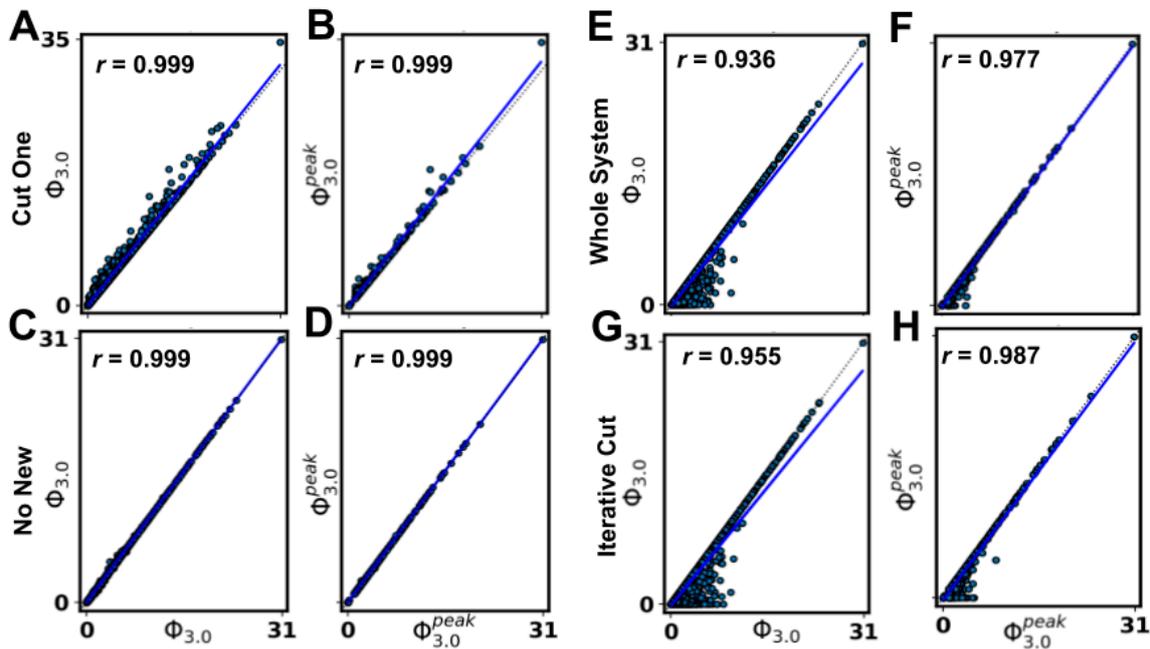
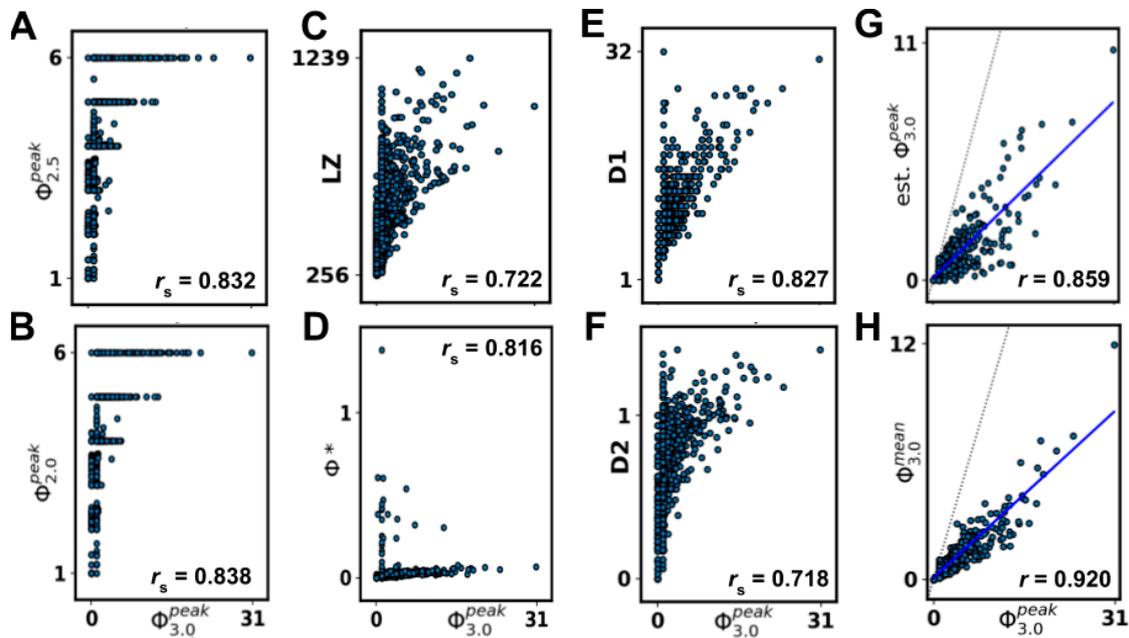


Figure 4. Results of the comparison between  $\Phi_{3.0}$  and approximations, with plotted linear fit (blue) and one-to-one relationship (dotted, gray); (A)  $\Phi_{3.0}$  of the state-dependent CO approximation, (B)  $\Phi_{3.0}^{peak}$  of the state-independent CO, (C)  $\Phi_{3.0}$  of the state-dependent NN approximation, (D)  $\Phi_{3.0}^{peak}$  of the state-independent NN. (E)  $\Phi_{3.0}$  of the state-dependent WS estimated main complex, (F)  $\Phi_{3.0}^{peak}$  of the state-independent WS, (G)  $\Phi_{3.0}$  of the state-dependent IC estimated main complex, (H)  $\Phi_{3.0}^{peak}$  of the state-independent IC.

In regard to estimating  $\Phi_{3.0}^{peak}$ , we took samples from  $n = 1, 2, \dots, 15$  states with results ranging from weak correlation ( $n = 1, r = 0.688$ ) to strong correlation ( $n = 15, r = 0.893$ ) as the number of samples

increased (for  $n = 5$ ;  $R^2 = 0.738$ ,  $SS\Phi_{3.0}^{peak} = 0.097 + 0.262\Phi_{3.0}$ ). This was in accordance with a very strong correlation between  $\Phi_{3.0}^{peak}$  and  $\Phi_{3.0}^{mean}$  ( $R^2 > 0.846$ ,  $\Phi_{3.0}^{mean} = 0.087 + 0.274\Phi_{3.0}^{peak}$ ). These strong correlations suggest that a network with a high value of  $\Phi_{3.0}^{peak}$  typically has several states with high  $\Phi_{3.0}$  values, not just a single state of high  $\Phi_{3.0}$ . See Figure 5g,h.



**Figure 5.** Results of comparison between state-independent  $\Phi_{3.0}^{peak}$  and heuristics and estimates of  $\Phi_{3.0}^{peak}$ . (A)  $\Phi_{2.5}$  modified from  $\Phi_{2.0}$ , (B)  $\Phi_{2.0}$  based on  $IIT_{2.0}$ , (C) LZ complexity (non-normalized), (D) decoder-based  $\Phi$ , based on  $\Phi_{2.0}$ , (E) state differentiation D1, (F) cumulative variance of system elements D, (G) estimated state-independent  $\Phi_{3.0}^{peak}$  using five randomly sampled states (H) state-independent  $\Phi_{3.0}^{mean}$ . G and H are plotted with linear fit (blue) and one-to-one relationship (dotted, gray).

