Accuracy and Performance of the State-based Φ and Liveliness Measures of Information Integration

David Gamez[a] and Igor Aleksander[b]

a. Department of Computing, Imperial College, London SW7 2BT, UK

dgamez@imperial.ac.uk, +44 7790 803 368

b. Department of Electrical and Electronic Engineering, Imperial College, London SW7 2BT, UK

i.aleksander@imperial.ac.uk, +44 7879 842168

NOTICE: this is the author's version of a work that was accepted for publication in Consciousness and Cognition. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version is forthcoming in Consciousness and Cognition (doi:10.1016/j.concog.2011.05.016).

Abstract

A number of people have suggested that there is a link between information integration and consciousness, and a number of algorithms for calculating information integration have been put forward. The most recent of these is Balduzzi and Tononi's state-based Φ algorithm, which has factorial dependencies that severely limit the number of neurons that can be analyzed. To address this issue an alternative state-based measure known as liveliness has been developed, which uses the causal relationships between neurons to identify the areas of maximum information integration. This paper outlines the state-based Φ and liveliness algorithms and sets out a number of test networks that were used to compare their accuracy and performance. The results show that liveliness is a reasonable approximation to state-based Φ for some network topologies, and it has a much more scalable performance than state-based Φ.

Keywords: information integration; consciousness; causation; neural networks, effective connectivity; Φ; liveliness

1. Introduction

Information integration is a property of systems of connected elements that expresses the extent to which they are capable of entering a large number of states that result from causal interactions among their elements (Tononi and Sporns, 2003). For example, a digital camera sensor with a million photodiodes has low information integration because none of its large number of states result from causal interactions among the photodiodes. A system consisting of a million lights controlled by a single switch has a high level of integration between its elements, but a low level of information integration because it can only enter a small number of states (all lights on; all lights off). The mammalian brain is an example of a system with high information integration because it can enter a large number of different states, and each of these states results from causal interactions between the neurons.

A number of people have suggested that there is a link between information integration and consciousness (Metzinger, 2003) or that information integration actually *is* consciousness (Tononi, 2004, 2008), and a number of algorithms for calculating information integration have been put forward. These include neural complexity (Tononi, Sporns and Edelman, 1994),[1] stateless Φ[2] (Tononi and Sporns, 2003), state-based Φ (Balduzzi and Tononi, 2008),[3]

[1] Neural complexity is not explicitly formulated as a measure of information integration – rather, it is claimed to be a measure of the balance between functional specialization at a small scale and global integration at a large scale. However, since its method and aims are very closely aligned with the later work of Tononi and Sporns (2003) and Balduzzi and Tononi (2008), it has been included as a measure of information integration in this paper.

[2] Tononi and Sporns' (2003) use of the symbol Φ is distinct from the many other quantities that Φ has been used to represent. Tononi and Sporns intended the "I" in Φ to represent the information that is integrated within an entity that is represented by the "O" in Φ.

causal density (Seth et al., 2006) and liveliness (Aleksander and Gamez, 2011) – see Seth et al. (2006) for a review of earlier work and a comparison. There have been a limited number of experiments suggesting a link between information integration and consciousness (Lee et al., 2009; Massimini et al., 2009, 2010), and in the longer term it is hoped that one of the measures of information integration could form part of a scientific theory of consciousness that could be used to make accurate predictions about conscious states.

The neural complexity and stateless Φ measures of information integration use equilibrium solutions to the network, which are insensitive to a network's changing weights and ignore alterations in the information integration over time. The causal density measure is based on Granger causality (Granger, 1969), which has the advantage that it works when the causal model of the system is unknown, but its measurement of the system across multiple points in time smears out the information integration and makes the algorithm insensitive to fluctuations within the sampling window. Causal density also has the disadvantage that it does not identify the regions of the system that have the highest 'density' of information integration, since it only provides a single number for the information integration of the entire system.[4] The state-based Φ algorithm put forward by Balduzzi and Tononi (2008) is the most temporally and spatially specific measure that has been put forward so far, since it can identify the regions of maximum information integration for each firing state of the system. Balduzzi and Tononi (2009) have also suggested how the structure of integrated information might be linked to the contents of consciousness, and a method for applying the state-based Φ algorithm to time series data has been put forward by Barrett and Seth (2011). The key limitation of Balduzzi and Tononi's (2008) algorithm is that its factorial dependencies severely limit the number of neurons that can be analyzed (see Section 5.7). To address this problem an alternative measure of information integration known as liveliness has been developed, based on earlier work by Aleksander (1973) and Aleksander and Atlas (1973). When the causal model of the system is known, both liveliness and state-based Φ can be used to measure changes in its information integration over time and make predictions about the moment to moment consciousness of the system.

If information integration is correlated with or claimed to *be* consciousness, then it must be a real property of the physical system, and different methods for measuring information integration should produce similar results. The accuracy of information integration algorithms could be evaluated by making the (problematic) assumption that information integration is correlated with consciousness, and carrying out experiments – for example, using fMRI or

[3] The information integration measure put forward by Tononi and Sporns (2003) will be referred to as "stateless Φ" to distinguish it from the related state-based measure of Φ of Balduzzi and Tononi (2008).

[4] This limitation could be addressed by applying a clustering algorithm to the weights identified by Granger causality, but we are not aware of any work that has done this so far.

EEG – that measure the correlation between the output of the information integration algorithms and reports of conscious states. If information integration is correlated with consciousness, then the algorithms that most accurately predict consciousness would be the most accurate measures of information integration. The main problem with this approach is that some or all of the algorithms might be measuring a property of the brain that is correlated with consciousness, but which has nothing to do with information integration. There is also the issue that our spatial and temporal access to the brain is severely limited, which makes it very difficult to measure information integration in humans. Another way of measuring the accuracy of information integration algorithms is to create simulated networks with well defined regions of high information integration. Different information integration algorithms could be run on these networks and their output compared with the areas of expected maximum information integration. This approach has the problem that our intuitions about the areas of maximum information integration might not be correct, but there does not appear to be a way of measuring the information integration of a network that does not depend on a particular algorithm.

While the simulated networks approach is problematic, until access to the brain improves it is the only method that is available for the comparison of different measures of information integration, and this paper proposes a number of test networks that could be used to benchmark different information integration algorithms. The first part of the paper outlines the state-based Φ and liveliness algorithms, and then a number of test networks are described, which were used to evaluate the accuracy and performance of the two measures. The results of the analyses are presented in Section 5, and the paper concludes with a discussion and some suggestions for testing the link between information integration and consciousness on real neural data.

2. Two State-based Measures of Information Integration

2.1 State-based Φ

Balduzzi and Tononi's (2008) state-based Φ algorithm uses relative entropy to measure the effective information that is generated by a subset of the network when it enters a particular state. The relative entropy, $H[\, p \parallel q]$, between probability distributions p and q is given by Equation 1:

$$H[\, p \parallel q] = \sum_i p_i \log_2 \frac{p_i}{q_i} \,. \tag{1}$$

When a network or subset of the network enters a particular state, x_1, at time t_1 there is a certain probability that each possible state at t_0 led to the current state at t_1. The set of these probabilities is called the *a posteriori* repertoire or $p(X_0 \rightarrow x_1)$. The entering of the network into state x_1 generates information because it defines the *a posteriori* repertoire of probabilities of the possible states that could have existed at t_0 and led to the current state. However, if the network is in a state of maximum entropy and its states are entirely random, then each state x_1 could have been caused by any other state, and the fact that you are in x_1 tells you nothing about the previous state of the network. In this case, the probability that each possible previous state of the network caused the current state is the same, and this set of probabilities is known as the *a priori* repertoire, or $p^{max}(X_0)$. According to Balduzzi and Tononi (2008), the amount of information generated by a particular state can be measured using the relative entropy between the *a posteriori* repertoire associated with x_1 and the *a priori* repertoire, as expressed in Equation 2:

$$ei(X_0 \rightarrow x_1) = H[p(X_0 \rightarrow x_1) \,\|\, p^{\max}(X_0)], \tag{2}$$

where $ei(X_0 \rightarrow x_1)$ is the effective information generated by the state x_1, $p(X_0 \rightarrow x_1)$ is the *a posteriori* repertoire and $p^{max}(X_0)$ is the *a priori* repertoire.

Equation 2 gives the effective information that is generated when the whole network or a subset enters a particular state, but it does not tell us whether this information was generated by causal interactions among the elements, or whether it is the sum of the information generated by the parts of the network acting independently. To answer this question Balduzzi and Tononi (2008) consider partitions of the network and calculate the relative entropy between the *a posteriori* repertoires of the parts considered independently and the *a posteriori* repertoire of the whole subset, as expressed in Equation 3:

$$ei(X_0 \rightarrow x_1 / P) = H\left[p(X_0 \rightarrow x_1) \,\middle\|\, \prod_{M^k \in P} p(M_0^k \rightarrow \mu_1^k) \right], \tag{3}$$

where $ei(X_0 \rightarrow x_1 / P)$ is the effective information of a particular partition, P, of the system into two or more parts, M^k is a part of the system, and μ^k is a state of M^k. To calculate $ei(X_0 \rightarrow x_1 / P)$, each part is considered as a system in its own right and the inputs from the other parts are treated as noise.

The *minimum information partition* is the subset or division of the subset across which the least information is integrated, and it is identified by comparing the normalized effective information values for each possible partition as

well as the normalized effective information for the subset as a whole (known as the total partition, whose effective information is calculated using Equation 2). Normalization is needed during this comparison process because the effective information across a partition between a single element and a number of elements is typically less than an equal bipartition, and the effective information across many partitions is typically higher than the effective information across few partitions.[5] The un-normalized value of effective information for the minimum information partition is the Φ value of the subset, which is calculated for every possible subset of the system.

