

Bidirectionally connected cores in a mouse connectome: towards extracting the brain subnetworks essential for consciousness

Jun Kitazono^{1,*}, Yuma Aoki², Masafumi Oizumi^{1,*}

¹Graduate School of Arts and Sciences, The University of Tokyo 3-8-1 Komaba, Meguro-ku, Tokyo, 153-8902, Japan,

²Graduate School of Information Science and Technology, The University of Tokyo 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

*Corresponding authors: Graduate School of Arts and Sciences, The University of Tokyo, Meguro-ku, Tokyo, Japan. Email: c-kitazono@g.ecc.u-tokyo.ac.jp, c-oizumi@g.ecc.u-tokyo.ac.jp

Where in the brain consciousness resides remains unclear. It has been suggested that the subnetworks supporting consciousness should be bidirectionally (recurrently) connected because both feed-forward and feedback processing are necessary for conscious experience. Accordingly, evaluating which subnetworks are bidirectionally connected and the strength of these connections would likely aid the identification of regions essential to consciousness. Here, we propose a method for hierarchically decomposing a network into cores with different strengths of bidirectional connection, as a means of revealing the structure of the complex brain network. We applied the method to a whole-brain mouse connectome. We found that cores with strong bidirectional connections consisted of regions presumably essential to consciousness (e.g. the isocortical and thalamic regions, and claustrum) and did not include regions presumably irrelevant to consciousness (e.g. cerebellum). Contrarily, we could not find such correspondence between cores and consciousness when we applied other simple methods that ignored bidirectionality. These findings suggest that our method provides a novel insight into the relation between bidirectional brain network structures and consciousness.

Key words: bidirectional; recurrent; network core; mouse connectome; consciousness.

Introduction

Where in the brain consciousness resides has been one of the biggest questions in science. Although we have not yet reached a conclusive answer, much empirical evidence has been accumulated in the course of searching for the minimal mechanisms sufficient for conscious experience, called neural correlates of consciousness (NCC) (Koch et al. 2016). Among the many problems that need to be solved in identifying NCC, we focus here on the problem of identifying the minimally sufficient subnetworks in the brain that support conscious experience. In this study, we simply refer to such subnetworks as “the locus of consciousness.” For example, it is commonly agreed that the retina is not included in the locus of consciousness because it has been empirically shown that neural activities in the retina do not directly correlate with what we perceive (Tononi and Laureys 2005). More importantly, a person who becomes retinally blind in adulthood continues to have vivid visual dreams (Tononi and Laureys 2005). As another notable example, the cerebellum is also not considered to be included in the locus of consciousness because lesions of the cerebellum do not much affect conscious experience, even though it has far more neurons than the cerebral

cortex and is densely connected to the rest of the brain (Lemon and Edgley 2010, Yu et al. 2015). On the other hand, which cortical areas or subcortical areas are essential for consciousness are still controversial (see Boly et al. 2017, Odegaard et al. 2017, Melloni et al. 2021 for general reviews, and Leopold 2012 for a review focusing on the primary visual cortex as an example).

In inferring the locus of consciousness in the brain, it is important to note suggestions that feed-forward processing alone is insufficient for subjects to consciously perceive stimuli; rather, feedback is also necessary, indicating the need for bidirectional (also called recurrent, reciprocal, or reentrant) processing (Cauller and Kulics 1991, Lamme et al. 1998, Supèr et al. 2001, Self et al. 2012, Auksztulewicz et al. 2012, Sachidhanandam et al. 2013, Koivisto et al. 2014, Tang et al. 2014, Manita et al. 2015). The feedback component disappears not only during the loss of specific contents of consciousness in awake states, but also during unconscious states, where conscious experiences are generally lost, such as general anesthesia (Lamme et al. 1998, Ku et al. 2011, Boly et al. 2012, Cohen et al. 2018), sleep (Cauller and Kulics 1988), and vegetative states (Boly et al. 2011). The importance of bidirectional processing is suggested to be

Received: October 6, 2021. **Revised:** February 28, 2022. **Accepted:** March 2, 2022

© The Author(s) 2022. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

independent of sensory modality (Dembski et al. 2021) (vision (Lamme et al. 1998, Supèr et al. 2001, Self et al. 2012, Koivisto et al. 2014, Tang et al. 2014), somatosensation (Cauller and Kulics 1991, Auksztulewicz et al. 2012, Sachidhanandam et al. 2013, Manita et al. 2015), and audition (Schlossmacher et al. 2021)) and species (humans (Boly et al. 2011, Auksztulewicz et al. 2012, Boly et al. 2012, Tang et al. 2014), monkeys (Cauller and Kulics 1988, Cauller and Kulics 1991, Lamme et al. 1998, Supèr et al. 2001, Self et al. 2012, Koivisto et al. 2014), rodents (Sachidhanandam et al. 2013, Manita et al. 2015), birds (Nieder et al. 2020), and even flies (Cohen et al. 2018)).

Given these findings, it appears reasonable that subnetworks in which brain areas are strongly bidirectionally connected would be included in the locus of consciousness. In fact, many major theories of consciousness have made similar predictions about the locus of consciousness in common, even though they differ in many other respects (Tononi and Edelman 1998, Dehaene and Naccache 2001, Edelman 2003, Lamme 2006, Dehaene and Changeux 2011, Oizumi et al. 2014, Tononi et al. 2016, Lamme 2018, Pennartz et al. 2019, Sikkens et al. 2019, Aru et al. 2020, Mashour et al. 2020). Under this criterion, the retina, for example, is evidently excluded from the locus of consciousness because it is connected to the other areas of the brain in a purely feed-forward manner. To examine the relation between subnetworks with strong bidirectional connections and consciousness, it is important to first identify such subnetworks and understand the bidirectional network structure of the brain. If we understand which subnetworks are strongly bidirectionally connected and which are only weakly so connected, we can quantitatively discuss the correspondence between these subnetworks and consciousness.

For this purpose, we propose a method for extracting subnetworks in which nodes are strongly connected in a bidirectional manner. We call such subnetworks “complexes”; this term and concept originate from integrated information theory, although the specific definition of complexes differs in the original theory (Tononi et al. 1994, Tononi 2004, Balduzzi and Tononi 2008, Oizumi et al. 2014, Tononi et al. 2016, Kitazono et al. 2020). To be specific, in this study, we first define a “main” complex as a subnetwork that has the local maximum of bidirectional connection strength. We evaluate the strength of bidirectional connections by a measure that takes a large value when nodes are connected by strong bidirectional edges. To reveal the network structure, we also define complexes as a weaker notion of a main complex. Complexes are less strongly bidirectionally connected than a main complex and form a nested structure. That is, a main complex is included in another less strongly connected complex; that complex is in turn included in yet another complex; and so on. In this hierarchical organization, a main complex, intuitively speaking, is a central core where there is no weakly connected parts and complexes are surrounding cores.

If we search for complexes by brute force, computation time grows exponentially with the number of nodes, because we need to take account of all possible subnetworks. To reduce computation time, we can use an algorithm proposed in our previous study (Kitazono et al. 2020). This algorithm, called hierarchical partitioning for complex search (HPC), enables the identification of complexes simply by hierarchically dividing the entire network. Because of the simplicity of this algorithm, the computation time increases only polynomially with the number of nodes. HPC allows us to find all complexes in a practical amount of time, even from large networks of thousands of elements, without omissions or misidentifications.

As a step in investigating the relationship between bidirectionally connected subnetworks—complexes—and consciousness, we applied the proposed method to a meso-scale, whole-brain mouse connectome (Oh et al. 2014) and identified the complexes. This connectome includes not only the cortical regions but also the subcortical, brainstem, and cerebellar regions, and has high spatial resolution. These characteristics make it suitable for discussing the relationship between brain regions and consciousness. We found that the extracted complexes with strong bidirectional connections consist of the brain regions that are thought to be essential to consciousness. In addition, to assess whether it is important to take account of the bidirectionality of connections, we examined how the results are affected if the bidirectionality of connections is ignored. We found that the complexes do not necessarily consist of the particular brain regions thought to be essential to consciousness, but rather of various brain regions that do not directly contribute to consciousness. We also applied a widely used method for extracting network cores, *s*-core decomposition, which does not consider bidirectionality. Interestingly, we could not find such correspondence between the extracted cores and the brain regions presumably essential to consciousness. In addition, we investigated the relationship between the complexes and the degree of nodes. We found that the complexes with strong bidirectional connections do not necessarily consist of high-degree nodes. This means that the core structures revealed by the complexes largely differ to the structures that are revealed by degree-based methods that ignore bidirectionality. These results indicate that the identification of bidirectional network structures will provide new insights into areas essential to consciousness.

The MATLAB codes of HPC are available at <https://github.com/JunKitazono/BidirectionallyConnectedCores>.

Results

Network cores with strong bidirectional connections: complexes

A simple example of a complex

In this study, we tried to extract the bidirectionally connected “cores” of the network, called “complexes.” Before

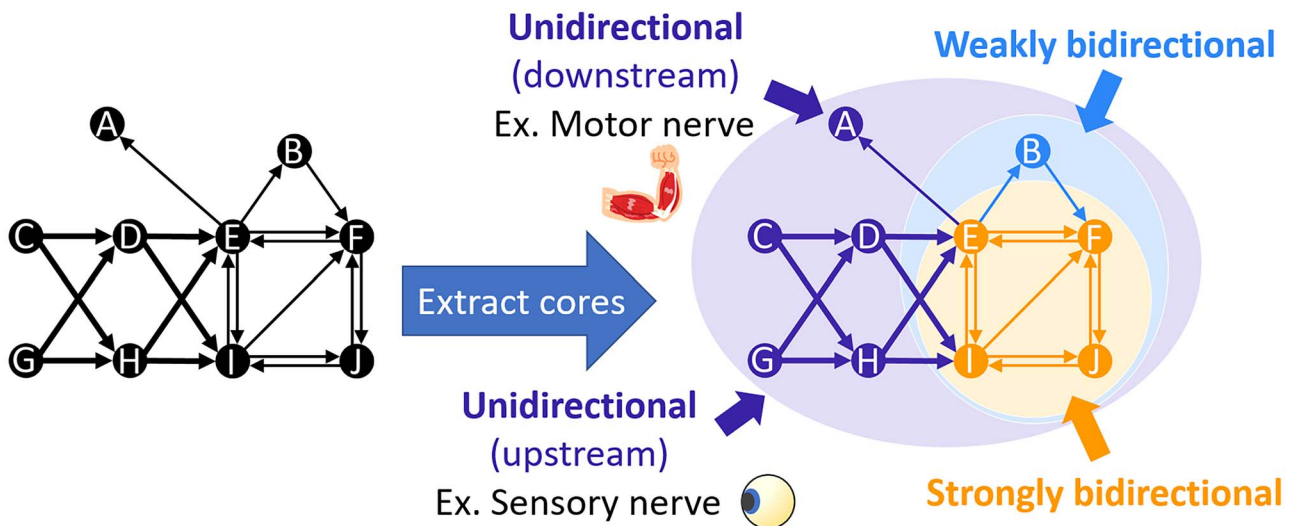


Fig. 1. Schematic of extracting network cores complexes and main complexes. A complex is a subnetwork that consists of stronger bidirectional connections than other subnetworks that include it. If we extract complexes from the network on the left, we obtain the result on the right. There are 2 complexes. One is the subnetwork $EFIJ$, which is colored orange. The other is the subnetwork $BEFIJ$, which is colored light blue. Comparing the 2 subnetworks, $EFIJ$ is more strongly bidirectionally connected than $BEFIJ$ because the nodes $EFIJ$ are all strongly connected to each other while the node B is only weakly connected to the nodes $EFIJ$ (i.e. there is only 1 edge in each direction, B to $EFIJ$ and $EFIJ$ to B). The node set $EFIJ$ turns out to be the most strongly connected complex in this network, which we call the main complex. The entire network is not a complex. The strength of the bidirectional connections in the entire network is zero because the entire network includes the nodes $CDGH$ and A , which are connected in a completely feed-forward manner. If we compare this network to the nervous system of the whole body of a mammal, we can consider the bidirectionally connected nodes $BEFIJ$ as the brain, the nodes $CDGH$ upstream as sensory nerves such as the retina, and the downstream node A as motor nerves.

we introduce the definition of complexes, let us first intuitively explain the concept of complexes taking the network shown in Fig. 1 as an example. In this example, the node A and the nodes $CDGH$ are only unidirectionally connected to $BEFIJ$, and therefore these nodes are not included in a complex. The node set $BEFIJ$ is a complex but only a weakly connected one, because the node B is only weakly connected to the nodes $EFIJ$ (i.e. there is only 1 edge in each direction, B to $EFIJ$ and $EFIJ$ to B). In contrast, the nodes $EFIJ$ are all strongly connected to each other, and the nodes $EFIJ$ therefore constitute a strongly connected complex. In this example, the node set $EFIJ$ turns out to be the most strongly connected complex, which we call the main complex. In general, complexes form a nested hierarchical structure, as do the node sets $EFIJ$ and $BEFIJ$. That is, a complex contains another complex that is smaller in size but more strongly connected. A complex smallest in size is the most strongly connected complex, which is a main complex.

We can consider this exemplar network as a toy network of the nervous system. For instance, we can consider the bidirectionally connected nodes $BEFIJ$ as the brain, the upstream nodes $CDGH$ as sensory nerves such as the retina (afferent nerves), and the downstream node A as motor nerves (efferent nerves). As we explain above, the node A (motor nerves for example) and the nodes $CDGH$ (the retina for example) are not included in the complexes. If we assume that bidirectional processing is essential for consciousness, the motor nerves and the retina would not be included in the locus of consciousness. In the mouse connectome network investigated in this study, there are no nodes that are only unidirectionally connected to the rest of the network. Thus, we

cannot evidently exclude some of the nodes because of the lack of bidirectional connections. Rather, we need to quantitatively investigate the degree of the bidirectional connections and look at the hierarchical structure of the complexes.