Finally, we tested whether the estimated MCs could predict  $\Phi_{3.0}$ .  $WS\Phi_{3.0}^{peak}$  was very strongly correlated with S.I.  $\Phi_{3.0}^{peak}$  ( $R^2 > 0.954$ , with  $WS\Phi_{3.0}^{peak} = -0.255 + 0.986\Phi_{3.0}^{peak}$ ) and with S.D.  $\Phi_{3.0}$  ( $R^2 > 0.876$ , with  $WS\Phi_{3.0} = -0.163 + 0.899\Phi_{3.0}$ ).  $IC\Phi_{3.0}$  was very strongly correlated with S.I.  $\Phi_{3.0}^{peak}$  ( $R^2 > 0.974$ , with  $IC\Phi_{3.0}^{peak} = -0.167 + 0.995\Phi_{3.0}^{peak}$ ) and very strongly correlated with  $\Phi_{3.0}$  ( $R^2 > 0.912$ , with  $IC\Phi_{3.0} = -0.119 + 0.927\Phi_{3.0}$ ). See Figure 4e–h.

Together, these results suggest that the tested approximations can be used as strong predictors of  $\Phi$ ; however, these approximations still require knowledge of the systems TPM, and their computational cost grows exponentially, leading to only a marginal increase in the size of networks that can be analyzed (see Appendix A.4).

### 3.3. Heuristics

The state differentiation measures D1 and D2 showed strong ( $r_s = 0.827$ ) and medium ( $r_s = 0.718$ ) rank order correlations with S.I.  $\Phi_{3.0}^{peak}$ , respectively (see Figure 5e,f).

S.D.  $\Phi_{2.0}$  and  $\Phi_{2.5}$  were weakly or less correlated with  $\Phi_{3.0}$  ( $r_s = 0.622$  and  $r_s = 0.473$ , respectively), while S.I. variants of  $\Phi_{2.0}$  and  $\Phi_{2.5}$  were strongly rank-order correlated with  $\Phi_{3.0}^{peak}$  ( $r_s = 0.838$  and  $r_s = 0.832$ , respectively) (Figure 5a,b).

The state-independent heuristic LZ and S were medium correlated with  $\Phi_{3.0}^{peak}$  ( $0.71 < r_s < 0.72$ ) (Figure 5c, only LZ shown). The state-independent measures SI and MI were weakly or less correlated with  $\Phi_{3.0}^{peak}$  ( $r_s < 0.54$ ), while  $\Phi^*$  was strongly rank-order correlated with  $\Phi_{3.0}^{peak}$ , ( $r_s = 0.82$ ) (Figure 5d,

only  $\Phi^*$  shown). For  $\Phi^*$ , the results showed two clusters of values, one seemingly linearly related to  $\Phi_{3,0}^{peak}$ , and one non-correlated cluster consisting of low  $\Phi_{3,0}^{peak}$ /high  $\Phi^*$  outliers. A post-hoc analysis removing outliers above two standard deviations of the mean negligibly influenced the results (see Appendix A.2).

Together, these results suggest that the tested heuristics might be accurate predictors of  $\Phi_{3,0}^{peak}$  on a group level however not necessarily for individual networks; they also drastically reduce computational demands (see Appendix A.4). In addition, all heuristics showed an increased variance of  $\Phi_{3,0}^{peak}$  with higher values, suggesting reduced correspondence for higher values.

### 3.4. Post-hoc Tests

For all measures, removing non-integrated ( $\Phi_{3,0}^{peak} = 0$ ) or irreducible circular networks ( $\Phi_{3,0}^{peak} = 1$ ) reduced the correlational values. This was true for all heuristics, while the approximations were minimally affected. After this adjustment, S.I. D1 and  $\Phi^*$  were the heuristics highest correlated with  $\Phi_{3,0}^{peak}$  ( $r_s = 0.703$  and  $r_s = 0.698$ , respectively), with LZ the third ( $r_s = 0.616$ ). This indicates that the results were influenced by a large cluster of non-integrated and circular networks and that the measures were sensitive to the difference between them (see Appendix A.3).

## 4. Discussion

We randomly generated a population of small networks (three to six nodes) with linear threshold logic and both excitatory and inhibitory connections. We evaluated several approximations and heuristic measures of integrated information based on how well they corresponded to the  $\Phi_{3,0}$ , according to the definition proposed by integrated information theory. The purpose of the work was to determine which methods, if any, might be used to test the theory. Since the accuracy of these methods cannot be evaluated for large networks of the size typically of interest for consciousness studies, we considered success in the current study—correspondence in small networks where  $\Phi_{3,0}$  can be computed—as a minimal requirement for any such measure. In summary, we observed that the computational approximations were strong predictors (as defined in Section 2.4) of both  $\Phi_{3,0}$  and  $\Phi_{3,0}^{peak}$ , while heuristic measures were only able to capture  $\Phi_{3,0}^{peak}$ . The approximation measures were still computationally intensive and required full knowledge of the systems TPM, meaning they only provided a marginal increase to the size of the systems that can be studied. Heuristic measures on the other hand, provided greater reductions in computation and knowledge requirements and can be applied to much larger systems, but only in a coarser state-independent manner.

### 4.1. Approximation Measures

The approximation measures we tested were developed by starting from the definition of  $\Phi_{3,0}$  and then making assumptions to simplify the computations. Although they did not reduce computation enough to substantially increase the applicability of  $\Phi_{3,0}$ , their success provides a blueprint for future approximations. We discuss two aspects of  $\Phi_{3,0}$  computation that should be investigated in future work: finding the MC of a network and finding the MIP of a mechanism–purview combination.

Regarding the estimates of the MC, the  $\Phi_{3,0}$  value of any subsystem within a network is a lower bound on the  $\Phi_{3,0}$  of the MC of that network. Moreover, the WS approximation (assuming the MC is the whole system) and the IC approximation (assuming the MC is the whole system after removing nodes without inputs or without outputs and inactive nodes) were both highly predictive of  $\Phi_{3,0}$  (and of  $\Phi_{3,0}^{peak}$ ). Estimating the MC provided computational savings by eliminating the need to compute  $\Phi_{3,0}$  for all possible subsets of elements. However, the computational cost of computing  $\Phi_{3,0}$  for an individual subsystem still grows exponentially with the size of the subsystem. Any MC estimate close to the full size of the network will still require substantial computation. Therefore, finding a minimal MC that still accurately estimates  $\Phi_{3,0}$  would be most efficient for reducing the computational demands. While this may limit the usability of MC estimates (for highly integrated systems, the MC is

more likely to be the whole system), such methods could be used to investigate questions regarding which part of a system is conscious (e.g., cortical location of consciousness [27]).

Using the CO approximation (assuming that at the system level, the MIP results from partitioning a single node), we observed very strong correlations with  $\Phi_{3,0}$  (and  $\Phi_{3,0}^{peak}$ ). Usually, the number of partitions to check grows exponentially with the number of nodes in the system, but with the CO approximation it grew linearly, providing a substantial computational savings. Extending the CO approximation (or some variant of it, see [28–30]) from the system-level MIP to the mechanism-level MIPs could provide even greater computational savings. While only a single system-level MIP needs to be found to compute  $\Phi_{3,0}$ , a mechanism-level MIP must be found for every mechanism–purview combination (the number of which grows exponentially with the system size).

As an aside, the IIT<sub>3,0</sub> formalism only considers bipartitions of nodes when searching for the MIP, presumably on the basis that further partitioning a mechanism (or system) could cause additional information loss (and, thus, never be a *minimum* information partition). To explore this, we employed an alternative definition of the MIP requiring a search over all partitions (AP, as opposed to bipartitions) for a subset of our networks. While we observed a very high correlation between all the partitions and bipartitions schemes (S.I.  $\Phi_{3,0}^{peak}$   $R^2 = 0.966$ ; S.D.  $\Phi_{3,0}$   $R^2 = 0.921$ ; see Appendix A.7), the correspondence was not exact. Note that the definition of a partition used for the ‘all partitions’ option is slightly different than the definition for ‘bipartitions’, so the set of partitions in the AP option is not strictly a superset of the set of bipartitions (see PyPhi v1.0 and its documentation [6] or Appendix A.7 for more details). Despite this difference, we saw a very strong correlation between the methods, suggesting that different rules for permissible cuts could be considered as potential approximations.