Balduzzi and Tononi (2008) define a *complex* as a subset that is not included in another subset with higher Φ. According to Balduzzi and Tononi, complexes are regions of the system where neurons integrate the most information, and the Φ value of each complex corresponds to the amount of information that is integrated. The *main* complexes of the system are the complexes whose subsets have strictly lower Φ, and Tononi (2004, 2008) claims that main complexes are the conscious parts of the system. A correlation-based interpretation of information integration would interpret the main complexes as the parts of the system that are most likely to be correlated with consciousness.

The state-based Φ analysis is computationally expensive because the calculations have to be run on every bipartition of every possible subset of the network.[6] The first part of the full analysis is the extraction of all the possible subsets of the network, with the number of ways of selecting *m* elements out of the *n* elements of the system being given by Equation 4:

$$\frac{n!}{m!(n-m)!}, \tag{4}$$

which has to be summed over all subset sizes from $m = 2$ to $m = n$. The next part is the calculation of effective information on every possible bipartition of each subset in order to find the minimum information partition. A bipartition is created by selecting *k* elements out of the *m* elements in the subset, where k ranges from 0 to $m / 2$. Putting the subset selections together with the bipartition selections gives Equation 5:

$$t_{analysis} = \sum_{m=2}^{\mathrm{n}} \sum_{k=0}^{m/2} \frac{n!}{m!(n-m)!} \frac{m!}{k!(m-k)!} f(m), \tag{5}$$

[5] Normalization can introduce instabilities in the Φ value of the subset – see Barrett and Seth (2011).
[6] This paper only considers bipartitions of the system, but the analysis can also be applied to other divisions of the system, such as tripartitions.

where $t_{analysis}$ is the full analysis time and $f(m)$ is the time taken to calculate the effective information on a single bipartition of a subset of size m. Equation 5 omits the fact that when the number of neurons in each half of the bipartition is exactly the same, the number of possible bipartitions has to be divided by two, because the selection of all possible combinations in one half results in the selection of all possible combinations in the other half. Experimental measures of the performance of the state-based Φ algorithm are given in Section 5.7.

2.2 Liveliness

The concept of liveliness emerged out of earlier work of Aleksander (1973) and Aleksander and Atlas (1973) as part of a logic- and state-based approach to neural networks. This work developed the concept of lively connections, neurons and functions and created a way of estimating the distribution of lively physical loops and consequent state cycles in a system. The earlier work was based on a static analysis of the system and it has recently been reworked into the state-based approach that is presented in this paper.

Liveliness or λ is a measure of the causal influence between two neurons for a particular state of the network. The causal influence measured by liveliness is different from anatomical connectivity because the connection weights between neurons do not tell you whether one neuron is influencing another at a particular point in time. For example, if neuron n_1 connects to neuron n_2 with a lively connection, then there is a high probability that a change in state of n_1 will result in a change in state of n_2 independent of the inputs to the other connections. With a low liveliness connection, the state of n_2 is largely independent of the state of n_1. In the network shown in Figure 1A, the connection between n_2 and n_3 has high liveliness because a change in the state of n_2 will be reflected in a change in the output of n_3. On the other hand, the connection from n_1 to n_3 in Figure 1A has zero liveliness because a change in the state of n_1 will not affect the output of n_2 in the next time step. The key hypothesis of the liveliness measure is that lively connections are responsible for the integration of the system and that the number of lively connections is linked to the system's capacity for differentiation.

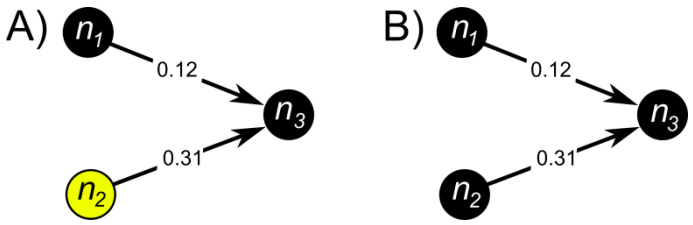

Figure 1. Simple network to illustrate liveliness. n_3 has a simple integrate and fire function with a threshold of 0.1; firing neurons are highlighted in yellow. A) In this state of the network the connection between n_1 and n_3 has zero liveliness because the next state of n_3 will be controlled entirely by the firing of n_2: a change in n_1 would not make any difference to the next state of n_3. B) Here the connection between n_1 and n_3 is lively because a change in n_1 will affect the next state of n_3.

The state-based liveliness algorithm is based on a transition between the current state of the network, s_0, and the next state of the network, s_1. The algorithm starts by using the state of the network s_0 to calculate the next state of the network, s_1. Then the state of each pre-synaptic neuron at s_0 is varied to determine whether it affects the state of the post-synaptic neurons at s_1. If a change in the state of a pre-synaptic neuron affects the next state of a post-synaptic neuron, then the connection is assigned a liveliness of 1. Otherwise, it is assigned a liveliness of 0. While this paper focuses on networks in which the connection liveliness is 1 or 0, this could become a graded value if the system was non-deterministic, if pre-synaptic neurons had continuous instead of binary output, or if pre-synaptic neurons were varied in combination – see the discussion in Section 6. Once the liveliness of the connections has been calculated, the liveliness of each neuron is given by the sum of the incoming liveliness of its connections. The neurons' liveliness can be plotted as a heat map to give a highly intuitive picture of the information integration of the network.

The final stage of the algorithm is the identification of clusters in the network. A cluster is defined as an area of the network that is informationally isolated because it does not have lively connections (i.e., it is not actively integrating information) with neurons in the rest of the network. Clusters can be identified by starting with a seed neuron and adding neurons with lively connections to this neuron until no more neurons can be added. The resulting group of neurons will be a cluster, and treating each neuron as a seed neuron enables all clusters to be identified. The cluster is a unit of information integration similar to Balduzzi and Tononi's (2008) complex and the total liveliness of a cluster is given by Equation 6:

$$\lambda_{cluster} = \lambda_{n-tot}\left(\frac{\lambda_{n-tot}}{n^2}\right), \tag{6}$$

where $\lambda_{n\text{-}tot}$ is the sum of the liveliness of the neurons in the cluster and n is the number of neurons in the cluster. In this equation, the maximum liveliness of the cluster is n^2, so $\lambda_{n\text{-}tot}/n^2$ is a measure of the density of lively connections within the cluster, which is similar to Seth's (2006) measure of causal density (with the difference that self

connections are taken into account). By itself, lively connection density is not an adequate measure of information integration because larger networks with low lively density could have many more lively connections (and therefore more differentiation) than a pair of neurons with a full set of lively connections. To address this issue the density of lively connections is used to scale the total amount of liveliness in the cluster, so that larger networks with greater total liveliness will have larger values of $\lambda_{cluster}$ than smaller networks with the same density of lively connections.[7]

The main computational costs of the liveliness algorithm are the cost of calculating the connection liveliness and the cost of expanding each seed to identify clusters, as shown in Equation 7:

$$t_{analysis} = c \times C_c + n \times C_{ex} \times \alpha, \tag{7}$$

where c is the number of connections, n is the number of neurons, C_c is the cost of calculating the liveliness of a single connection, C_{ex} is the cost of expanding a single connection from a seed neuron to identify a cluster, and α is the average cluster size. In biologically-inspired networks with 1,000 – 10,000 connections per neuron the first part of the equation will dominate the calculation cost, and the analysis time should vary approximately linearly with the number of connections. Experimental measures of the performance of the liveliness algorithm are given in Section 5.7.

3. Test Networks

3.1 Introduction

This section outlines six networks that were designed to compare the areas of maximum information integration identified by the state-based Φ and liveliness algorithms, and a set of simpler test networks that were used to measure the performance. The architectures of these networks was selected to illustrate potential links between information integration in the brain and consciousness, and the discussion in the following sections makes some loose comparisons between each network's structure and structures in the brain. However, these comparisons must be taken with a great deal of caution because the small size of these networks makes any comparison between their structure and brain structures *extremely* preliminary.

[7] In practice, $\lambda_{n\text{-}tot}$ will typically scale with the size of the network, so Equation 6 could potentially be written as $\lambda_{cluster} = n(\lambda_{n\text{-}tot}/n^2)$. However, this reformulation would not express our intuition that the differentiation of the network is linked to the total number of lively connections, and not to the total number of elements.

3.2 Neurons

The state-based Φ and liveliness algorithms are applicable to any set of elements that have causal links between them, such as neurons based on logic functions, weightless neurons (Aleksander, 2005), and weighted neurons, such as Izhikevich (2003). In the analyses described in this paper the test networks were created using weightless neurons, which can be used to implement logic-based networks and to approximate weighted neural models. To understand how a weightless neuron can approximate a weighted neuron, consider the network shown in Figure 2. If n_5 has a threshold of 0.5, then it will fire whenever *one* of the other neurons are firing, which corresponds to a weightless neuron with function f_4 in Table 1. On the other hand, if n_5 has a threshold of 2, then it will only fire when *all* of the other neurons are firing, which corresponds to a weightless neuron with function f_1 in Table 1. Weightless functions can also be created to model networks in which each connection has a different weight.

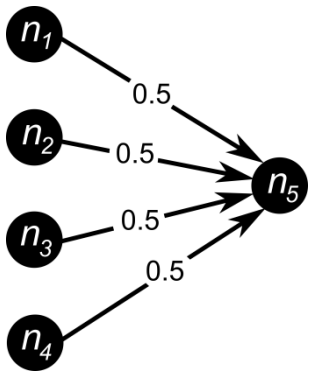

Figure 2. Four neurons connect to a fifth with connection weight 0.5. n_5 has an integrate and fire function that can have different thresholds. There is a long time step and no delay, so only spikes received during the current time step influence whether n_5 fires or not.