Outline of complexes and related concepts

The mathematical definition of a “complex” is rather complicated. To get the gist of it, we first outline 2 important concepts, namely strength of bidirectional connections and minimum cut, and then outline complexes. Please see Methods for mathematically formal explanations.

Strength of bidirectional connections To define complexes, i.e. bidirectionally connected cores of a network, we first need to have a measure that quantifies how strongly the 2 divided parts of a network are bidirectionally connected. We propose a measure that is low when the connections in 1 direction are weak even though those in the other direction are strong (Fig. 2a and 2b) and that is high when the connections in both directions are strong (Fig. 2c). Specifically, we define the strength of bidirectional connections as the minimum value of the sum of the weights of connections going from one part to the other and the sum in the opposite direction (Eq. (6) in Methods). The strength of the bidirectional connections defined this way is zero when the connections are completely unidirectional as in Fig. 2a, low when the connections in 1 direction are weak as in Fig. 2b, and high when the connections are strong in both directions as in Fig. 2c.

Minimum cut A complex is a network core whose parts are strongly connected to each other in a bidirectional

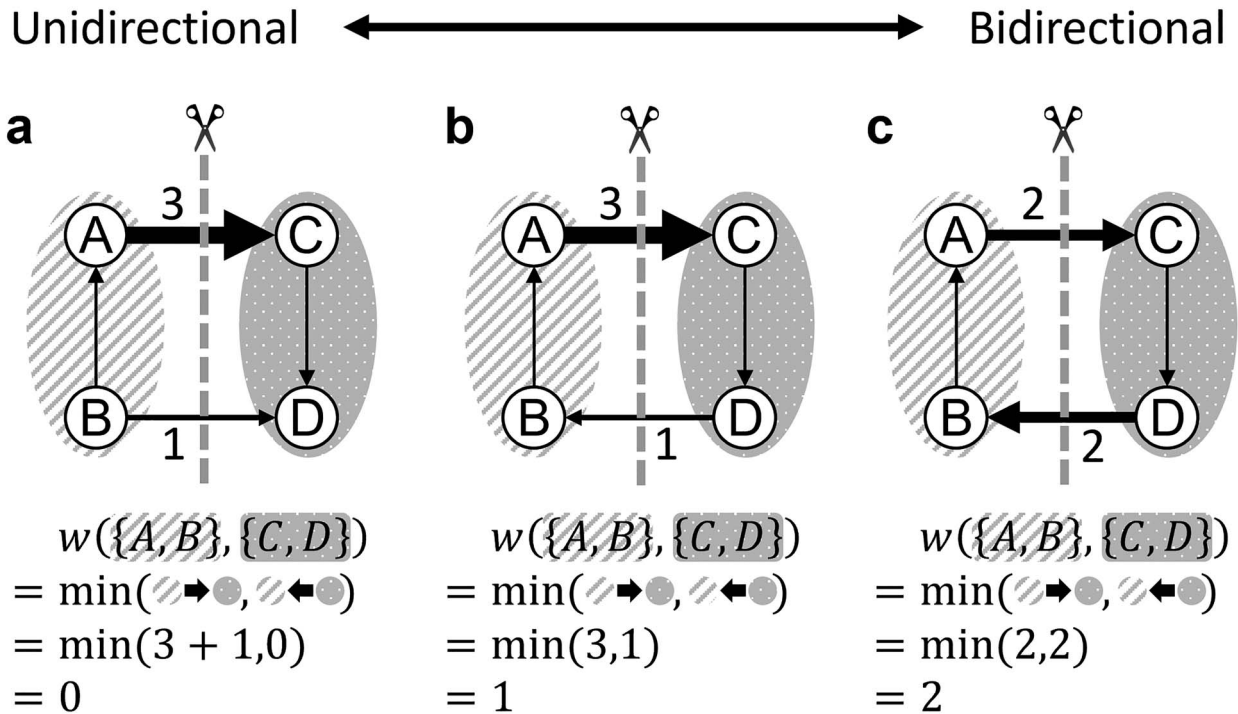


Fig. 2. Strength of bidirectional connections. To measure the strength of bidirectional connections, we take the minimum value of the sum of the weights of connections going from one part to the other and the sum in the opposite direction. In the examples a–c, the strength of bidirectional connections between AB and CD is measured. a) The connection is completely unidirectional: there are connections from AB to CD but there are no connections in the opposite direction. In this case, the strength of bidirectional connections is 0. b) The connection from AB to CD is strong but that from CD to AB is weak. In this case, the strength of bidirectional connections is high, which equals to 1. c) The connections in both directions are strong and the strength of bidirectional connections is high, which equals to 2.

manner. In other words, a complex cannot be “cut” into 2 parts without losing many strong edges no matter how it is cut. To measure such “inseparability” of a network, we consider the bi-partition of the network for which the strength of bidirectional connections is minimum among those for all possible bi-partitions, which we call a minimum cut (or a min-cut). We call the strength of bidirectional connections for a min-cut the “min-cut weight” and represent it by w^{mc} . As the value of a min-cut weight w^{mc} gives the lower bound of the strength of bidirectional connections for any possible bi-partitions of the network, any part of the network is “bidirectionally” connected to its complement part with a strength greater than or equal to w^{mc} .

If a network consists of disconnected parts as shown in Fig. 3a, the min-cut is the partition that cuts the network into the 2 disconnected parts and w^{mc} is 0. On the other hand, in a network where all the parts are strongly connected and cannot be separated without many edges being cut as in Fig. 3b, w^{mc} is large. As illustrated in these examples, a larger min-cut weight w^{mc} indicates a network that is more inseparable.

Please note that w^{mc} can either grow or shrink as network size increases, depending on structures of a network. This means that the largest subnetwork (the entire network) or the smallest subnetwork (a subnetwork consisting of 2 nodes) is not necessarily the most inseparable.

Complex Complexes and main complexes are defined using the min-cut weight w^{mc} introduced above. A main complex is a subnetwork that has “locally” maximal w^{mc} . Local maximum means that w^{mc} in a main complex is larger than that in any other subnetwork containing it and any other smaller subnetwork contained within it (both the left and right inequality in Fig. 4 hold). In general, a network can have multiple main complexes. In addition to main complexes, the notion of complexes is also useful for revealing the structure of a network. Briefly, a complex is a weaker notion of a main complex, i.e. a subnetwork such that its w^{mc} is greater than w^{mc} of any other subnetwork containing it (only the right inequality in Fig. 4 holds). Complexes form a hierarchical structure: a main complex is included in a complex larger in size but with smaller w^{mc} , and the complex is included in yet another complex even larger in size but with smaller w^{mc} . Metaphorically speaking, if we consider a network as a mountain whose height is determined by the min-cut weight w^{mc} , a main complex tells us the peak of the mountain and the surrounding complexes tell us the contour lines of the mountain, as illustrated in Fig. 1.

A schematic explanation of complexes is shown in Fig. 4. We consider a network, which is the same as that in Fig. 1. For example, the subnetwork consisting of the 4 nodes $\{E, F, I, J\}$ is a complex because its min-cut weight w^{mc} is greater than any larger subnetworks

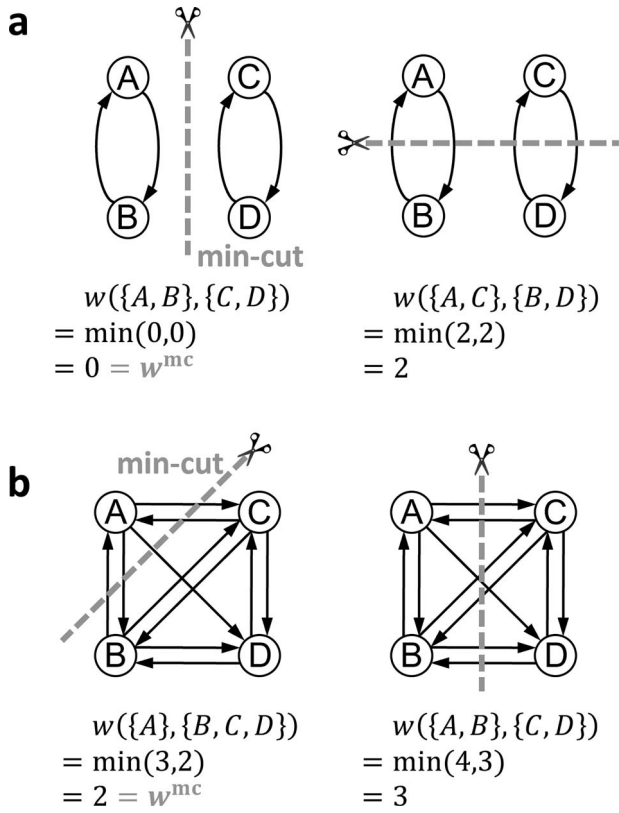


Fig. 3. Schematic of minimum cut. The minimum cut (min-cut) is the cut for which the strength of bidirectional connections is minimum among all possible cuts. In this figure, we assume that all edge weights are 1. a) A network consisting of 2 mutually disconnected groups AB and CD. The min-cut partitions this network into the 2 parts AB and CD, and its weight is zero ($w^{mc} = 0$). On the other hand, if the network is cut into AC and BD, the strength of bidirectional connections becomes nonzero. b) A network where all the parts are strongly connected in a bidirectional manner. This network cannot be separated without cutting many edges. The strength of bidirectional connections is therefore high, even for its min-cut ($w^{mc} = 2$).

containing it ($\{B, E, F, I, J\}$, $\{D, E, F, H, I, J\}$, etc.). In addition, the subnetwork $\{E, F, I, J\}$ is also a main complex because its w^{mc} is greater than those of not only larger subnetworks but also of any smaller subnetworks contained within it ($\{F, J\}$, $\{E, F, I\}$, etc.).

Extracting complexes

Fast and exact algorithm to search for complexes

If complexes are searched for by brute force, the computation time increases exponentially with the number of nodes N . This is because it is necessary to compute the min-cut weight w^{mc} for all of the $O(2^N)$ subnetworks and then compare these values. On the other hand, by using a fast and exact method we proposed in our previous study, HPC (Kitazono et al. 2020), we need to compute w^{mc} for only $N - 1$ subnetworks. As a result, the overall computation time increases only polynomially with N , and it is possible to analyze networks with several thousands of nodes in a practical time (see Fig. 5 for an actual computation time evaluated by a simulation).

We illustrate how HPC works using the example shown in Fig. 6. In the following, for simplicity of notation, we write a subnetwork consisting of a node set S simply as S . In HPC, a network is hierarchically partitioned by min-cuts until it is decomposed into single nodes. First, the whole network $V = \{A, B, C, D, E, F, G\}$ is divided by its min-cut (indicated by a dashed line) into $V_L = \{A, B, E, F\}$ and $V_R = \{C, D, G\}$. Then, V_L is divided into V_{LL} and V_{LR} , and V_R into V_{RL} and $\{G\}$. Finally, the whole network V is decomposed into 7 single nodes. After this process, we obtain the set of hierarchically partitioned subnetworks $V, V_L, V_R, V_{LL}, V_{LR}, V_{RL}$. We consider all the set of subnetworks $\mathcal{V} = \{V, V_L, V_R, V_{LL}, V_{LR}, V_{RL}\}$, excluding single nodes. We can then mathematically prove that any complex in the network belongs to \mathcal{V} . The proof is based on the mathematical property “monotonicity” and its satisfaction by the strength of bidirectional connections (Eq. (6) in Methods). Thus, we can consider \mathcal{V} as the candidate complexes. We can select complexes from these \mathcal{V} candidates without omissions or misidentifications. See Methods for more details.

In this process, we need to evaluate w^{mc} of only $N - 1$ ($= 6$) subnetworks, which are the subnetworks in \mathcal{V} . This number is much smaller than the number of subnetworks evaluated in the brute force method, $2^N - N - 1$ ($= 57$), which is the number of subnetworks consisting of more than 1 node.

Complexes in a network form a hierarchical structure

As we mention above, we can find complexes from among the candidate subnetworks (\mathcal{V}) appearing in the hierarchical partitioning process. Since the candidate subnetworks form a nested hierarchical structure, as we can see in Fig. 6, complexes in a network consequently form a nested hierarchical structure as in Fig. 1. That is, a complex contains another complex that is smaller in size but has a greater w^{mc} . A complex that is locally the smallest in size has a locally maximum w^{mc} , which is a main complex. See Methods for mathematical details.

Please note that a nested hierarchical structure is not necessarily a single peak structure, but can have multiple peaks (i.e. there can be multiple main complexes in a network). For example, in Fig. S4, there are 2 main complexes and the complexes form a nested hierarchical structure with the 2 main complexes as peaks.

