

Effects of parity, frustration, and stochastic fluctuations on integrated information of consciousness for networks with two small-sized loops

Tadaaki Hosaka (✉ hosaka-t@rs.tus.ac.jp)

Tokyo University of Science

Research

Keywords: integrated information theory 3.0, split brain, major complex, bridge, pyphi

Posted Date: April 9th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-385815/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

RESEARCH

Effects of parity, frustration, and stochastic fluctuations on integrated information of consciousness for networks with two small-sized loops

Tadaaki Hosaka

Correspondence:

hosaka.tadaaki@gmail.com
 School of Management, Tokyo
 University of Science, Tokyo,
 Japan
 Full list of author information is
 available at the end of the article

Abstract

Background: Integrated Information Theory (IIT) has been attracting attention as a theory of consciousness. The latest version, IIT3.0, is still at the stage of accumulating knowledge concerning fundamental networks. This paper presents an evaluation of the system-level integrated conceptual information of a major complex, Φ^{Max} , associated with the center of consciousness for a small-scale network containing two small loops in accordance with the IIT3.0 framework. We focus on the following parameters characterizing the system model: 1) number of nodes in the loop, 2) frustration of the loop, and 3) temperature controlling the stochastic fluctuation of the state transition. Specifically, assuming that the two loops are coupled systems, such as cerebral hemispheres, the effect of these parameters on the values of Φ^{Max} and conditions for major complexes formed by a single loop, rather than the entire network, is investigated.

Results: Our first finding is that parity of the number of nodes forming a loop has a strong effect on the integrated conceptual information Φ^{Max} . For loops with an even number of nodes, the number of concepts tends to decrease, and Φ^{Max} becomes smaller. When the loop is formed with an odd number of nodes, the system without frustration and the system with two frustrated loops can have exactly the same Φ^{Max} . It is also shown that, although counterintuitive, the value of Φ^{Max} can be maximized in the presence of stochastic fluctuations. Our second finding is that a major complex is more likely to be formed by a small number of nodes under small stochastic fluctuations. In particular, this tendency is enhanced for larger numbers of nodes constituting a loop. On the other hand, the entire network can easily become a major complex under larger stochastic fluctuations, and this tendency can be reinforced by frustration.

Conclusions: Our results indicating that the entire network dominates and maintains a high level of consciousness in the presence of a certain degree of fluctuation and frustration may qualitatively correspond to actual neural behaviors. The results of this study are expected to contribute to the verification of the consistency of IIT with the actual nervous system in the future.

Keywords: integrated information theory 3.0; split brain; major complex; bridge; pyphi

Background

Integrated Information Theory (IIT), proposed by Giulio Tononi, has been attracting attention as a theory that can mathematically describe the quality and quantity

of consciousness generated in causal systems such as neural networks. Essentially, in IIT, the amount of information obtained by integrating the components of a system into one, rather than dividing them into several parts, is regarded as the quantity of subjective consciousness emerging in the system. Since IIT was first proposed[1], the computational framework has been continually improved[2, 3], and the latest version is IIT3.0[4, 5]. IIT has not been constructed in a bottom-up manner to be consistent with neurophysiological experiments but has been proposed through careful observation of phenomena related to subjective consciousness to represent necessary properties of consciousness in the framework of information theory.

A brief review of IIT3.0[4] is provided in the Appendix to facilitate the understanding of the present work. In IIT research, the integrated conceptual information, Φ , and that of complexes represented as Φ^{Max} , in system-level integration are frequently referred to simply as integrated information. However, to avoid confusion, in this paper, the term “integrated information” is used only for mechanism-level integrated information φ or φ^{Max} , and the term “minimum information partition (MIP)” is also used for mechanism-level integration while the expression “integrated conceptual information” is always used as a reference to Φ or Φ^{Max} . Furthermore, we do not deal with the formation of minor complexes, and the notation Φ^{Max} is utilized only to express integrated conceptual information of a major complex.

Following the release of newer versions of IIT, improvements were reported from various points of view. The metric of integrated information used in IIT2.0[6] can have a negative value, which is a disadvantage. Thus, several improved metrics have been proposed [7, 8, 9]. Oizumi et al. [10] also interpreted IIT from the perspective of information geometry and proposed another metric of integrated information. Mediano et al. [11] applied these metrics to some common networks and compared their usefulness in a unified manner. Another major problem in the calculation of integrated information is the extremely large computational power required. This is an obstacle in verifying the validity of IIT in large-scale systems such as the human brain. For example, the number of node partitioning patterns increases exponentially with system size. Some approximate computational methods have been proposed to quickly find the optimal partitioning pattern [12, 13], which can be applied to the framework of IIT2.0, but not IIT3.0, owing to its complicated hierarchical structure. In another direction including IIT3.0, an attempt to evaluate the quantity of integrated information in the limit of an infinite number of nodes was reported using a technique called mean field approximation in statistical physics[14, 15]. Krohn and Ostwald [16] expressed the integrated conceptual information Φ of the IIT3.0 framework using probabilistic models.

In one application of IIT, an attempt to determine the integrated information from EEG data was reported[17]. Recently, IIT research has grown beyond the original purpose of evaluating the quality and quantity of consciousness. Niizato et al. [18] utilized integrated conceptual information to characterize the behavior of schools of fish.

IIT also deals with the quality of consciousness (qualia) [19]. IIT3.0[4] assumes that a constellation of concepts associated with the value of mechanism-level integrated information φ^{Max} and the core cause and core effect (past and future probability distributions), correspond to the quality of consciousness. The validity

of this assumption needs to be confirmed in the future, and it has recently been suggested that category theory, which mathematically deals with relationships between multiple mathematical structures, can be a powerful tool for this purpose [20, 21].

IIT has influenced other fields and theories as well. IIT claims that consciousness emerges as intrinsic information independent of physical entities, however, Barrett [22] discussed the hypothesis of how consciousness can be related to the fundamental fields of physics. Safron [23] recently proposed a theory that unifies IIT with two other prominent theories of consciousness, the free energy principle [24] and the global neuronal workspace theory [25]. On the other hand, despite the prominence of IIT research, problems related to the formulation of IIT and problems from a philosophical perspective have been pointed out, partly because of the lack of evidence supported by experiments [26, 27, 28]. For IIT to develop in the future, it will be important to verify its consistency with clinical findings.

As mentioned above, many metrics and approximation methods have been proposed for IIT2.0, and much knowledge has been accumulated so far. The most significant development from IIT2.0 to IIT3.0 is the introduction of system-level integrated information. As a result, IIT3.0 has become more complicated with hierarchical computational procedures and increased computational complexity, and no practical approximation methods have been proposed. Analyses for several networks were described in the original article of IIT3.0[4], however, since then, only a few papers have discussed the value of Φ for specific networks. Popiel et al. [29] examined the trend of Φ for a traditional fully connected Ising network of five nodes. In their research, a parameter of temperature was introduced for controlling stochastic fluctuations in the state transitions, and the variation of Φ with respect to the temperature was investigated in the context of phase transitions for the magnetization and magnetic susceptibility of the statistical mechanics. Furthermore, few insights into how major complexes are generated have yet been described. We believe that IIT3.0 is still at the stage of accumulating knowledge concerning fundamental networks.

In this study, we evaluate the integrated conceptual information of a major complex, Φ^{Max} , in networks significant for validating IIT. We consider the well-known clinical finding that patients with a split brain appear to have independent consciousness in their left and right hemispheres[30, 31], and then consider small networks in which two loops corresponding to both hemispheres are connected by a bridge. To enhance the diversity of the system model, incorporating the concept of frustration for loops and stochastic fluctuations for state transitions, we examine multiple topologies with different numbers of nodes in the loops. In particular, we describe how often a small loop becomes a major complex, or conversely, how often the entire network becomes a major complex. Owing to the computational complexity, the maximum number of nodes is kept at eight in our experiments, and it is not possible to derive strict relationships between the split brain and IIT in terms of neurophysiology. Our goal is to provide, in the realm of IIT, rich insight into the factors that make a part, rather than the whole, of the network a major complex.

Methods

In this study, we examine the integrated conceptual information of a major complex, Φ^{Max} , and the corresponding major complex constitution for networks with the five topologies shown in Fig. 1. In the network shown in Fig. 1(a), the loop consisting of nodes 1, 2, and 3 and the loop consisting of nodes 4, 5, and 6 are connected by the edge between nodes 3 and 4. Similarly, in networks (b) and (c), two loops consisting of three or four nodes are connected by an edge. Such an edge, when removed, divides the entire network into two subsystems and is termed a bridge. Bridges are important paths in network efficiency because they are frequently passed through during traffic between two arbitrary nodes. From this perspective, although the number of nodes is extremely small compared with a human brain, the two loops correspond to the two brain hemispheres and the bridge to the corpus callosum. In addition, the networks in Fig. 1(b) are expected to have a stronger tendency to form a major complex in only one loop than in the network (a) because more nodes are included in each loop, and intuitively, the loop is likely to be isolated. For comparison, we also examine the topology shown in Fig. 1(d) and (e). In Fig. 1(d), adding a single edge to the network in Fig. 1(a) makes the edge between the loops non-bridging, and we expect the entire network to have a strong tendency to form a major complex. In the fully connected network shown in Fig. 1(e), where edges are set between all nodes, it appears that the entire network is likely to be a major complex.

To calculate the integrated information according to the IIT framework, we first need to define the transition probability matrix (TPM) of the target network. In this study, the state of node $i (= 1, 2, \dots, N)$ is represented as $S_i (= \pm 1)$, corresponding to firing and non-firing, although the index representing time is not explicit. The input to node i at a given time is represented as

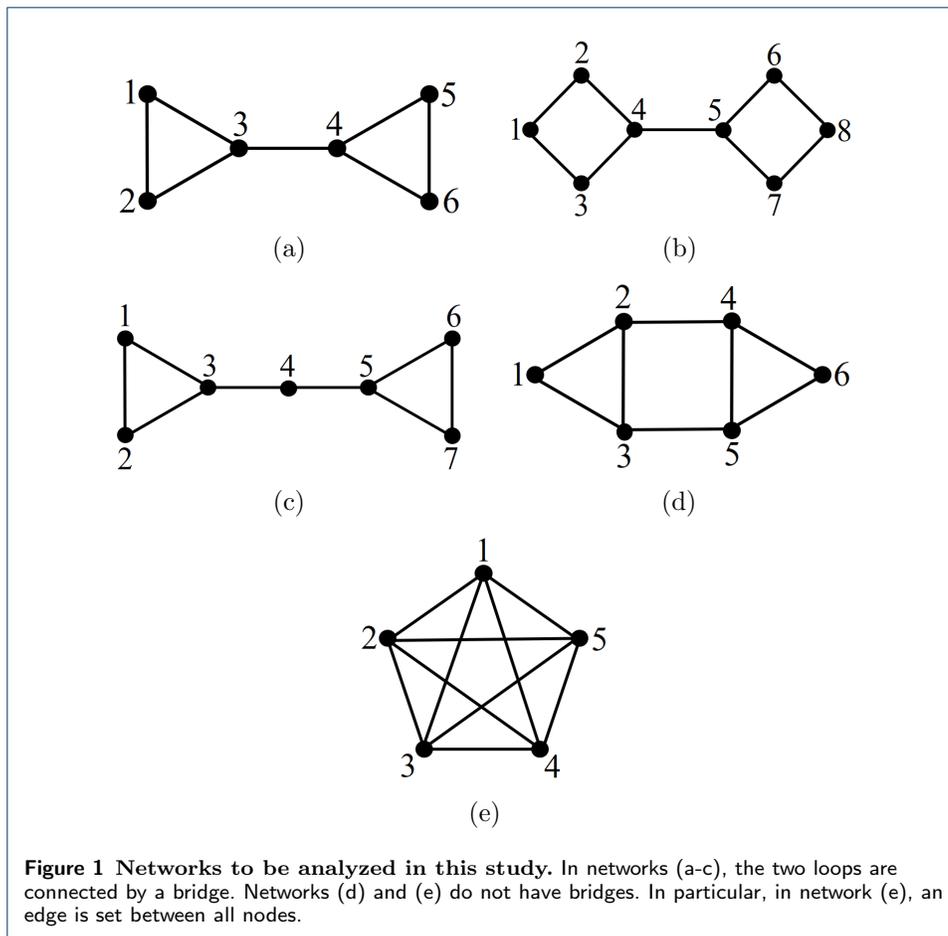
$$\sigma_i = \sum_{j=1}^N J_{ij} S_j, \quad (1)$$

where J_{ij} is assumed to be zero if there is no edge between nodes i and j , and 1 or -1 if there is an edge between nodes. Also, for self-couplings, $J_{ii} = 0$ is assumed. $J_{ij} = 1$ means that nodes i and j tend to assume the same state, and $J_{ij} = -1$ means that nodes i and j tend to assume different states. Using σ_i , the probability $p(S_i)$ that node i will assume $S_i (= \pm 1)$ at the next time step is represented as

$$p(S_i) = \frac{1}{1 + \exp(-2\sigma_i S_i / T)}, \quad (2)$$

where T is a parameter called temperature which controls the accuracy of neuronal behavior, as introduced in the traditional Hopfield model and the above-mentioned work[29]. In this study, we also examine the dependence of Φ^{Max} on temperature T .

All 2^N network configurations are subject to our investigation; however, states which give completely equivalent computational processes of the integrated conceptual information are excluded as duplicates. For example, in the network of Fig. 1(a) with all edges $J_{ij} = 1$, the configurations $S_1 = -1, S_i = 1$ ($i = 2, 3, 4, 5, 6$), and



another configuration $S_2 = -1, S_i = 1$ ($i = 1, 3, 4, 5, 6$) are completely equivalent from the perspective of computing the integrated conceptual information. In such a case, only one of the multiple equivalent states should be considered.

Stable states satisfy the condition of $J_{ij}S_iS_j = 1$ for any edge and its related nodes, and the existence of such a stable state in a loop can be determined by whether the product of all edges J_{ij} included in the loop becomes 1. If the product is -1 , there is no stable state for the loop. Within the field of magnetism in statistical physics, this kind of loop characterized by a negative product is referred to as “frustration”. In this study, we also investigate the effect of the existence of frustrated loops on the value of Φ^{Max} because actual neural circuits are expected to be frustrated with a mixture of inhibitory and excitatory connections. The amount of mechanism-level integrated information φ^{Max} tends to be higher when past and future states are definitively determined based on the current state. Therefore, we presume that integrated conceptual information Φ^{Max} is also smaller in networks with frustrated loops.

