

Criticality as a determinant of integrated information Φ in human brain networks

Hyoungkyu Kim^{1,2}, Anthony G. Hudetz^{1,2,3}, George A. Mashour^{1,2,3}, UnCheol Lee^{1,2*}

¹*Department of Anesthesiology, University of Michigan Medical School, Ann Arbor, MI, USA*

²*Center for Consciousness Science, University of Michigan Medical School, Domino's Farms, P.O. Box 385, Ann Arbor, MI, USA*

³*Neuroscience Graduate Program, University of Michigan, Ann Arbor, MI, USA*

* Correspondence:

UnCheol Lee
uclee@med.umich.edu

Keywords: criticality, integrated information, human consciousness, brain network

Abstract

Integrated information theory (IIT) postulates that consciousness arises from the cause-effect structure of a system but the optimal network conditions for this structure have not been elucidated. In the study, we test the hypothesis that network criticality, a dynamically balanced state between a large variation of functional network configurations and a large constraint of structural network configurations, is a necessary condition for the emergence of a cause-effect structure that results in a large Φ , a surrogate of integrated information. We also hypothesized that if the brain deviates from criticality, the cause-effect structure is obscured and Φ diminishes. We tested these hypotheses with a large-scale brain network model and high-density electroencephalography (EEG) acquired during various levels of human consciousness during general anesthesia. In the modeling study, maximal criticality coincided with maximal Φ . The constraint of the structural network on the functional network is maximized in the maximal criticality. The EEG study demonstrated an explicit relationship between Φ , criticality, and level

of consciousness. Functional brain network significantly correlated with structural brain network only in conscious states. The results support the hypothesis that network criticality maximizes Φ .

Introduction

Integrated information theory (IIT) proposes that consciousness equates with integrated information in a system and the integrated information is maximized when integration and differentiation of the systems' components are balanced. IIT proposes algorithmic methods to identify the differentiated parts of a system and quantify the integrated information across the parts [1–6]. Φ is a measure of complexity of the cause-effect structure of the minimum information partition among all possible partitions. However, in a dynamic system such as the brain, the optimal conditions under which the cause-effect structure—that is, the basis of integrated information—arises has not been elucidated. In this study, we hypothesized that network criticality, a balanced state between a large variation of functional network configurations and a large constraint of structural network configurations, may be the basis of the high Φ in conscious brains.

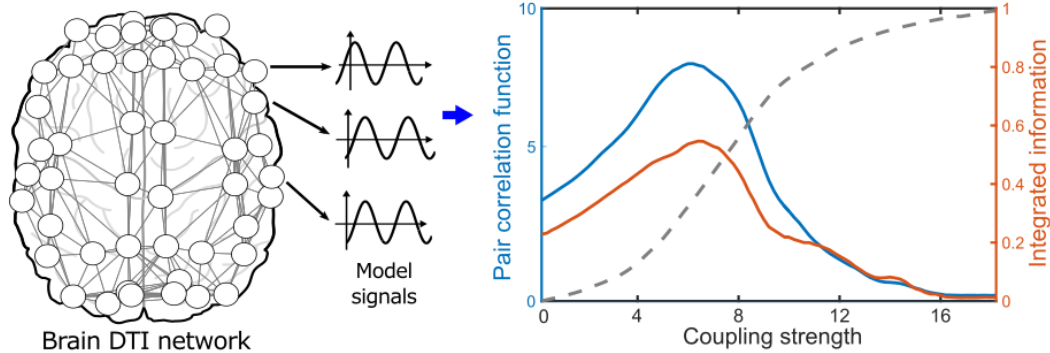
Criticality was originally introduced for studying phase transition in physics, which was simply defined as a balanced state between order and disorder in the activities of the elements that make up a system [7,8]. This property has been observed broadly in physical and non-physical systems and has been suggested as an optimal state for information storage, transmission, and integration with high susceptibility to external stimuli [9–13]. In particular, several computational modeling and empirical studies suggest that the brain dynamics associated with consciousness reside near a critical state [14–19]. Furthermore, recent studies have attempted not only to identify whether the conscious brain resides near the critical state but also to understand systematically how various types of brain perturbations (sleep, anesthesia, and traumatic injuries) lead to a deviation from criticality [9,20–23]. Such approaches introduce the criticality hypothesis as a theoretical framework to study pharmacological and pathological states of unconsciousness, such as anesthesia and coma. Regarding the characteristics of a critical state such as highly informative, highly susceptible, highly efficient, and highly integrative, criticality shares many commonalities with the brain state that IIT proposes as conducive to consciousness. However, the relationship between criticality, Φ , and human consciousness has not been demonstrated explicitly.