Demonstration of the proposed method in a toy example

In this subsection, we demonstrate with a simple example how we can understand the structure of a network by extracting the complexes. We will also explain how to visualize the results, which will be used in showing the results of the mouse connectome analysis in the next subsection. In addition, to show the significance of considering bidirectionality, we illustrate using the same example how the results are affected if bidirectionality is not considered. Finally, to highlight the characteristics of

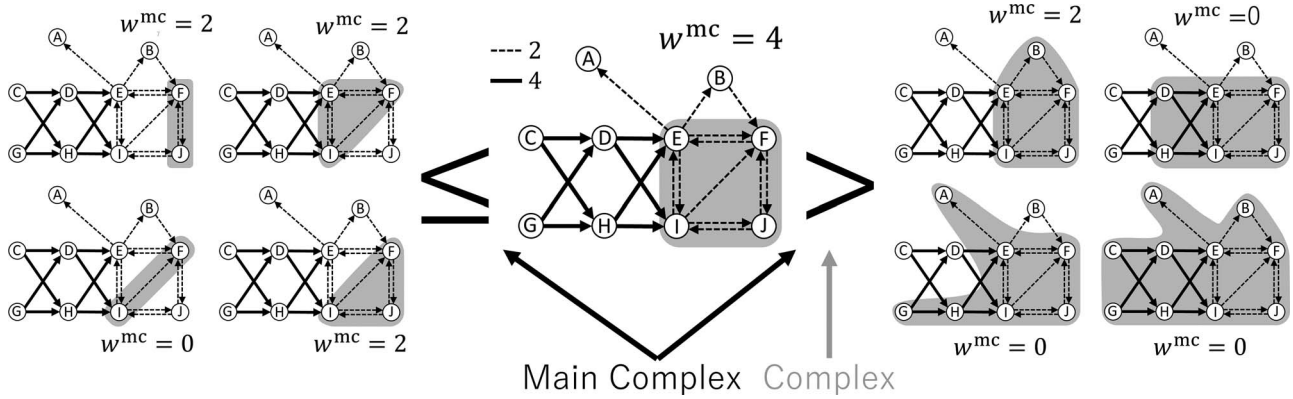


Fig. 4. Schematic of the definition of complexes and main complexes. The subnetwork $\{E, F, I, J\}$ is a complex because it has a greater min-cut weight w^{mc} than any larger subnetworks that contain it, namely, $\{B, E, F, I, J\}$, $\{D, H, E, F, I, J\}$, $\{A, E, F, G, H, I, J\}$, and so on. In addition, the subnetwork $\{E, F, I, J\}$ is a main complex because it has a greater min-cut weight than not only the larger subnetworks but also any smaller subnetworks within it, namely, $\{E, J\}$, $\{F, I\}$, $\{E, F, I\}$, and so on.

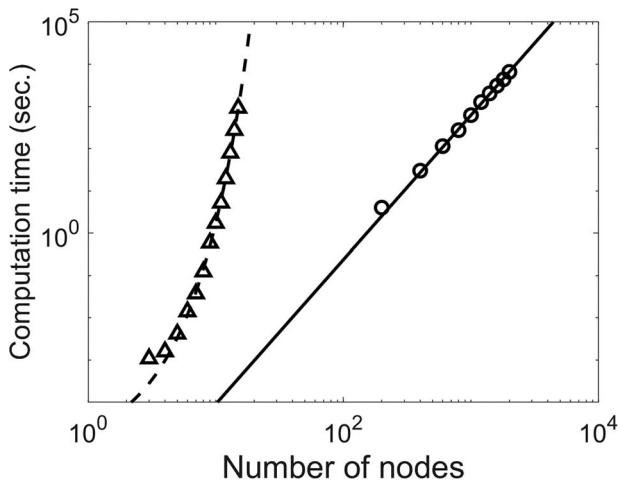


Fig. 5. Computation time of hierarchical partitioning for complex search. Computation time was evaluated by a simulation. In the simulation, networks with different numbers of nodes were randomly generated. The weight of each edge was sampled from a uniform distribution in the interval $(0, 1)$. The circles and the solid line indicate the computation time of hierarchical partitioning for complex search and a fitted linear function ($\log_{10} T = 3.419 \log_{10} N - 7.463$ ($T \propto N^{3.419}$)). The triangles and dashed line indicate the computation time of the exhaustive search and a fitted exponential function ($\log_{10} T = 0.5421N - 5.179$). The simulation was done on a machine with an Intel Xeon Gold 5220 processor at 2.20 GHz. All the calculations were implemented in MATLAB 2019a.

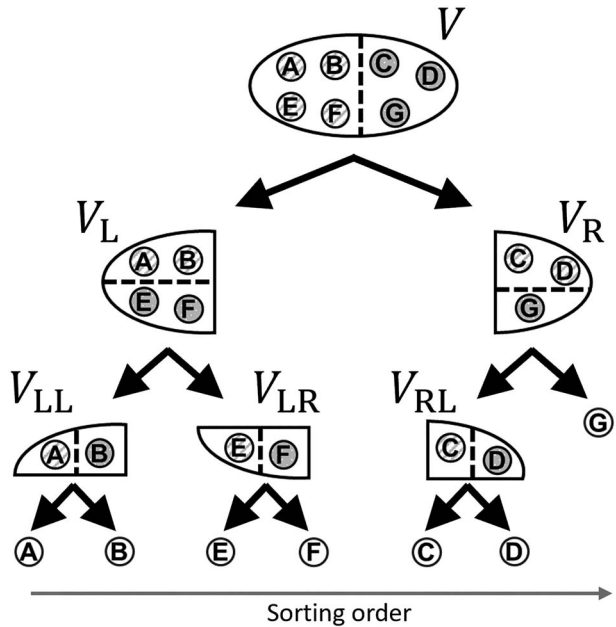


Fig. 6. Schematic of hierarchical partitioning for complex search. In HPC, a network is hierarchically partitioned by min-cuts until the network is decomposed into single nodes. In this example, the whole network $V = \{A, B, C, D, E, F, G\}$ is divided by its min-cut (indicated by a dashed line) into $V_L = \{A, B, E, F\}$ and $V_R = \{C, D, G\}$. Then, V_L is divided into V_{LL} and V_{LR} , and V_R into V_{RL} and $\{G\}$. Finally, the whole network V is decomposed into seven single nodes. In this process, we only need to evaluate the w^{mc} of $N - 1$ ($= 6$) subnetworks. This number is much smaller than the number of subnetworks evaluated in the brute force method, $2^N - N - 1$ ($= 57$), which is the number of subnetworks consisting of more than 1 node. The subnetworks appearing in this hierarchical partitioning process are candidate complexes. The bottom arrow indicates how we sort rows (columns) of the connection matrices in Figs. 7d and 7h and 8b and 8e (See Methods).

the proposed method, we compare it with a representative method for extracting network cores, s-core decomposition, which does not consider the bidirectionality of connections.

Understanding a network structure based on complexes

We consider the network shown in Fig. 7a, which is the same as that in Fig. 1. We visualize the complexes in this network in Fig. 7b. As mentioned in the description of Fig. 1, there are 2 complexes in this network. One is the node set $\{E, F, I, J\}$ (indicated by orange), and the other is the node set $\{B, E, F, I, J\}$ (indicated by light blue), and their min-cut weight w^{mc} values are 2 and 1, respectively. The node set $EFIJ$ is the main complex. The whole network with nonzero w^{mc} is always a complex because it is not

contained in a larger subnetwork. However, in this case, the min-cut weight w^{mc} of the entire network is 0. Thus, the entire network is not a complex because we do not call a network a complex when its w^{mc} is 0, i.e. it is completely separable. From this figure, we can see that the 2 complexes are nested. That is, the complex $\{E, F, I, J\}$, which has a larger w^{mc} , is contained in $\{B, E, F, I, J\}$, which has a smaller w^{mc} .

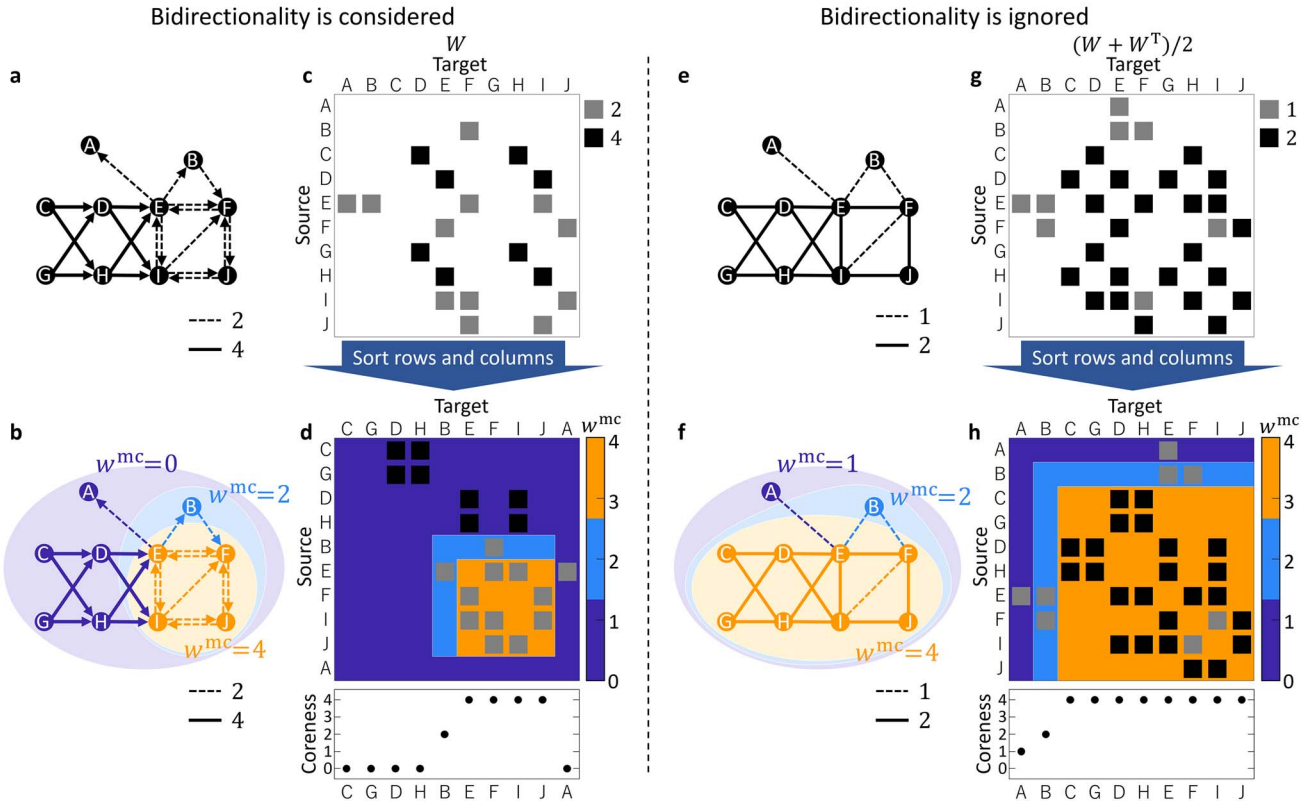


Fig. 7. Complexes in a toy network. Bidirectionality is considered in a–d and ignored in e–h. a) A toy network, which is the same as the network in Fig. 1. b) The structure of complexes. Each complex is indicated by a color representing the min-cut weight w^{mc} . c) The connection matrix of the network in a. Edge weight is shown in grayscale: white, gray, and black indicate 0, 1, and 2, respectively. d) The rows and columns of the connection matrix are sorted according to the hierarchical structure of the complexes. The color map changing from blue to orange indicates the min-cut weight w^{mc} of the complexes. Square areas correspond to the complexes and they are superimposed in the ascending order of w^{mc} . The plot at the bottom shows coreness values. e) The same network as in a except that the direction of connections are ignored, which corresponds to ignoring bidirectionality of connections (see Methods). f–h) The connection matrix, the complexes, and the sorted connection matrix and the coreness as in b–d.

We can also visualize complexes using connection matrices (Fig. 7c and 7d). In Fig. 7d, the rows and columns of the connection matrix (Fig. 7c) are sorted according to the hierarchical structure of the complexes (see Methods for a detailed description of the sorting process). In Fig. 7d, the color map indicates the min-cut weight w^{mc} of the complexes. Square areas correspond to the complexes and are superimposed in ascending order of w^{mc} . We can see that the colored square areas in the sorted connection matrix in Fig. 7d corresponds to the colored areas in Fig. 7b.

By using the complexes and their w^{mc} , we can also see how each node is distributed in complexes with different strength (w^{mc}). For example, node E is included in the strongest complex, i.e. the complex with the highest w^{mc} , and node B is included in the weaker complex, and so on. To quantify the strength of the complexes that each node is included in, we use an index called “coreness.” We define the coreness of node v as k_v if node v is included in a complex with $w^{mc} = k_v$ but not included in any complex with $w^{mc} > k_v$ (Eq. (11) in Methods). The coreness values correspond to the color of the nodes in Fig. 7b, and in the same way, to the color of the diagonal elements in Fig. 7d. From this figure, we can see, for example, that nodes E, F, I, and J have the largest value of coreness, which means

that they are included in a complex with the largest value of w^{mc} . On the other hand, the nodes A, C, D, G, and H have a value of 0 for coreness, indicating that they are not included in a complex with $w^{mc} > 0$.

Effect of considering bidirectionality

To illustrate the significance of considering bidirectionality, we compare the complexes extracted when considering bidirectionality with those extracted when ignoring bidirectionality. When we ignore bidirectionality, we quantify the strength of connections by the sum of all the edge weights between 2 parts (divided by 2 for consistency with the case when considering bidirectionality) as in Eq. (3) in Methods. Quantifying the strength of connections with this simple measure is equivalent to quantifying the strength of bidirectional connections with the original measure (Eq. (6) in Methods) after symmetrizing a network (i.e. taking the mean of the original connection matrix W and its transpose W^T) to make it virtually undirected. See Methods for more details.

Figure 7e represents the same network as that in Fig. 7a but the direction of the connection is ignored. The symmetrized connection matrix $(W + W^T)/2$ is shown in Fig. 7g. Figure 7f and 7h shows the results of the extracted complexes in this undirected network. Unlike

the case when bidirectionality is considered (Fig. 7c and d), the main complex contains not only *E, F, I*, and *J* but also *C, D, G*, and *H*. This is because the nodes *C, D, G*, and *H* are strongly unidirectionally connected to other nodes but not bidirectionally connected. Also, w^{mc} for the entire network is nonzero but is zero in the original directed network. Reflecting the structure of the complexes, the coreness values are highest for the nodes *C, D, E, F, G, H, I*, and *J*, and the coreness of every node is nonzero (Fig. 7h). As can be seen in this example, if the bidirectionality of connections is ignored, i.e. only the summed strength of connections is considered, the structure of the complexes is substantially changed.