Let us define the specific edge settings in our experiments. As a network with no frustration, all edges are set to $J_{ij} = 1$. As a network with frustration in one loop, we set $J_{13} = -1$ in the topologies of Fig. 1(a) and (c), $J_{24} = -1$ in (b), $J_{12} = -1$ in (d) and (e). The other edges are assumed to be 1. For a network with frustration in both loops, we also set the edge located in the symmetric position of the right

loop to -1 . However, for network (e), we set $J_{45} = -1$. Note that in network (d), the loop formed by nodes 2, 3, 5, and 4 is not frustrated in any case.

As described above, for the five topologies shown in Fig. 1, we consider the following cases: a) no frustration, b) frustration in one loop, and c) frustration in both loops, resulting in a total of 15 networks. For each network, we find the major complex and obtain the integrated conceptual information Φ^{Max} for all states except for degenerations which are equivalent from a computational point of view, while changing the temperature parameter T . For our experiments, we use the Python library called PyPhi, published by Mayner et al.[32] for computing integrated conceptual information.

Results

In this section, we present the results of our simulations. The changes in Φ^{Max} with respect to T are shown, followed by the indices related to the formation of the major complex. A deeper discussion of results is presented in the next section.

Integrated conceptual information Φ^{Max} with respect to temperature T

Fig. 2 shows the value of the integrated conceptual information of a major complex Φ^{Max} versus the temperature parameter T for the network in Fig. 1(a). Fig. 2(a) illustrates the result of no frustration, (b) illustrates the result of frustration in one loop, and (c) illustrates the result of frustration in both loops. The legend describes the representative configurations of nodes 1 through 6 as 0/1 in order, which is handled as ± 1 in the calculation of Eqs. (1) and (2) to derive the TPM. The results for configurations that do not appear in the legend are equal or fairly close to one of the graphs in the figure.

For each T in the network without frustration, a particularly large value of Φ^{Max} is observed when the states of all nodes are equal, that is, $S_i = 1$ or $S_i = -1$, which corresponds to stable states in the limit of $T \rightarrow 0$. A series of changes in Φ^{Max} for other configurations is relatively small and tends to have a maximum value in the range of $T = 1$ to $T = 2$, instead of monotonically decreasing with T . Because the mechanism-level integrated information φ can be qualitatively regarded as the extent to which the current state can define the past and future states, it might be intuitively inferred that the value of φ is larger when $T \rightarrow 0$, which has no probabilistic fluctuations. However, the computation of the system-level integrated conceptual information Φ is too complicated to be understood intuitively because it involves the hierarchical computation of unidirectionally partitioning the subsystem, calculating the mechanism-level integrated information, and measuring the distance between conceptual structures. It is not incomprehensible that Φ^{Max} may be larger in the presence of stochastic fluctuations, which is further investigated in the Discussion section. On the other hand, when the value of T is high, the value of Φ^{Max} approaches zero regardless of node configuration. For $T \rightarrow \infty$, the transition probability between any two configurations approaches $1/2^N$. This also holds for the partitioned system, and it is obvious that there is no difference between the two systems before and after partitioning in the calculation process of Φ , resulting in $\Phi \rightarrow 0$.

When a single loop is frustrated, no such prominently large Φ^{Max} is observed, unlike in the network where no frustration is present. One of the configurations with

a moderately large Φ^{Max} value is $\{S_i\} = \{1, -1, 1, 1, 1, 1\}$. Considering $J_{13} = -1$, we see that the state at the next time step for this configuration is uniquely determined in the limit of $T \rightarrow 0$. We believe that this causal uniqueness leads to large values of Φ^{Max} . As in the case of the network without frustration, in some series the maximum value of Φ^{Max} is observed within the range of $T = 1$ to $T = 2$.

A series of changes in Φ^{Max} for the network with frustration in the two loops is identical to that of the network without frustration, indicating that both networks have causally equivalent configurations. We examine this in more detail in the Discussion section. It should be noted that the current IIT version cannot handle the time evolution of states. Even if two distinct networks have the same value of Φ^{Max} at any instant, either one may transition to another state with a different value of Φ^{Max} at the next time point or may move between multiple states while maintaining a constant value of Φ^{Max} . As an example of one of the configurations that exhibits a pronouncedly large Φ^{Max} , let us consider $\{S_i\} = \{1, -1, 1, -1, -1, 1\}$ in a network with two frustrated loops. Considering $J_{13} = J_{45} = -1$ in our setting, we can see that this configuration transitions at the next time point with probability 1 to a flipped version of the current states in $T \rightarrow 0$. Thus, the system maintains a high value of Φ^{Max} switching between two mirror states. This is quite different from the situation in which a network without frustration maintains only one stable state of $S_i = 1$, as described above.

The integrated conceptual information Φ^{Max} for the networks in Fig. 1(b, c, d, e) is shown in the Additional file 1. For networks (c, d), the tendency is qualitatively similar to the case of network (a). Namely, the following are observed: 1) In the network without frustration, Φ^{Max} is remarkably high in the stable state of $T \rightarrow 0$, that is, all the nodes are 1 or -1 , and the integrated conceptual information is relatively low in other states while reaching a maximum value in the range of $T = 1$ to $T = 2$. 2) In networks with a single frustrated loop, no state with a remarkably high value of Φ^{Max} is observed. 3) The graph illustrating the change in Φ^{Max} in networks with two frustrated loops is identical with that of the network without frustration. On the other hand, networks (b, e) show a different trend from networks (a, c, d). Specifically, we find that 1) Φ^{Max} has a maximum value near $T = 0$ and overall decreases as T increases; 2) the stable state in $T \rightarrow 0$ without frustration has a relatively large Φ^{Max} , but the degree of protrusion is limited; and 3) a series of changes in Φ^{Max} for the network with two frustrated loops is not identical with that of the network without frustration.

Moreover, considering all results, we can see that the scale of the Φ^{Max} differs greatly depending on network topology. In comparing network (c) with network (a), it has only one more node added, but Φ^{Max} is roughly doubled. On the other hand, although network (b) has the largest number of nodes, Φ^{Max} is at most approximately 1.6, which is notably lower than the other networks. The number of nodes constituting a loop in this network is even (four), unlike networks (a, c, d), and the parity of the number of nodes has a significant influence on the number of concepts in conceptual structure, directly affecting the value of Φ^{Max} . We explore this in more depth in the Discussion section.

Formation of major complexes

We examine the effect of network topology on the tendency of a subsystem to become a major complex. Two kinds of ratios of the following events among the 2^N configurations are investigated for each network:

Ratio1 Major complex is formed by less than or equal to half of the total number of nodes N ,

Ratio2 A major complex is formed by all nodes.

However, in practice, configurations that are essentially equivalent, such as mirror states, are excluded instead of duplicates being counted, thus reducing the total number of configurations to less than 2^N .

Two types of statistics are obtained for each ratio by setting different ranges of target T . First, when calculating these indices, any configuration that satisfies each event at any one of the simulation temperatures T is counted. Second, in order to clarify the dependence of the ratios on the parameter T , only cases satisfying the requirement at any point among $T \geq 1.5$ are counted for ratio1, and only cases satisfying the requirement at any point among $T \leq 0.1$ are counted for ratio2.

Ratio1, the rate of major complexes formed by less than or equal to half of the total number of nodes N , is shown in Fig. 3. The horizontal axis represents the network topology with the number of frustrated loops, and the vertical axis represents the ratio. The bar graph shows statistics for the entire temperature range, and the line graph shows statistics for $T \geq 1.5$. As expected, this ratio is larger for networks containing bridges in Fig. 1(a, b, c) than for networks without bridges in Fig. 1(d, e). The formation of major complexes with such a small number of nodes is frequently observed when values of T are relatively small. As shown by the green lines, the ratio significantly decreases in the presence of relatively large stochastic fluctuation, $T \geq 1.5$, and in particular, becomes zero for networks (d, e). In the next section, we discuss the reason why major complexes with a small number of nodes are likely to occur under low stochastic fluctuation. Furthermore, this ratio is especially large for network (b), which has more nodes in the loops on both sides of the bridge. However, in network (b), it has been observed that nodes $\{1, 2, 3, 4\}$ or $\{5, 6, 7, 8\}$ rarely form a major complex even for low T , and a major complex is often composed of three or fewer nodes.

Ratio2, the rate of the major complexes formed by all nodes, is shown in Fig. 4. The bar graph depicts statistics for the entire temperature range, and the line graph depicts statistics for $T \leq 0.1$. For the entire range of T , this index is almost 1 except for network (b), which also implies that it is common for the entire network to be integrated into a major complex. However, this ratio significantly decreases in the absence of stochastic fluctuation, $T \leq 0.1$, for networks (a, b, c) with bridges. On the other hand, no significant temperature dependence is observed in networks (d, e) without a bridge. Another feature of this index is that the effect of frustration is significant in network (b). In this network, the presence of frustration reduces the number of completely random nodes with $p(S_i = \pm 1) = 0.5$, encouraging the entire network to become a major complex, and conversely, preventing a small number of nodes from becoming a major complex. The role of frustration is explained in detail in the Discussion section.

These results indicate that in a system with a bridge, each loop tends to dominantly determine the integrated conceptual information, especially when the

stochastic fluctuations of state transitions are small, while the entire network tends to become a major complex when the stochastic fluctuations are large or when the network is densely connected.

Discussion

Congruence of Φ^{Max} in cases of no frustration and two frustrations

We consider what causes the congruence of Φ^{Max} in the case of no frustration and of two frustrations for the network in Fig. 1(a, c, d). Fig. 5(a) shows two different networks, where the white and black circled nodes represent $S_i = +1$ and $S_i = -1$, respectively, the solid and dashed edges represent $J_{ij} = +1$ and $J_{ij} = -1$, respectively. The left network is free of frustration, and the right network has frustration in each loop. The numbers beside each node are values of σ_i , which are derived by Eq. 1. In these networks, the absolute value of σ_i for each node is equal. For nodes with different signs of σ_i between the two networks, the values of $p(S_i = +1)$ and $p(S_i = -1)$ are also flipped. However, this probability flip has no effect on the distance between the non-partitioned and partitioned repertoires measured by the earth mover's distance (EMD), because the ground metric of EMD both for calculating φ and Φ is defined by the Hamming distance, that is, the discrepancy of each node state. It can be said that $1/-1$ representing the state of each node is simply a label. Thus, these two networks form an equivalence causality between two adjacent times, resulting in equal values of Φ^{Max} . Fig. 5(b) depicts another example, in which interchanging the positions of the two rightmost nodes (node 5 and node 6 in Fig. 1(a)) does not affect causality. When a loop is formed with three nodes, it is always possible to set the configuration such that the value of Φ^{Max} is equal for both networks with and without frustration in the two loops. On the other hand, such a counterpart cannot exist in a network with only one frustrated loop. Moreover, when the loop is formed by four nodes, it is not always possible to determine configurations that give values of Φ^{Max} common to systems with different numbers of frustrations.

In IIT3.0 [4, 5], the constellation of concepts corresponds to the qualia which the network experiences as subjective consciousness. Although the two networks mentioned above have an equal value of Φ^{Max} , their constellations are different because the sign flip of σ_i alters the constellation, but not EMD. Furthermore, as mentioned in the Results section, considering time evolution, the two networks do not necessarily maintain common values of Φ^{Max} as they transition from the current states. Therefore, these two networks can have the same value of Φ^{Max} simultaneously, but the qualia occurring in them are different. However, in the exceptional case where the sign flip is present in all nodes, IIT considers the same qualia occur in these two networks and designates such constellations isomorphism [19]. The quest for conditions that allow two different networks to maintain equal values of Φ^{Max} during time evolution, is a future challenge.

Maximum of Φ^{Max} in the range of $T = 1$ to $T = 2$

We examine why Φ^{Max} has a maximum in the range of $T = 1$ to $T = 2$, depending on the network topology and node configuration. The value of φ^{Max} associated with each concept comprising the conceptual structure can be interpreted as the

degree to which the current state defines past or future states. Temperature T is a parameter that determines the accuracy of neuronal response, and a larger T makes the neuron's behavior more stochastic. Thus, we might intuitively expect that as T increases, the value of φ^{Max} will decrease, and consequently, Φ^{Max} will also decrease. Indeed, it is obvious that in the limit of $T \rightarrow \infty$, $\Phi^{\text{Max}} \rightarrow 0$, but the experimental results suggest that Φ^{Max} does not necessarily show a monotonic decrease.

This non-monotonicity for Φ^{Max} is explained using a simple network of four nodes, as shown in Fig. 6(a). Let us assume that states of all nodes are identical, that is, $S_i = 1$ or $S_i = -1$. The dependence of Φ^{Max} on T for this network is shown by the blue line in Fig. 6(b), which also shows a maximum at $T = 1.5$. For $T \geq 0.4$, a major complex is formed by all nodes, and the conceptual structure consists of 14 concepts as only the combination of nodes 3 and 4 in the power set is not a concept.

Let us take as an example a mechanism consisting of node 1 and node 4 with all nodes being a subsystem, for which the optimal purview is a combination of node 2 and node 3 at any temperature. Representing the mechanism-level integrated information obtained for the optimal past and future purviews as $\varphi_{\text{cause}}^{\text{Max}}$ and $\varphi_{\text{effect}}^{\text{Max}}$, respectively (see the Appendix), in this case, the relation $\varphi_{\text{cause}}^{\text{Max}} < \varphi_{\text{effect}}^{\text{Max}}$ holds, and φ^{Max} is determined by $\varphi_{\text{cause}}^{\text{Max}}$ indicated by the orange line in Fig. 6(b). Fig. 6(c) shows the change in the probability of each bin of the cause repertoire with respect to T . Note that only the probabilities for the four configurations determined by the states of nodes 2 and 3 in the purview are shown. The solid lines show the values of the core cause, that is, the cause repertoire for the non-partitioned purview. The dashed lines show the values for the purview partitioned by MIP. Within the limit of $T \rightarrow \infty$, the probability of each bin in both non-partitioned and partitioned repertoires approaches the value of a uniform distribution (in this case, $1/4$). However, the speed of change as T rises is not constant, and it is probable that the magnitudes of probabilities for the non-partitioned repertoire and the partitioned repertoire are reversed somewhere in the middle, as in the bin of $\{11\}$. Additionally, as shown in the bin of $\{01\}$ in the partitioned repertoire, the change may not be monotonous and rather a mixture of decrease and increase.

These can be seen in the repertoire for three reasons. First, the value of σ_i (Eq. 1) is different depending on the node, and consequently, the derivative of Eq. 2 for a specific value of T differs for each node. Second, the operation of partitioning is computationally realized by the marginalization of TPM, which also makes the differential coefficients inconsistent between the non-partitioned and partitioned mechanism/purview. Third, especially for the cause repertoire, the value of each bin indirectly affects the others because normalization is performed during the calculation based on the Bayes' formula.