The number of neurons in each test network was fixed to ensure that only the functions and topology of the network influenced the results. A size of 12 neurons was selected because it is roughly the maximum number that can be analyzed by the state-based Φ algorithm in a reasonable time (see Section 5.7). To study the effect of different function and weight combinations on information integration, five of the test networks were created in four different versions in which all of the neurons implemented one of the functions listed in Table 2. These versions of the networks with different functions will be referred to as 100f, 75f, 50f and 25f. The next few sections outline the connections and firing patterns of each test network in more detail.

n_1	n_2	n_3	n_4	f_1	f_2	f_3	f_4
0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	1
0	0	1	0	0	0	0	1
0	1	0	0	0	0	0	1
1	0	0	0	0	0	0	1
0	0	1	1	0	0	1	1
0	1	0	1	0	0	1	1
1	0	0	1	0	0	1	1
0	1	1	0	0	0	1	1
1	0	1	0	0	0	1	1
1	1	0	0	0	0	1	1
0	1	1	1	0	1	1	1
1	0	1	1	0	1	1	1
1	1	0	1	0	1	1	1
1	1	1	0	0	1	1	1
1	1	1	1	1	1	1	1

Table 1. Examples of weightless functions that could be implemented by neuron n_5 in Figure 2. f_1 corresponds to a weighted neuron with a threshold of 2; f_2 corresponds to a weighted neuron with a threshold of 1.5; f_3 corresponds to a weighted neuron with a threshold of 1 and f_4 corresponds to a weighted neuron with a threshold of 0.5.

	Condition	**Output**
100f	100 % of inputs = 1	1
75f	75% of inputs = 1	1
50f	50% of inputs = 1	1
25f	25% of inputs = 1	1

Table 2. Neuron functions. Output is zero unless the condition is matched. The 100%, 75%, 50% and 25% functions correspond to f_1, f_2, f_3 and f_4 in Table 1.

3.3 Uniform Test Network

The uniform test network is similar to Balduzzi and Tononi (2008), Figure 12 and it was intended to loosely correspond to densely connected cortex. Each neuron receives six connections from randomly selected neurons without any self connections, as shown in Figure 3. This network was created in four different versions (100f, 75f, 50f, 25f), in which all of the neurons implemented one of the functions in Table 2. The following firing patterns were used for the analysis of the network:

- No neurons firing.
- 25% randomly selected neurons firing.
- 50% randomly selected neurons firing.
- 75% randomly selected neurons firing.
- 100% randomly selected neurons firing.

These firing patterns are shown in the first rows of figures 10 and 11.

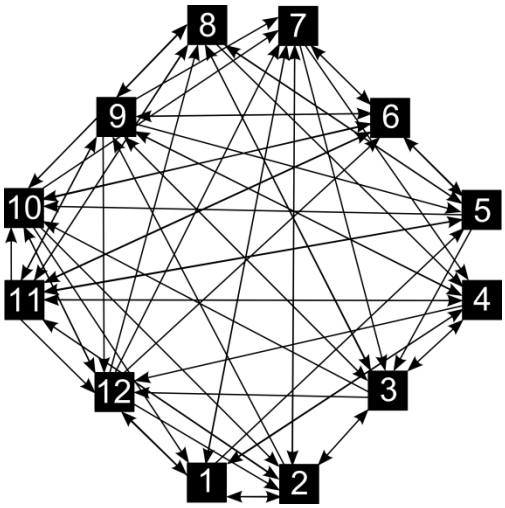

Figure 3. Uniform test network. The six connections to each neuron are not always visible in this diagram because of overlap between them.

3.4 Partitioned Test Network

A critical test of any algorithm for consciousness is the ability to identify two separate consciousness within the same system – for example, in split brain patients (Gazzaniga, 1970). The partitioned test network shown in Figure 4 consists of two independent clusters, and each neuron receives three connections from randomly selected neurons in the same cluster. This network was created in four different versions (100f, 75f, 50f, 25f) in which all of the neurons implemented one of the functions in Table 2. The following firing patterns were used for the analysis of the network:

- No neurons firing
- 25% neurons firing
- 50% neurons firing
- 75% neurons firing
- 100% neurons firing

These firing patterns are shown in the first rows of figures 13 and 14.

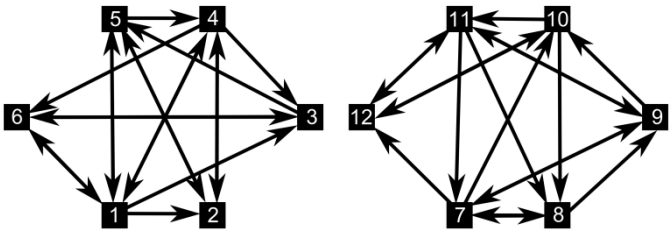

Figure 4. Partitioned test network

3.5 Modular Test Network

The modular test network loosely models the division of the brain into highly integrated nuclei with low amounts of connectivity between them, and it is similar to Balduzzi and Tononi (2008), Figure 13. The network is divided into three fully connected ‘modules’ with a single one way connection between each module, as shown in Figure 5. This network was created in four different versions (100f, 75f, 50f, 25f) in which all of the neurons implemented one of the functions in Table 2. The following firing patterns were used for the analysis of the network:

- No neurons firing.
- 1 randomly selected neuron firing in each module.
- 2 randomly selected neurons firing in each module.
- 3 randomly selected neurons firing in each module.
- 4 randomly selected neurons firing in each module.

These firing patterns are shown in the first rows of figures 16 and 17.

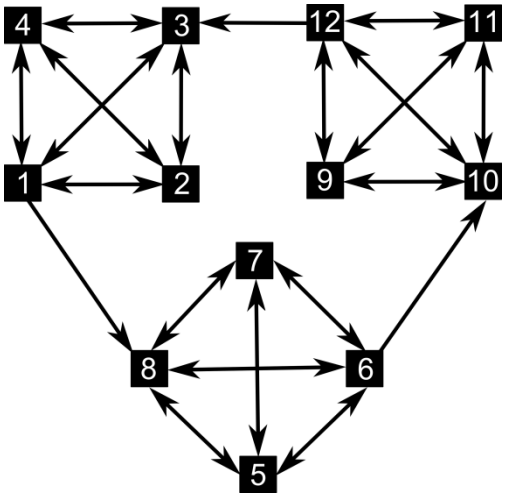

Figure 5. Modular test network

3.6 Sensory Test Network

The sensory test network loosely corresponds to a highly integrated part of the brain that receives sensory information from the periphery – what Balduzzi and Tononi (2008) refer to as ports-in – and it is also similar to the right half of Balduzzi and Tononi (2008), Figure 5. In this network, six of the neurons are fully connected and they each receive a single connection from one of the other six neurons, as shown in Figure 6. This network was created in four different

versions (100f, 75f, 50f, 25f) in which all of the neurons implemented one of the functions described in Table 2. The following firing patterns were used for the analysis of the network:

- No neurons firing.
- 50% of the centre neurons firing; peripheral neurons quiescent.
- Centre neurons quiescent; 50% of the peripheral neurons firing.
- All centre neurons firing; peripheral neurons quiescent.
- All centre neurons quiescent; peripheral neurons firing.
- 50% of all neurons firing.

These firing patterns are shown in the first rows of figures 19 and 20.

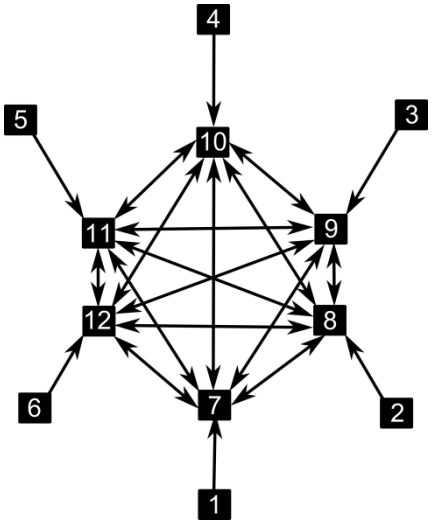

Figure 6. Sensory test network

3.7 Motor Test Network

The motor test network loosely corresponds to a highly integrated part of the brain passing motor control information to the periphery – what Balduzzi and Tononi (2008) refer to as ports-out. It is similar to the left half of Balduzzi and Tononi (2008), Figure 5, and to Balduzzi and Tononi (2008), Figure 6. In this network, six of the neurons are fully connected and each of the central neurons is connected to one of the six neurons at the periphery, as shown in Figure 7. This network was created in four different versions (100f, 75f, 50f, 25f) in which all of the neurons implemented one of the functions described in Table 2. The following firing patterns were used for the analysis of the network:

- No neurons firing.
- 50% of the centre neurons firing; peripheral neurons quiescent.

- Centre neurons quiescent; 50% of the peripheral neurons firing.
- All centre neurons firing; peripheral neurons quiescent.
- All centre neurons quiescent; peripheral neurons firing.
- 50% of all neurons firing.

These firing patterns are shown in the first rows of figures 22 and 23.