A representative existing method for extracting cores of a network is s-core decomposition (Chatterjee and Sinha 2007, Hagmann et al. 2008, Heuvel and Sporns 2011, Harriger et al. 2012), which does not consider the bidirectionality of connections (Text S1). When s-core decomposition is applied to the network in Fig. 7, the obtained s-cores are identical to the complexes when bidirectionality is ignored. In general, it can be mathematically proven that s-cores are identical to complexes when bidirectionality is ignored under a certain condition (see Text S1 for details). In this example, the condition holds, and accordingly the extracted s-cores and the complexes are exactly the same.

Complexes in a mouse connectome

To demonstrate whether our method is able to extract meaningful bidirectionally connected cores in a brain network, we applied it to a mouse connectome (Oh et al. 2014) and extracted complexes. We consider this mouse connectome to be highly suitable for this purpose because it includes not only the cortical regions but also the subcortical, brainstem, and cerebellar regions, and has high spatial resolution.

The connectome (Supplementary Table 3 of Oh et al. 2014) consists of 213 brain regions in each hemisphere, giving 426 nodes in total. Each brain region is at a mid-ontology level and is classified into one of the major brain regions such as the isocortex, thalamus, and cerebellar cortex. The connection matrix is shown in Fig. 8a. The color coding at the top and left of the connection matrix indicates the major brain regions. The color of each entry in the matrix indicates the edge weight between the brain regions that can be considered to be proportional to the average out-degree of neurons projecting from one region to the other. See Oh et al. 2014 for a detailed description.

Brain regions included in complexes

We extracted complexes in the mouse connectome using the proposed method. The extracted complexes are visualized in Fig. 8b. In Fig. 8b, the rows and columns of the connection matrix (Fig. 8a) are sorted according to the structure of the complexes in the same way as in Fig. 7. The color coding, which changes from blue to yellow, represents the value of the min-cut weight w^{mc} of the complexes. See Table S1 for specific region names and

the detail values of w^{mc} . On the left side of Fig. 8b, the brain regions included in the complexes with high w^{mc} values are enlarged. Specifically, the regions included in the complexes with the highest (a main complex) to the 11th highest w^{mc} are extracted. The 11th highest w^{mc} corresponds to the upper quartile of the coreness value for all regions.

We first found that the extracted complexes were symmetric between the left and right hemispheres. That is, when a region on 1 side was included in a complex, the corresponding region on the opposite side was also included in the complex. This means that the connections between the left and right sides were strong enough so that each complex straddled the left and right sides symmetrically (in Text S2, we analyze how strongly the structure of the complexes depended on the left-right connections.) In the following, since the extracted complexes were symmetric, we do not distinguish between the left and right brains.

We now describe the regions that constitute the complexes with high w^{mc} . We observed that many regions in the cerebral cortex are included in top complexes (complexes with high w^{mc}). In particular, mainly the isocortical regions constitute the first through third complexes. The only exceptions are the claustrum (CLA) and the basolateral amygdalar nucleus (BLA) in the cortical subplate, which are included in the third complex. The 4th to 9th complexes consist of the regions listed above plus other isocortical and thalamic regions, and the lateral parts of the entorhinal cortex (ENTl) in the hippocampal formation. The 10th and 11th complexes further includes some regions in the isocortex, olfactory areas, cortical subplate, and pallidum. The regions in the other major regions are not included in the 1st to 11th complexes.

Thus, the regions included in the complexes with the highest w^{mc} are not evenly distributed among all major regions, but are rather concentrated in the cortical (particularly isocortical) regions and thalamic regions. We can confirm the unevenness among the major regions from the coreness values (Fig. 8d, Table S2). Regions in the isocortex have particularly high coreness values (i.e. they are included in complexes with high w^{mc}). Also, regions in the thalamus have high coreness values. Other regions with high coreness values are the CLA and BLA in the cortical subplate, followed by ENTl in the hippocampal formation, and some regions in the olfactory areas, cortical subplate and pallidum. On the other hand, regions in the other major regions have low coreness values. In particular, regions in the cerebellar cortex and cerebellar nuclei have much lower coreness values.

These results suggest that there appears to be a good correspondence between whether or not a region is included in complexes with high w^{mc} and whether or not a region is considered important for consciousness. For example, the isocortex and thalamus are considered essential to consciousness, whereas the cerebellar cortex and cerebellar nuclei do not directly contribute to consciousness (Boly et al. 2012, Koch et al. 2016,

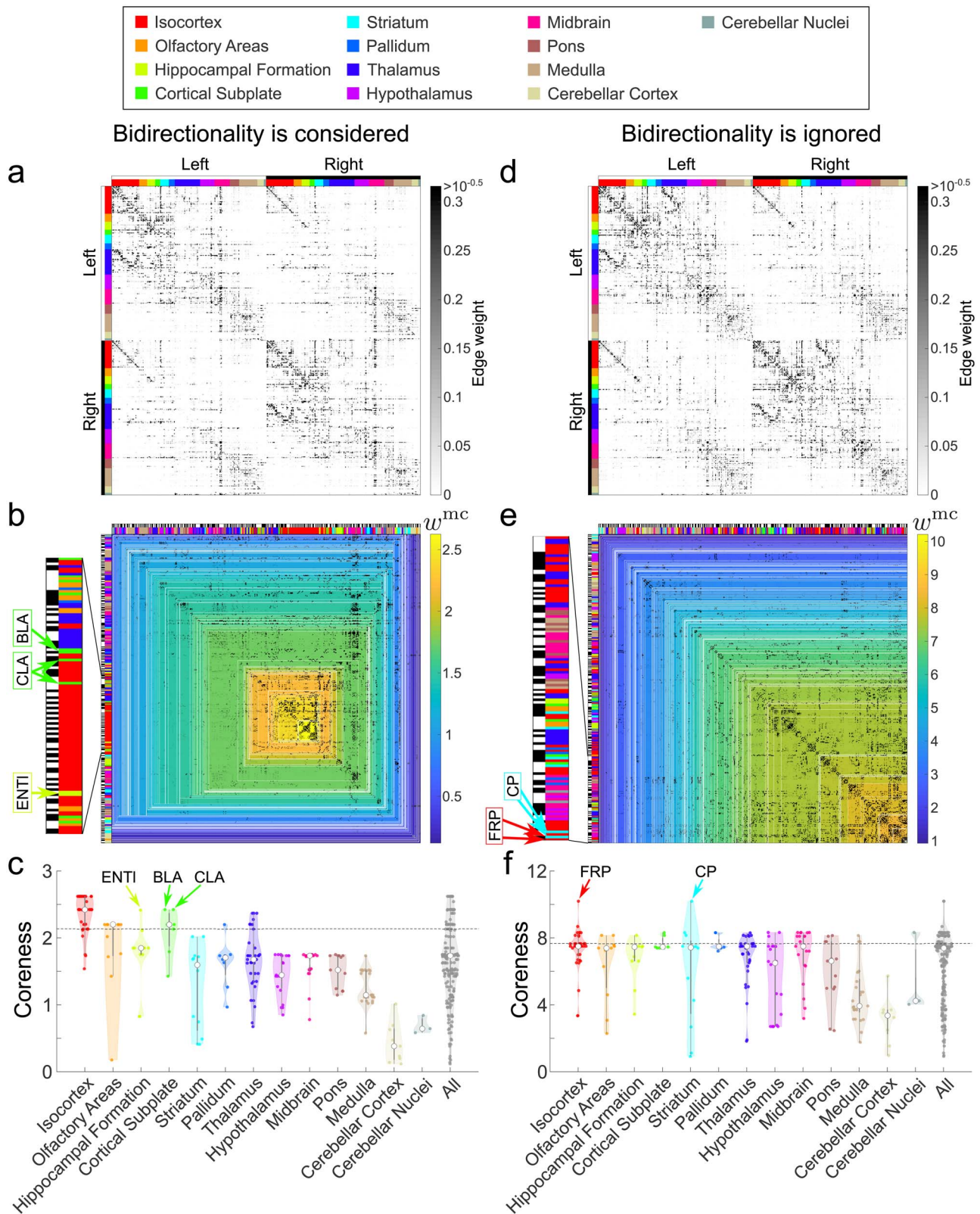


Fig. 8. Complexes in a mouse connectome. Bidirectionality is considered in panels a–c and ignored in panels d–f. a) The inter-region connection matrix of the mouse connectome. The color bars at the left and top of the matrix represent major brain regions and whether they are in the left or right brain. b) Structure of complexes. A connection matrix in which the rows and columns are sorted according to the hierarchical structure of complexes as in Fig. 7. The change in color map from blue to yellow indicates the min-cut weight w^{mc} of the complexes. Square areas correspond to complexes and are superimposed in ascending order of w^{mc} . At the left, the brain regions included in the complexes with high w^{mc} values (top 1, i.e. the main complex, to top 11) are enlarged. c) The coreness values are plotted for each major brain region. The regions above the dashed line, which indicates the upper quartile of coreness values for all regions, correspond to the enlarged regions in b. d) The inter-region connection matrix when bidirectionality is ignored, i.e. the mean of the original connection matrix W and its transpose W^T , $(W + W^T)/2$. e) Structure of complexes. At the left, the brain regions included in the complexes with high w^{mc} values (top 1, i.e. the main complex, to top 14) are enlarged. f) The coreness values are plotted for each major brain region. The regions above the dashed line, which indicates the upper quartile of coreness values for all regions, correspond to the enlarged regions in e.

Tononi et al. 2016). Other than the isocortical and thalamic regions, the CLA in the cortical subplate has long been associated with consciousness (Crick and Koch 2005). We discuss the relationship between consciousness and the regions included or not included in the top complexes in detail in Discussion.

Large difference in complexes when bidirectionality is ignored

Next, we investigated how the results change when the bidirectionality of the connections is ignored, i.e. the direction of connections is ignored and only the summed strength of connections is considered, as is in the example in Fig. 7d–f. Figure 8d shows the symmetrized connection matrix, based on which the complexes are extracted.

Let us first mention that similar to the case when considering bidirectionality, the results are symmetric between the left and right brains. That is, one of the following 2 conditions is satisfied: (1) as is in the case when considering bidirectionality, if a region on 1 side was included in a complex, the corresponding region on the opposite side was also included in the complex; or (2) if a region on 1 side was included in a complex S , the corresponding region on the opposite side was included in another complex with the same strength of bidirectional connections as that of S . We therefore do not distinguish between the left and right brains in the following.

Figure 8e represents the extracted complexes. See Table S1 for specific region names and the detail values of w^{mc} . In the left side of Fig. 8e, the brain regions included in the complexes with the highest to 14th values of w^{mc} is enlarged. The 14th highest w^{mc} corresponds to the upper quartile of the coreness value for all regions. By comparing the brain regions in the top complexes shown in Fig. 8e with Fig. 8b, we can see that these are largely different in the sense that the brain regions in the top complexes are evenly distributed in almost all of the major brain regions when bidirectionality is ignored but are included in particular major regions such as the isocortex or thalamus when bidirectionality is considered. In fact, regions in all major regions except the cerebellar cortex are included in the complex with second highest w^{mc} when bidirectionality is ignored. We then investigate the change by ignoring bidirectionality using the coreness values (Fig. 8f, Table S2). We observed that the difference among the major brain regions becomes small when bidirectionality is ignored. The maximum values of coreness are equal for many major regions. This reflects the fact that the regions included in the complex with large w^{mc} are evenly distributed in many major regions. Also, the median coreness values (represented by white circles in Fig. 8f) are equal for many major regions.

From Fig. 8f, we can see that there are 2 regions that have a particularly high coreness value, namely the frontal pole (FRP) in the isocortex and the caudoputamen

(CP) in the striatum. The high coreness of these 2 regions is due to the strong connection from CP to the FRP. On the other hand, when bidirectionality is considered, the coreness value of CP is low because the connection in the opposite direction, FRP to CP, is weak.

To directly compare the 2 cases, namely when bidirectionality is considered or ignored, we made a scatter plot of coreness in Fig. 9, (1, 2) or (2, 1) panel (see also the network diagram that compares the 2 cases in Fig. S5). We can see that the distributions in the 2 cases are very different: regions with high coreness values when bidirectionality is considered do not necessarily have high coreness values when bidirectionality is ignored, and vice versa. We can also see that the points lie completely above the identity line ($y = x$) in the panel (2, 1) in Fig. 9. This is because the coreness for complexes is always larger when bidirectionality is ignored than when it is considered (See Text S3 for a proof).

Thus, if we ignore the bidirectionality of connections, the results change drastically; the complexes no longer necessarily consist of regions presumably essential to consciousness. This suggests that considering the bidirectionality of connections is important in associating the network core complexes with consciousness.

Comparison with other existing methods

To further assess the significance of considering bidirectionality, we compare the proposed method with other existing methods that do not take account of bidirectionality. We first consider s-core decomposition (Chatterjee and Sinha 2007, Hagmann et al. 2008, Heuvel and Sporns 2011, Harriger et al. 2012), one of the most popular methods for extracting network cores. As we mentioned in the toy network analysis, s-core decomposition does not consider bidirectionality of connections and s-cores become identical to complexes when bidirectionality is ignored under a certain mathematical condition (see Text S1 for details). In the mouse connectome case, this condition does not hold exactly, but almost does, and the obtained s-cores are almost the same as the complexes when bidirectionality is ignored. We can see that the coreness values for s-core decomposition (coreness for s-core decomposition is defined in the same way as for complexes; see Methods) and those for the complexes when bidirectionality is ignored are almost identical (Fig. 9, (2,3) or (3,2) panel). Since for complexes the difference in coreness values among the major regions is small when bidirectionality is ignored (Fig. 8f), the difference is accordingly also small for s-cores. This means that the s-cores with a high s do not necessarily consist of regions in particular major regions, and therefore do not consist mainly of regions considered essential to consciousness.