φ is obtained by EMD between non-partitioned and partitioned repertoires, and due to factors mentioned above, even φ^{Max} can exhibit non-monotonicity with respect to T . It is natural that Φ^{Max} , which is calculated using the set of φ^{Max} , can show non-monotonicity. In the network of this study, σ_i is likely to take on a value of one or two, and the probability defined in Eq. 2 has a large slope around $T = 1.0$. Therefore, it is speculated that the above-mentioned effects are most apparent at these temperatures.

It is believed that stochastic uncertainties, such as thermal and spontaneous fluctuations, exist in actual neuronal behaviors. These neuronal fluctuations are expected to be well modeled by Eq. 2 and the value of T giving the maximum value of Φ^{Max} , rather than larger values of T which make the behavior of the neuron too imprecise. Further work is needed to link the results of the present study to real neuronal activity and the amount of consciousness.

Relationship between network topologies and magnitudes of Φ^{Max}

We investigated the value of Φ^{Max} versus T for the five types of network topologies shown in Fig. 1. From Fig. 2 and the Additional file 1, it can be seen that the scale of the value of Φ^{Max} differs greatly depending on network topology. We discuss the causes of these differences.

Fig. 7 shows the number of concepts included in the major complex for the five network topologies investigated in this study, under conditions in which loops are not frustrated and states of all nodes are identical. In general, the larger the number of concepts in a conceptual structure, the larger the Φ^{Max} , as was also mentioned in [4]. The figure includes the number of concepts in a non-partitioned system (solid line) and that in a system partitioned by the optimal unidirectional cut (dashed line). The value of Φ^{Max} is significantly affected by the difference in the number of concepts between the non-partitioned and partitioned systems, and it can be confirmed that these differences roughly correspond to the order of magnitude of Φ^{Max} for the five topologies.

Even though networks (a, b, c) in Fig. 1 have similar topologies with two loops connected by a bridge, a significant difference in the number of concepts occurs. We believe that this is due to the parity of the number of nodes comprising the loop. Here, we consider a network made up of a single loop, as shown in Fig. 8(a). The state of each node is probabilistically determined by the state of the nodes connected by edges and the type of edge J_{ij} , using Eqs. (1) and (2). Therefore, the state S_i at time t can affect the state of only its neighboring nodes connected by the edge at time $t + 1$, and not the states of itself nor of the non-adjacent nodes. Similarly, the state S_i at time t is affected only by the state at time $t - 1$ of its neighboring nodes.

Thus, in the single-loop network shown in Fig. 8(a), two neighboring nodes cannot be a concept, except when the number of nodes constituting a loop, M , is three. For example, in the loop of $M = 4$, when nodes 1 and 2 constitute a mechanism with all nodes a subsystem, partitioning any purview into a group of nodes that are causally related to node 1 (i.e., nodes 2 and 4) and a group of nodes causally related to node 2 (i.e., nodes 1 and 3) will result in $\varphi^{\text{Max}} = 0$ (Fig. 8(b)). Note that a similar argument does not hold for the case of $M = 3$.

Extending this idea, we see that in a loop with M of even numbers, only a combination of every other node can be a concept. For example, in a loop of $M = 6$, nodes $\{1, 3, 5\}$ and nodes $\{2, 4, 6\}$, etc., can be a concept ($\varphi^{\text{Max}} > 0$), while nodes $\{1, 3, 4, 5\}$ lead to $\varphi^{\text{Max}} = 0$ by breaking any purview into a group that has causal relations with nodes $\{1, 3, 5\}$ (i.e., nodes $\{2, 4, 6\}$) and a group that has causal relations with nodes $\{2, 4, 6\}$ (i.e., nodes $\{1, 3, 5\}$), as in the case of Fig. 8(b). It is also clear that a concept consisting of the entire network can only be generated when M is an odd number.

Table 1 Φ^{Max} and the number of concepts in the network consisting of a single loop. The parity of the number of nodes M affects the number of concepts and consequently the value of Φ^{Max} .

Number of nodes M	Φ^{Max}	Number of concepts	
		Non-partitioned	Partitioned
3	1.846	7	2
4	0.302	6	6
5	8.254	21	6
6	2.200	14	10

From the above discussion, it can be expected that the number of concepts is significantly reduced for loops of even M . For the four networks shown in Fig. 8(a), the value of Φ^{Max} and the number of concepts in non-partitioned and unidirectionally partitioned systems are shown in Table 1 under the conditions of $T = 1.0$ and identical node states. It can be seen that the value of Φ^{Max} and the number of concepts repeatedly increase or decrease as M changes. Note that the value of Φ^{Max} is extremely low for $M = 4$ because the same number of concepts are obtained in partitioned and non-partitioned systems, and that the value of Φ^{Max} for $M = 6$ is also close to that of $M = 3$. Although our network of interest, shown in Fig. 1, does not consist of a single loop, we believe that the main cause of the extremely low Φ^{Max} value of network (b) is the parity of nodes composing the loop. On the other hand, we speculate that the increase in the number of node combination patterns which can directly influence the concept generation contributes to the high Φ^{Max} of network (c).

A real neuronal network contains loops of various sizes. Since no analysis focusing on parity has been conducted thus far, new findings may be obtained by paying special attention to parity when examining IIT with actual brain data in the future.

Generation of small-sized major complexes with small T

Ratio1 is the fraction of major complexes formed by less than or equal to $N/2$ nodes and tends to be larger when parameter T is smaller. In this section, we investigate possible reasons for this trend. In particular, for $T < 0.04$, $p(S_i = 1)$ can be approximated as 1 or 0, which corresponds to the positive and negative of σ_i . This means that for $T \approx 0$, the state of each node at the next time step is determined by a majority vote of the current states of its neighboring nodes connected by edges. Additionally, in the case of $\sigma_i = 0$, the state of the next time step for such a node is determined at random, because $p(S_i = \pm 1) = 0.5$, regardless of T . These two effects of majority voting and complete randomness can reduce the number of concepts in the entire network, and consequently make the subsystem composed of a small number of nodes more likely to become a major complex, as explained below.

Firstly, to illustrate the effect of majority voting, we consider the network shown in Fig. 9(a), in which impact of voting is pronounced. Representation for nodes and edges are the same as in Fig. 5. In this network, when all the nodes are set as a subsystem, no concept spanning nodes 3 and 4 can be generated in the case of $T \approx 0$. This is because for any given purview, if a cut is set up to divide all nodes into two groups separated by an edge between nodes 3 and 4, while allowing one group to be an empty set, it is always possible to achieve $\varphi^{\text{Max}} = 0$. Note that this

can be done only if $p(S_i = \pm 1)$ is approximated as 1 or 0. As a result, only concepts consisting of combinations of elements from the node set $\{1, 2, 3\}$ or the node set $\{4, 5, 6\}$ can exist.

Furthermore, for a partitioned subsystem with the unidirectional cut separating nodes $\{1, 2, 3\}$ and $\{4, 5, 6\}$, the obtained concepts are exactly the same as for the non-partitioned system, considering this unidirectional cut has no effect on the majority decision in $T \approx 0$. For a mechanism that includes node 3 or node 4, $\varphi_{\text{cause}}^{\text{Max}}$ or $\varphi_{\text{effect}}^{\text{Max}}$ is slightly different from the non-partitioned system, so the value of Φ^{Max} does not vanish, although it is expected to be sufficiently small. In this network, the value of Φ^{Max} is larger when, for example, a node set $\{4, 5, 6\}$ is a subsystem and a unidirectional cut is placed in it, than when the entire network is a subsystem.

On the other hand, as the value of T increases, the situation of strict majority voting terminates, and middle values, $0 < p(S_i = 1) < 1$, need to be handled. Unlike the situation with $T \approx 0$, partitions can change the value of transition probability, and then concepts spanning nodes 3 and 4, such as a node set $\{1, 2, 4, 5\}$, can be generated. In fact, in the network shown in Fig. 9(a), the entire network has 29 concepts at $T = 1.5$, compared to 14 concepts at $T = 0.03$. As a result, a major complex is formed by all nodes in the region of large T . The number of concepts for networks other than the one discussed here are also likely reduced due to “insensitivity” to σ_i caused by $T \approx 0$, resulting in a tendency to generate small major complexes.

Secondly, to illustrate the effect of complete randomness, we consider the network shown in Fig. 9(b). In this case, regardless of T , nodes 1 and 2 have $p(S_i = 1) = 0.5$. Concepts involving nodes adjacent to such random nodes are less likely to occur. For example, consider a mechanism consisting of nodes 1 and 3, adjacent to node 2 of a random variable. In this case, to avoid trivial vanishing φ^{Max} , node 2, which has a causal relationship with both nodes 1 and 3, must be included in the purview. However, by setting the partition to isolate only node 2, $\varphi^{\text{Max}} = 0$ is eventually obtained for any purviews. This is because node 2 originally has $p(S_i = 1) = 0.5$, thus, isolating it does not change its substantive causal relationship, resulting in at least $\varphi_{\text{effect}}^{\text{Max}} = 0$. A subsystem with many complete random nodes has a smaller number of concepts, and as a result, is less likely to become a major complex. Additionally, to the effect of majority voting, for the network in Fig. 9(b), when T is small, the subsystem consisting of nodes $\{4, 5, 6\}$ is a major complex. As another example, for the network shown on the left of Fig. 9(c), where only one node has $p(S_i = 1) = 0.5$, all nodes form the major complex even if T is small. On the other hand, the network on the right side of Fig. 9(c) has four nodes with $p(S_i = 1) = 0.5$, and the size of the major complex remains small even if T is large.

As described in the previous section, the value of Φ^{Max} can be larger in the range of $1 \leq T \leq 2$ than in $T \approx 0$, and we optimistically expected that this temperature range with Eq. (2) was able to express the uncertainty of real neurons. In this section, we argued that within this temperature range, the entire network has a strong tendency to become a major complex. This may be consistent with the fact that our consciousness is usually an integration of information from the left and right hemispheres. Although these ideas are only self-serving conjectures from the results of the present study, they may also become important in the future in verifying the consistency of IIT with a real human brain.

Relation between ratio1 and the size of loops

As described in the previous subsection, more nodes with $p(S_i = \pm 1) = 0.5$ tend to generate fewer concepts and reduce the size of the major complex. A necessary condition for complete randomness is that the number of edges coming into a node be an even number, and it can be said that $\sigma_i = 0$ is most likely to occur when two nodes are connected with each other, which is satisfied by the loop in our study. Therefore, it can be inferred that by increasing the size of the loop, the major complex is more likely to be formed by a small number of nodes, rather than by all nodes.

In particular, because the loop in the network shown in Fig. 1(b) consists of an even number (four) of nodes, the entire four nodes in the loop cannot become a concept, as mentioned earlier, further reducing the size of the major complex. For network (b), not only is the ratio1 larger than that of networks (a, c), but a huge proportion of the major complex is formed by only two nodes, such as nodes 2 and 4.

The computation of Φ^{Max} based on IIT3.0 requires operations of exponential order on N , and it is not practical to increase N unless an approximate method is used. However, it can be inferred that the value of ratio1 becomes large with increasing N , especially when loop size is even. Since the human cerebral hemisphere is not a single loop, it is necessary to analyze networks with complex structures, such as those with a mixture of multiple loop sizes.

Effect of frustration on major complexes

Finally, let us discuss the effect of frustration on the formation of major complexes. As illustrated in Fig. 4, the difference caused by the presence of frustration is pronounced in ratio2, the percentage of major complexes formed by the entire network, especially for the network in Fig. 1(b). The rate for network (b) is the lowest in the absence of frustration and increases with the number of frustrated loops.

To investigate the rationale for this, we focus on two networks consisting of four nodes, as shown in Fig. 10. The representation for nodes and edges are the same as in Fig. 5. The network on the left side of Fig. 10(a) has no frustration, and all nodes have $p(S_i = \pm 1) = 0.5$, where even a pair of nodes cannot become a concept, as shown in Fig. 9(b). As a result, the major complex is not formed by all nodes; in this case, two nodes, such as nodes 3 and 4, become a major complex, and its conceptual structure consists of only several single nodes.

On the other hand, in the network of Fig. 10(b) with one edge inverted, the presence of frustration reduces the number of random nodes with $p(S_i = \pm 1) = 0.5$ to only two, and node sets $\{1, 3\}$ and $\{2, 4\}$ can be a concept. As a result, all four nodes can form a major complex. However, it should be noted that a concept with all nodes cannot exist. For a loop consisting of four nodes, not all nodes can have $p(S_i = \pm 1) = 0.5$ in the presence of frustration, regardless of the node configuration. Therefore, we can state that frustration plays a role in mitigating randomness characterized by $p(S_i = \pm 1) = 0.5$, facilitating the entire network to become a major complex.

Network (b) has two loops consisting of four nodes, and because of the effect described above, ratio1 decreases and ratio2 increases with the number of frustrations. However, the overall trend described in the earlier section, that two nodes

often form a major complex when T is small, and that the entire network is likely to be a major complex, especially when T is large, is preserved.

When three nodes comprise a loop, not all nodes can have $p(S_i = \pm 1) = 0.5$, regardless of the presence of frustration. We speculate that this is a primal reason why no significant relationships between the number of frustrations and the two ratios were observed for networks (a, c). In the case of a loop consisting of five or more nodes, if not all, but four or five neighboring nodes have $p(S_i = \pm 1) = 0.5$ in a cluster, the entire loop is not expected to be a major complex, although it may depend on the value of T . In the actual neural networks, a single neuron belongs to many loops. It is interesting to investigate relations between frustration and formation of major complexes in such complicated network structures.

Conclusions

In this study, we evaluated the integrated conceptual information Φ^{Max} for small networks with two loops according to the IIT3.0 framework. The experimental results indicated that the parity of the number of nodes, the presence of frustration, and the stochastic fluctuation of state transitions strongly affected the magnitude of Φ^{Max} and the formation of major complexes. As for the value of Φ^{Max} , the following trends were often observed: 1) Φ^{Max} did not monotonically decrease with respect to the magnitude of the fluctuation, but reached its maximum in the presence of some degree of stochastic fluctuation, and 2) even numbers of nodes in the loop reduced the number of concepts, resulting in a smaller value of Φ^{Max} . As for the formation of major complexes, the following trends were often observed: 1) major complexes were easily formed by less than half of the nodes in networks with bridges, as compared to networks without bridges, especially in regions where stochastic fluctuations were small, and 2) frustration in some networks reduced the number of random nodes, facilitating the formation of major complexes in the entire network.

Actual neuronal behavior fluctuates and frustrations exist in real neuronal circuits due to the mixture of excitatory and inhibitory connections. Under these circumstances, we usually maintain a high level of consciousness with integrated information from the left and right hemispheres. This fact corresponds to our results in $1 \leq T \leq 2$ on an analogy level at this stage. It is expected that the results of this study will help to establish the consistency of IIT with the real brain and consciousness in the future.