66 To identify a relationship between criticality and Φ , we analyzed computationally a large-scale
 67 human brain network model adjusting criticality with a control parameter. The criticality was
 68 defined using the pair correlation function (PCF), a surrogate measure of susceptibility [24]. The
 69 relative change of Φ across states was estimated with $\bar{\Phi}$, a surrogate measure of Φ that we
 70 developed for high-density EEG. Empirically, we modulated the level of human consciousness
 71 in a stepwise manner with an anesthetic and calculated both $\bar{\Phi}$ and PCF from continuous EEG
 72 data. From the modeling study and EEG analysis, we were able to quantitatively study the
 73 relationship between Φ , criticality, and consciousness, and suggested network criticality as a
 74 necessary condition for a high Φ in human brain networks.

A. Oscillator model in brain network

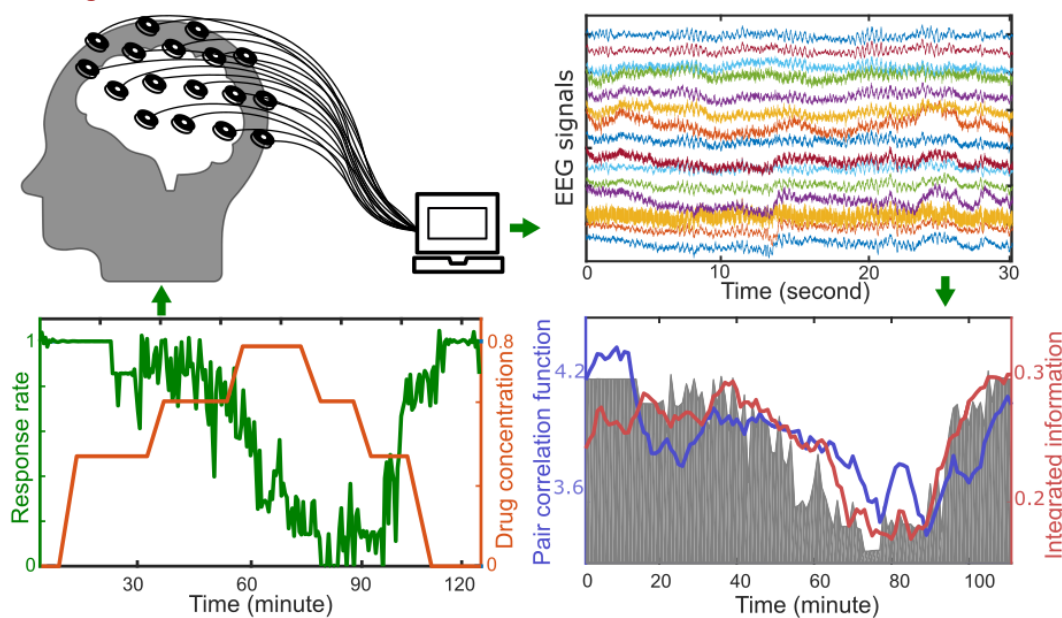
Equation of Kuramoto model

$$\dot{\theta}_j(t) = \omega_j + S \sum_{k=1}^N A_{jk} \sin\{\theta_k(t - \tau_{jk}) - \theta_j(t)\}, \quad j = 1, 2, \dots, N$$