Next, we investigated whether the complexes with strong bidirectional connections simply consist of the brain regions with high degree, i.e. network hubs (Heuvel and Sporns 2013, Fulcher and Fornito 2016). The degree of a node is the sum of weights of edges connecting to it, irrespective of direction (Eq. (13) in Methods), and

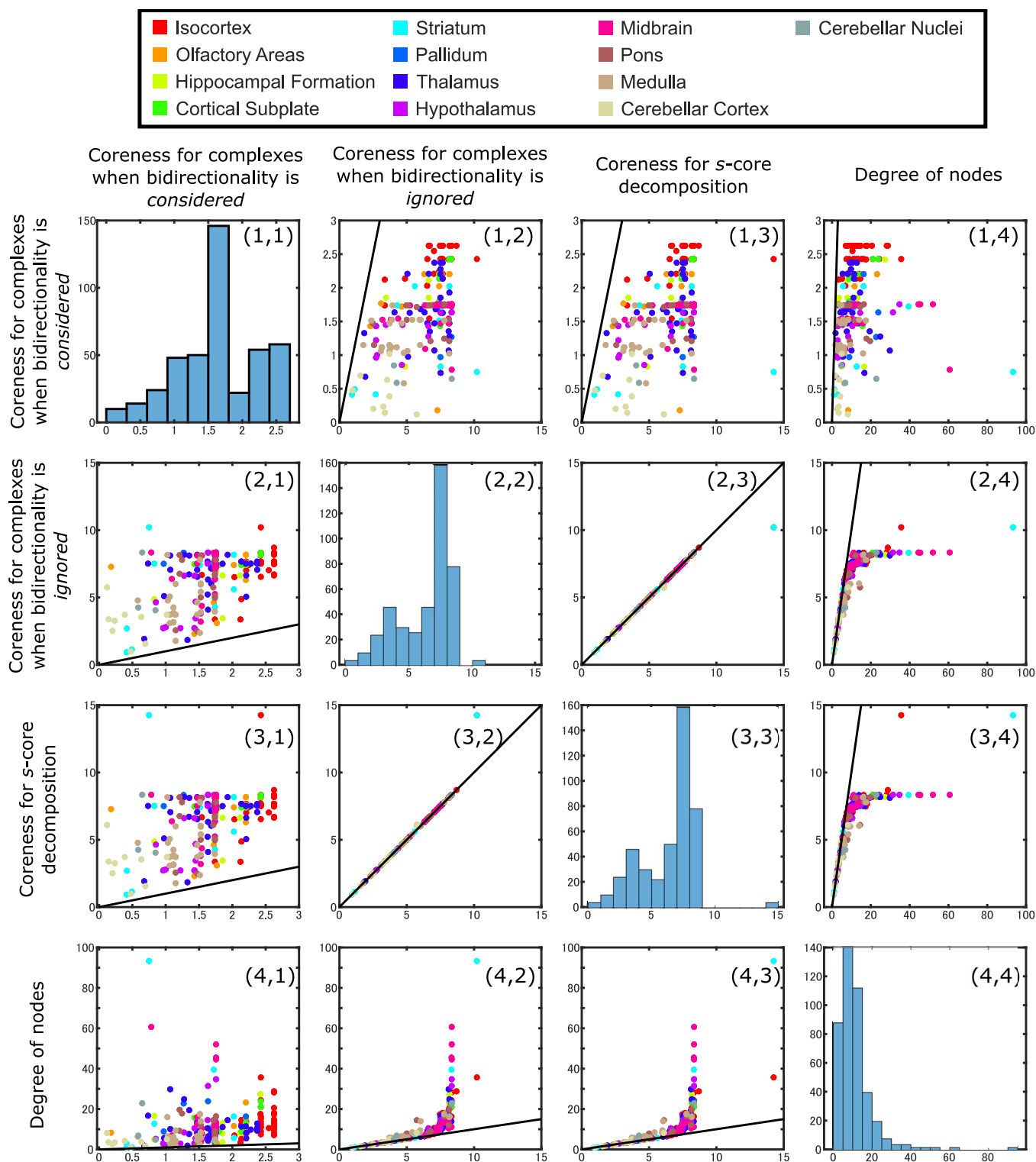


Fig. 9. Histograms and scatter plots of coreness values and degree of nodes. Histograms of the coreness values and degree of nodes appear along the diagonal. Scatter plots of pairs of coreness value or degree of nodes appear in the off diagonal. Color of the points indicates major brain regions. The line in each scatter plot is the identity line ($y = x$).

the network hubs are nodes with high degree. If the complexes with strong bidirectional connections consist of the hub regions, this means that bidirectionality does not matter to the extraction of complexes. We observed that the complexes when bidirectionality is considered do not necessarily consist of regions with high

degree. In Fig. 9, (1,4) or (4,1) panel, we can see that the coreness values for the complexes when bidirectionality is considered and the degree are only weakly correlated: many brain regions with high coreness values have low degree. In contrast, the coreness for the complexes when bidirectionality is ignored (and the coreness for

s-core decomposition) corresponds well to the degree, especially around the lower degree range (Fig. 9, (2,4) or (4,2) panel). Thus, these results indicate that the consideration of bidirectionality in the proposed method enabled us to extract core structures in the mouse connectome that cannot be extracted by simple degree-based methods. We also compared the proposed method with a motif analysis (Milo et al. 2002, Sporns and Kötter 2004, Onnela et al. 2005, Rubinov and Sporns 2010, Harriger et al. 2012). We found that nodes in strong complexes tended to participate more in motifs with cycles than in motifs without cycles. See Text S4 for the details of the results.

Discussion

In this study, we proposed a method to find the network cores, called “complexes,” which consist of strong bidirectional connections. If we search for complexes by brute force, computation time grows exponentially with the number of nodes. To solve this problem, we introduced a fast and exact algorithm proposed in our previous study, HPC (Kitazono et al. 2020). The HPC algorithm reduces the computation time to polynomial time and enables the analysis of large networks consisting of up to several thousand nodes in a practical amount of time. By utilizing HPC, we extracted complexes in a mesoscale, whole-brain mouse connectome consisting of 426 regions (Oh et al. 2014), with the aim of identifying subnetworks in the brain relevant for consciousness. We found that complexes with strong bidirectional connections include many brain regions that have been considered essential for consciousness in previous studies. We also found that if bidirectionality is ignored, the brain regions included in the complexes with strong connections are evenly distributed in major brain regions regardless of whether or not they are relevant for consciousness. These results indicate that bidirectionality may be the key that characterizes the regions essential for consciousness.

Correspondence between complexes and essential regions for consciousness

In this subsection, we discuss the relationship between complexes and the regions essential for consciousness. We first discuss in detail the brain regions with high coreness, i.e. the regions included in the complexes with strong bidirectional connections, and then the brain regions that are not included in such strong complexes.

First, many regions in the cerebral cortex, especially the isocortical regions, have high coreness. Previous studies suggested that bidirectional interaction among isocortical regions is essential for consciousness (Koch et al. 2016, Lamme 2018, Mashour et al. 2020). In addition to the isocortical regions, the claustrum (CLA) in the cortical subplate also has high coreness. Francis Crick speculated that the CLA is the seat of consciousness, and that metaphorically speaking it plays the role of the

conductor that orchestrates the brain (Crick and Koch 2005). In fact, recent studies in mice suggest that the CLA is involved in the control of arousal and sleep levels (Narikiyo et al. 2020, Norimoto et al. 2020). The CLA is also suggested to have a role in salience processing and attention control (Mathur 2014, Goll et al. 2015, Atlan et al. 2018, Jackson 2018, Smith et al. 2019, Smith et al. 2020), and might therefore be involved in selecting what comes to one’s conscious perception.

As for the subcortical regions, many thalamic regions also have high coreness. It is suggested that the thalamo-cortical loop—a circuit composed of the thalamus and cortical regions—is important for consciousness (Mumford 1991, Llinás 2003, Ward 2011, Aru et al. 2019).

As we discuss above, the brain regions with high coreness seem to correspond well with the regions that are considered essential to consciousness. However, the brain regions with high coreness also include some regions which have not yet been shown to be relevant to consciousness. A notable example is the basolateral amygdalar nucleus (BLA) in the cortical subplate. The BLA has the same coreness as the CLA, which is highest excluding the isocortical regions. The BLA is thought to be critical for emotion (positive and negative valence) and to mediate conditioning both for reward and fear (O’Neill et al. 2018). To our knowledge, however, the relationship between the BLA and consciousness is little understood (e.g. whether the BLA directly contributes to subjective experience of emotions). Further investigation of such brain regions would be useful.

In addition to investigating whether the regions with high coreness are relevant to consciousness, it is also important to investigate the converse, namely whether the regions with low coreness, which are very weakly bidirectionally connected, are presumably irrelevant to consciousness. Notably, we found that all nodes in the cerebellar cortex and cerebellar nuclei have low coreness. It is well known that the cerebellum does not directly contribute to consciousness (Lemon and Edgley 2010, Yu et al. 2015) even though it has much more neurons than the cerebrum. We also found that the regions in the midbrain, medulla, and pons—the major regions which constitute the brainstem—have low coreness. Although the brainstem is important for enabling consciousness, it is not thought to contribute directly to conscious experience, in the same way that the heart is important for enabling consciousness but does not contribute directly to conscious experience. These are called background conditions (Koch et al. 2016).

Taking our results together, we have found that (1) brain regions presumably essential to consciousness have high coreness—that is, they are included in complexes with strong bidirectional connections; and that (2) brain regions presumably irrelevant to consciousness have low coreness, meaning that the regions are only weakly bidirectionally connected to other regions.

Significance of considering bidirectionality

We here discuss how considering bidirectionality affects the results of the complexes and its importance in relating the complexes with the locus of consciousness. As we have seen in Results, when bidirectionality is ignored, the structure of the complexes largely differs from that when bidirectionality is considered. One large difference is that the difference in coreness between major regions is smaller when bidirectionality is ignored. Regions in the cerebellum (cerebellar cortex and cerebellar nuclei) and brainstem (midbrain, pons, and medulla), which have smaller coreness than other regions such as the isocortex and thalamus when bidirectionality is considered, have similarly high coreness to these regions when bidirectionality is ignored. As mentioned in the previous section, these regions have not been so far considered to directly contribute to consciousness (Lemon and Edgley 2010, Yu et al. 2015, Koch et al. 2016). Another large difference is that the caudoputamen (CP) in the striatum, which is not included among complexes with large w^{mc} when bidirectionality is considered, forms the main complex when bidirectionality is ignored. The striatum, more broadly the basal ganglia, is not thought to contribute directly to consciousness (Tononi 2004, Boly et al. 2017) (but see Slagter et al. 2017, Afrasiabi et al. 2021).

Thus, when bidirectionality is ignored, regions both relevant and irrelevant to consciousness are evenly included in the strong complexes. This means that the seemingly good correspondence between complexes and regions relevant to consciousness we identified when considering bidirectionality is largely lost.

The operation of ignoring bidirectionality is equivalent to symmetrizing a network and treating it as practically an undirected network, as described in Results. Therefore, the fact that ignoring bidirectionality leads to large changes in the complex can be rephrased as saying that ignoring “directionality” leads to large changes in the complex. In a previous study, Kale et al. studied how ignoring directionality affects graph-theoretic measures of connectomes and showed that hubs are particularly affected (Kale et al. 2018). The concept of hubs is based on the local structure of a network, connectivity of a single node. The concept of complexes, on the other hand, is based on the global structure of a network, which is how strongly subnetworks are connected. Thus, the present study can be said to show that not only local structures (e.g. hubs) but also global structures, i.e. complexes, can be greatly affected by ignoring directionality.

Comparison with other core extraction methods in terms of bidirectionality

In the literature, a variety of methods have been applied to connectomes to extract network cores in which elements are densely connected to each other. In what follows, in terms of bidirectionality of connections, we compare complexes with 3 representative methods for core extraction, namely *s*-core decomposition, network hubs, and modularity maximization.

In this study, we first compared *s*-core decomposition with the complexes. *s*-core decomposition is a representative method that has been widely applied to connectomes of various species (Chatterjee and Sinha 2007, Hagmann et al. 2008, Harriger et al. 2012), and the extracted cores are related to certain functions. We showed that the *s*-cores extracted from the mouse connectome largely differ to the complexes when bidirectionality is considered, but are almost identical to the complexes when bidirectionality is ignored. This means that the consideration of bidirectionality enabled us to reveal core structures that cannot be revealed by *s*-core decomposition.

We next compared network hubs (Heuvel and Sporns 2013, Fulcher and Fornito 2016) with the complexes. Previous studies showed that the brain network contains cores in which hubs (high-degree nodes) are densely connected to each other (called “rich-clubs”) (Heuvel and Sporns 2011, Harriger et al. 2012, Fulcher and Fornito 2016). We showed that in the mouse connectome the complexes with strong bidirectional connections included not only high-degree nodes but many low-degree nodes (Fig. 9). This means that the core structures revealed by the complexes largely differ to the structures that can be revealed by hub-based methods.

Finally, we discuss modularity maximization, which is also widely used in connectome analysis (Sporns and Betzel 2016). Similar to the proposed method, modularity maximization is a method used to extract subnetworks with dense connections. Its objective is, however, qualitatively different from that of the proposed method. The objective of modularity maximization is to partition a network into non-overlapping cores (called modules or communities) with dense internal connections, and not to decompose a network hierarchically as for complexes. This difference in objectives hampers direct quantitative comparison of the 2 methods by experiments. We therefore confined ourselves here to a qualitative comparison in terms of bidirectionality. The mathematical formulation of modularity maximization is suitable for undirected networks (Sporns and Betzel 2016). It is therefore impossible to consider the direction of connections and hence bidirectionality. However, a variant of the modularity maximization methods considers the direction of connections when defining the density of connections (Leicht and Newman 2008). This variant does not consider bidirectionality, however, and extracted modules do not therefore necessarily consist of bidirectional connections, i.e. modules have fully feed-forward structures.

As exemplified above, the core extraction methods in wide current use for connectome analysis do not consider the bidirectionality of connections. Thus, we conclude that the main result of the present paper, which has revealed the correspondence between the network cores of the brain and consciousness, can only be achieved by methods such as the proposed method which take account of the bidirectionality of connections.