Appendix

Brief review of IIT3.0

Suppose that the entire network consists of N nodes. It is assumed that the states of the nodes are probabilistically updated as a discrete Markov process and that the states of the nodes in the network at the current time and the transition probability matrix (TPM) between the states are known. IIT postulates a hierarchical structure: the consciousness of the entire network is dominated by the subsystem with the largest integrated conceptual information, which is specified as a set of integrated information of smaller subsets termed mechanisms.

Mechanism-level integrated information

Based on the current state of the nodes in the mechanism and the Bayes' formula, the probability distributions of the states for a specific set of nodes one step before and after, termed the cause and effect repertoires, respectively, are evaluated with nodes outside the subsystem fixed values. The set of nodes for which the probability distributions are computed is termed a purview and may differ from the mechanism, although it must be a subset of the subsystem. In addition, the purviews of the cause and effect repertoires are defined separately.

The nodes in the mechanism (current) and the purview (past/future) are independently partitioned into two groups, assuming that one group of the mechanism is not causally related to one group of purviews, and the other group of the mechanism is not causally related to the other group of purviews. Realizing the assumption by marginalizing the TPM, the cause and effect repertoires are also calculated for this partitioned mechanism and purview. Then, the distance between the repertoires for non-partitioned and partitioned purviews is evaluated by the earth mover's distance (EMD) with the ground metric defined as the Hamming distance. The partition minimizing the EMD is termed the minimum information partition (MIP), and this minimum value is defined as the integrated information φ .

The integrated information is evaluated for all possible purviews in the same way as described above. The maximal value among the past is termed the maximally irreducible cause $\varphi_{\text{cause}}^{\text{Max}}$ and the maximal value among the future is termed the maximally irreducible effect $\varphi_{\text{effect}}^{\text{Max}}$. The cause and effect repertoires obtained for the optimal non-partitioned purviews are termed the core cause and core effect, respectively. The lower value between $\varphi_{\text{cause}}^{\text{Max}}$ and $\varphi_{\text{effect}}^{\text{Max}}$ is defined as the quantity of consciousness for the target mechanism φ^{Max} , which is also often referred to simply as the integrated information.

System-level integrated information

Next, the quantity of consciousness in an upper level, that is, a subsystem and the entire network, is defined on the basis of the mechanism-level integrated information. For any possible mechanism, that is, any combination of nodes in the subsystem, the computation described in the previous section is performed. Then, a set of mechanisms satisfying $\varphi^{\text{Max}} > 0$ is obtained. A mechanism with $\varphi^{\text{Max}} > 0$ is termed a concept, and the set of concepts is called the conceptual structure.

We consider a unidirectionally partitioned system in which the nodes in the subsystem are divided into two groups; however, the causality between the different groups is unidirectionally removed. For all patterns of unidirectional cuts, the conceptual structure is derived according to the previously described procedures. Then, the distance between the conceptual structures of the non-partitioned subsystem and each partitioned subsystem is calculated using an extended version of the EMD. The extended EMD is defined as the sum of the transportation costs of moving φ^{Max} of each concept in the non-partitioned subsystem to the corresponding concept in the partitioned subsystem where the excess or unmatched portion within the value of φ^{Max} , if any, is replaced by the transportation cost to the repertoire obtained under the unconstrained condition. The unidirectional cut that minimizes this distance is also called MIP, as in the case of the mechanism level, and the minimum

value is termed the integrated conceptual information Φ , which is also often referred to simply as the integrated information.

Finally, the procedure described thus far is performed for all possible subsystems in the network. The subsystem that gives the maximum value of the integrated conceptual information is termed a major complex. Then, amongst the subsystems that contain nodes not included in the major complex, the node which gives the maximum value of Φ is termed a minor complex. Subsequently, the next minor complex is obtained in the same way, focusing only on nodes that are not included in any of the previously selected complexes. In short, a complex can be interpreted as a subsystem whose Φ is a local maximum. The integrated conceptual information of a complex is denoted by Φ^{Max} , and its conceptual structure is termed a maximally irreducible conceptual structure. IIT3.0 concludes that the value of Φ^{Max} for complexes, especially for a major complex, is regarded as the quantity of consciousness in the network, analogous to the consciousness of the entire brain.

Acknowledgements

Not applicable.

Funding

Not applicable.

Abbreviations

IIT: Integrated information theory; MIP: Minimum information partition; TPM: Transition probability matrix; EMD: Earth movers' distance

Availability of data and materials

All supporting data are submitted in Additional files.

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Authors' contributions

TH carried out all aspects of this study.

Authors' information

Not applicable.

Author details

School of Management, Tokyo University of Science, Tokyo, Japan.

References

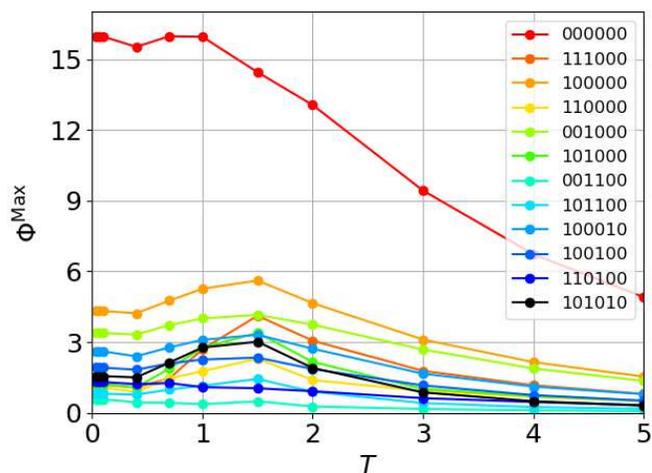
1. Tononi, G.: An information integration theory of consciousness. *BMC neuroscience* **5**, 42 (2004). doi:10.1186/1471-2202-5-42
2. Tononi, G.: Consciousness as integrated information: A provisional manifesto. *The Biological Bulletin* **215**, 216–42 (2009). doi:10.2307/25470707
3. Tononi, G.: Integrated information theory of consciousness: An updated account. *Archives Italiennes de Biologie* **150**, 56–90 (2012). doi:10.4449/aib.v149i5.1388
4. Oizumi, M., Albantakis, L., Tononi, G.: From the phenomenology to the mechanisms of consciousness: Integrated information theory 3.0. *PLOS Computational Biology* **10**(5), 1–25 (2014). doi:10.1371/journal.pcbi.1003588
5. Tononi, G., Boly, M., Massimini, M., Koch, C.: Integrated information theory: From consciousness to its physical substrate. *Nature Reviews Neuroscience* **17** (2016). doi:10.1038/nrn.2016.44
6. Balduzzi, D., Tononi, G.: Integrated information in discrete dynamical systems: Motivation and theoretical framework. *PLOS Computational Biology* **4**(6), 1–18 (2008). doi:10.1371/journal.pcbi.1000091
7. Barrett, A.B., Seth, A.K.: Practical measures of integrated information for time-series data. *PLOS Computational Biology* **7**(1), 1–18 (2011). doi:10.1371/journal.pcbi.1001052
8. Seth, A.K., Barrett, A.B., Barnett, L.: Causal density and integrated information as measures of conscious level. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **369**(1952), 3748–3767 (2011). doi:10.1098/rsta.2011.0079

9. Oizumi, M., Amari, S., Yanagawa, T., Fujii, N., Tsuchiya, N.: Measuring integrated information from the decoding perspective. *PLOS Computational Biology* **12**(1), 1–18 (2016). doi:10.1371/journal.pcbi.1004654
10. Oizumi, M., Tsuchiya, N., Amari, S.: Unified framework for information integration based on information geometry. *Proceedings of the National Academy of Sciences* **113**(51), 14817–14822 (2016). doi:10.1073/pnas.1603583113
11. Mediano, P., Seth, A., Barrett, A.: Measuring integrated information: Comparison of candidate measures in theory and simulation. *Entropy* **21**, 17 (2018). doi:10.3390/e21010017
12. Kitazono, J., Kanai, R., Oizumi, M.: Efficient algorithms for searching the minimum information partition in integrated information theory. *Entropy* **20**, 173 (2017). doi:10.3390/e20030173
13. Toker, D., Sommer, F.T.: Information integration in large brain networks. *PLOS Computational Biology* **15**(2), 1–26 (2019). doi:10.1371/journal.pcbi.1006807
14. Aguilera, M., A. Di Paolo, E.: Integrated information in the thermodynamic limit. *Neural Networks* **114**, 136–146 (2019). doi:10.1016/j.neunet.2019.03.001
15. Aguilera, M.: Scaling behaviour and critical phase transitions in integrated information theory. *Entropy* **21**(12), 1–24 (2019). doi:10.3390/e21121198
16. Krohn, S., Ostwald, D.: Computing integrated information. *Neuroscience of Consciousness* **2017**(1) (2017). doi:10.1093/nc/nix017
17. Kim, H., Hudetz, A.G., Lee, J., Mashour, G.A., Lee, U.: Estimating the integrated information measure phi from high-density electroencephalography during states of consciousness in humans. *Frontiers in Human Neuroscience* **12**, 42 (2018). doi:10.3389/fnhum.2018.00042
18. Niizato, T., Sakamoto, K., Mototake, Y., Murakami, H., Tomaru, T., Hoshika, T., Fukushima, T.: Finding continuity and discontinuity in fish schools via integrated information theory. *PLOS ONE* **15**(2), 1–29 (2020). doi:10.1371/journal.pone.0229573
19. Balduzzi, D., Tononi, G.: Qualia: The geometry of integrated information. *PLOS Computational Biology* **5**(8), 1–24 (2009). doi:10.1371/journal.pcbi.1000462
20. Tsuchiya, N.: “what is it like to be a bat?”—a pathway to the answer from the integrated information theory. *Philosophy Compass* **12**(3), 12407 (2017). doi:10.1111/phc3.12407
21. Northoff, G., Tsuchiya, N., Saigo, H.: Mathematics and the brain: A category theoretical approach to go beyond the neural correlates of consciousness. *Entropy* **21**(12), 1234–1234 (2019). doi:10.3390/e21121234
22. Barrett, A.: An integration of integrated information theory with fundamental physics. *Frontiers in Psychology* **5**, 63 (2014). doi:10.3389/fpsyg.2014.00063
23. Safron, A.: An integrated world modeling theory (iwmt) of consciousness: Combining integrated information and global neuronal workspace theories with the free energy principle and active inference framework; toward solving the hard problem and characterizing agentic causation. *Frontiers in Artificial Intelligence* **3**, 30 (2020). doi:10.3389/frai.2020.00030
24. Friston, K.: The free-energy principle: A unified brain theory? *Nature Reviews. Neuroscience* **11**, 127–38 (2010). doi:10.1038/nrn2787
25. Dehaene, S., Sergent, C., Changeux, J.-P.: A neuronal network model linking subjective reports and objective physiological data during conscious perception **100**(14), 8520–8525 (2003). doi:10.1073/pnas.1332574100
26. Cerullo, M.A.: The problem with phi: A critique of integrated information theory. *PLOS Computational Biology* **11**(9), 1–12 (2015). doi:10.1371/journal.pcbi.1004286
27. Barrett, A., Mediano, P.: The phi measure of integrated information is not well-defined for general physical systems. *Journal of Consciousness Studies* **26**(1-2), 11–20 (2019)
28. Pautz, A.: What is the integrated information theory of consciousness? *Journal of Consciousness Studies* **26**(1-2), 188–215 (2019)
29. Popiel, N.J.M., Khajehabdollahi, S., Abeyasinghe, P.M., Riganello, F., Nichols, E.S., Owen, A.M., Soddu, A.: The emergence of integrated information, complexity, and ‘consciousness’ at criticality. *Entropy* **22**(3), 1–12 (2020). doi:10.3390/e22030339
30. Sperry, R.W.: Cerebral organization and behavior. *Science* **133**(3466), 1749–1757 (1961). doi:10.1126/science.133.3466.1749
31. Gazzaniga, M.S., Bogen, J.E., Sperry, R.W.: Some functional effects of sectioning the cerebral commissures in man. *Proceedings of the National Academy of Sciences* **48**(10), 1765–1769 (1962). doi:10.1073/pnas.48.10.1765
32. Mayner, W.G.P., Marshall, W., Albantakis, L., Findlay, G., Marchman, R., Tononi, G.: Pyphi: A toolbox for integrated information theory. *PLOS Computational Biology* **14**(7), 1–21 (2018). doi:10.1371/journal.pcbi.1006343

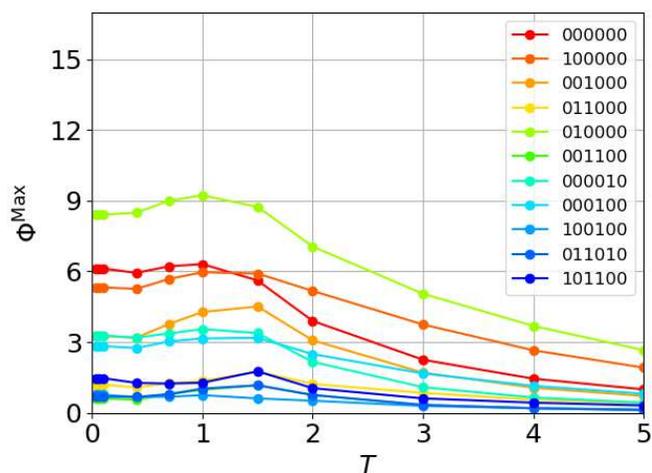
Additional Files

Additional file 1 — Integrated conceptual information Φ^{Max} with respect to temperature T for the networks in Fig. 1(b-e)

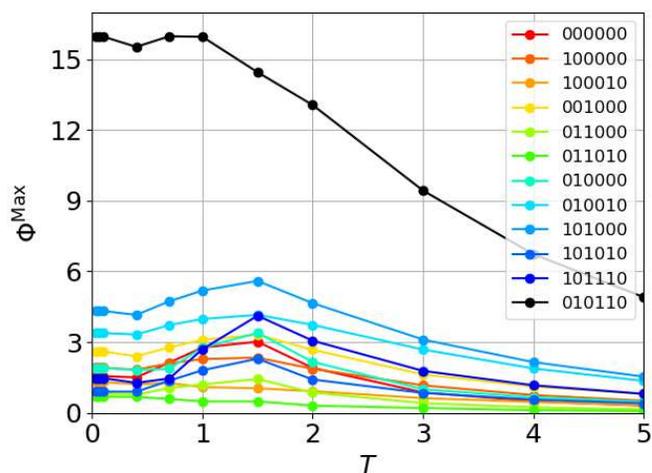
The view of these graphs is the same as in Fig. 2.