B. Analysis in human EEG

EEG signals from sevoflurane-induced unconsciousness



75

76 Figure 1. Schematic flow diagram for studying the relationship between criticality (defined as pair
77 correlation function, PCF), integrated information (Φ) and human consciousness. (A) Model study: a
78 simple coupled oscillator model (Kuramoto model) was implemented on an anatomically informed human
79 brain network structure constructed from diffusion tensor imaging (DTI). The Φ was calculated while
80 modulating the level of criticality with a control parameter (coupling strength). The level of criticality was
81 defined using PCF, a susceptibility measure, of the simulated brain activity. (B) Empirical study: 64-
82 channel EEG data derived from seven healthy volunteers were recorded while gradually increasing
83 sevoflurane concentration from 0.4% to 0.6% to 0.8%, then decreasing it from 0.8% to 0.6% to 0.4%. The
84 changes of PCF and Φ were compared with the response rate to verbal commands.

Methods

Ethics Statement: This study was conducted at the University of Michigan Medical School and approved by the Institutional Review Board (HUM00061087); after careful discussion, written informed consent was obtained from all participants.

Human brain network model: Many recent studies have successfully applied Kuramoto/Stuart-Landau models to the brain in order to understand the organizational principles of multiscale brain function, surrogates of information flow, and complex dynamics at the whole brain network level [25–27]. Similarly, we hypothesized that the application of simple oscillatory models, which can modulate the criticality with a control parameter, to an anatomically informed brain network structure could inform the relationship between brain network criticality and integrated information.

We used a large-scale brain network model that implements a coupled simple oscillator model on the scaffold of an anatomically informed human brain network structure. The human brain network consists of 78 parcels of the cerebral cortex constructed from diffusion tensor imaging (DTI) of 80 young adults [28].

$$\dot{\theta}_j(t) = \omega_j + S \sum_{k=1}^N A_{jk} \sin \left(\theta_k(t - \tau_{jk}) - \theta_j(t) \right), \quad j = 1, 2, \dots, N. \quad (1)$$

Here S is the coupling strength between oscillators and A_{jk} denotes the anatomical connections between oscillator j and k , yielding 1 if a connection exists and 0 otherwise. τ_{jk} is the time delay between node j and k . $\theta_j(t)$ is the phase of oscillator j at time t . ω_j is the intrinsic frequency of oscillator j .

Simulation procedures: All parameters for the models were set to simulate alpha oscillations in the brain. Alpha oscillation models have successfully explained empirically observed brain network behaviors such as functional connectivity, traveling waves, and network state transitions based on electroencephalography (EEG) and magnetoencephalography (MEG) [25,26,29–37]. Thus, alpha oscillations were analyzed to understand the behaviors of Φ and the criticality at the brain network level. The natural frequencies of the oscillators in our simulation were given as a Gaussian distribution around 10 Hz with a standard deviation of 0.5 Hz. Time delay was set proportional to the physical distances between nodes with a propagation speed of 8.6m/s [38]. The coupling strength between the oscillators was continuously changed from 0 to

18. In each parameter set, 100 configurations were simulated and the results were averaged over all configurations.

Experimental procedures: We conducted a secondary analysis of high-density EEG data from a study of sevoflurane-induced unconsciousness in humans; the details of the experiment can be found in the supplementary text of this article and the previously published Blain-Moraes et al [39]. The current study tested different hypotheses and, unlike the original study, included computational analyses of brain network models.

Sevoflurane dataset: 64-channel EEG were recorded from seven healthy volunteers as sevoflurane concentrations in high-flow oxygen (8 L/min) were gradually increased from 0.4% to 0.6% to 0.8% (the average range at which unconsciousness was induced) or beyond, then decreased from 0.8% to 0.6% to 0.4%. The EEG was recorded with eyes closed. The loss and recovery of consciousness were defined as the loss and recovery of response to the verbal command ‘squeeze your left [or right] hand twice,’ on a recorded loop every 30 seconds, with right/left hand commands randomized.