Number of main complexes

In general, since main complexes are “local” maxima in terms of w^{mc} , there can be multiple main complexes in a network as we mentioned in Results. An extreme example occurs when a network consists of 2 mutually disconnected modules: in this case, there will be 2 (or more) main complexes. The presence of multiple main complexes in a network indicates that the network consists of multiple weakly coupled modules.

In the mouse connectome, there are 5 main complexes when bidirectionality is considered, albeit that we only mentioned one of them in Results. The main complex we mentioned has the highest w^{mc} of all main complexes (and of all subnetworks by definition) and is largest in size among all main complexes. The other main complexes, which we did not mention, have low w^{mc} and consist of only 2 regions, i.e. are minimum in size. This means that the mouse connectome can be almost considered to consist of 1 large module.

Thus, we can ascertain modular structure using complexes. In contrast, when we use s -core decomposition, this cannot be ascertained. Consider a network consisting of 2 densely connected parts, as shown in Fig. S6. In this case, s -core decomposition extracts the entire network as a single core and does not extract the modular structure in this network. The proposed method, on the other hand, extracts 2 modules as 2 main complexes. This is because s -core decomposition uses only local information, i.e. degree of nodes, and cannot consider the global structure of a network as a whole. On the other hand, the proposed method uses the global information of networks, min-cuts, which allows us to extract the modular structure.

Limitations of this study and future direction

We searched for complexes using a method proposed in our previous study, hierarchical partitioning for complex search (HPC) (Kitazono et al. 2020). The computation time of HPC increases only polynomially with the number of the nodes, which is much smaller than the exponential increase in brute force search. This enables us to find complexes in a network with several thousand nodes in a practical amount of time. However, to find complexes in networks with more nodes ($N \gg 2,000-3,000$), a further speeding-up is required. One possible solution is the use of approximation algorithms for min-cut search (Kirkpatrick et al. 1983) instead of the exact algorithm, the Hao-Orlin algorithm (Hao and Orlin 1994), which we used in this study.

In this study, we discussed how complexes extracted from the mouse connectome, which consists of anatomical connections, are related to consciousness. We should note, however, that it is not the anatomical connections themselves that are directly responsible for conscious experiences at a particular time, but rather interactions between brain regions that result from the brain activity (Cauller and Kulics 1991, Lamme et al. 1998, Supèr et al. 2001, Auzstulewicz et al. 2012, Self et al. 2012,

Sachidhanandam et al. 2013, Koivisto et al. 2014, Tang et al. 2014, Manita et al. 2015). The location of bidirectional interactions changes from time to time, and the brain regions that mediate consciousness can also change accordingly (Koch et al. 2016, Tononi et al. 2016). To capture such dynamic change in consciousness, future research should therefore aim to extract complexes from functional or causal networks constructed by quantifying interactions using brain activity data. The relationship between anatomical and functional networks is not as simple as a one-to-one correspondence. It is empirically known, however, that there are some similarities between them (Mišić et al. 2016, Grandjean et al. 2017, Sethi et al. 2017), as would be naturally expected from the fact that anatomical connections are the basis for interactions between brain regions. In addition, bidirectional connections are suggested to play a crucial role in synchronizing brain regions (Gollo et al. 2014). We therefore expect that complexes extracted from functional networks could be similar to the complexes extracted from the anatomical network.

Methods

Strength of bidirectional connections

Here, we propose a way of quantifying how strongly 2 parts of a graph are bidirectionally connected. We consider a directed graph $G(V, E)$, where V and E are the node set and the edge set, respectively. For a bi-partition of the node set V , (V_L, V_R), there are 2 types of edges that connect V_L and V_R depending on its direction. One is the set of edges outgoing from V_L to V_R :

$$E(V_L \rightarrow V_R) = \{(u, v) \in E | u \in V_L, v \in V_R\}. \quad (1)$$

The other is the set of the edges incoming to V_L from V_R (or equivalently, outgoing from V_R to V_L):

$$E(V_R \rightarrow V_L) = \{(u, v) \in A | u \in V_R, v \in V_L\}. \quad (2)$$

When we ignore the directions of the connections between V_L and V_R , the simplest way of quantifying the strength of the connections is to add up all the weights of the edges that connect V_L and V_R regardless of their directions:

$$w_{\text{simple sum}}(V_L; V_R) = \frac{1}{2} \sum_{e \in E(V_L \rightarrow V_R) \cup E(V_R \rightarrow V_L)} w_e, \quad (3)$$

where w_e represent the weight of the edge e . The factor 2 in the denominator is for consistency with the strength of bidirectional connections, as explained later.

On the other hand, when we consider the bidirectionality of connections between V_L and V_R , we first separately

add up the weight of the edges for each direction,

$$w(V_L \rightarrow V_R) = \sum_{e \in E(V_L \rightarrow V_R)} w_a, \quad (4)$$

$$w(V_R \rightarrow V_L) = \sum_{e \in E(V_R \rightarrow V_L)} w_a, \quad (5)$$

and then define the strength of bidirectional connections as their minimum:

$$w(V_L; V_R) = \min(w(V_L \rightarrow V_R), w(V_R \rightarrow V_L)). \quad (6)$$

With this definition, if 2 parts of a network are only connected unidirectionally, as in Fig. 2a, the strength of bidirectional connections $w(V_L; V_R)$ is 0, which means that the 2 parts are considered to be “disconnected” bidirectionally. In Fig. 2b, the connection from one part to the other part is strong (3) but that in the other direction is weak (1). Consequently, the strength of bidirectional connections is low ($w(V_L; V_R) = 1$). In Fig. 2c, the connections in both directions are strong (2) and the strength of bidirectional connections is high ($w(V_L; V_R) = 2$). If we ignore the directionality of connections and add up the edge weights in the 2 directions ($w_{\text{simple sum}}$), the strength of connections is evaluated as 2 in all 3 cases.

The 2 measures $w(V_L; V_R)$ and $w_{\text{simple sum}}$, with and without considering bidirectionality, are equal to each other when connections are symmetric ($w_{(u,v)} = w_{(v,u)}$).

Minimum cut

Definition of minimum cut

A cut of a graph $G(V, E)$ is called a minimum cut if the strength of the connections for the cut is not higher than that of any other cut. More formally, a minimum cut $(V_L^{\text{mc}}, V_R^{\text{mc}})$ is defined as follows:

$$(V_L^{\text{mc}}, V_R^{\text{mc}}) = \arg \min_{(V_L, V_R) \in \mathcal{P}_V} w(V_L; V_R), \quad (7)$$

where \mathcal{P}_V denotes the set of all bi-partitions of V . We denote the weight of the minimum cut $(V_L^{\text{mc}}, V_R^{\text{mc}})$ of a graph G as

$$w_G^{\text{mc}} := w(V_L^{\text{mc}}; V_R^{\text{mc}}). \quad (8)$$

Fast and exact algorithms for searching for min-cuts

We defined a measure of strength of bidirectional connections as in Eq. (6). Although this definition is different from the canonical definition of a graph cut weight for directed graphs, the minimum cut problems for the 2 definitions are equivalent (Text S5). Therefore, we can use a well-established algorithm to solve the minimum cut problem. In this study, we utilize the Hao-Orlin algorithm (Hao and Orlin 1994). Its time complexity is $O(|V||E| \log(|V|^2/|E|))$, where $|V|$ and $|E|$ are the number of nodes and edges, respectively.

Complex

In this section, we introduce the definition of a complex (Balduzzi and Tononi 2008, Tononi 2008, Kitazono et al. 2020). We also introduce a main complex, which is a stronger definition of a complex (Balduzzi and Tononi 2008, Tononi 2008, Kitazono et al. 2020).

To formally define complexes, we need to introduce the concept of an induced subgraph. Let G be a graph consisting of node set V and edge set E , and let $S \subseteq V$ be a subset of nodes. Then, an induced subgraph $G[S]$ is the graph consisting of all the nodes in S and all the edges connecting the nodes in S . The min-cut weight of $G[S]$ is denoted by $w_{G[S]}^{\text{mc}}$. We are now ready to define complexes.

Definition 1.1 (Complex). An induced subgraph $G[S]$ ($S \subseteq V$) is called a complex if it satisfies $w_{G[S]}^{\text{mc}} > 0$ and $w_{G[S]}^{\text{mc}} > w_{G[T]}^{\text{mc}}$ for any subset T that is a superset of S ($T \supset S$ and $T \subseteq V$).

A schematic explanation of the definition of a complex is shown in Fig. 4. In this schematic, we consider induced subgraphs of a graph G consisting of 10 nodes $\{A, B, \dots, J\}$. An induced subgraph $G[\{E, F, I, J\}]$ is a complex because it has greater w^{mc} than any induced subgraph of G that is its supergraph (e.g. $G[\{B, E, F, I, J\}]$ and $G[\{D, E, F, H, I, J\}]$).

The whole graph G is a complex if it satisfies $w_G^{\text{mc}} > 0$ by definition. We define $w^{\text{mc}} = 0$ for single nodes because we cannot consider partitions of a single node. Therefore, single nodes cannot be complexes.

An induced subgraph is called a main complex if its min-cut weight w^{mc} is larger than those of any induced subgraphs that are its supergraphs, and is also larger than or equal to those of any induced subgraphs that are its subgraphs. That is, a complex is called a main complex if its min-cut weight w^{mc} is larger than or equal to those of any induced subgraphs that are its subgraphs.

Definition 1.2 (Main complex). A complex is called a main complex if it satisfies $w_{G[S]}^{\text{mc}} \geq w_{G[R]}^{\text{mc}}$ for any subset R of S ($R \subset S$).

A schematic explanation of the definition of main complexes is shown in Fig. 4. An induced subgraph $G[\{E, F, I, J\}]$ is a main complex because it is a complex and has greater w^{mc} than any induced subgraph that is its subgraph (e.g. $G[\{F, J\}]$ and $G[\{E, F, I\}]$).

At this point some readers might wonder if, instead of Definition 1.1, we could define a core, which we call here a sub-complex, in the following form:

Definition 1.3 (Sub-complex). An induced subgraph $G[S]$ ($S \subseteq V$) is called a sub-complex if it satisfies $w_{G[S]}^{\text{mc}} > 0$ and $w_{G[S]}^{\text{mc}} > w_{G[R]}^{\text{mc}}$ for any subset R that is a subset of S ($R \subset S$).

The difference between Definitions 1.1 and 1.3 is whether we compare the w^{mc} of a core with those of its supersets or subsets. We explain here why we define a core complex in accordance with Definition 1.1, and not with Definition 1.3. The reason lies in 2 properties of complexes. The first property is the hierarchical nature of complexes. As mentioned in the Results, complexes form a nested hierarchical structure. That is, if there are 2 complexes, they satisfy either the property that

one contains the other or that they do not overlap with each other. On the other hand, sub-complexes do not necessarily form such a hierarchical structure: when there are 2 sub-complexes, the 2 may partially overlap. As an example, consider a fully connected network where the weights of the edges are all equal. In this case, every subnetwork is a sub-complex, and extracting sub-complexes does not make sense. The second property is that complexes can be said to be representative of other subnetworks. The coreness of a node v is defined as “the largest of the w^{mc} of all complexes containing v (Eq. (11)).” As we will see later in this section, the value of coreness is equal to “the largest of the w^{mc} of all subnetworks (not necessarily complexes) that contain v (Eq. (12)).” This means that the w^{mc} of complexes is representative of the w^{mc} of all subnetworks. On the other hand, if we define coreness for sub-complexes in the same way, it can be smaller than Eq. (12). From these 2 properties, we consider that it is reasonable to use Definition 1.1.

Relation to integrated information theory

As we mention in Introduction, the term and concept of complexes in the proposed method originate from the integrated information theory (IIT) of consciousness (Tononi 2004, Balduzzi and Tononi 2008, Oizumi et al. 2014, Tononi et al. 2016). We here clarify how the proposed method is related to IIT.

The definition of complexes in this study basically follows that in IIT. However, there are some important differences between the two. First, the most major difference is in systems to be analyzed: the proposed method targets static graphs (e.g. connectomes) while IIT targets stochastic dynamical systems (e.g. brain dynamics). Since the targeted systems are different, the measures of coupling strength between subsystems are also different: this study uses strength of bidirectional connections, which is based on weight of edges in a graph, while IIT uses a measure called integrated information (also called Φ) (Tononi 2004, Balduzzi and Tononi 2008, Oizumi et al. 2014, Tononi et al. 2016). Integrated information Φ measures how much information is lost when interactions between subsystems are removed (Amari et al. 2018, Oizumi et al. 2016). Apart from these differences in the targeted systems and the measures of coupling strength, complexes in this study are defined in the same way as those in IIT 2.0, which is a previous version of IIT (Balduzzi and Tononi 2008). Note that the equivalent to the minimum cut is called the minimum information partition in IIT.

Although the measures of coupling strength in this study and IIT differ, as stated above, a mathematical idea to take account of bidirectionality is borrowed from IIT. In this study, to quantify the strength of bidirectional connections, we consider the minimum of the sum of weights in 1 direction and the sum of weights in the opposite direction (Eq. (6)). This definition, i.e. taking the minimum of the 2 directions, is from the operation in

computing Φ in the latest version of IIT (IIT 3.0, Oizumi et al. 2014, Tononi et al. 2016), where the amount of information is computed in 2 directions for time: forward (past to future) and backward (future to past), and the minimum value is taken.

Using graph weights to quantify coupling strengths has advantages and disadvantages over integrated information Φ . The advantages include that we can apply the HPC algorithm to find complexes rapidly, whereas we cannot use HPC for Φ because Φ does not have the mathematical property required by HPC to be valid (Kitazono et al. 2020). The disadvantages include that since a graph is fully described by one-to-one relations between nodes, graph-based analysis inevitably omits the effects of many-to-many interactions. In contrast, when Φ is used, such many-to-many interactions can be considered (Kitazono et al. 2020).