(a) no frustration

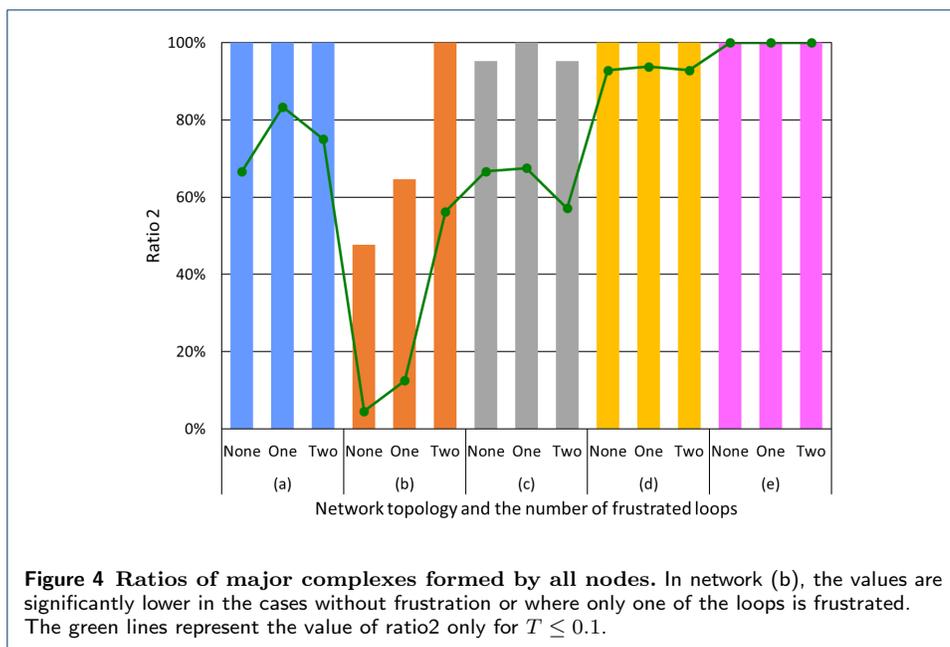
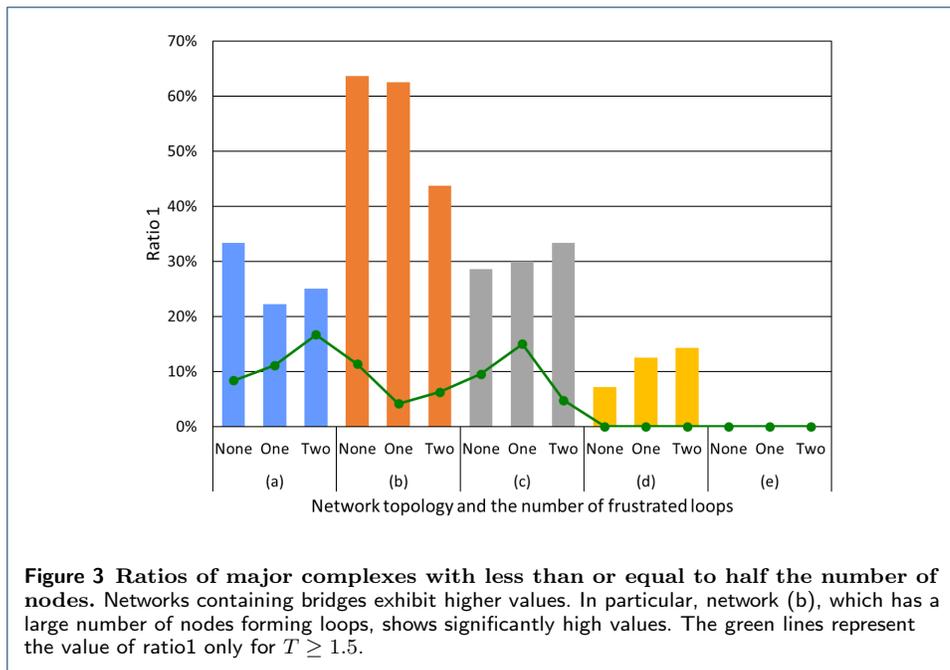


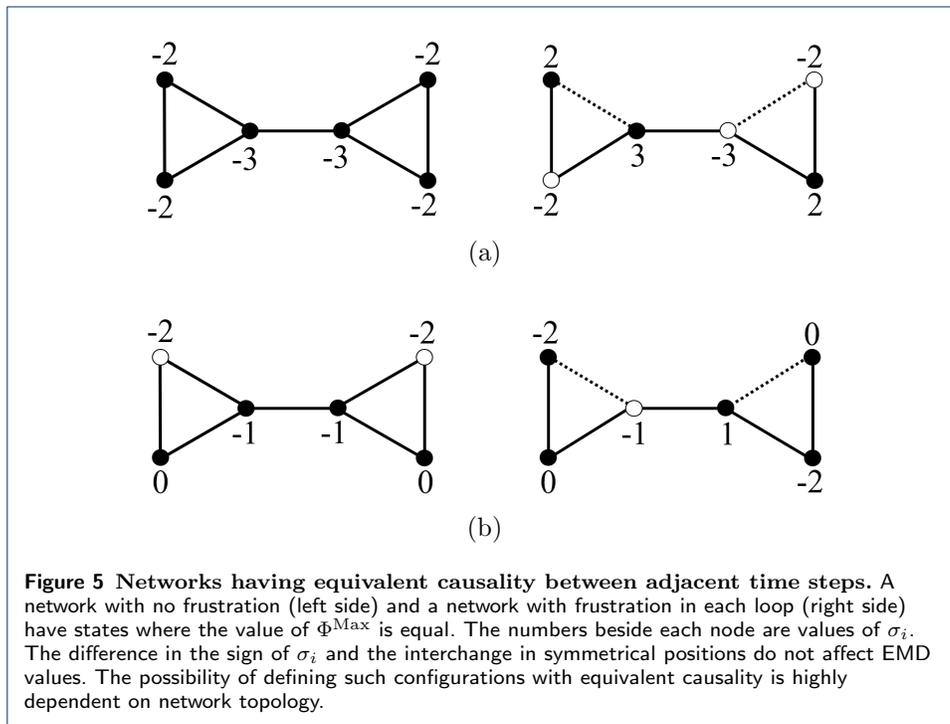
(b) frustration in one loop

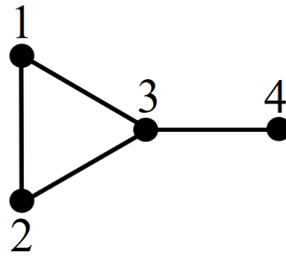


(c) frustration in both two loops

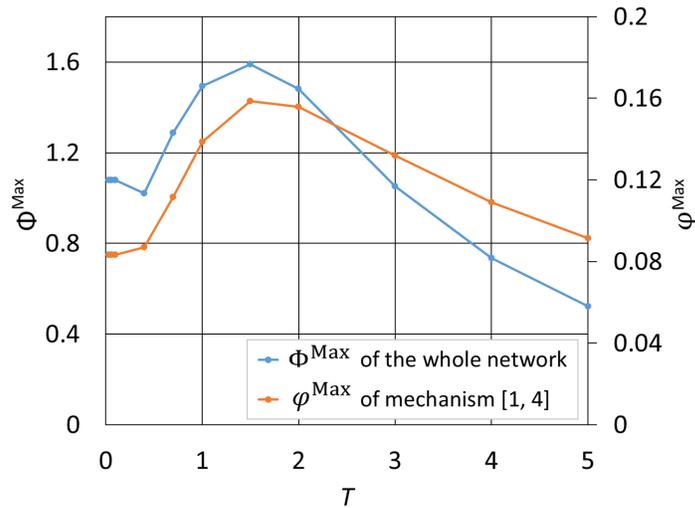
Figure 2 Integrated conceptual information Φ^{Max} with respect to temperature T for the network in Fig. 1(a). Only typical patterns are shown, and configurations showing the same or remarkably similar changes to one of the representatives are omitted. The results in no frustration and two frustrated loops are identical.



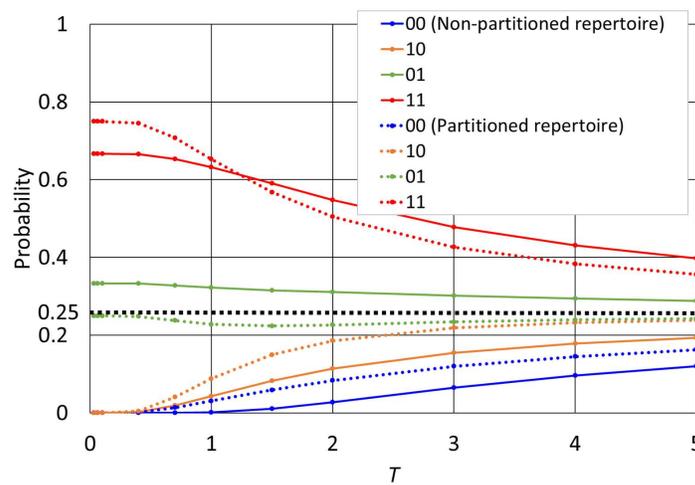




(a) Networks consisting of four nodes

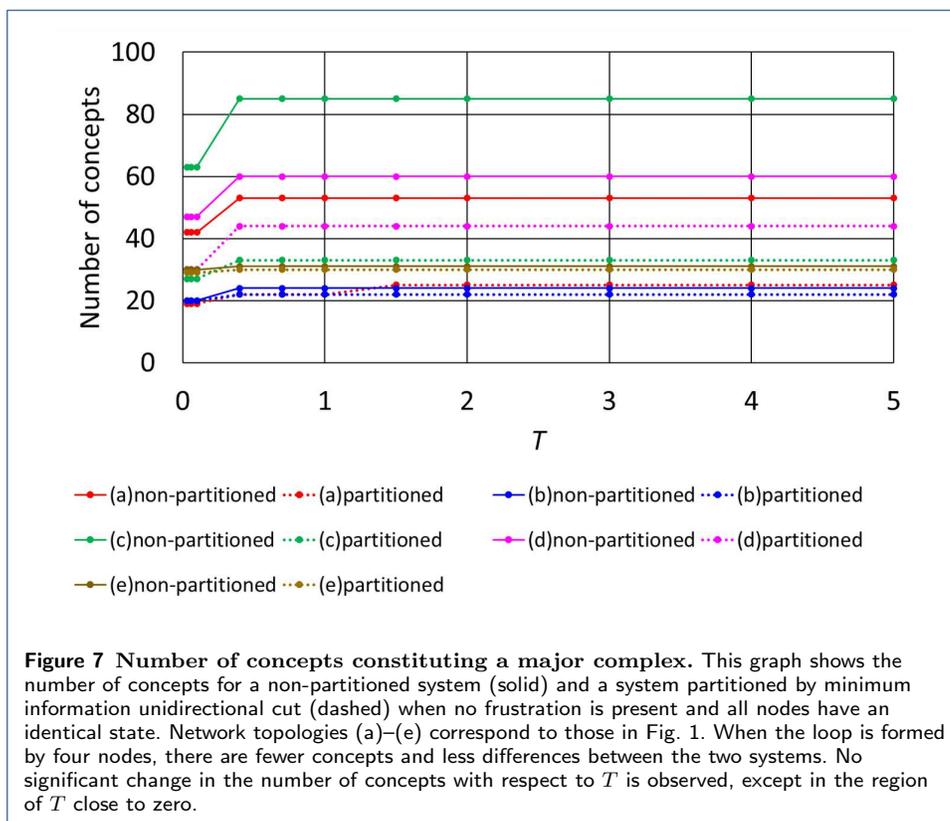


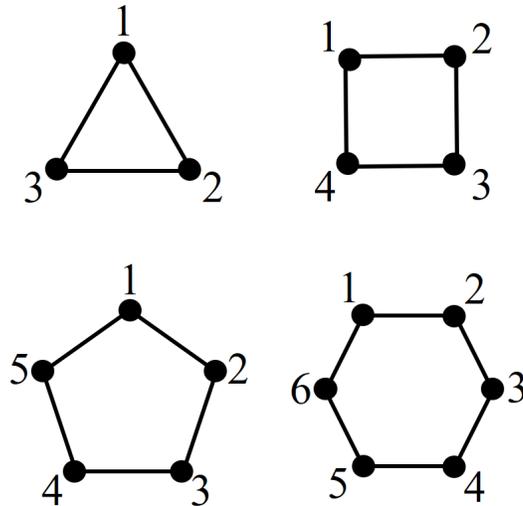
(b) Φ^{Max} and φ^{Max} with respect to T



(c) Probabilities of each bin in the non-partitioned and partitioned cause repertoires

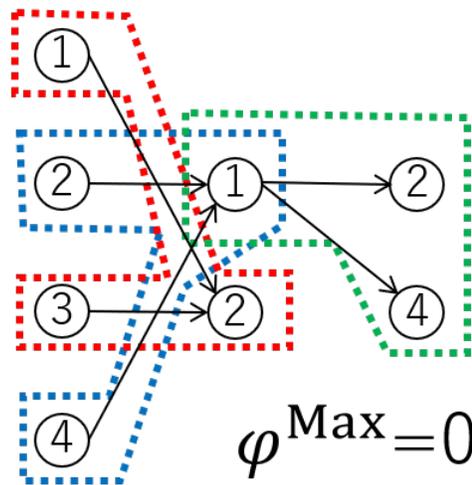
Figure 6 Non-monotonicity of the integrated information. (a) System of four nodes corresponds to a part of the network analyzed in this study. (b) Contrary to the intuitive prediction of monotonic decrease with temperature, both Φ^{Max} and φ^{Max} can have maxima at $T = 1.5$. (c) The values of each bin of the cause repertoire are illustrated when nodes 1 and 4 are taken as the mechanism. The monotonicity and speed of approaching the uniform distribution when T is increased depends on each bin.