#### 4.2. Heuristic Measures

Although heuristic measures did not capture state-dependent  $\Phi_{3,0}$ , most were rank-correlated with state-independent  $\Phi_{3,0}^{peak}$ . However, all heuristic measures were negatively impacted by removing networks with  $\Phi_{3,0}^{peak} = 0||1$ , indicating that reducible ( $\Phi_{3,0}^{peak} = 0$ ) or circular ( $\Phi_{3,0}^{peak} = 1$ ) networks can confound comparisons, as a majority of networks fall in this range. The heuristics that showed the strongest correlation after removal of  $\Phi_{3,0}^{peak} = 0||1$  networks were measures of state differentiation (D1), integrated information ( $\Phi^*$ ), and complexity (LZ). Together, these results suggest that D1,  $\Phi^*$ , and, to a lesser degree, LZ could be useful heuristics for  $\Phi_{3,0}^{peak}$  at the group level, although unreliable at the individual level.

The heuristic D1 measures the number of states accessible by a system [15], and the strong correlation we observed indicates that systems with a large repertoire of available states are also likely to have high  $\Phi_{3,0}^{peak}$  (assuming the systems are irreducible, i.e.,  $\Phi_{3,0}^{peak} > 0$ ). This finding is interesting because clinical results also corroborate state differentiation as a factor in unconsciousness, where it has been observed that the state repertoire of the brain is reduced during anesthesia [31]. While D1 is computationally tractable, it requires full knowledge of the system (i.e., a TPM with  $2^{2n}$  bits of information), that the system is integrated, and that transitions are relatively noise-free. As such, unfortunately, D1 cannot be applied to larger artificial or biological systems of interest (such as the brain). The second measure that correlated well with  $\Phi_{3,0}^{peak}$  can also be seen to quantify state differentiation to some extent. LZ is a measure of signal complexity [32], offering a concrete algorithm to quantify the number of unique patterns in a signal. While LZ has been used to differentiate conscious and unconscious states [13,33], it cannot distinguish between a noisy system and an integrated but complex one from observed data alone. Thus, some knowledge of the structure of the system in question is required for its interpretation. In addition, while LZ allows for analysis of real systems based on time-series data, it is also the measure that is the furthest removed from IIT (but see [14]). It is highly dependent on the size of the input and is hard to interpret without normalization, which makes it difficult to compare systems of varying size. Finally, the measure  $\Phi^*$  is aimed at providing a tractable measure of integrated information using mismatched decoding and is applicable to time-series data,

both discrete and continuous [10].  $\Phi^*$  is relatively fast to compute and can also be applied to continuous time series like EEG. However, while we observed a high correlation with  $\Phi_{3,0}^{peak}$ , a cluster of high  $\Phi^*$  values with corresponding low  $\Phi_{3,0}^{peak}$  values limited the interpretation. This suggests that  $\Phi^*$  might not be reliable for low  $\Phi_{3,0}^{peak}$  networks, but the analysis of larger networks is needed to draw a conclusion. While the results did not suggest a clear tractable alternative to  $\Phi_{3,0}$ , several of the measures could be useful in statistical comparisons of groups of networks.

Prior work directly comparing  $\Phi_{3,0}$  with measures of differentiation (e.g., D1, LZ) reported lower correlations than those observed here for  $\Phi_{3,0}$  [15]. There are at least three possible reasons for this: (a) the current work considered only linear nodes instead of nodes implementing general logic, (b) we compared against  $\Phi_{3,0}^{peak}$  and not  $\Phi_{3,0}^{mean}$ , and (c) we considered only the whole system as a basis for the heuristics, and not the subset of elements that constitutes the MC. For (b), we reran the analysis replacing  $\Phi_{3,0}^{peak}$  with  $\Phi_{3,0}^{mean}$ , producing negligible deviances in the results (see Appendix A.5). For (c), the results of the WS (whole-system approximation) suggested that using the whole system to approximate the MC does not make a substantial difference (at least for networks of this size). This leaves (a), the types of network studied, as the likely reason for the differences in the strength of the correlations.

All heuristic measures' rank correlations with  $\Phi_{3,0}^{peak}$  were negatively impacted by removing networks with  $\Phi_{3,0}^{peak} = 0||1$ . This suggests that such networks are indeed relevant to consider and that finding a tractable measure that separates  $\Phi_{3,0}^{peak} = 0$  and  $\Phi_{3,0}^{peak} \geq 0$  networks would be useful in its own right. Evident in the results was that all heuristics, except S, SI, and MI, showed an inverse predictability with  $\Phi_{3,0}^{peak}$ , i.e., low scores on a given heuristic corresponded to a low score on  $\Phi_{3,0}^{peak}$ , but the higher the scores, the larger the spread of  $\Phi_{3,0}^{peak}$  (see Figure 5). This could explain why the correlations drop when removing networks with  $\Phi_{3,0}^{peak} = 0||1$ . This inverse predictability indicates two things. First, that the tested measures could be useful as negative markers, that is, low scores on measures can indicate low  $\Phi_{3,0}^{peak}$  networks, but not the converse. Secondly, it suggests  $\Phi_{3,0}^{peak}$  has dependencies on aspects of the underlying network that are not captured by any of the heuristic measures.

### 4.3. Future Outlook

Finally, we discuss several topics that we consider to be relevant for future work. First, there are several conceptual aspects of  $\Phi_{3,0}$  that are worth considering when developing future methods. *Composition*: One of the major changes in  $\text{IIT}_{3,0}$  from previous iterations of the theory is the role of all possible mechanisms (subsets of nodes) in the integration of the system as a whole. To our knowledge, all existing heuristic measures of integrated information are wholistic, always looking at the system as a whole. Future heuristics could take a compositional approach, combining integration values from subsets of measurements, rather than using all measurements at once. *State dependence*: We report that heuristic measures do not correlate with state-dependent  $\Phi_{3,0}$  (see Appendix A.6 for a perturbation-based approach), but a more accurate statement is that there are no (data-based) state-dependent heuristics; the nature of heuristic measures does not naturally accommodate state-dependence. *Cut directionality*:  $\Phi_{3,0}$  uses unidirectional cuts, i.e., separating one directed connection, while other heuristics use bidirectional cuts ( $\Phi_{2,0}$ ,  $\Phi_{2,5}$ ) or even total cuts, separating system elements ( $\Phi^*$ , SI, MI). This leads, in effect, to an overestimation of integrated information, even for feedforward and ring-shaped networks (see Figure 2). This could potentially partially explain the inverse predictability noted above.

Secondly, there are differences in the data used for the different measures. Only the approximations (and D1/D2/ $\Phi_{2,0}$ / $\Phi_{2,5}$ ) were calculated on the full TPM, the other heuristics were calculated on the basis of the generated time-series data. However, while deterministic networks such as those considered here can be fully described by both time-series data and TPM, given that the system was initialized to all possible states at least once, data from deterministic systems might be "insufficient" as a time

series, as they often converge on a few cyclical states and, as such, need to be regularly perturbed. One solution to this could be to add noise to the system to avoid fixed points. In addition, as all heuristics considered here (except D1/D2/ $\Phi_{2.0}$ / $\Phi_{2.5}$ ) were dependent on the size of the generated time series (see Appendix A.1), future work should control for the number of samples and discuss the impact of non-self-sustainable activity (convergence on a set of attractor states).