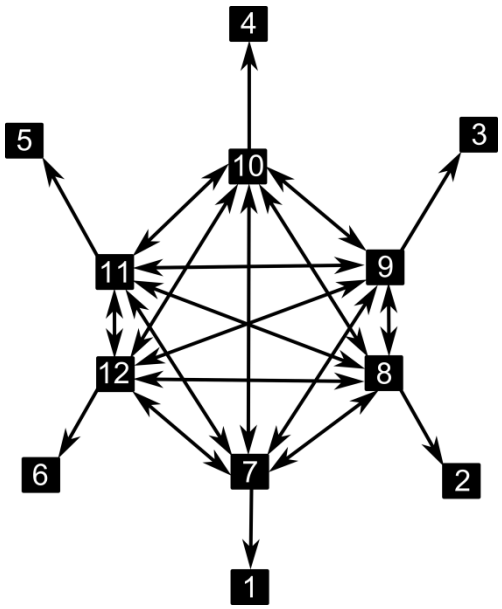

Figure 7. Motor test network

3.7 Sensory-motor Test Network

The sensory-motor test network is intended to illustrate a system that receives information from the world, processes it internally and then produces an output that could correspond to a motor command. The three neurons at the bottom of Figure 8 are the network's input, the middle ring of neurons holds the network's internal states and the top three neurons are the output of the network. This network was created in a single version in which the neurons' truth tables were hand coded to implement the input→internal-processing→output behaviour. The following firing patterns were used for the analysis of the network:

- No neurons firing.
- Two input neurons firing.
- Three 'internal' neurons firing in response to the input signal.
- Three 'internal' neurons firing in response to the previous pattern.
- Output neurons firing in response to the state of the 'internal' neurons.

These firing patterns are shown in the first rows of figures 25 and 26.

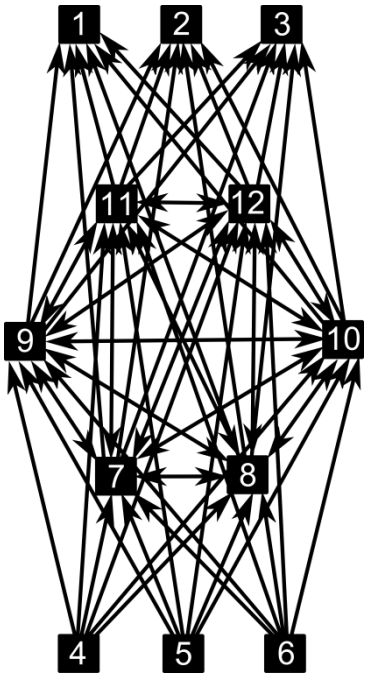

Figure 8. Sensory-motor test network

4. Materials and Methods

The test networks were created in the SpikeStream neural simulator and analyzed using state-based Φ and liveliness plugins written in C++ for SpikeStream.[8] These plugins were unit tested and validated on some of the networks from Balduzzi and Tononi (2008). A screenshot of the liveliness plugin is shown in Figure 9.

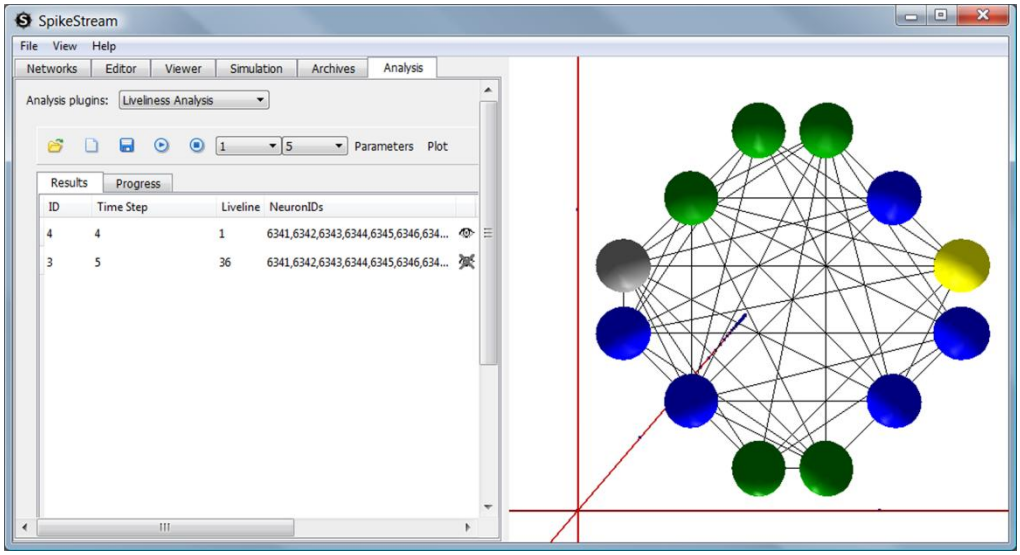

Figure 9. SpikeStream plugin for liveliness analysis. The colours in the network on the right correspond to the amount of liveliness that is associated with each neuron for a particular state.

[8] More information about SpikeStream can be obtained at: http://spikestream.sf.net. An open source release of SpikeStream with the liveliness and state-based Φ plugins is available to download.

The state-based Φ results were difficult to interpret because a large number of complexes were often found for each time step and there were often many overlapping complexes with the same value of Φ. To alleviate this difficulty only main complexes with the maximum value of Φ were included in the results and complexes with lower Φ were discarded. Complexes or clusters with Φ or liveliness less than 1 were also discarded.

To provide a simple visual comparison between the two algorithms, the results were plotted as spectrograms showing the normalized maximum integration between each pair of neurons averaged over all time steps. These spectrograms were calculated by recording the highest Φ or liveliness of a complex or cluster that each pair of neurons was involved in. The results were then normalized between 0 and 1, averaged over all time steps, and plotted using the colour scale shown on the right of each spectrogram. The neuron IDs on the X and Y axes correspond to the neuron numbers in the network diagrams; the colour shows the average normalized integration between each pair of neurons according to the state-based Φ and liveliness algorithms.

To measure the performance of the two algorithms, networks with different numbers of weightless neurons were created, with each neuron receiving connections from another five randomly selected neurons. These networks were trained on 5 different patterns and the time that each algorithm took per time step was averaged over 25 runs on a Pentium IV 3.2 GHz single core computer. The performance of the state-based Φ algorithm was measured for a full analysis and for an analysis in which the Φ of the entire network was calculated without examining subsets.

5. Results

5.1 Uniform Network Results

The results for the liveliness and state-based Φ analyses of the uniform network are given in figures 10-12. Both algorithms predicted a reasonably uniform information integration distribution across the network, which matched expectations based on the connectivity. The deviations from complete uniformity were most likely due to randomness in the connections and firing patterns.

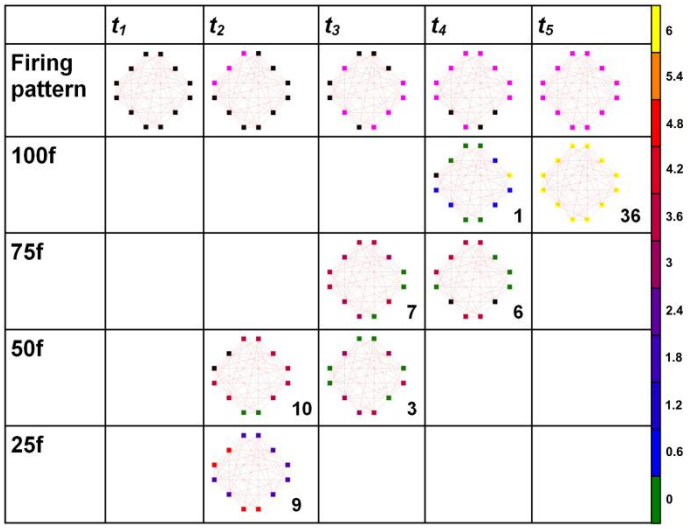

Figure 10. Liveliness results for the uniform network. The first row shows the firing pattern for each network at time steps t_1-t_5; subsequent rows show the clusters that were identified for each function and time step. An empty cell indicates that no clusters with λ>=1 were found. The numbers are the cluster liveliness calculated according to Equation 6 and rounded to 0 decimal places. The amount of liveliness of each neuron is indicated using the colour scale on the right of the figure. Neurons coloured black are not included in the cluster.

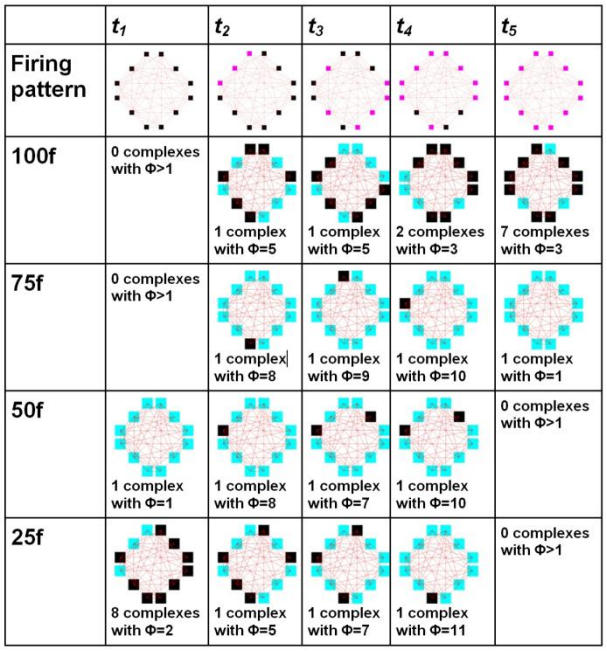

Figure 11. State-based Φ results for the uniform network showing some of the main complex(es) for each time step and function. The first row shows the firing pattern for each network at time steps t_1-t_5; subsequent rows show the complexes that were identified for each function and time step. When more than one complex was discovered with the same Φ, only a typical example is shown. The Φ values were rounded to 0 decimal places after the main complex(es) had been identified.