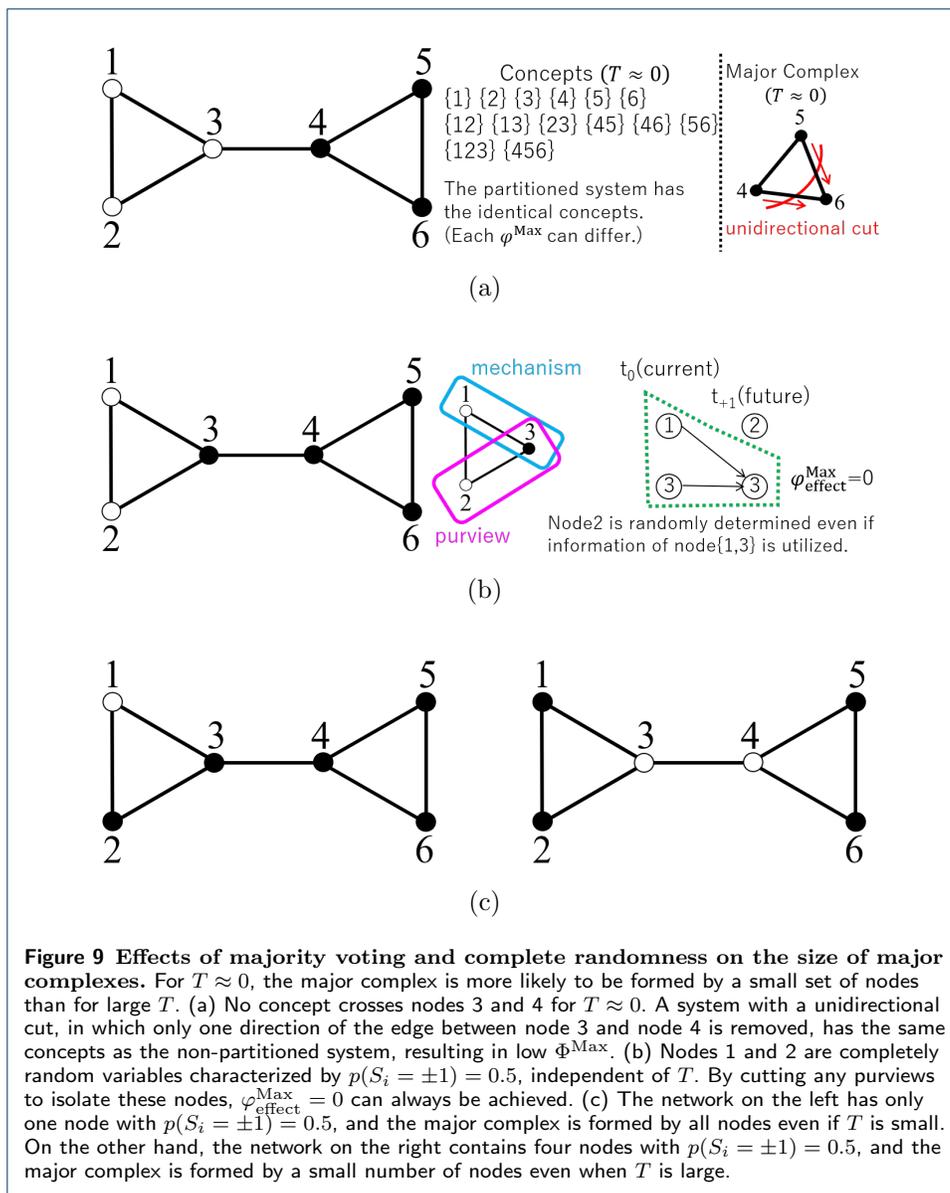
(a) Networks consisting of a single loop

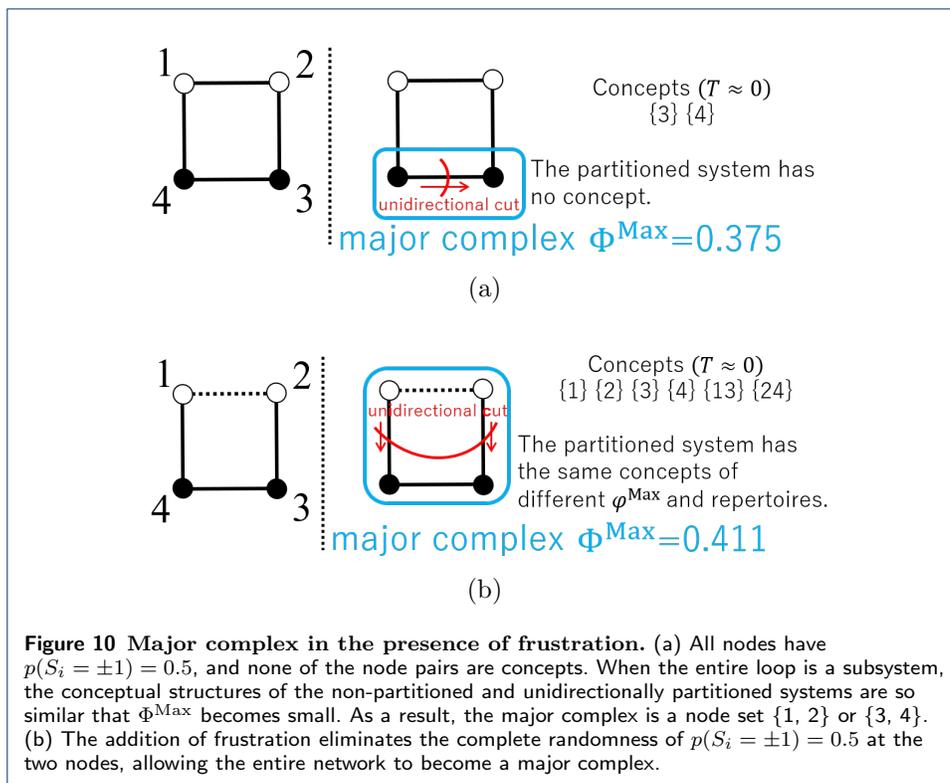
t_{-1} (past) t_0 (current) t_{+1} (future)



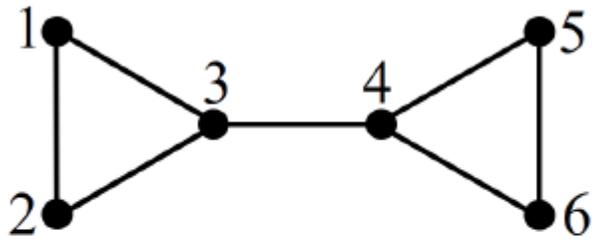
(b) Example of the partition leading to $\varphi^{Max} = 0$ in the case of a loop formed by four nodes

Figure 8 Effect of parity on the number of concepts in a network of a single loop. When the number of nodes in a loop is even, the number of combinations of nodes with $\varphi^{Max} > 0$ is significantly reduced because the nodes can be always divided into two causal groups without overlap. (b) Nodes 2 and 4 have a causal relationship with node 1, but not with node 2. Therefore, the same repertoire is obtained in the partitioned system as in the non-partitioned system, resulting in $\varphi^{Max} = 0$. This does not hold true when the number of nodes constituting a loop, M , is an odd number.

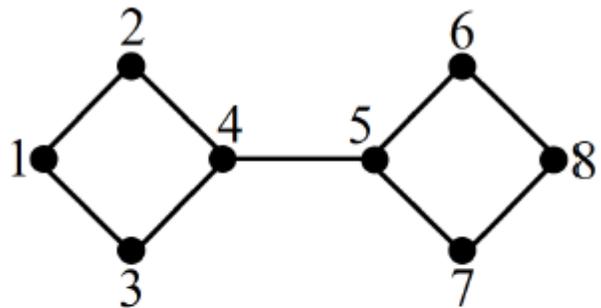




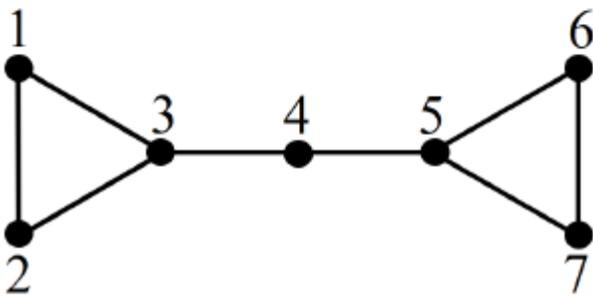
Figures



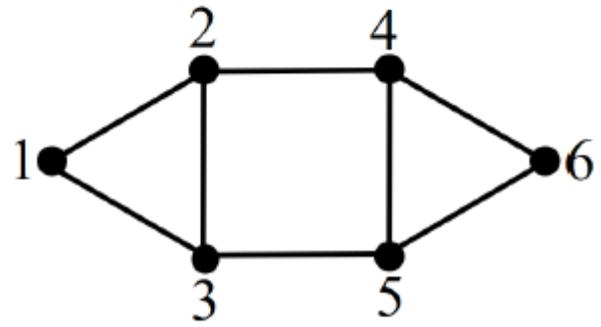
(a)



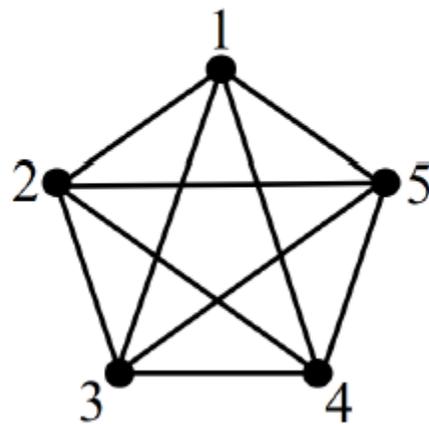
(b)



(c)



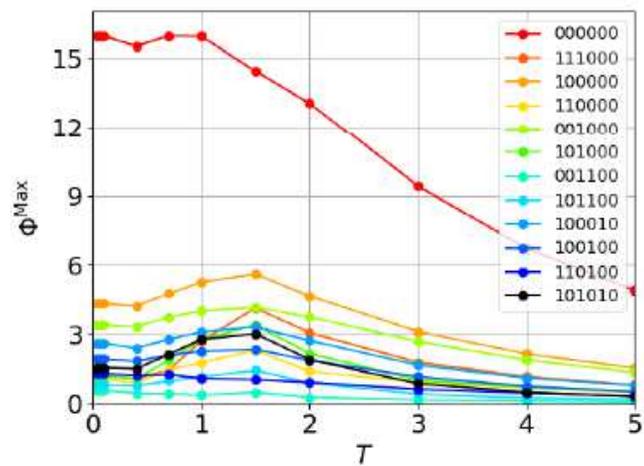
(d)



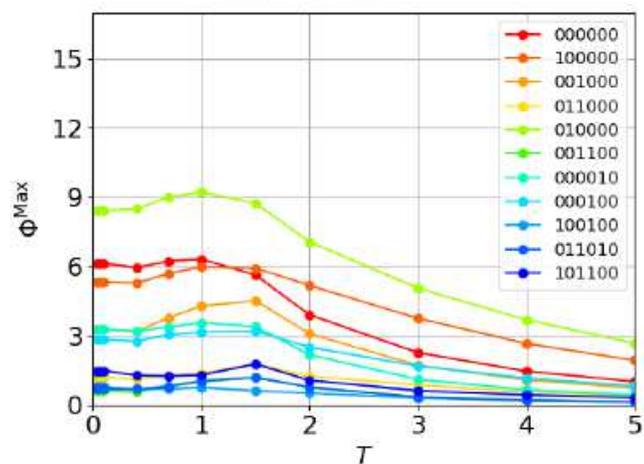
(e)

Figure 1

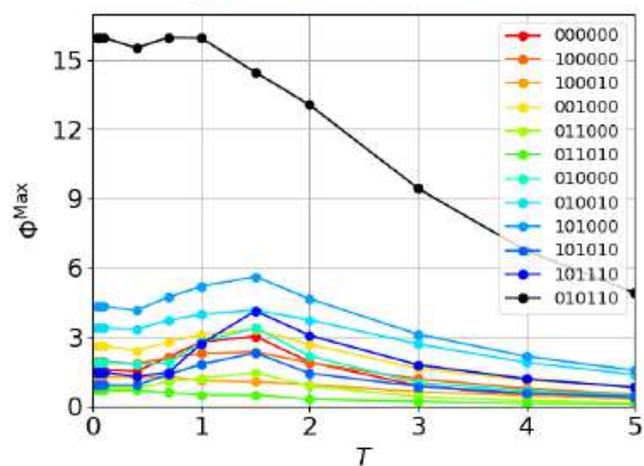
Networks to be analyzed in this study. In networks (a-c), the two loops are connected by a bridge. Networks (d) and (e) do not have bridges. In particular, in network (e), an edge is set between all nodes.



(a) no frustration



(b) frustration in one loop



(c) frustration in both two loops

Figure 2

Integrated conceptual information Φ^{Max} with respect to temperature T for the network in Fig. 1(a). Only typical patterns are shown, and configurations showing the same or remarkably similar changes to one of the representatives are omitted. The results in no frustration and two frustrated loops are identical.

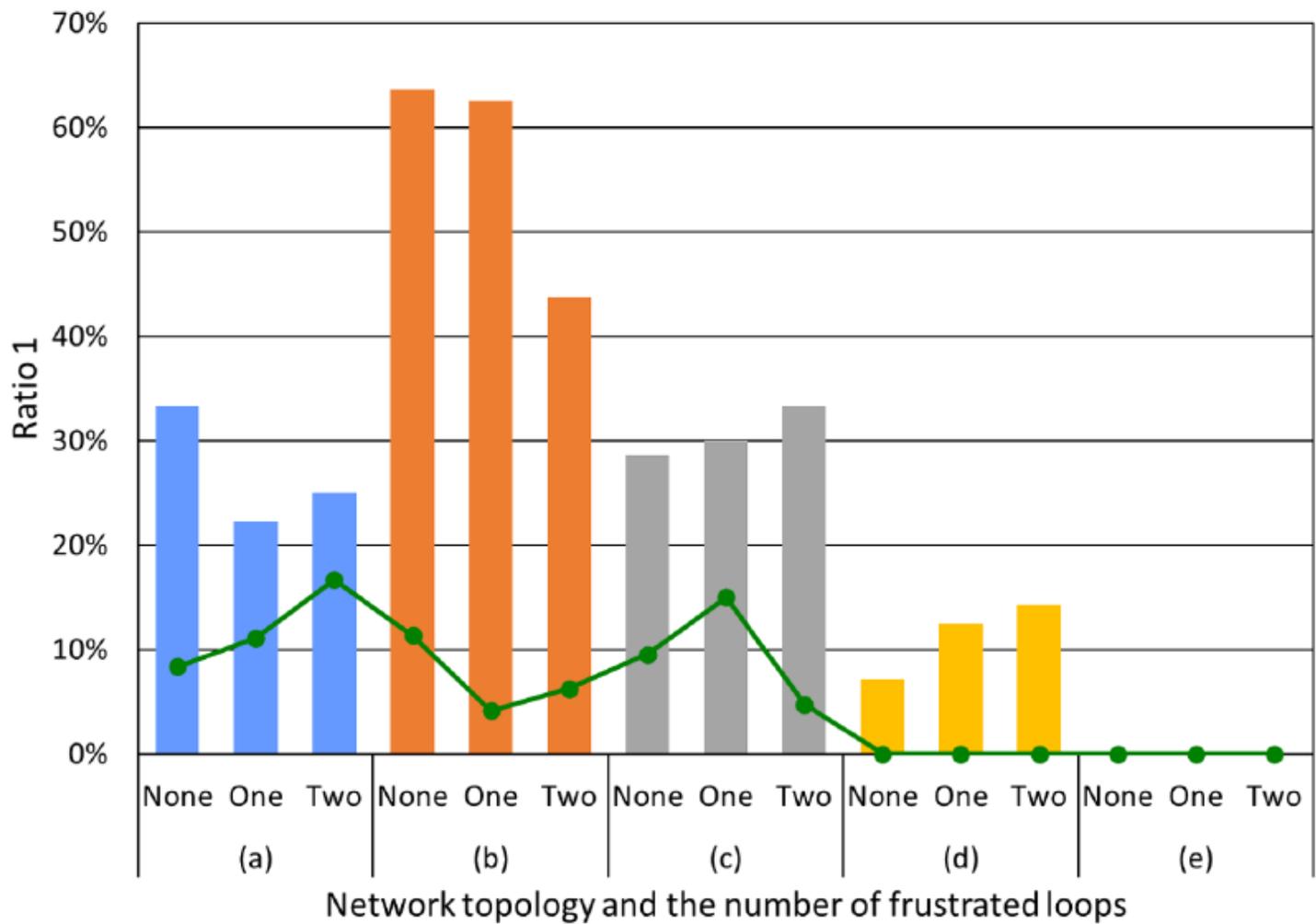


Figure 3

Ratios of major complexes with less than or equal to half the number of nodes. Networks containing bridges exhibit higher values. In particular, network (b), which has a large number of nodes forming loops, shows significantly high values. The green lines represent the value of ratio1 only for $T \geq 1:5$.