Hierarchical partitioning for complex search

If we search for complexes by brute force, computation time increases exponentially with the number of nodes. Therefore, when the number of nodes in the network exceeds several tens, it becomes practically impossible to identify the complexes. On the other hand, using the algorithm HPC, which we proposed in a previous study (Kitazono et al. 2020), the computation time increases only polynomially with the number of nodes. HPC is an exact method that does not use approximations and can extract all complexes without any omissions or misidentifications. This makes it possible to extract all complexes from a network consisting of several thousand nodes in a practical computation time. An actual computation time evaluated by a simulation is shown in Fig. 5.

In what follows in this subsection, we write the induced subgraph $G[S]$ for a node subset S as S for simplicity of notation.

HPC primarily consists of 2 steps. The first is listing candidates of (main) complexes. HPC narrows down candidates for (main) complexes by hierarchically partitioning a network. The second step is screening the candidates to find (main) complexes.

In the first step, HPC hierarchically partitions a network with min-cuts (Fig. 6). HPC starts by dividing the whole network with its min-cut, and then repeatedly divides the subnetworks with their min-cuts until the entire network is completely decomposed into single nodes. This procedure in HPC is summarized as follows:

- 1) Find the min-cut (V_L, V_R) of the whole network V and divide the whole network V into the 2 subnetworks V_L and V_R .
- 2) Find the min-cuts of the subnetworks found in the previous step, V_L and V_R , and divide them into (V_{LL}, V_{LR}) and (V_{RL}, V_{RR}), respectively.
- 3) Repeat this division until the whole network is decomposed into single nodes.

After the procedure above, we obtain the set of hierarchically partitioned subnetworks, that is, $V, V_L, V_R, V_{LL}, V_{LR}, V_{RL}, V_{RR}$, and so on. We consider all the set of subnetworks

$$\mathcal{V} = \{V, V_L, V_R, V_{LL}, V_{LR}, V_{RL}, V_{RR}, \dots\}, \quad (9)$$

excluding single nodes. Then, the following theorem holds.

Theorem 1.4. Any complex $S \subseteq V$ belongs to \mathcal{V} ($S \in \mathcal{V}$).

Thus from this theorem, \mathcal{V} can be seen as the set of candidates of complexes. The theorem is based on satisfaction of a mathematical property “monotonicity” by the strength of bidirectional connections (Eq. (6)). Let us consider the strength of bidirectional connections $w(S; T)$ between 2 subsets of nodes S and T . If we then add another set of nodes U to S , the strength of bidirectional connections does not decrease. That is,

$$w(S \cup U; T) \geq w(S; T). \quad (10)$$

Also, if we add U to T , $w(S; T \cup U) \geq w(S; T)$. This inequality means that the strength of bidirectional connections monotonically increases as nodes are added. We call this property “monotonicity.” By using monotonicity, we can easily show that a subnetwork cannot be a complex if it straddles the boundary of a min-cut of a subnetwork that contains it, and can prove Theorem 1.4 (see our previous work (Kitazono et al. 2020) for the proof).

After the hierarchical partitioning procedure described above, in the second step, we need to check whether each candidate of complexes belonging to \mathcal{V} is actually a (main) complex or not in accordance with Def. 1.1. We can efficiently check this by taking advantage of the hierarchical (tree) structure. For more detail please, see our previous work (Kitazono et al. 2020).

In general, a network can have multiple min-cuts. If this is the case, depending on which min-cut is used to divide a network in the hierarchical partitioning process, the candidate set of complexes \mathcal{V} can vary. However, even though \mathcal{V} varies, the resulting complexes (and also main complexes) do not vary. This is because any of the candidate sets contains all (main) complexes independent of which min-cut is used. Therefore we do not have to care which of multiple min-cuts we select.

Coreness of each node

Using the complexes and their w^{mc} , we define a “coreness” of each node. When a node is included in complexes with high w^{mc} , the coreness of that node is high, and conversely, when a node is included only in complexes with low w^{mc} , the coreness of that node is low. Specifically, we define the coreness of a node v as k_v if the node v is included in a complex with $w^{\text{mc}} = k_v$ but not included in any complex with $w^{\text{mc}} > k_v$. Equivalently, we can define the coreness of a node v as the largest of the w^{mc} of all

complexes containing the node v :

$$k_v = \max_{C \in \mathcal{G}_{\text{complex}} | v \in V(C)} w_C^{\text{mc}}, \quad (11)$$

where $\mathcal{G}_{\text{complex}}$ denotes the set of all complexes in the graph G and $V(C)$ denotes the set of all nodes in the complex C . The coreness Eq. (11) is equal to the largest of the w^{mc} of all subnetworks containing node v (see Text S6 for a proof):

$$k_v = \max_{S | v \in V(S)} w_{G[S]}^{\text{mc}}. \quad (12)$$

Coreness for s -cores

We can define a coreness for s -cores in the same way as for complexes: we define the coreness of a node v as s if node v is included in s -core but is not included in any s' -core with $s' > s$.

Degree of a node

We define the degree of a node v as the sum of the weights of all edges connecting v and other nodes, irrespective of the direction of edges:

$$\begin{aligned} \text{deg}(v) &= \frac{1}{2} \sum_{e \in E(v, V)} w_e, \\ E(v, V) &= \{(v, u) \in E | u \in V\}. \end{aligned} \quad (13)$$

The factor 2 in the denominator is for consistency with the strength of connections: when we measure the strength of connections (Eq. (3)) between a node and the nodes connecting to it, it becomes equal to the degree of the node. This degree can be also regarded as the mean of in-degree and out-degree.

Sorting rows and columns of a connection matrix according to the structures of complexes

In Figs. 7b and 7h and 8d and 8h, we sorted rows and columns of a connection matrix according to the hierarchical structures of the complexes. Here we explain this sorting process in detail.

To start, we sort the rows and columns in the order of the leaf nodes of the hierarchical structure obtained by hierarchical partitioning (Fig. 6). In the case of Fig. 6, the rows (columns) are sorted in the order of A, B, E, F, C, D, and G. We now explain in detail. At each step of the hierarchical partitioning process, we sort the nodes according to which of the 2 subnetworks (e.g. V_L or V_R) they are classified in. Therefore, at the end of the process, nodes classified into the same groups until a late stage of the process are placed close to each other, whereas those classified into different groups at an early stage are placed away from each other. Since the hierarchical structure obtained by the hierarchical partitioning is the

basis of the hierarchical structure formed by the complexes, the result is that nodes in the same complex with high w^{mc} are placed close to each other.

This is the rough flow of how the order of nodes is determined. This alone, however, is not enough to uniquely determine the order. There is still arbitrariness with regard to which of the 2 subnetworks (e.g. V_L or V_R) comes first at each step of the process. To eliminate this arbitrariness, we chose to place the upstream subnetwork first. That is, for example, when the strength of the connections from V_L to V_R (Eq. (4)) is higher than that in the opposite direction (Eq. (5)), i.e. when V_L is located relatively upstream to V_R , V_L is placed ahead of V_R . If the strengths of the connections in the 2 directions are equal to each other, we arranged the rows (columns) so that their original order is maintained.

Supplementary material

Supplementary material is available at *Cerebral Cortex Journal* online.

Funding

This work was partially supported by Japan Science and Technology Agency ACT-X (Grant Number JPMJAX20A6), Japan Science and Technology Agency CREST (Grant Numbers JPMJCR1864 and JPMJCR15E2), including AIP challenge program, Japan Science and Technology Agency Moonshot R&D (Grant Number JPMJMS2012), and Japan Society for the Promotion of Science KAKENHI (Grant Numbers 18H02713 and 20H05712), Japan.

Conflict of interest statement: The authors declare no competing financial interests or potential conflicts of interest.

Data and code availability

The MATLAB codes of HPC are available at <https://github.com/JunKitazono/BidirectionallyConnectedCores>.

References

- Afrasiabi M, Redinbaugh MJ, Phillips JM, Kambi NA, Mohanta S, Raz A, Haun AM, Saalman YB. Consciousness depends on integration between parietal cortex, striatum, and thalamus. *Cell Systems*. 2021;12(4):363–373.e11. <https://doi.org/10.1016/j.cels.2021.02.003>.
- Amari S, Tsuchiya N, Oizumi M. Geometry of information integration. In: Ay, N., Gibilisco, P., Matúš, F, (eds) *Information geometry and its applications*. IGAIA IV 2016. Springer Proceedings in Mathematics & Statistics. Vol. 252. Cham: Springer; 2018. https://doi.org/10.1007/978-3-319-97798-0_1.
- Aru J, Suzuki M, Larkum ME. Cellular mechanisms of conscious processing. *Trends Cogn Sci*. 2020;24(10):814–825. <https://doi.org/10.1016/j.tics.2020.07.006>.
- Aru J, Suzuki M, Rutiku R, Larkum ME, Bachmann T. Coupling the state and contents of consciousness. *Front Syst Neurosci*. 2019;13:43. <https://doi.org/10.3389/fnsys.2019.00043>.
- Atlan G, Terem A, Peretz-Rivlin N, Sehwat K, Gonzales BJ, Pozner G, Tasaka GI, Goll Y, Refaeli R, Zviran O. The claustrum supports resilience to distraction. *Curr Biol*. 2018;28(17):2752–2762.e7. <https://doi.org/10.1016/j.cub.2018.06.068>.
- Auksztulewicz R, Spitzer B, Blankenburg F. Recurrent neural processing and somatosensory awareness. *J Neurosci*. 2012;32(3):799–805. <https://doi.org/10.1523/JNEUROSCI.3974-11.2012>.
- Balduzzi D, Tononi G. Integrated information in discrete dynamical systems: motivation and theoretical framework. *PLoS Comput Biol*. 2008;4(6):1–18. <https://doi.org/10.1371/journal.pcbi.1000091>.
- Boly M, Garrido MI, Gosseries O, Bruno MA, Boveroux P, Schnakers C, Massimini M, Litvak V, Laureys S, Friston K. Preserved feedforward but impaired top-down processes in the vegetative state. *Science*. 2011;332(6031):858–862. <https://doi.org/10.1126/science.1202043>.
- Boly M, Massimini M, Tsuchiya N. Are the neural correlates of consciousness in the front or in the back of the cerebral cortex. *Clin Neuroimag Evid Neurosci*. 2017;37(40):9603–9613. <https://doi.org/10.1523/JNEUROSCI.3218-16.2017>.
- Boly M, Moran R, Murphy M, Boveroux P, Bruno MA, Noirhomme Q, Ledoux D, Bonhomme V, Bricchant JF, Tononi G et al. Connectivity changes underlying spectral EEG changes during propofol-induced loss of consciousness. *J Neurosci*. 2012;32(20):7082–7090. <https://doi.org/10.1523/JNEUROSCI.3769-11.2012>.
- Caulier LJ, Kulics AT. A comparison of awake and sleeping cortical states by analysis of the somatosensory-evoked response of postcentral area 1 in rhesus monkey. *Exp Brain Res*. 1988;72(3):584–592. <https://doi.org/10.1007/BF00250603>.
- Caulier LJ, Kulics AT. The neural basis of the behaviorally relevant N1 component of the somatosensory-evoked potential in SI cortex of awake monkeys: evidence that backward cortical projections signal conscious touch sensation. *Exp Brain Res*. 1991;84(3):607–619. <https://doi.org/10.1007/BF00230973>.
- Chatterjee N, Sinha S. 2007. Understanding the mind of a worm: hierarchical network structure underlying nervous system function in *c. elegans*. In: Banerjee R, Chakrabarti BK, editors. *Progress in brain research*. Vol. 168. Elsevier. p. 145–153. [https://doi.org/10.1016/S0079-6123\(07\)68012-1](https://doi.org/10.1016/S0079-6123(07)68012-1).
- Cohen D, van Swinderen B, Tsuchiya N. Isoflurane impairs low-frequency feedback but leaves high-frequency feedforward connectivity intact in the fly brain. *eNeuro*. 2018;5(1). <https://doi.org/10.1523/ENEURO.0329-17.2018>.
- Crick FC, Koch C. What is the function of the claustrum? *Philos Trans R Soc Lond B Biol Sci*. 2005;360(1458):1271–1279. <https://doi.org/10.1098/rstb.2005.1661>.
- de Haan EHF, Corballis PM, Hillyard SA, Marzi CA, Seth A, Lamme VAF, Volz L, Fabri M, Schechter E, Bayne T et al. Split-brain: what we know now and why this is important for understanding consciousness. *Neuropsychol Rev*. 2020;30(2):224–233. <https://doi.org/10.1007/s11065-020-09439-3>.
- Dehaene S, Changeux JP. Experimental and theoretical approaches to conscious processing. *Neuron*. 2011;70(2):200–227. <https://doi.org/10.1016/j.neuron.2011.03.018>.
- Dehaene S, Naccache L. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*. 2001;79(1–2):1–37. [https://doi.org/10.1016/S0010-0277\(00\)00123-2](https://doi.org/10.1016/S0010-0277(00)00123-2).
- Dembski C, Koch C, Pitts M. Perceptual awareness negativity: a physiological correlate of sensory consciousness. *Trends Cogn Sci*. 2021;25(8):660–670. <https://doi.org/10.1016/j.tics.2021.05.009>.
- Edelman GM. Naturalizing consciousness: a theoretical framework. *Proc Natl Acad Sci*. 2003;100:5520–5524. <https://doi.org/10.1073/pnas.0931349100>.