Thirdly, studies comparing measures of information integration, differentiation, and complexity, have also observed both qualitative and quantitative differences between the measures, even for simple systems [19,20]. Thus, there might be a large number of networks where the tested heuristics would correspond to  $\Phi_{3.0}$  if only certain prerequisites are met, such as a certain degree of irreducibility or small-worldness. One could, for example, imagine systems that have evolved to become highly integrated through interacting with an environment [34]. Such evolved networks might have further qualities than being integrated, such as state differentiation that serves distinctive roles for the system, i.e., differences that make a behavioral difference to an organism, which is an important concept in IIT (although considered from an internal perspective in the theory) [5]. While it is still an open question what  $\Phi_{3.0}$  captures of the underlying network above that of the heuristics considered here, investigation into structural and functional aspects that lead to systems with high  $\Phi_{3.0}$  could point to avenues for developing new measures inspired by IIT. Further, while estimates of the upper bound of  $\Phi_{3.0}$ , given a system size, have been proposed (e.g., see [15]), not much is known about the actual distribution of  $\Phi_{3.0}$  over different network types and topologies. Here, we explored a variety of network topologies, but the system properties, such as weight, noise, thresholds, element types, and so on, were omitted because of the limited scope of the paper. Investigating the relation between such network properties and  $\Phi_{3.0}$  would be an interesting research project moving forward. This could be useful as a testbed for future IIT-inspired measures and be informative about what kind of properties could be important for high  $\Phi_{3.0}$  in biological systems and the properties to aim for in artificial systems to produce “consciousness”.

Finally, there are several approximations and heuristics not included in the present study [11,12,19,28,35–40], some of which are specifically applicable to time-series data [10–12,19,21,28,40]. Accordingly, the present work should not be considered an exhaustive exploration of  $\Phi_{3.0}$  correlates.

**Author Contributions:** Conceptualization, A.S.N.; methodology, A.S.N.; software, A.S.N.; validation, A.S.N.; formal analysis, A.S.N.; investigation, A.S.N., B.E.J., and W.M.; writing—original draft preparation, A.S.N.; writing—review and editing, A.S.N., B.E.J., and W.M.; visualization, A.S.N.; supervision, W.M.

**Funding:** This study was supported by the European Union’s Horizon 2020 research and innovation programme under grant agreement 7202070 (Human Brain Project (HBP)) and the Norwegian Research Council (NRC: 262950/F20 and 214079/F20)

**Acknowledgments:** We appreciate the discussions, theoretical contributions, project administration, and funding acquisition of Johan F. Storm and discussions and consulting with Larissa Albantakis, Benjamin Thürer, and Arnfinn Aamodt.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## Appendix A.

### Appendix A.1. Input Size

For each network  $N$  with  $n \in \{3, 4, 5, 6\}$  elements, we generated an observed time series as a matrix  $A_N$ , consisting of  $n$  columns and  $m$  rows. To cover the full state space of  $N$ , we perturbed each  $N$  into  $2^n$  possible initial conditions  $S_i$ . For each initial condition  $S_i$  we simulated  $2^n + 1$  observations (referred to as an epoch) to ensure that we explored the full behavior of the network. Thus,  $A_N$  was a matrix of at least size  $n \times m(n)$ , where

$$m(n) = (2^n + 1)2^n$$

However, as the LZ compression is dependent on the amount of data to compress, we wanted the size of  $A_N$  to be equal for all  $n$ . Hence, we needed to adjust the number of timesteps that we ran the simulation for, so that the size of  $A_N$  would always be the same as the largest network in the set,  $\check{n}$ . Thus, for the specific case of  $\check{n} = 6$ , the size of  $A_N$  is given by

$$\check{n} \times m(\check{n}) = 6 \times m(6) = 24,960$$

To get the same size of  $A_N$  for a network  $N$  of size  $n \in \{3, 4, 5, 6\}$ , we needed an adjusted number of timesteps  $m'(n) \approx \alpha(n) \times m(n)$  (rounded to the nearest integer) where the adjustment factor  $\alpha(n)$  is given by

$$n \times \alpha(n) \times m(n) = 24960$$

$$\alpha(n) = 24960/n(2^n + 1)2^n$$

For the general case, the shape of  $A_N$  is  $n$  - by -  $m'(n)$  where

$$m'(n) \approx \alpha(n) \times m(n)$$

$$m(n) = (2^n + 1)2^n$$

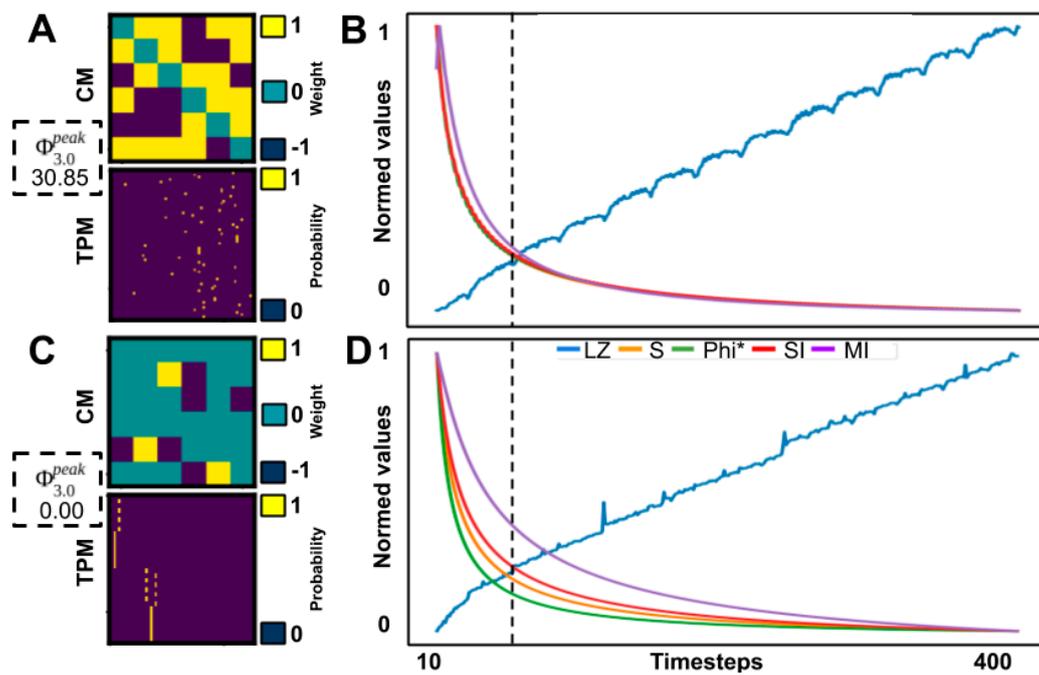
$$\alpha(n) = \check{n}(2^{\check{n}} + 1)2^{\check{n}}/n(2^n + 1)2^n$$

where  $n \in \{a, a+1, \dots, \check{n}\}$ , for some  $a, \check{n} \in \mathbb{N}$ , with  $\check{n} > a$ .

To test the effect of varying the amount of data, i.e., the size of  $A_N$ , we generated data based on two networks with  $n = 6$ : one with high  $\Phi_{3,0}^{peak}$ , and one with low  $\Phi_{3,0}^{peak}$ , with number of timesteps in one epoch,  $E \approx \alpha(n)(2^n + 1) = \{10, 11, \dots, 425\}$ . See Figure A1 for results.

Evident in the results is that all heuristics on generated data were affected by the number of timesteps (i.e., the size of  $A_N$ ). This indicated that various measures were dependent on the amount of data they were calculated on.

However, as we forced each generated time series to have the same size, the networks with fewer elements generated longer time series, i.e., fewer columns required more rows. As this could potentially confound the observed results, we reanalyzed Spearman's  $r_s$  between the heuristics on the observed data and  $\Phi_{3,0}$  for each network size class separately. Except for the heuristic SI, which increased relative to the results presented in the main text, the other measures were less affected (See Table A1).