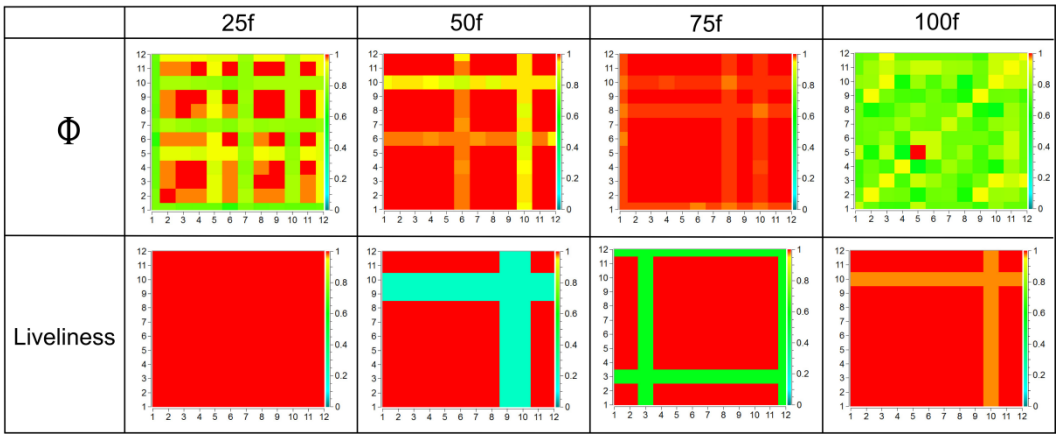

Figure 12. Integration spectrograms for the uniform network showing the normalized maximum integration between each pair of neurons for different functions averaged over all time steps

5.2 Partitioned Network Results

The results for the partitioned network are given in figures 13-15. Both algorithms correctly identified the partition in the network; the key difference between them was that the state-based Φ algorithm typically identified more integration in the right group of neurons with IDs 7-12, whereas the liveliness algorithm typically identified more integration in the left group of neurons with IDs 1-6.

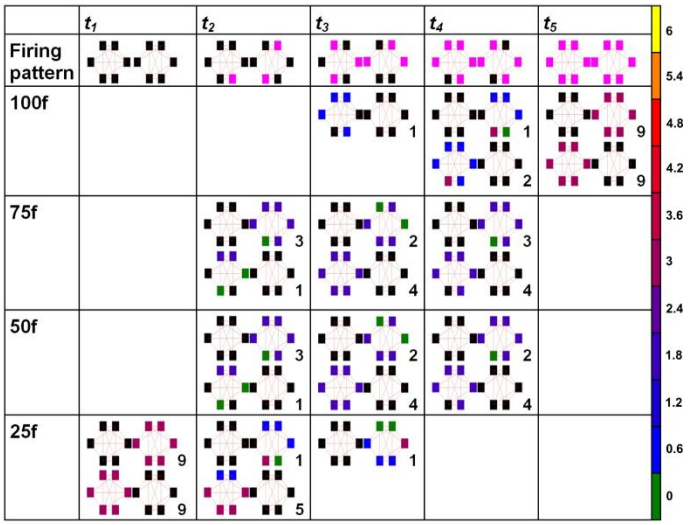

Figure 13. Liveliness results for the partitioned network. The first row shows the firing pattern for each network at time steps t_1-t_5; subsequent rows show the clusters that were identified for each function and time step. An empty cell indicates that no clusters with λ>=1 were found. The numbers indicate the cluster liveliness calculated according to Equation 6 and rounded to 0 decimal places. The amount of liveliness of each neuron is indicated using the colour scale on the right of the figure. Neurons coloured black are not included in the cluster.

	t_1	t_2	t_3	t_4	t_5
Firing pattern					
100f	0 complexes with Φ>1	3 complexes with Φ=4	1 complex with Φ=4	1 complex with Φ=3	3 complexes with Φ=3
75f	2 complexes with Φ=1	1 complex with Φ=5	1 complex with Φ=5	1 complex with Φ=5	2 complexes with Φ=1
50f	2 complexes with Φ=1	1 complex with Φ=5	1 complex with Φ=5	1 complex with Φ=5	2 complexes with Φ=1
25f	2 complexes with Φ=3	1 complex with Φ=3	2 complexes with Φ=4	2 complexes with Φ=4	0 complexes with Φ>1

Figure 14. State-based Φ results for the partitioned network showing some of the main complex(es) for each time step and function. The first row shows the firing pattern for each network at time steps t_1-t_5; subsequent rows show the complexes that were identified for each function and time step. When more than one complex was discovered with the same Φ, only a typical example is shown. The Φ values were rounded to 0 decimal places after the main complex(es) had been identified.

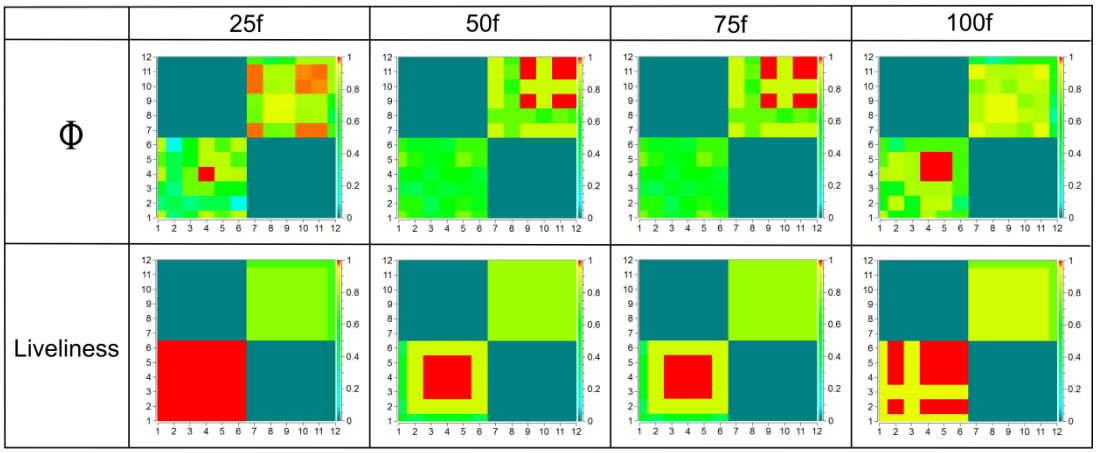

Figure 15. Integration spectrograms for the partitioned network showing the normalized maximum integration between each pair of neurons for different functions averaged over all time steps

5.3 Modular Network Results

The liveliness and state-based Φ results for the modular network are shown in figures 16-18. These integration spectrograms are quite messy, but they do demonstrate that the state-based Φ and liveliness algorithms identified modularity in all of the networks. The state-based Φ results are a bit cleaner for the 75f, 50f and 25f networks, showing better defined areas of high integration between neurons 1-4, 5-8 and 9-12.

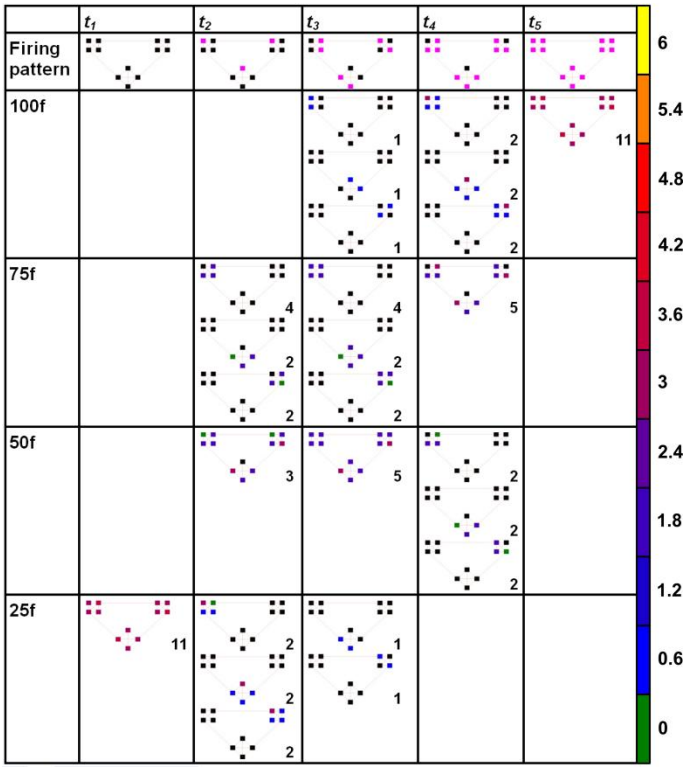

Figure 16. Liveliness results for the modular network. The first row shows the firing pattern for each network at time steps t_1-t_5; subsequent rows show the clusters that were identified for each function and time step. An empty cell indicates that no clusters with λ>=1 were found. The numbers indicate the cluster liveliness calculated according to Equation 6 and rounded to 0 decimal places. The amount of liveliness of each neuron is indicated using the colour scale on the right of the figure. Neurons coloured black are not included in the cluster.

	t_1	t_2	t_3	t_4	t_5
Firing pattern					
100f	0 complexes with Φ>1	3 complexes with Φ=4	2 complexes with Φ=4	3 complexes with Φ=2	18 complexes with Φ=2
75f	3 complexes with Φ=1	3 complexes with Φ=4	9 complexes with Φ=2	6 complexes with Φ=2	9 complexes with Φ=1
50f	7 complexes with Φ=1	6 complexes with Φ=2	3 complexes with Φ=2	3 complexes with Φ=4	3 complexes with Φ=1
25f	18 complexes with Φ=2	3 complexes with Φ=2	6 complexes with Φ=2	3 complexes with Φ=4	0 complexes with Φ>1

Figure 17. State-based Φ results for the modular network showing some of the main complex(es) for each time step and function. The first row shows the firing pattern for each network at time steps t_1-t_5; subsequent rows show the complexes that were identified for each function and time step. When more than one complex was discovered with the same Φ, only a typical example is shown. The Φ values were rounded to 0 decimal places after the main complex(es) had been identified.