- Fulcher BD, Fornito A. A transcriptional signature of hub connectivity in the mouse connectome. *Proc Natl Acad Sci U S A*. 2016;113:1435–1440. <https://doi.org/10.1073/pnas.1513302113>.
- Goll Y, Atlan G, Citri A. Attention: the claustrum. *Trends Neurosci*. 2015;38(8):486–495. <https://doi.org/10.1016/j.tins.2015.05.006>.
- Gollo LL, Mirasso C, Sporns O, Breakspear M. Mechanisms of zero-lag synchronization in cortical motifs. *PLoS Comput Biol*. 2014;10(4):e1003548. <https://doi.org/10.1371/journal.pcbi.1003548>.
- Grandjean J, Zerbi V, Balsters JH, Wenderoth N, Rudin M. Structural basis of large-scale functional connectivity in the mouse. *J Neurosci*. 2017;37(34):8092–8101. <https://doi.org/10.1523/JNEUROSCI.0438-17.2017>.
- Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, Sporns O. Mapping the structural core of human cerebral cortex. *PLoS Biol*. 2008;6(7):e159. <https://doi.org/10.1371/journal.pbio.0060159>.
- Hao JX, Orlin JB. A faster algorithm for finding the minimum cut in a directed graph. *J Algorithm Comput Technol*. 1994;17(3):424–446. <https://doi.org/10.1006/jagm.1994.1043>.
- Harriger L, van den Heuvel MP, Sporns O. Rich club organization of macaque cerebral cortex and its role in network communication. *PLoS One*. 2012;7(9):e46497. <https://doi.org/10.1371/journal.pone.0046497>.
- Jackson J. Attention: noisy networks are tuned out by the claustrum. *Curr Biol*. 2018;28(17):R937–R939. <https://doi.org/10.1016/j.cub.2018.07.031>.
- Kale P, Zalesky A, Gollo LL. Estimating the impact of structural directionality: how reliable are undirected connectomes? *Netw Neurosci*. 2018;2(2):259–284. https://doi.org/10.1162/netn_a_00040.
- Kirkpatrick S, Gelatt CD, Vecchi MP. Optimization by simulated annealing. *Science*. 1983;220(4598):671–680. <https://doi.org/10.1126/science.220.4598.671>.
- Kitazono J, Kanai R, Oizumi M. Efficient search for informational cores in complex systems: application to brain networks. *Neural Netw*. 2020;132:232–244. <https://doi.org/10.1016/j.neunet.2020.08.020>.
- Koch C, Massimini M, Boly M, Tononi G. Neural correlates of consciousness: progress and problems. *Nat Rev Neurosci*. 2016;17(5):307–321. <https://doi.org/10.1038/nrn.2016.22>.
- Koivisto M, Kastrati G, Revonsuo A. Recurrent processing enhances visual awareness but is not necessary for fast categorization of natural scenes. *J Cogn Neurosci*. 2014;26(2):223–231. https://doi.org/10.1162/jocn_a_00486.
- Ku SW, Lee U, Noh GJ, Jun IG, Mashour GA. Preferential inhibition of frontal-to-parietal feedback connectivity is a neurophysiologic correlate of general anesthesia in surgical patients. *PLoS One*. 2011;6(10):e25155. <https://doi.org/10.1371/journal.pone.0025155>.
- Lamme VAF, Zipser K, Spekreijse H. Figure-ground activity in primary visual cortex is suppressed by anesthesia. *Proc Natl Acad Sci U S A*. 1998;95(6):3263–3268. <https://doi.org/10.1073/pnas.95.6.3263>.
- Lamme VAF. Towards a true neural stance on consciousness. *Trends Cogn Sci*. 2006;10(11):494–501. <https://doi.org/10.1016/j.tics.2006.09.001>.
- Lamme VAF. Challenges for theories of consciousness: seeing or knowing, the missing ingredient and how to deal with panpsychism. *Philos Trans R Soc Lond Ser B Biol Sci*. 2018;373(1755). <https://doi.org/10.1098/rstb.2017.0344>.
- Leicht EA, Newman MEJ. Community structure in directed networks. *Phys Rev Lett*. 2008;100(11):118703. <https://doi.org/10.1103/PhysRevLett.100.118703>.
- Lemon RN, Edgley SA. Life without a cerebellum. *Brain*. 2010;133(3):652–654. <https://doi.org/10.1093/brain/awq030>.
- Leopold DA. Primary visual cortex: awareness and blindsight. *Annu Rev Neurosci*. 2012;35:91–109. <https://doi.org/10.1146/annurev-neuro-062111-150356>.
- Llinás R. Consciousness and the thalamocortical loop. *Int Congr Ser*. 2003;1250:409–416. [https://doi.org/10.1016/S0531-5131\(03\)01067-7](https://doi.org/10.1016/S0531-5131(03)01067-7).
- Manita S, Suzuki T, Homma C, Matsumoto T, Odagawa M, Yamada K, Ota K, Matsubara C, Inutsuka A, Sato M et al. A top-down cortical circuit for accurate sensory perception. *Neuron*. 2015;86(5):1304–1316. <https://doi.org/10.1016/j.neuron.2015.05.006>.
- Mashour GA, Roelfsema P, Changeux JP, Dehaene S. Conscious processing and the global neuronal workspace hypothesis. *Neuron*. 2020;105(5):776–798. <https://doi.org/10.1016/j.neuron.2020.01.026>.
- Mathur B. The claustrum in review. *Front Syst Neurosci*. 2014;8:48. <https://doi.org/10.3389/fnsys.2014.00048>.
- Melloni L, Mudrik L, Pitts M, Koch C. Making the hard problem of consciousness easier. *Science*. 2021;372(6545):911–912. <https://doi.org/10.1126/science.abj3259>.
- Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U. Network motifs: simple building blocks of complex networks. *Science*. 2002;298(5594):824–827. <https://doi.org/10.1126/science.298.5594.824>.
- Mišić B, Betzel RF, de Reus MA, van den Heuvel MP, Berman MG, McIntosh AR, Sporns O. Network-level structure-function relationships in human neocortex. *Cereb Cortex*. 2016;26(7):3285–3296. <https://doi.org/10.1093/cercor/bhw089>.
- Mumford D. On the computational architecture of the neocortex. *Biol Cybern*. 1991;65(2):135–145. <https://doi.org/10.1007/BF00198477>.
- Narikiyo K, Mizuguchi R, Ajima A, Shiozaki M, Hamanaka H, Johansen JP, Mori K, Yoshihara Y. The claustrum coordinates cortical slow-wave activity. *Nat Neurosci*. 2020;23(6):741–753. <https://doi.org/10.1038/s41593-020-0625-7>.
- Nieder A, Wagener L, Rinnert P. A neural correlate of sensory consciousness in a corvid bird. *Science*. 2020;369(6511):1626–1629. <https://doi.org/10.1126/science.abb1447>.
- Norimoto H, Fenk LA, Li HH, Tosches MA, Gallego-Flores T, Hain D, Reiter S, Kobayashi R, Macias A, Arends A et al. A claustrum in reptiles and its role in slow-wave sleep. *Nature*. 2020;578(7795):413–418. <https://doi.org/10.1038/s41586-020-1993-6>.
- Odegaard B, Knight RT, Lau H. Should a few null findings falsify prefrontal theories of conscious perception? *J Neurosci*. 2017;37(40):9593–9602. <https://doi.org/10.1523/JNEUROSCI.3217-16.2017>.
- Oh SW, Harris JA, Ng L, Winslow B, Cain N, Mihalas S, Wang Q, Lau C, Kuan L, Henry AM. A mesoscale connectome of the mouse brain. *Nature*. 2014;508(7495):207–214. <https://doi.org/10.1038/nature13186>.
- Oizumi M, Albantakis L, Tononi G. From the phenomenology to the mechanisms of consciousness: integrated information theory 3.0. *PLoS Comput Biol*. 2014;10(5):e1003588. <https://doi.org/10.1371/journal.pcbi.1003588>.
- Oizumi M, Tsuchiya N, Amari S-I. Unified framework for information integration based on information geometry. *Proc Natl Acad Sci U S A*. 2016;113(51):14817–14822. <https://doi.org/10.1073/pnas.1603583113>.
- Onnela J-P, Saramäki J, Kertész J, Kaski K. Intensity and coherence of motifs in weighted complex networks. *Phys Rev E Stat Nonlinear Soft Matter Phys*. 2005;71(6):065103. <https://doi.org/10.1103/PhysRevE.71.065103>.

- O'Neill PK, Gore F, Salzman CD. Basolateral amygdala circuitry in positive and negative valence. *Curr Opin Neurobiol.* 2018;49:175–183. <https://doi.org/10.1016/j.conb.2018.02.012>.
- Pennartz CMA, Dora S, Muckli L, Lorteije JAM. Towards a unified view on pathways and functions of neural recurrent processing. *Trends Neurosci.* 2019;42(9):589–603. <https://doi.org/10.1016/j.tins.2019.07.005>.
- Pinto Y, Neville DA, Otten M, Corballis PM, Lamme VAF, de Haan EHF, Foschi N, Fabri M. Split brain: divided perception but undivided consciousness. *Brain.* 2017;140(5):1231–1237. <https://doi.org/10.1093/brain/aww358>.
- Rubinov M, Sporns O. Complex network measures of brain connectivity: uses and interpretations. *NeuroImage.* 2010;52(3):1059–1069. <https://doi.org/10.1016/j.neuroimage.2009.10.003>.
- Sachidhanandam S, Sreenivasan V, Kyriakatos A, Kremer Y, Petersen CCH. Membrane potential correlates of sensory perception in mouse barrel cortex. *Nat Neurosci.* 2013;16(11):1671–1677. <https://doi.org/10.1038/nn.3532>.
- Schlossmacher I, Dellert T, Bruchmann M, Straube T. Dissociating neural correlates of consciousness and task relevance during auditory processing. *NeuroImage.* 2021;228:117712. <https://doi.org/10.1016/j.neuroimage.2020.117712>.
- Self MW, Kooijmans RN, Supèr H, Lamme VA, Roelfsema PR. Different glutamate receptors convey feedforward and recurrent processing in macaque V1. *Proc Natl Acad Sci U S A.* 2012;109(27):11031–11036. <https://doi.org/10.1073/pnas.1119527109>.
- Sethi SS, Zerbi V, Wenderoth N, Fornito A, Fulcher BD. Structural connectome topology relates to regional BOLD signal dynamics in the mouse brain. *Chaos.* 2017;27(4):047405. <https://doi.org/10.1063/1.4979281>.
- Sikkens T, Bosman CA, Olcese U. The role of top-down modulation in shaping sensory processing across brain states: implications for consciousness. *Front Syst Neurosci.* 2019;13:31. <https://doi.org/10.3389/fnsys.2019.00031>.
- Slagter HA, Mazaheri A, Reteig LC, Smolders R, Figeë M, Mantione M, Schuurman PR, Denys D. Contributions of the ventral striatum to conscious perception: an intracranial EEG study of the attentional blink. *J Neurosci.* 2017;37(5):1081–1089. <https://doi.org/10.1523/JNEUROSCI.2282-16.2016>.
- Smith JB, Lee AK, Jackson J. The claustrum. *Curr Biol.* 2020;30(23):R1401–R1406. <https://doi.org/10.1016/j.cub.2020.09.069>.
- Smith JB, Watson GDR, Liang Z, Liu Y, Zhang N, Alloway KD. A role for the claustrum in salience processing? *Front Neuroanat.* 2019;13:64. <https://doi.org/10.3389/fnana.2019.00064>.
- Sporns O, Betzel RF. Modular brain networks. *Annu Rev Psychol.* 2016;67:613–640. <https://doi.org/10.1146/annurev-psych-122414-033634>.
- Sporns O, Kötter R. Motifs in brain networks. *PLoS Biol.* 2004;2(11):e369. <https://doi.org/10.1371/journal.pbio.0020369>.
- Supèr H, Spekreijse H, Lamme VA. Two distinct modes of sensory processing observed in monkey primary visual cortex (v1). *Nat Neurosci.* 2001;4(3):304–310. <https://doi.org/10.1038/85170>.
- Tang H, Buia C, Madhavan R, Crone NE, Madsen JR, Anderson WS, Kreiman G. Spatiotemporal dynamics underlying object completion in human ventral visual cortex. *Neuron.* 2014;83(3):736–748. <https://doi.org/10.1016/j.neuron.2014.06.017>.
- Tononi G. An information integration theory of consciousness. *BMC Neurosci.* 2004;5:42. <https://doi.org/10.1186/1471-2202-5-42>.
- Tononi G. Consciousness, information integration, and the brain. In: Laureys S, editors. *Progress in brain research.* Vol. 150. Elsevier; 2005. pp. 109–126. [https://doi.org/10.1016/S0079-6123\(05\)50009-8](https://doi.org/10.1016/S0079-6123(05)50009-8).
- Tononi G. Consciousness as integrated information: a provisional manifesto. *Biol Bull.* 2008;215(3):216–242. <https://doi.org/10.2307/25470707>.
- Tononi G, Boly M, Massimini M, Koch C. Integrated information theory: from consciousness to its physical substrate. *Nat Rev Neurosci.* 2016;17(7):450–461. <https://doi.org/10.1038/nrn.2016.44>.
- Tononi G, Edelman GM. Consciousness and complexity. *Science.* 1998;282(5395):1846–1851. <https://doi.org/10.1126/science.282.5395.1846>.
- Tononi G, Sporns O, Edelman GM. A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proc Natl Acad Sci U S A.* 1994;91(11):5033–5037. <https://doi.org/10.1073/pnas.91.11.5033>.
- van den Heuvel MP, Sporns O. Rich-club organization of the human connectome. *J Neurosci.* 2011;31(44):15775–15786. <https://doi.org/10.1523/JNEUROSCI.3539-11.2011>.
- van den Heuvel MP, Sporns O. Network hubs in the human brain. *Trends Cogn Sci.* 2013;17(12):683–696. <https://doi.org/10.1016/j.tics.2013.09.012>.
- Ward LM. The thalamic dynamic core theory of conscious experience. *Conscious Cogn.* 2011;20(2):464–486. <https://doi.org/10.1016/j.concog.2011.01.007>.
- Yu F, Jiang QJ, Sun XY, Zhang RW. A new case of complete primary cerebellar agenesis: clinical and imaging findings in a living patient. *Brain.* 2015;138(6):e353. <https://doi.org/10.1093/brain/awu239>.