**Figure A1.** Heuristics on generated time-series data, over varying number of timesteps, for two different six-node networks. (A) a high  $\Phi_{3.0}^{peak}$  network with highly connected CM and complex TPM, (B) a low  $\Phi_{3.0}^{peak}$  network with sparse connected CM and simple TPM, (C) normed values for different heuristics over timesteps between sampled timepoints for network A, (D) normed values for network B. The values for the two networks (B,D) were normalized between 0 and 1. In the original analysis, 64 timesteps were used (dashed line).

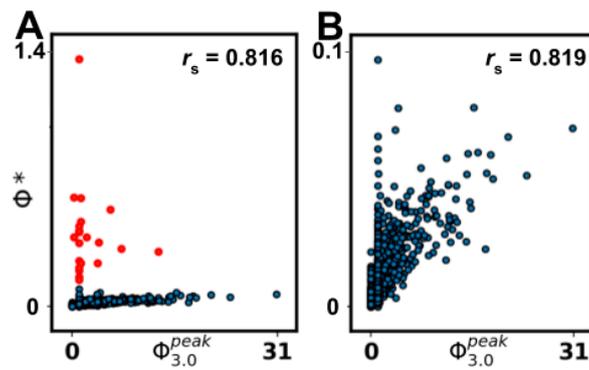
**Table A1.**  $r_s$  between  $\Phi_{3.0}^{peak}$  and heuristics on the observed data, for different network sizes.

$n$	LZ	S	$\Phi^*$	SI	MI
3	0.776	0.747	0.696	0.625	0.032
4	0.799	0.778	0.794	0.717	0.078
5	0.786	0.753	0.825	0.772	0.162
6	0.756	0.668	0.848	0.833	0.276
$\bar{x}$	0.777	0.743	0.791	0.738	0.137

Note: Each correlation value reflects the  $r_s$  between the state-independent heuristic and  $\Phi_{3.0}^{peak}$ .  $n$ : size of the network in number of nodes.

Appendix A.2.  $\Phi^*$  Post-hoc Analysis

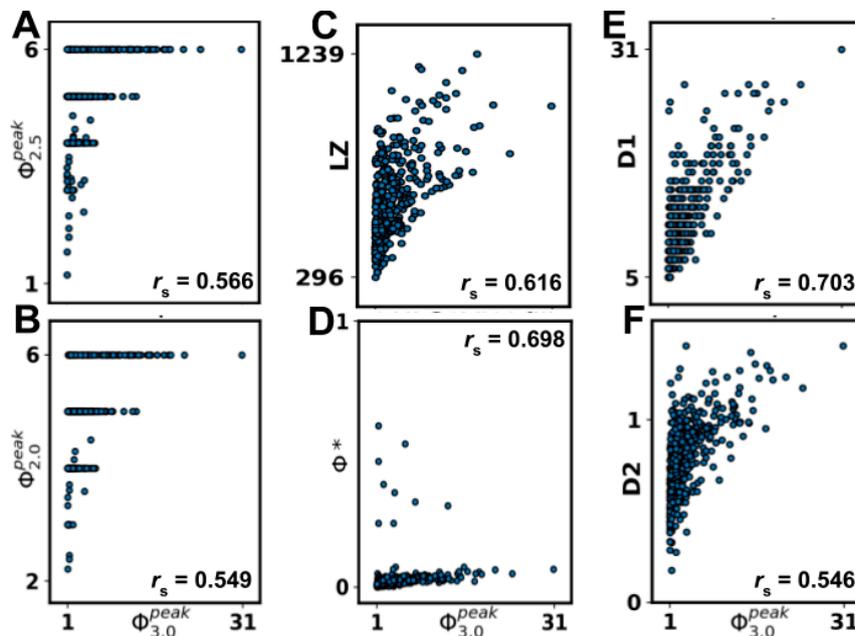
To investigate the distribution of  $\Phi^*$  relative to  $\Phi_{3.0}^{peak}$  after removing a cluster of high  $\Phi^*$  and low  $\Phi_{3.0}^{peak}$  values, we removed the outliers above two standard deviations of the mean. This did not improve the results drastically, as the bulk of observations lay within a narrow band of low  $\Phi^*$  values (see Figure A2).



**Figure A2.** Results of the comparison between state-independent  $\Phi^*$  and  $\Phi_{3,0}^{peak}$  with and without outliers;  $r_s = 0.816$  and  $r_s = 0.819$ , respectively. (A) scatter with outliers > two standard deviations marked in red, (B) scatter with outliers > two standard deviations removed.

*Appendix A.3. Post-hoc Analysis of Networks not Totally Reducible or Reducible to Circular Systems*

Systems that are completely reducible ( $\Phi_{3,0}^{peak} = 0$ ) or reducible to a circular or ring-shaped (sub)network ( $\Phi_{3,0}^{peak} = 1$ ) might not be representative for candidate heuristics, as these networks can be considered “special” cases in terms of  $IIT_{3,0}$ . The absolute difference in correlation values can be seen in Table A2, and the corresponding scatter plots of some select measures are shown in Figure A3. Note that we have here included the bipartitioning versus the all-partitioning comparison (AP) (see Appendix A.7). Most measures dropped in correlational values, while those that increased were low to begin with. Only measures A to D had an  $r > 0.8$ , while measure H and L stayed close to  $r_s = 0.7$ . The other measures had  $r_s < 0.65$ . This suggests that the reported correlational values for most heuristics (F to N) were primarily driven by a cluster of non or trivially integrated networks ( $\Phi_{3,0}^{peak} = 0||1$ ). For measures F to N, Spearman’s rank order correlation was used, Pearson’s correlation otherwise.



**Figure A3.** Results of the comparison between state-independent  $\Phi_{3,0}^{peak}$  and heuristics of  $\Phi_{3,0}^{peak}$  after all networks with  $\Phi_{3,0}^{peak} = 0||1$  were removed; (A)  $\Phi$  based on  $\Phi_{2,5}$ , (B)  $\Phi$  based on  $\Phi_{2,0}$ , (C) LZ complexity (non-normalized), (D)  $\Phi^*$ , (E) state differentiation D1, (F) cumulative variance of system elements D2.

**Table A2.** Difference in the results after removing networks and states with  $\Phi_{3.0}^{peak} = 0||1$ .

#	S.D. Measure	$\Delta r$	S.I. Measure	$\Delta r$
	$\Phi_{3.0}$		$\Phi_{3.0}^{peak}$	
A	CO $\Phi_{3.0}$	-0.004	CO $\Phi_{3.0}^{peak}$	-0.004
B	NN $\Phi_{3.0}$	0.000	NN $\Phi_{3.0}^{peak}$	0.000
C	WS $\Phi_{3.0}$	0.027	WS $\Phi_{3.0}^{peak}$	0.006
D	IC $\Phi_{3.0}$	0.021	IC $\Phi_{3.0}^{peak}$	0.006
E			Est <sub>5</sub> $\Phi_{3.0}$	-0.080
F	$\Phi_{2.0}$	-0.396	$\Phi_{2.0}^{peak}$	-0.289
G	$\Phi_{2.5}$	-0.491	$\Phi_{2.5}^{peak}$	-0.266
H			D1	-0.124
I			D2	-0.172
J			S	-0.405
K			LZ	-0.106
L			$\Phi^*$	-0.118
M			SI	0.063
N			MI	0.010
O	AP $\Phi_{3.0}$	-0.029	AP $\Phi_{3.0}^{peak}$	0.014

**Abbreviations:**  $\Delta r$ : change in correlation values with measures A–F using Pearson’s  $r$ , and G–O using Spearman’s  $r_s$ ; SEst<sub>5</sub>:  $\Phi_{3.0}^{peak}$  estimated from five sample states.