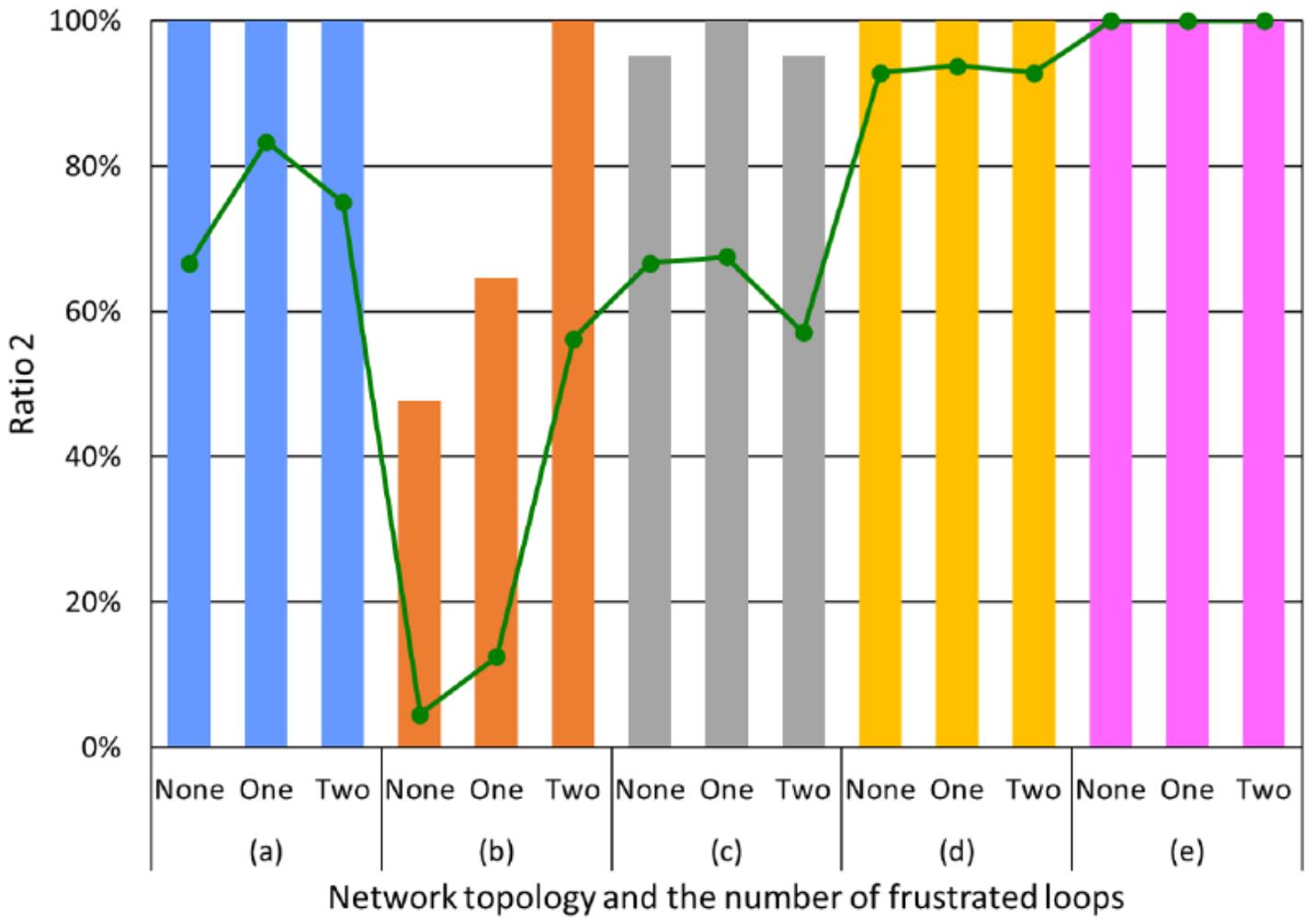
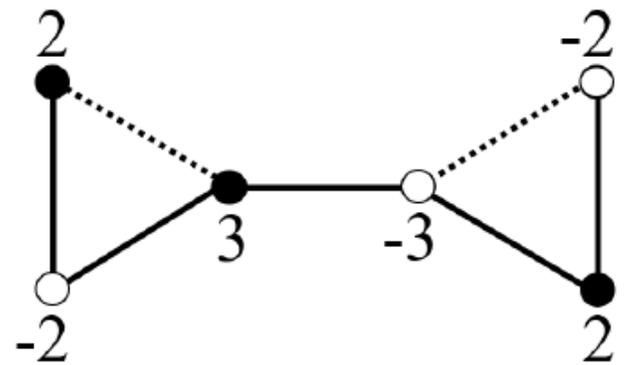
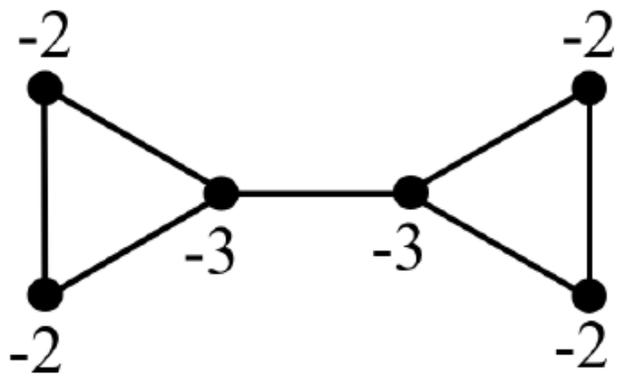
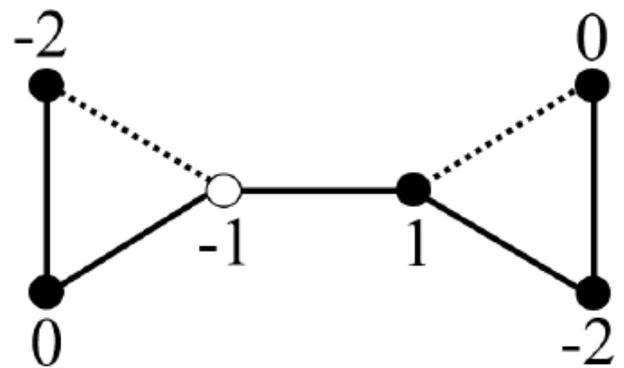
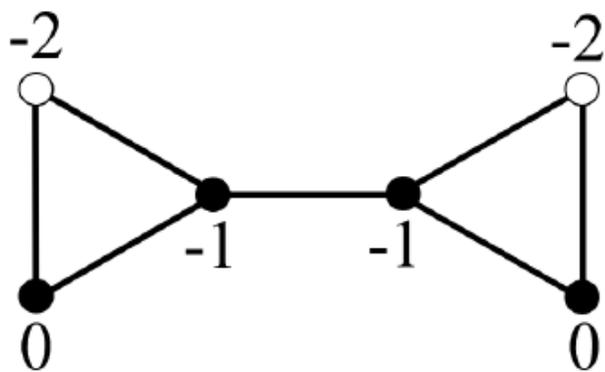


Figure 4

Ratios of major complexes formed by all nodes. In network (b), the values are significantly lower in the cases without frustration or where only one of the loops is frustrated. The green lines represent the value of ratio2 only for $T \leq 0:1$.



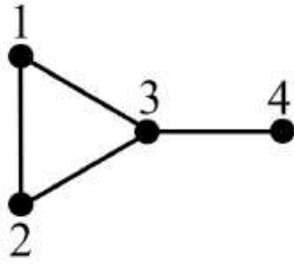
(a)



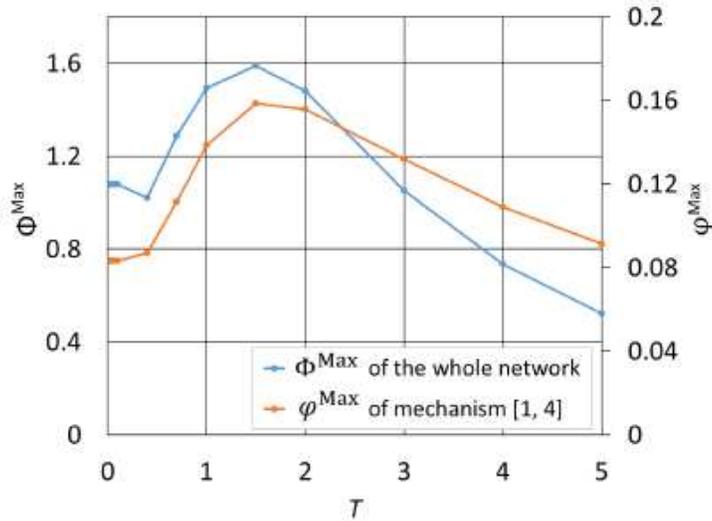
(b)

Figure 5

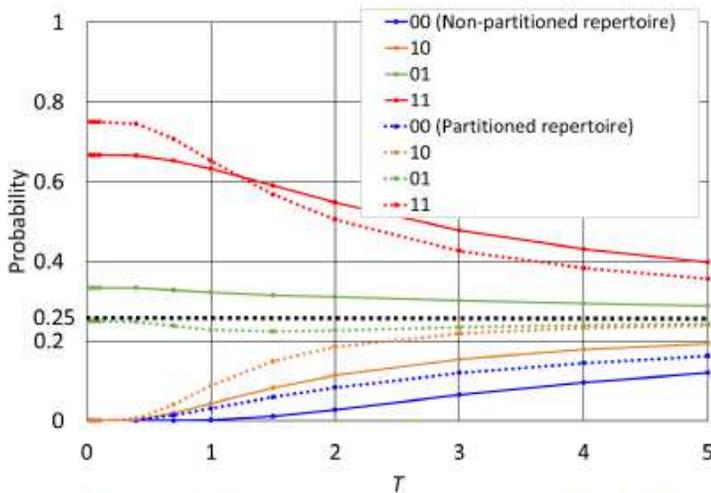
Networks having equivalent causality between adjacent time steps. A network with no frustration (left side) and a network with frustration in each loop (right side) have states where the value of Φ_{Max} is equal. The numbers beside each node are values of o_i . The difference in the sign of o_i and the interchange in symmetrical positions do not affect EMD values. The possibility of defining such configurations with equivalent causality is highly dependent on network topology.



(a) Networks consisting of four nodes



(b) Φ^{Max} and φ^{Max} with respect to T



(c) Probabilities of each bin in the non-partitioned and partitioned cause repertoires

Figure 6

Non-monotonicity of the integrated information. (a) System of four nodes corresponds to a part of the network analyzed in this study. (b) Contrary to the intuitive prediction of monotonic decrease with temperature, both Φ^{Max} and φ^{Max} can have maxima at $T = 1.5$. (c) The values of each bin of the cause repertoire are illustrated when nodes 1 and 4 are taken as the mechanism. The monotonicity and speed of approaching the uniform distribution when T is increased depends on each bin.