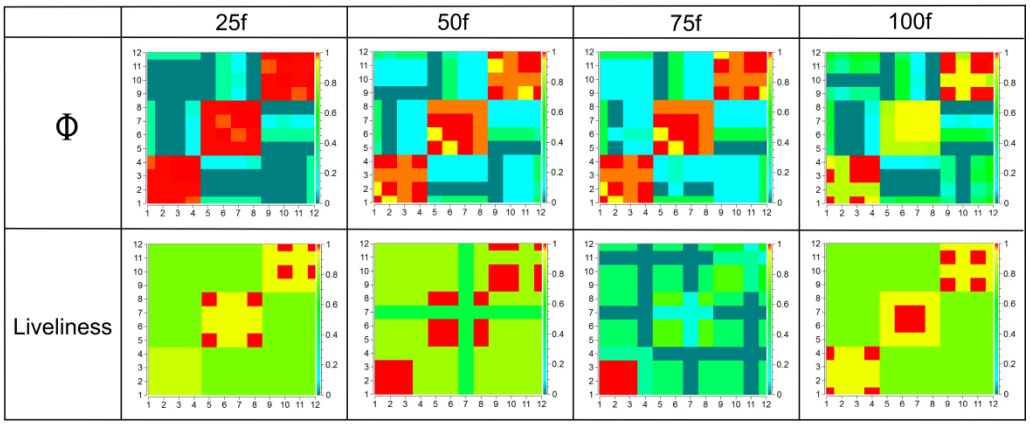

Figure 18. Integration spectrograms for the modular network showing the normalized maximum integration for different functions averaged over all time steps

5.4 Sensory Network Results

The results for the sensory network in figures 19-21 show that the state-based Φ algorithm was generally better at identifying the expected high integration between the central group of six neurons. However, neither algorithm

appears to have identified patterns of integration between the peripheral neurons and the centre, which would be expected to appear as diagonal lines starting at (1,7) and finishing at (6,12) and starting at (7,1) and finishing at (12,6). Traces of these diagonal lines can be seen in the state-based Φ spectrograms for the 100f and 75f motor networks.

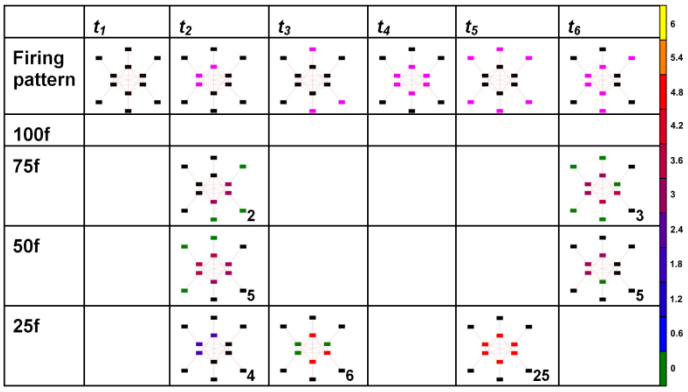

Figure 19. Liveliness results for the sensory network. The first row shows the firing pattern for each network at time steps t_1-t_6; subsequent rows show the clusters that were identified for each function and time step. An empty cell indicates that no clusters with λ>=1 were found. The numbers indicate the cluster liveliness calculated according to Equation 6 and rounded to 0 decimal places. The amount of liveliness of each neuron is indicated using the colour scale on the right of the figure. Neurons coloured black are not included in the cluster.

	t_1	t_2	t_3	t_4	t_5	t_6
Firing pattern						
100f	0 complexes with Φ>1	3 complexes with Φ=3	0 complexes with Φ>1	15 complexes with Φ=2	0 complexes with Φ>1	1 complex with Φ=3
75f	0 complexes with Φ>1	3 complexes with Φ=2	0 complexes with Φ=0	20 complexes with Φ=1	0 complexes with Φ=0	1 complex with Φ=4
50f	20 complexes with Φ=1	3 complexes with Φ=2	20 complexes with Φ=1	0 complexes with Φ>1	20 complexes with Φ=1	1 complex with Φ=5
25f	20 complexes with Φ=2	1 complex with Φ=2	20 complexes with Φ=2	0 complexes with Φ>1	20 complexes with Φ=2	1 complex with Φ=3

Figure 20. State-based Φ results for the sensory network showing some of the main complex(es) for each time step and function. The first row shows the firing pattern for each network at time steps t_1-t_6; subsequent rows show the complexes that were identified for each function and time step. When more than one complex was discovered with the same Φ, only a typical example is shown. The Φ values were rounded to 0 decimal places after the main complex(es) had been identified.

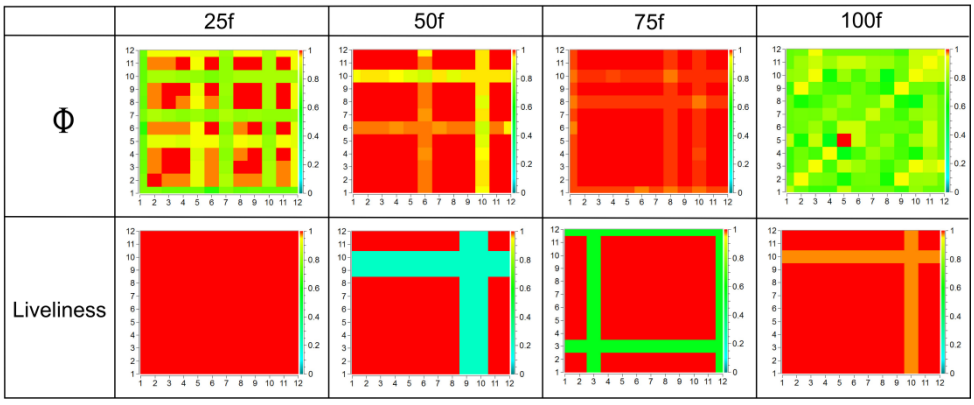

Figure 21. Integration spectrograms for the sensory network showing the normalized maximum integration for different functions averaged over all time steps

5.5 Motor Network Results

The motor network liveliness and state-based Φ results in figures 22-24 show generally poor results for the liveliness algorithm, which identified few clusters for most of the firing patterns in the 100f and 75f networks. The state-based Φ

results for the 100f and 75f networks were much closer to expectations, showing integration within the centre cluster of neurons 6-12 and integration between the central neurons and the peripheral neurons, which appears as diagonal lines starting at (1,7) and finishing at (6,12) and starting at (7,1) and finishing at (12,6).

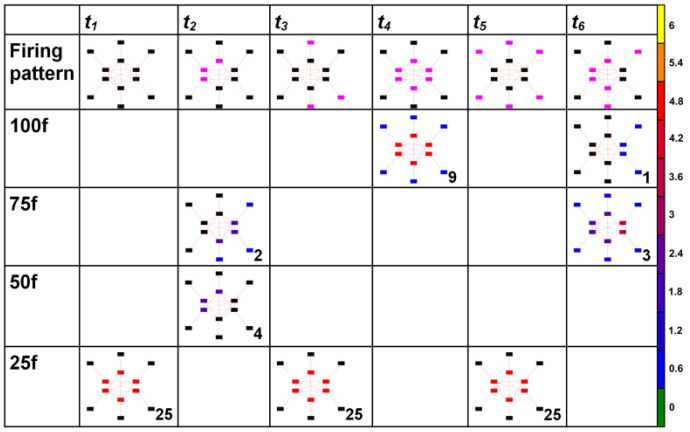

Figure 22. Liveliness results for the motor network. The first row shows the firing pattern for each network at time steps t_1-t_6; subsequent rows show the clusters that were identified for each function and time step. An empty cell indicates that no clusters with λ>=1 were found. The numbers indicate the cluster liveliness calculated according to Equation 6 and rounded to 0 decimal places. The amount of liveliness of each neuron is indicated using the colour scale on the right of the figure. Neurons coloured black are not included in the cluster.

	t_1	t_2	t_3	t_4	t_5	t_6
Firing pattern						
100f	6 complexes with Φ=1	3 complexes with Φ=3	6 complexes with Φ=1	15 complexes with Φ=2	15 complexes with Φ=2	4 complexes with Φ=3
75f	6 complexes with Φ=1	3 complexes with Φ=3	1 complex with Φ=2	20 complexes with Φ=2	20 complexes with Φ=2	4 complexes with Φ=3
50f	20 complexes with Φ=2	3 complexes with Φ=3	20 complexes with Φ=2	0 complexes with Φ>1	20 complexes with Φ=2	2 complexes with Φ=3
25f	15 complexes with Φ=2	3 complexes with Φ=3	15 complexes with Φ=2	0 complexes with Φ>1	15 complexes with Φ=2	2 complexes with Φ=5

Figure 23. State-based Φ results for the motor network showing some of the main complex(es) for each time step and function. The first row shows the firing pattern for each network at time steps t_1-t_6; subsequent rows show the complexes that were identified for each function and time step. When more than one complex was discovered with the same Φ, only a typical example is shown. The Φ values were rounded to 0 decimal places after the main complex(es) had been identified.