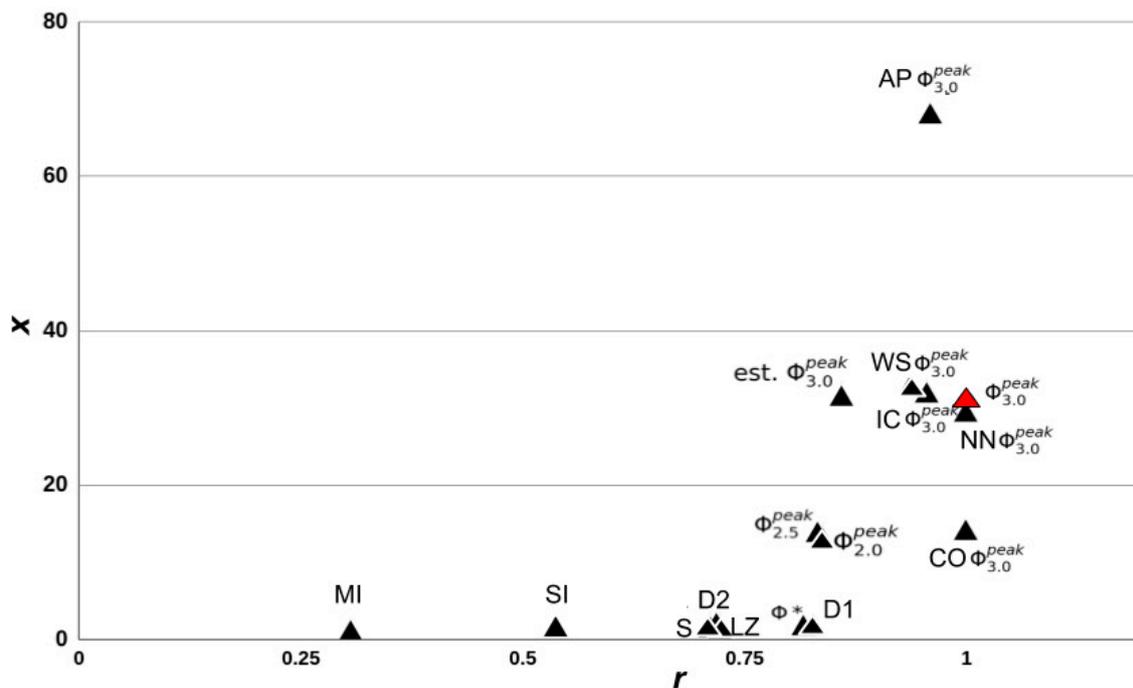
#### Appendix A.4. Estimated Computational Demands

To estimate the computational demands, seven networks of each  $n \in \{3, 4, 5, 6\}$  were randomly generated, with  $p(W_{ij} = 1) \in \{0.7, 0.8, 0.9, 1.0\}$ , and  $p(W_{ij} = -1) \in \{0.3, 0.4, 0.5\}$ . The average times were recorded for each measure, then fitted to a logarithmic regression with reported exponent  $x$ , in the form of  $time = bn^x$ , where  $b$  is a constant, and  $n$  is the system size in nodes. In essence,  $x > 1$  indicates an exponential (more than linear) increase, while  $x < 1$  indicates a less than linear increase. The reported exponents, especially for the measures of  $\Phi$ , were likely underestimated. However, these estimates were highly dependent on underlying computational power, parallelization, efficiency of algorithmic implementation, as well as utilization of shortcuts. As such, the estimated computational demands are guiding at best. Here, we used a 32 gb, 16-core (Intel Xeon E5-1660 v4 @ 3.20 GHz, 20480 KB), parallelized on the level of states for  $\Phi_{2.0/2.5/3.0}$ , at the level of partitions (MIP search) for  $\Phi^*$ , SI, and MI1, and non-parallelized for LZ, S, D1, and D2. See Table A3 for the average time taken to compute the measures (in seconds) for each network size and fitted logarithmic regression, and Figure A4 for an overview of the relationship between computational time and correlation with  $\Phi_{3.0}^{peak}$ . Note that we have here included the all-partitioning (AP) “approximation” (see Appendix A.7).

**Table A3.** Estimated computational demands.

#	Measure	$t(n = 3)$	$t(n = 4)$	$t(n = 5)$	$t(n = 6)$	$x$
	$\Phi_{3.0}^{peak}$	0.40	1.51	102.67	9397.08	31.13
A	CO $\Phi_{3.0}^{peak}$	0.39	1.22	26.61	874.54	13.74
B	NN $\Phi_{3.0}^{peak}$	0.35	1.27	68.61	7070.00	29.06
C	WS $\Phi_{3.0}^{peak}$	0.32	1.41	91.57	8379.66	32.33
D	IC $\Phi_{3.0}^{peak}$	0.31	1.18	74.64	8175.50	31.56
E	Est <sub>5</sub> $\Phi_{3.0}$	0.08	0.30	20.53	1879.41	31.13
F	$\Phi_{2.0}^{peak}$	0.39	1.49	27.60	691.91	12.59
G	$\Phi_{2.5}^{peak}$	0.37	1.90	33.12	850.35	13.47
H	D1	0.00003	0.00003	0.00004	0.0001	1.43
I	D2	0.000	0.001	0.001	0.002	1.64
J	S	0.005	0.005	0.006	0.007	1.14
K	LZ	0.03	0.02	0.02	0.02	0.99
L	$\Phi^*$	0.23	0.29	0.43	0.60	1.38
M	SI	0.20	0.29	0.38	0.38	1.24
N	MI	0.17	0.22	0.18	0.06	0.71
O	AP $\Phi_{3.0}^{peak}$	0.44	3.98	2021.79	-	67.73

**Abbreviations:**  $t(n = i)$ : time in seconds to calculate the relevant measure for a system of size  $n = 3, 4, 5, 6$ ;  $x$ : exponent in a logarithmic regression fit of the form  $time = bn^x$ , where  $n$  is the system size in nodes, and  $b$  is a constant (not reported); Est<sub>5</sub>:  $\Phi_{3.0}^{peak}$  estimated from five sample states.



**Figure A4.** Overview of computational times recorded, for each measure ( $\Phi_{3.0}^{peak}$  marked red), over correlation ( $r/r_s$ ) with  $\Phi_{3.0}^{peak}$ . The y-axis corresponds to the exponent  $x$  of the logarithmic fit between times (in seconds) for networks of size  $n = 3, 4, 5, 6$ , in the form of  $y = bn^x$ , where  $b$  is a constant.

*Appendix A.5. Comparisons versus  $\Phi_{3.0}^{mean}$*

We investigated to what extent replacing  $\Phi_{3.0}^{peak}$  with  $\Phi_{3.0}^{mean}$  (similarly for measures A–D, F, G, O) affected the overall results. Analysis and statistical comparisons were performed as in Sections 2.3 and 2.4. All the approximations and heuristics were negligibly affected, suggesting that for small networks of  $n \in \{3, 4, 5, 6\}$ , the mean and peak state-dependent  $\Phi_{3.0}$  were estimated with similar

accuracy (Table A4). Note that we have here included the bipartitioning versus the all-partitioning (AP) comparison (see Appendix A.7).

**Table A4.** Results after replacing  $\Phi_{3,0}^{peak}$  with  $\Phi_{3,0}^{mean}$ .

#	S.I. Measure	$r$	$\Delta r$
A	CO $\Phi_{3,0}^{mean}$	0.999	.
B	NN $\Phi_{3,0}^{mean}$	0.999	.
C	WS $\Phi_{3,0}^{mean}$	0.947	−0.03
D	IC $\Phi_{3,0}^{mean}$	0.946	−0.041
E	Est <sub>5</sub> $\Phi_{3,0}$	0.922	0.063
F	$\Phi_{2,0}^{mean}$	0.839	0.001
G	$\Phi_{2,5}^{mean}$	0.787	−0.045
H	D1	0.812	−0.015
I	D2	0.774	0.056
J	S	0.744	0.033
K	LZ	0.716	−0.006
L	$\Phi^*$	0.801	−0.015
M	SI	0.499	−0.038
N	MI	0.304	−0.002
O	AP $\Phi_{3,0}^{mean}$	0.979	0.02

**Abbreviations:**  $r$ : correlation values with measures;  $\Delta r$ : change in correlation values; A–F using Pearson’s  $r$ , and G–O using Spearman’s  $r_s$ .