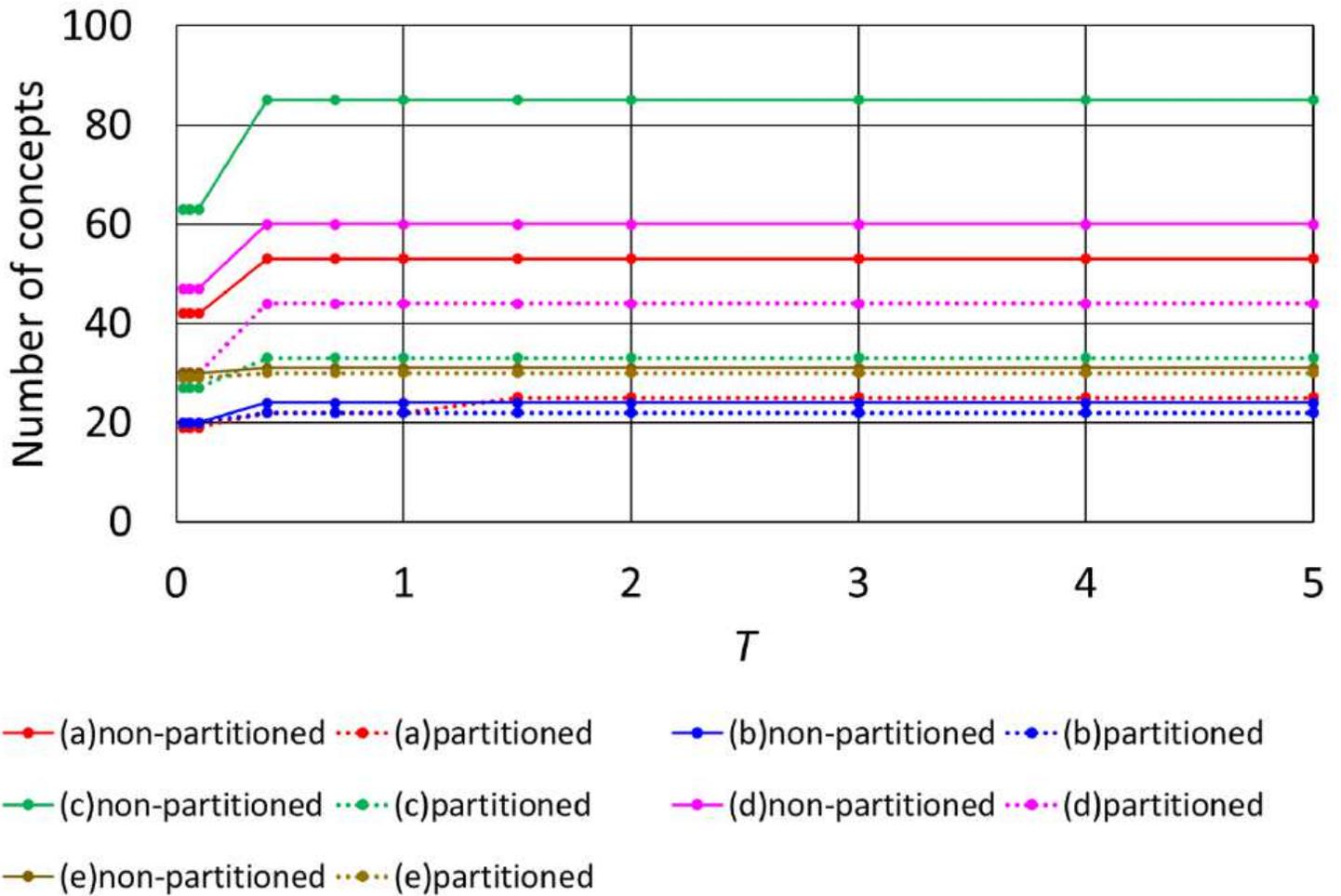
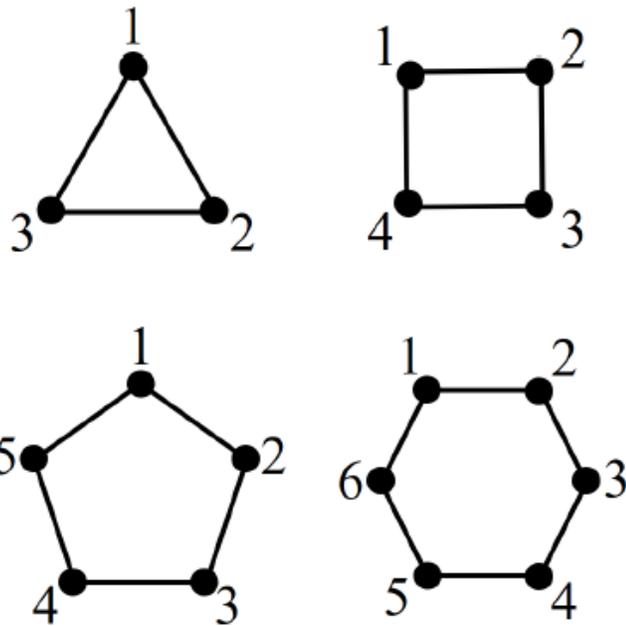


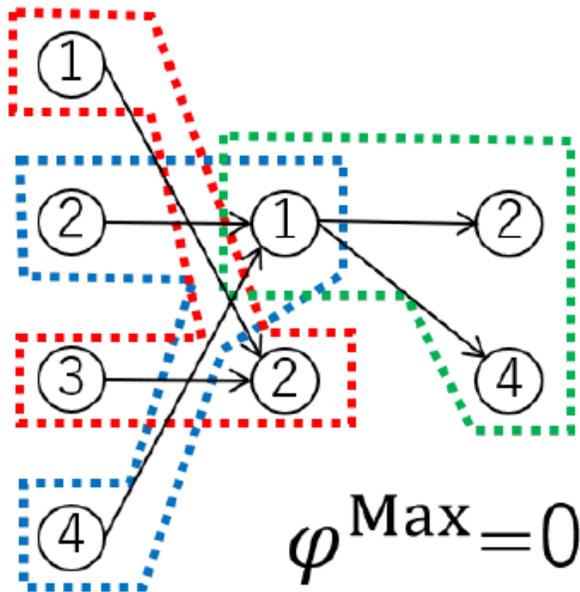
Figure 7

Number of concepts constituting a major complex. This graph shows the number of concepts for a non-partitioned system (solid) and a system partitioned by minimum information unidirectional cut (dashed) when no frustration is present and all nodes have an identical state. Network topologies (a)–(e) correspond to those in Fig. 1. When the loop is formed by four nodes, there are fewer concepts and less differences between the two systems. No significant change in the number of concepts with respect to T is observed, except in the region of T close to zero.



(a) Networks consisting of a single loop

t_{-1} (past) t_0 (current) t_{+1} (future)

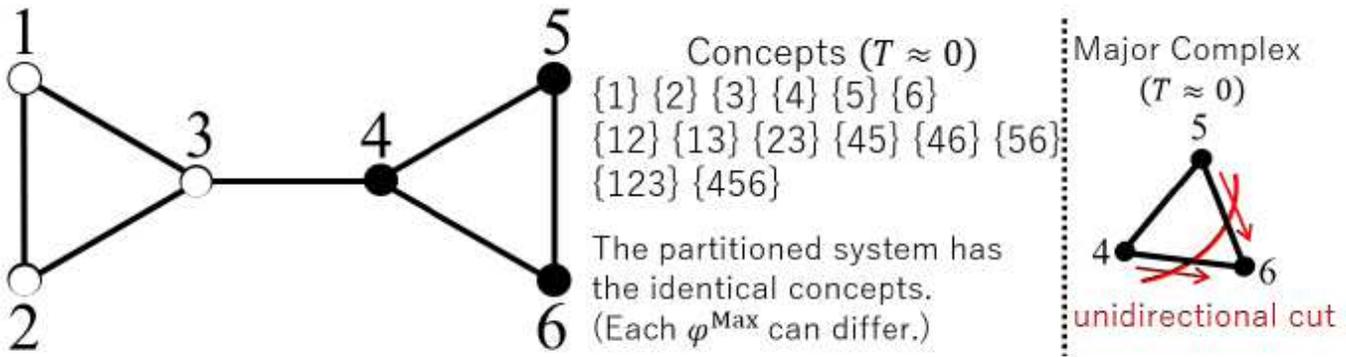


(b) Example of the partition leading to $\varphi^{\text{Max}} = 0$ in the case of a loop formed by four nodes

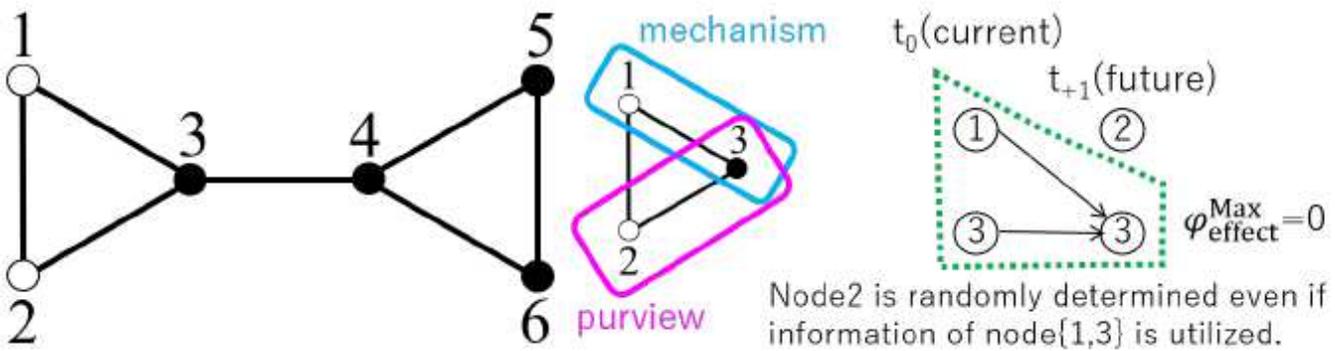
Figure 8

Effect of parity on the number of concepts in a network of a single loop. When the number of nodes in a loop is even, the number of combinations of nodes with $\varphi^{\text{Max}} > 0$ is significantly reduced because the nodes can be always divided into two causal groups without overlap. (b) Nodes 2 and 4 have a causal relationship with node 1, but not with node 2. Therefore, the same repertoire is obtained in the partitioned

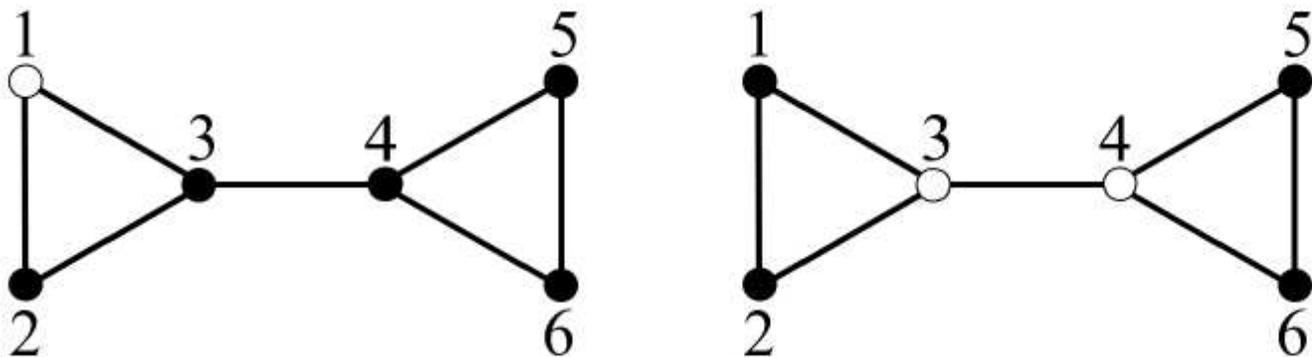
system as in the non-partitioned system, resulting in $\varphi^{\text{Max}} = 0$. This does not hold true when the number of nodes constituting a loop, M , is an odd number.



(a)



(b)



(c)

Figure 9

Effects of majority voting and complete randomness on the size of major complexes. For $T \approx 0$, the major complex is more likely to be formed by a small set of nodes than for large T . (a) No concept crosses nodes 3 and 4 for $T \approx 0$. A system with a unidirectional cut, in which only one direction of the edge

between node 3 and node 4 is removed, has the same concepts as the non-partitioned system, resulting in low Φ_{Max} . (b) Nodes 1 and 2 are completely random variables characterized by $p(S_i = \pm 1) = 0.5$, independent of T . By cutting any purviews to isolate these nodes, $\Phi_{\text{Max}} = 0$ can always be achieved. (c) The network on the left has only one node with $p(S_i = \pm 1) = 0.5$, and the major complex is formed by all nodes even if T is small. On the other hand, the network on the right contains four nodes with $p(S_i = \pm 1) = 0.5$, and the major complex is formed by a small number of nodes even when T is large.

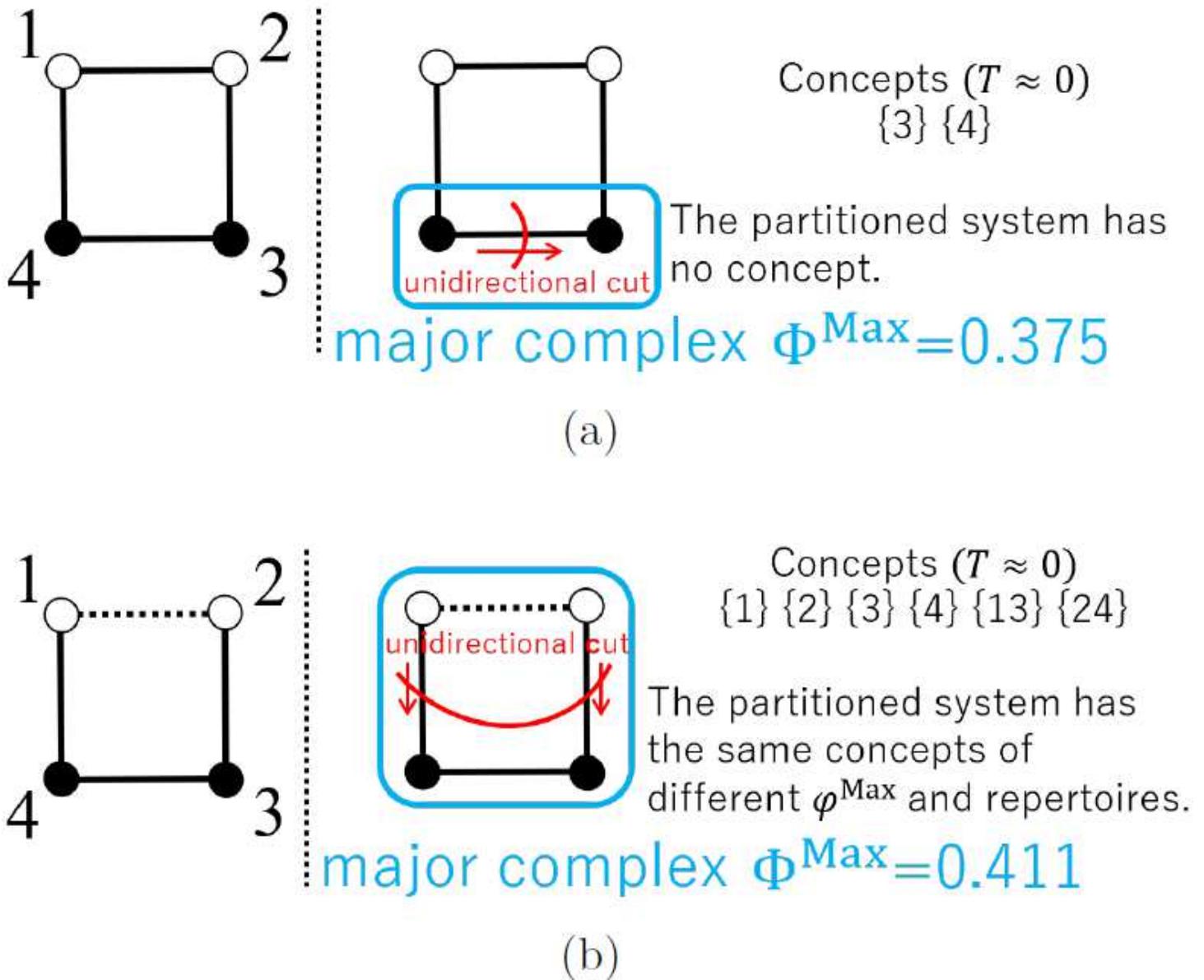


Figure 10

Major complex in the presence of frustration. (a) All nodes have $p(S_i = \pm 1) = 0.5$, and none of the node pairs are concepts. When the entire loop is a subsystem, the conceptual structures of the non-partitioned and unidirectionally partitioned systems are so similar that Φ_{Max} becomes small. As a result, the major complex is a node set $\{1, 2\}$ or $\{3, 4\}$. (b) The addition of frustration eliminates the complete randomness of $p(S_i = \pm 1) = 0.5$ at the two nodes, allowing the entire network to become a major complex.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [additionalfile1hosakatadaaki.pdf](#)