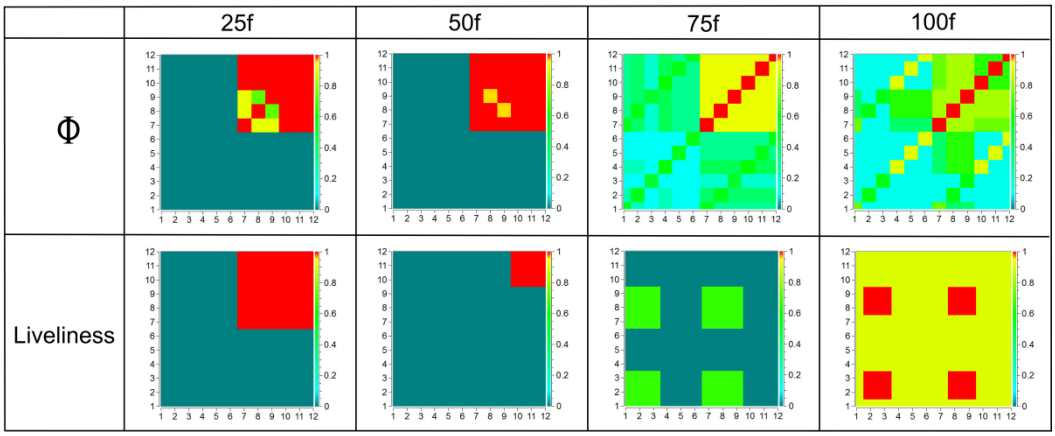

Figure 24. Integration spectrograms for the motor network showing the normalized maximum integration for different functions averaged over all time steps

5.6 Sensory-motor Network Results

The sensory-motor liveliness and state-based Φ results in figures 25-27 show broadly similar results for both algorithms with the expected high integration within the central cluster and reduced integration with peripheral

neurons. The liveliness algorithms indicates higher levels of information integration when information is being passed from the input to the centre neurons and during the processing of the centre neurons, whereas the state-based Φ algorithm indicates higher levels of information integration during processing of input information and when information is being passed to the output neurons. This timing difference is likely to be due to the fact that the liveliness algorithm is attempting to decide which parts of the current state of the network affect the next state of the network, whereas the state-based Φ algorithm is examining the extent to which previous states of the network are responsible for the network's current state.

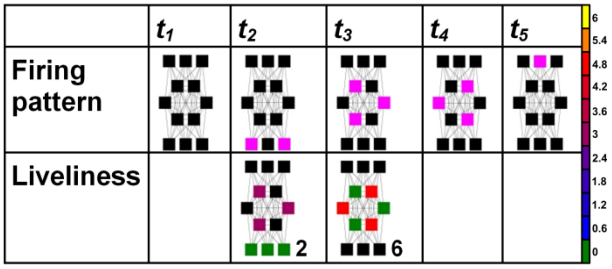

Figure 25. Liveliness results for the sensory motor network. The first row shows the firing pattern for each network at time steps t_1-t_5; the second row shows the clusters that were identified for each time step. An empty cell indicates that no clusters with λ>=1 were found. The numbers indicate the cluster liveliness calculated according to Equation 6 and rounded to 0 decimal places. The amount of liveliness of each neuron is indicated using the colour scale on the right of the figure. Neurons coloured black are not included in the cluster.

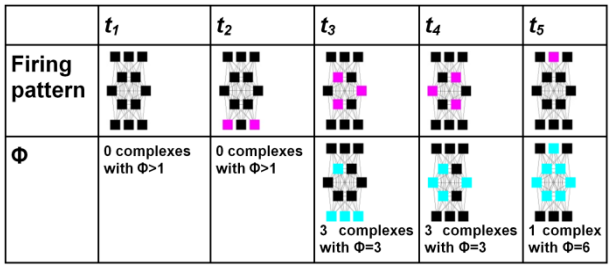

Figure 26. State-based Φ results for the sensory motor network. The first row shows the firing pattern for each network at time steps t_1-t_6; the second row shows the complexes that were identified for each time step. When more than one complex was discovered with the same Φ, only a typical example is shown. The Φ values were rounded to 0 decimal places after the main complex(es) had been identified.

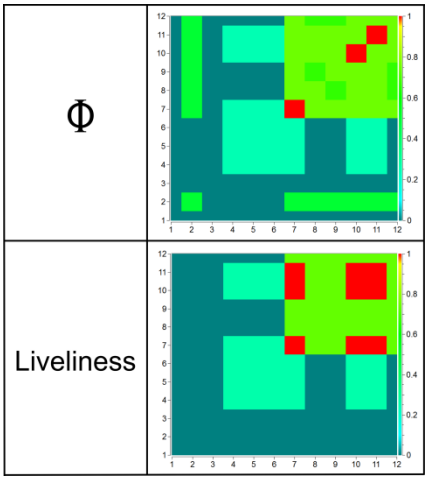

Figure 27. Integration spectrogram for the sensory-motor network showing the normalized maximum integration averaged over all time steps

5.7 Performance Results

The measured and predicted performance of the state-based Φ algorithm is shown in Figure 28. The factorial dependencies caused the full analysis time of this algorithm to increase very rapidly, and the fitted exponential

function, shown as a red dotted line, suggest that it would take approximately 10^9 years to analyze a network of 30 neurons using a desktop computer. The analysis times without subsets also increased exponentially with the number of neurons due to the factorial dependency on the number of bipartitions of the network. Even with heavy optimization and supercomputing power, the performance of the state-based Φ algorithm is likely to remain a major limitation for the foreseeable future.

The measured performance times for the liveliness algorithm in Figure 29 show a linear dependency on the number of neurons, which is also a linear dependency on the number of connections because there was a fixed number of five connections per neuron in all of the networks. It took around 13 seconds to analyze a 100 neuron network for liveliness.

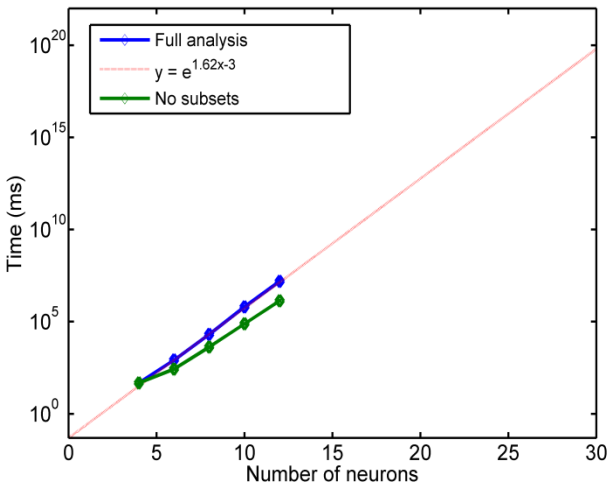

Figure 28. Measured and predicted analysis times for the state-based Φ algorithm showing results for full analyses and analyses without subsets. The full analysis times approximate the function $y=e^{1.62x-3}$. Error bars of +/- 1 standard deviation were plotted, but are not visible on the logarithmic scale.

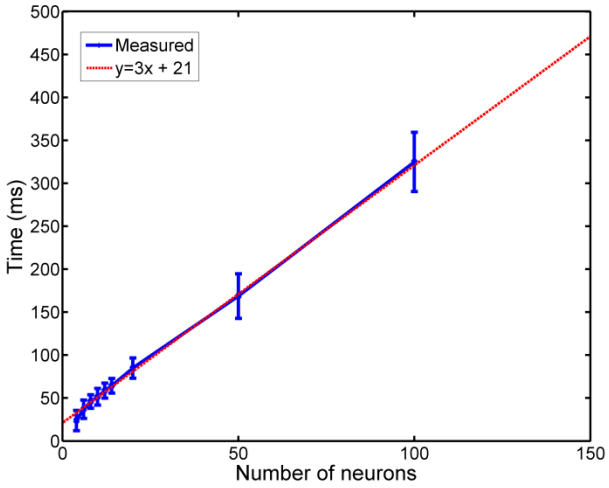

Figure 29. Measured analysis times for the liveliness algorithm, which approximate the function $y=3x+21$. Error bars are +/- 1 standard deviation.

6. Discussion

One of the main problems with the state-based Φ algorithm is that the analysis of networks with random connections and firing patterns can result in a large number of overlapping main complexes. If Tononi's (2004, 2008) claim that the main complexes are associated with consciousness is correct, then many of the test networks would be attributed up to 20 separate overlapping consciousnesses. If the number of main complexes increases with the size of the network, then it seems likely that the brain will contain an extremely large number of main complexes, which sits poorly with our intuition that the normal waking brain is associated with a single consciousness. One way of

addressing this issue would be to merge overlapping complexes with similar values of Φ, perhaps using the rate of change of Φ to identify boundaries.

While the integration spectrograms produced by the state-based Φ algorithm approximately matched our expectations about the information integration of the network, the output from the liveliness algorithm was patchy, with no results being returned for some of the networks and firing patterns – a problem that was particularly apparent on the sensory and motor networks. One possible explanation of these results is that the liveliness algorithm did not take combinations of neurons into account. To understand why this might be taking place, consider a version of the network shown in Figure 2 in which the threshold of n_5 is 1 and all of the neurons are quiescent. In this case a change in state of any two of the presynaptic neurons will cause the postsynaptic neuron to fire, but the current liveliness algorithm would assign a liveliness of 0 to all connections - ignoring the fact that neurons are lively in combination with each other. In future work this problem could be addressed by measuring the liveliness of different combinations of neurons and taking the average of each connection. It might also be possible to use the *a posteriori* repertoire to measure the extent to which combinations of neurons affect subsequent states. Whichever approach is chosen, care would have to be taken to avoid introducing factorial dependencies into the liveliness algorithm, which would make its performance similar to state-based Φ.