#### Appendix A.6. Initial-state-dependent Heuristics

All heuristics on generated time-series data were calculated in a state-independent manner, using the time-series data generated for the whole network. However, while generating the time-series data, we periodically perturbed the network into a new state, ensuring that our data fully explored the state space of the network. This procedure resembles the perturbations applied by TMS during empirical studies of consciousness [14]. As such, we aimed to test whether the sequence of states following an initial state (i.e., an epoch) can be useful when estimating  $\Phi_{3,0}$ . In other words, we calculated LZ, S,  $\Phi^*$ , SI, and MI on the basis of each of  $2^n$  epochs separately (e.g.,  $t^1$  to  $t^{257}$  in Figure 1c) rather than for all epochs appended together. Since each epoch varied in length based on  $n$ , we correlated the “initial-state-dependent heuristics” with  $\Phi_{3,0}$  for each network size separately. Statistical comparisons were otherwise performed as in Section 2.4. See Table A5 for results. As each epoch was  $\alpha(n)$ -timesteps long, they contained large swathes of cyclical state repetitions (one network could at most visit  $2^n$  unique states before repeating); one should be careful in drawing conclusions from this approach. However, further tests exploring this topic in particular could be informative for the future use of a perturbational approach [14,16].

**Table A5.**  $r_s$  between  $\Phi_{3,0}$  and “initial-state-dependent heuristics” for varying network sizes.

$n$	LZ	S	$\Phi^*$	SI	MI
3	0.391	−0.032	0.146	0.232	0.248
4	0.405	−0.072	−0.126	0.182	0.080
5	0.442	−0.079	−0.101	0.192	0.099
6	0.453	−0.040	−0.092	0.129	0.252
$\bar{x}$	0.423	−0.056	−0.043	0.184	0.169

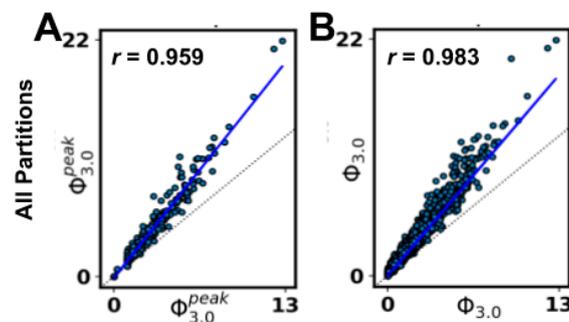
Note: Each correlation value reflects the  $r_s$  between the initial-state-dependent heuristics and  $\Phi_{3,0}$ ;  $n$ : size of the network in number of nodes.

#### Appendix A.7. All Partitions

Bipartitioning (BP) when finding the MIP is the only partitioning scheme considered in IIT<sub>3,0</sub>. However, it is not clear how partitioning a mechanism in more than two parts would affect  $\Phi_{3,0}$ ,

nor how it would be affected by different rules for cutting. While  $IIT_{3.0}$  is defined using BP, a criticism against the theory is that one could use tripartitioning, or more, and that BP should be considered an approximation in its own right with respect to more extensive partitioning schemes. As such, we tested the default BP versus that of all possible partitions (AP) [6] to investigate how well they corresponded (on networks with  $n \in \{3, 4, 5\}$ ). While a superset of BP should result in less than or equal  $\Phi_{3.0}$  due to usually increased information loss with increased number of partitions, the way AP is implemented in PyPhi v1.0 [6] requires that any partition includes at least a mechanism element. As such, AP is not a superset of BP, but the results might be informative in terms of other more expedient partitioning schemes based on other requirements for permissible cuts. Statistical comparisons were performed as defined in Section 2.4.

Bipartitioning was a very strong linear predictor of all partitioning, both with S.I.  $\Phi_{3.0}^{peak}$  ( $R^2 = 0.966$ ,  $AP\Phi_{3.0}^{peak} = -0.134 + 1.438BP\Phi_{3.0}^{peak}$ ) and with S.D.  $\Phi_{3.0}$  ( $R^2 = 0.921$ ,  $AP\Phi_{3.0} = -0.033 + 1.541BP\Phi_{3.0}$ ). We also observed significantly higher  $\Phi_{3.0}$  values for all partitioning, with relative increases of S.D. ( $M = 32.29 \pm 150.11\%$ ,  $t = -5.77$ ,  $p < 0.0001$ ) and S.I.: ( $M = 16.21 \pm 28.60$ ,  $t = -21.16$ ,  $p < 0.0001$ ) (Figure A5).



**Figure A5.** Results of the comparison between  $\Phi_{3.0}$  and approximations, with plotted linear fit (blue) and one-to-one relationship (dotted, gray); (A)  $\Phi_{3.0}^{peak}$  of the state-independent all-partitioning (AP) approximation, (B)  $\Phi_{3.0}$  of the state-dependent AP.

## References

1. Crick, F.; Koch, C. Towards a neurobiological theory of consciousness. *Semin. Neurosci.* **1990**, *2*, 263–275.
2. Chalmers, D.J. Facing up to the problem of consciousness. *J. Conscious. Stud.* **1995**, *2*, 200–219.
3. Balduzzi, D.; Tononi, G. Integrated information in discrete dynamical systems: Motivation and theoretical framework. *PLoS Comput. Biol.* **2008**, *4*, e1000091. [[CrossRef](#)]
4. Tononi, G. An information integration theory of consciousness. *BMC Neurosci.* **2004**, *5*, 42. [[CrossRef](#)] [[PubMed](#)]
5. Oizumi, M.; Albantakis, L.; Tononi, G. From the phenomenology to the mechanisms of consciousness: Integrated information theory 3.0. *PLoS Comput. Biol.* **2014**, *10*, e1003588. [[CrossRef](#)]
6. Mayner, W.G.P.; Marshall, W.; Albantakis, L.; Findlay, G.; Marchman, R.; Tononi, G. PyPhi: A toolbox for integrated information theory. *PLoS Comput. Biol.* **2018**, *14*, e1006343. [[CrossRef](#)] [[PubMed](#)]
7. Marshall, W.; Albantakis, L.; Tononi, G. Black-boxing and cause-effect power. *PLoS Comput. Biol.* **2018**, *14*, e1006114. [[CrossRef](#)] [[PubMed](#)]
8. Marshall, W.; Kim, H.; Walker, S.I.; Tononi, G.; Albantakis, L. How causal analysis can reveal autonomy in models of biological systems. *Philos. Trans. A Math. Phys. Eng. Sci.* **2017**, *375*, 20160358.
9. Albantakis, L.; Tononi, G. The Intrinsic Cause-Effect Power of Discrete Dynamical Systems—From Elementary Cellular Automata to Adapting Animats. *Entropy* **2015**, *17*, 5472–5502. [[CrossRef](#)]
10. Oizumi, M.; Amari, S.-I.; Yanagawa, T.; Fujii, N.; Tsuchiya, N. Measuring Integrated Information from the Decoding Perspective. *PLoS Comput. Biol.* **2016**, *12*, e1004654. [[CrossRef](#)] [[PubMed](#)]
11. Barrett, A.B.; Seth, A.K. Practical measures of integrated information for time-series data. *PLoS Comput. Biol.* **2011**, *7*, e1001052. [[CrossRef](#)] [[PubMed](#)]