A second limitation of the liveliness algorithm is that its simple clustering method is only capable of identifying completely isolated parts of the network. While the liveliness algorithm's ability to generate heat maps make this problem less serious, it could be addressed by applying more sophisticated clustering algorithms, such as Feldt et al. (2009), to the lively weights, or by looking for contour lines in the neuron heat maps using visual processing techniques. Clustering algorithms that could generate overlapping clusters should be avoided because this would lead to the problems with overlapping consciousnesses that were discussed earlier. Clusters contained inside clusters with higher liveliness could be discarded in a similar way to the processing of subsets described in Balduzzi and Tononi (2008).

The detailed results show that maximum values of liveliness were typically present when all or none of the neurons were firing, whereas higher values of Φ were associated with intermediate levels of firing. Since an intermediate firing state can occur in more possible ways, it might be thought that the differentiation of the network would be higher in these intermediate firing states, which would make state-based Φ a more accurate measure of information integration. A second way of interpreting the differentiation criteria is that it should be higher for inhomogeneous architectures. In these experiments both state-based Φ and liveliness had their maximum values for

the uniform network, which had a substantial degree of randomness in its connectivity, but more systematic work would be required to establish whether there is a systematic link between the information integration measures and the degree of homogeneity of the network. State-based Φ and liveliness also had significant differences in the timing of their maximum values, which was particularly apparent in the sensory-motor network, where the cluster liveliness peaked two time steps before the maximum Φ. One way of explaining this temporal difference is that state-based Φ is identifying causal influences that *previous* states of the system exert on its current state, whereas liveliness is measuring the degree to which each element and connection influences the *next* state of the network.

While some of the differences in the information integration that was identified by the state-based Φ and liveliness algorithms can be explained by limitations of the liveliness algorithm, the question about which algorithm is a *correct* measure of information integration is much more difficult to address. In this paper the accuracy of information integration algorithms was evaluated by benchmarking them on simulated networks whose areas of information integration were reasonably intuitively obvious. However, even with simple networks it rapidly becomes unclear what the 'actual' information integration of the network should be once different functions and noisy connections and firing patterns are introduced. While there does not appear to be any easy way around this problem, further analyses of simulated and real biological networks might go some way towards addressing it.

7. Measuring the Link between Information Integration and Consciousness on Real Neural Data

At the present stage of research most of the work on information integration has been largely theoretical and based on simulated networks, with very few attempts to experimentally test the link between information integration and consciousness on humans or animals. One of the reasons is that Balduzzi and Tononi's (2008) measure depends on knowledge about the causal structure of the system and it was only defined for discrete systems – problems that have been addressed by Barrett and Seth's (2011) method for calculating information integration on time series data. A second key issue has been the factorial dependency of the state-based Φ algorithm, which severely limits the number of elements that can be analyzed and was one motivation for the development of the liveliness algorithm described in this paper.

If it could be shown that different ways of measuring information integration produced the same results on simulated systems, then the current set of information integration algorithms could be viewed as a toolbox of methods,

whose use would depend on the type of data that is available from the system. The benchmarking work on simulated networks would provide grounds for believing that the algorithms in the toolbox are equivalent ways of measuring information integration – much as one would expect digital and mercury thermometers to produce similar values for the temperature of the system.

The specific characteristics of the liveliness algorithm suggest several ways in which it could be used to test the link between information integration and consciousness. Liveliness is a perturbational approach in which the liveliness of a connection is determined by changing the state of the source element to see if it alters the state of the destination element, and in animals this type of perturbation would be relatively easy to perform by distributing electrodes throughout the animal's brain and injecting noise into one electrode while measuring from the other electrodes. The degree of change in the other electrodes could be used as a measure of the liveliness of their connection with the source electrode – with the connection liveliness taking continuous, rather than discrete values. By carrying out this procedure when the animal was asleep and awake it would be possible to establish whether there was a link between liveliness and consciousness and identify areas of the system with maximum liveliness. In humans a perturbational approach could be employed by using a similar method to Massimini et al. (2009, 2010), who injected noise into the human brain using transcranial magnetic stimulation (TMS) and recorded the resulting pattern of activity using EEG when the subjects were conscious and unconscious. By systematically using TMS to inject noise at different points across the brain and measuring the change of activity it would be possible to measure the liveliness of connections between different parts of the brain. This experiment could be performed on conscious and unconscious subjects to establish whether there is a link between liveliness and consciousness. This approach could be extended in both animal and human models by asking subjects to carry out a particular task and repeating the measurements during the task. Similar experiments with greater spatial resolution could be carried out by combining fMRI with TMS.

A second way of testing the link between information integration and consciousness would be to construct a large scale neural model based on the brain and match the behaviour of this model to the behaviour of a subject's brain by using the model's local field potential to predict the subject's fMRI patterns. Such a calibrated model should be able to produce approximately the same fMRI patterns as the subject's brain under different conditions. This calibrated model could then be analyzed using the liveliness algorithm and used to make predictions about the regions of the subject's fMRI scan that were linked to consciousness. Separate 'brain reading' techniques, such as Kay et al.

(2008), could be used to identify the contents of the predicted consciousness within the lively clusters, which could be compared with first person reports.

8. Conclusions

Information integration is a property of a network that expresses the extent to which it is capable of entering a large number of states that result from causal interactions among its elements. Tononi (2004, 2008) has made an interesting claim about the link between information integration and consciousness, but the poor performance of Balduzzi and Tononi's (2008) state-based Φ algorithm has made this very difficult to test. To address this problem this paper has set out a new algorithm for measuring information integration and compared the accuracy and performance of this algorithm with state-based Φ on a number of test networks.

The results of these experiments demonstrated that the state-based Φ algorithm matched our expectations about the areas of maximum information integration better than the liveliness algorithm, particularly for the sensory and motor networks. However, the state-based Φ algorithm had very poor performance and produced results that were difficult to interpret with large numbers of overlapping main complexes. The liveliness algorithm has a much cleaner output and scales linearly with the number of neurons and connections - with further work it might be possible to improve it, so that it can be used to approximate the state-based Φ algorithm on larger networks.

The question about which algorithm (if any) is a *correct* measure of information integration is a difficult topic that needs to be investigated in future work, and further experiments are needed on the link between information integration and consciousness. The development of efficient ways of measuring information integration will play a key role in the testing of information integration on more realistic systems and this paper has made some suggestions about how the liveliness algorithm could be applied to real neural data.

Acknowledgements

This work was supported by the *Association for Information Technology Trust*.

References

Aleksander, I. (1973). Random Logic Nets: Stability and Adaptation. *International Journal of Man-Machine Studies* 5: 115-31.

Aleksander, I. (2005). *The World in My Mind, My Mind in the World: Key Mechanisms of Consciousness in People, Animals and Machines*. Exeter: Imprint Academic.

Aleksander, I. and Atlas, P. (1973). Cyclic Activity in Nature: Causes of Stability. *International Journal of Neuroscience* 6: 45-50.

Aleksander, I. and Gamez, D. (2011). Informational Theories of Consciousness: A Review and Extension. In *From Brains to Systems: Brain-Inspired Cognitive Systems 2010*, edited by C. Hernández, R. Sanz, J. Gomez, L.S. Smith, A. Hussain, A. Chella and I. Aleksander. Berlin: Springer.

Balduzzi, D. and Tononi, G. (2008). Integrated Information in Discrete Dynamical Systems: Motivation and Theoretical Framework. *PLoS Computational Biology* 4(6): e1000091.

Balduzzi, D. and Tononi, G. (2009). Qualia: The Geometry of Integrated Information. *PLoS Computational Biology* 5(8): e1000462.

Barrett, A. B. and Seth, A. K. (2011). Practical Measures of Integrated Information for Time-Series Data. *PLoS Computational Biology* 7(1): e1001052.

Feldt, S., Waddell, J., Hetrick, V. L., Berke, J. D. and Żochowski, M. (2009). Functional clustering algorithm for the analysis of dynamic network data. *Physics Review E* 79(5): 056104.

Gazzaniga, M. S. (1970). *The Bisected Brain*. New York: Appleton-Century-Crofts.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37: 424-438.

Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Transactions on Neural Networks* 14(6): 1569-1572.

Kay, K.N., Naselaris, T., Prenger, R.J. and Gallant, J.L. (2008). Identifying natural images from human brain activity. *Nature* 452: 352-55.

Lee U., Mashour G. A., Kim S., Noh G. -J., Choi B. –M. (2009). Propofol induction reduces the capacity for neural information integration: Implications for the mechanism of consciousness and general anesthesia. *Consciousness and Cognition* 18(1): 56-64.

Massimini M, Boly M, Casali A, Rosanova M, Tononi G. (2009). A perturbational approach for evaluating the brain's capacity for consciousness. *Progress in Brain Research* 177:201-214.

Massimini M, Ferrarelli F, Murphy M, Huber R, Riedner B, Casarotto S, and Tononi G. (2010). Cortical reactivity and effective connectivity during REM sleep in humans. *Cognitive Neuroscience* 1(3):176-183.

Metzinger, T. (2003). *Being No One*. Cambridge, Massachusetts: The MIT Press.

Seth, A.K., Izhikevich, E., Reeke, G.N. and Edelman, G.M. (2006). Theories and measures of consciousness: An extended framework. *PNAS* 103(28): 10799–804.

Tononi, G. (2004). An Information Integration Theory of Consciousness. *BMC Neuroscience* 5:42.

Tononi, G. (2008). Consciousness and Integrated Information: a Provisional Manifesto. *Biological Bulletin* 215: 216–242.

Tononi, G. and Sporns, O. (2003). Measuring information integration. *BMC Neuroscience* 4:31.

Tononi, G., Sporns, O. and Edelman, G.M. (1994). A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Proc. Natl. Acad. Sci. USA* 91: 5033-7.