12. Tegmark, M. Improved Measures of Integrated Information. *PLoS Comput. Biol.* **2016**, *12*, e1005123. [[CrossRef](#)] [[PubMed](#)]
13. Schartner, M.M.; Seth, A.K.; Noirhomme, Q.; Boly, M.; Bruno, M.-A.; Laureys, S.; Barrett, A. Complexity of Multi-Dimensional Spontaneous EEG Decreases during Propofol Induced General Anaesthesia. *PLoS ONE* **2015**, *10*, e0133532. [[CrossRef](#)] [[PubMed](#)]
14. Casali, A.G.; Gosseries, O.; Rosanova, M.; Boly, M.; Sarasso, S.; Casali, K.R.; Casarotto, S.; Bruno, M.-A.; Laureys, S.; Tononi, G.; Massimini, M. A theoretically based index of consciousness independent of sensory processing and behavior. *Sci. Transl. Med.* **2013**, *5*, 198ra105. [[CrossRef](#)]
15. Marshall, W.; Gomez-Ramirez, J.; Tononi, G. Integrated Information and State Differentiation. *Front. Psychol.* **2016**, *7*, 926. [[CrossRef](#)]
16. Haun, A.M.; Oizumi, M.; Kovach, C.K.; Kawasaki, H.; Oya, H.; Howard, M.A.; Adolphs, R.; Tsuchiya, N. Conscious Perception as Integrated Information Patterns in Human Electroencephalography. *eNeuro* **2017**, *4*. [[CrossRef](#)]
17. Kim, H.; Hudetz, A.G.; Lee, J.; Mashour, G.A.; Lee, U. ReCCognition Study Group Estimating the Integrated Information Measure Phi from High-Density Electroencephalography during States of Consciousness in Humans. *Front. Hum. Neurosci.* **2018**, *12*, 42. [[CrossRef](#)]
18. Hudetz, A.G.; Liu, X.; Pillay, S. Dynamic repertoire of intrinsic brain states is reduced in propofol-induced unconsciousness. *Brain Connect.* **2015**, *5*, 10–22. [[CrossRef](#)]
19. Mediano, P.A.M.; Seth, A.K.; Barrett, A.B. Measuring Integrated Information: Comparison of Candidate Measures in Theory and Simulation. *Entropy* **2018**, *21*, 17. [[CrossRef](#)]
20. Kanwal, M.; Grochow, J.; Ay, N. Comparing information-theoretic measures of complexity in Boltzmann machines. *Entropy* **2017**, *19*, 310. [[CrossRef](#)]
21. Oizumi, M.; Tsuchiya, N.; Amari, S.-I. Unified framework for information integration based on information geometry. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 14817–14822. [[CrossRef](#)] [[PubMed](#)]
22. Ferenets, R.; Lipping, T.; Anier, A.; Jäntti, V.; Melto, S.; Hovilehto, S. Comparison of entropy and complexity measures for the assessment of depth of sedation. *IEEE Trans. Biomed. Eng.* **2006**, *53*, 1067–1077. [[CrossRef](#)]
23. Gosseries, O.; Schnakers, C.; Ledoux, D.; Vanhauzenhuyse, A.; Bruno, M.-A.; Demertzi, A.; Noirhomme, Q.; Lehembre, R.; Damas, P.; Goldman, S.; Peeters, E.; Moonen, G.; Laureys, S. Automated EEG entropy measurements in coma, vegetative state/unresponsive wakefulness syndrome and minimally conscious state. *Funct. Neurol.* **2011**, *26*, 25–30. [[PubMed](#)]
24. Schartner, M.M.; Carhart-Harris, R.L.; Barrett, A.B.; Seth, A.K.; Muthukumaraswamy, S.D. Increased spontaneous MEG signal diversity for psychoactive doses of ketamine, LSD and psilocybin. *Sci. Rep.* **2017**, *7*, 46421. [[CrossRef](#)] [[PubMed](#)]
25. Amari, S.; Tsuchiya, N.; Oizumi, M. Geometry of information integration. *arXiv* **2017**, arXiv:1709.02050.
26. Kitazono, J.; Oizumi, M. Figshare—Practical PHI Toolbox for Integrated Information Analysis. Available online: [https://figshare.com/articles/phi\\_toolbox\\_zip/3203326](https://figshare.com/articles/phi_toolbox_zip/3203326) (accessed on 10 December 2018).
27. Boly, M.; Massimini, M.; Tsuchiya, N.; Postle, B.R.; Koch, C.; Tononi, G. Are the Neural Correlates of Consciousness in the Front or in the Back of the Cerebral Cortex? Clinical and Neuroimaging Evidence. *J. Neurosci.* **2017**, *37*, 9603–9613. [[CrossRef](#)] [[PubMed](#)]
28. Kitazono, J.; Kanai, R.; Oizumi, M. Efficient Algorithms for Searching the Minimum Information Partition in Integrated Information Theory. *Entropy* **2018**, *20*, 173. [[CrossRef](#)]
29. Hidaka, S.; Oizumi, M. Fast and exact search for the partition with minimal information loss. *PLoS ONE* **2018**, *13*, e0201126. [[CrossRef](#)]
30. Arsiwalla, X.D.; Verschure, P.F.M.J. Integrated information for large complex networks. In Proceedings of the 2013 International Joint Conference on Neural Networks (IJCNN 2013), Dallas, TX, USA, 4–9 August 2013; pp. 1–7.
31. Hudetz, A.G.; Mashour, G.A. Disconnecting Consciousness: Is There a Common Anesthetic End Point? *Anesth. Analg.* **2016**, *123*, 1228–1240. [[CrossRef](#)]
32. Lempel, A.; Ziv, J. On the Complexity of Finite Sequences. *IEEE Trans. Inf. Theory* **1976**, *22*, 75–81. [[CrossRef](#)]
33. Schartner, M.M.; Pigorini, A.; Gibbs, S.A.; Arnulfo, G.; Sarasso, S.; Barnett, L.; Nobili, L.; Massimini, M.; Seth, A.K.; Barrett, A.B. Global and local complexity of intracranial EEG decreases during NREM sleep. *Neurosci. Conscious* **2017**, *2017*, niw022. [[CrossRef](#)] [[PubMed](#)]

34. Albantakis, L.; Hintze, A.; Koch, C.; Adami, C.; Tononi, G. Evolution of integrated causal structures in animats exposed to environments of increasing complexity. *PLoS Comput. Biol.* **2014**, *10*, e1003966. [[CrossRef](#)]
35. Toker, D.; Sommer, F. Moving Past the Minimum Information Partition: How To Quickly and Accurately Calculate Integrated Information. *arXiv* **2016**, arXiv:1605.01096 [q-bio.NC].
36. Khajehabdollahi, S.; Abeyasinghe, P.; Owen, A.; Soddu, A. The emergence of integrated information, complexity, and consciousness at criticality. *bioRxiv* **2019**, 521567. [[CrossRef](#)]
37. Esteban, F.J.; Galadí, J.; Langa, J.A.; Portillo, J.R.; Soler-Toscano, F. Informational structures: A dynamical system approach for integrated information. *PLoS Comput. Biol.* **2018**, *14*, e1006154. [[CrossRef](#)]
38. Virmani, M.; Nagaraj, N. A novel perturbation based compression complexity measure for networks. *Heliyon* **2019**, *5*, e01181. [[CrossRef](#)]
39. Toker, D.; Sommer, F.T. Information integration in large brain networks. *PLoS Comput. Biol.* **2019**, *15*, e1006807. [[CrossRef](#)]
40. Mori, H.; Oizumi, M. Information integration in a globally coupled chaotic system. In Proceedings of the 2018 Conference on Artificial Life: A Hybrid of the European Conference on Artificial Life (ECAL) and the International Conference on the Synthesis and Simulation of Living Systems (ALIFE 2018), Cambridge, MA, USA, 23–27 July 2018; pp. 384–385.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).