Average reference was used for re-referencing and the windowed sinc-FIR filter (in the MATLAB toolbox from EEGLAB) was used to avoid a possible shifting of the signal. We analyzed 12-second-long EEG epochs with 1-minute-long moving windows.

Criticality: Criticality, a boundary state between order and disorder, has long been proposed to play an important role in neural dynamics and brain function. Empirical evidence supports the hypothesis that the brain operates at or near the critical point, not only at the neuronal network level [22,40,41], but also at the large-scale or global network level [9,12,42–44]. Until now, most studies have focused on scale-free behavior, showing power law distribution of empirically observed variables. It has also been recently proposed that high correlation between functional and structural brain networks [14,23,25,26,45] and a large pair correlation function (PCF) [24,46] is evidence of criticality. In both the brain network model and empirical EEG data, we estimated criticality with PCF, which is the variance of global phase synchronization and is equivalent to susceptibility in statistical physics

$$PCF = N \{ \langle Re^2[z(t)] \rangle_t - \langle Re[z(t)] \rangle_t^2 \} \quad (2)$$

where $Re[z(t)]$ is the real part of the $z(t)$ in Eq. (3).

$$z = r e^{i\psi} = \frac{1}{N} \sum_{j=1}^N e^{i\theta_j} \quad (3)$$

where Ψ is the order parameter phase. The absolute value $r = |z|$ represents the degree of synchronization. The r is equal to zero when the phases of nodes are uniformly distributed and one when all the nodes have the same phase.

Calculation of $\bar{\Phi}$: IIT defines integrated information (Φ) as the effective information (φ) of the minimum information partition (MIP) in a system [1–3,5,47]. The MIP is also defined as the partition having minimum effective information among all possible partitions.

$$\Phi [X;x] = : \varphi [X;x, \text{MIP}(x)] \quad (4)$$

$$\text{MIP} (x) = : \arg \min\{ \varphi (X;x , P) \} \quad (5)$$

where X is the system, x is a state, and P is a partition $P=\{M^1, \dots, M^r \}$.

Identifying the MIP requires searching all possible partitions and comparing their effective information φ . This is the most time-consuming and computationally demanding process in the application to high-density EEG. Furthermore, considering the fact that EEG data recorded during anesthetic state transitions are non-Gaussian and continuous time series, we used the $\tilde{\Phi}_{AR}$ as a measure of integrated information [48]. The computed $\tilde{\Phi}_{AR}$ values from original signals were compared with the $\tilde{\Phi}_{AR}$ values of surrogate data sets, and non-significant $\tilde{\Phi}_{AR}$ values were set as zero.

Although many improvements have been made in the algorithms of Φ over the last decade, the computation time is still unrealistic because of the need to search an enormous number of partitions to identify the MIP. In our previous study, we proposed a surrogate measure $\bar{\Phi}$ to circumvent the explosive computational time of Φ for high-density EEGs. $\bar{\Phi}$ estimates the relative change of Φ across states by considering the average feature of many small sample units rather than trying to identify the MIP and its effective information for all EEG channels. A sample unit consists of a small number of EEG channels randomly selected from 64 channels and the total number of sample units is taken as large enough to represent the behavior of the entire high-density EEG montage. In this study, we limited our interest only to the relative changes of Φ values across states, rather than attempting to calculate the exact Φ value for each brain state, which would be impossible to measure using the superficial and spatially imprecise brain activity recorded by EEGs.

For each sample unit, we were able to calculate the MIP and its effective information—that is, the Φ of the sample unit. For instance, in this study, we selected 8 random EEG channels for a sample unit and acquired 200 sample units that were randomly selected from the baseline

states. The same 200 sample units determined in the baseline were then compared across EEG windows to investigate the increase or decrease of Φ values. Since the number of possible bipartitions of 8 channels is $\sum_{k=1}^4 C(8, k) = 162$, where C stands for the combination of k unique elements chosen from eight possible elements, calculating Φ for all EEG windows of seven subjects during state transitions is possible within a relatively short computational time.

The average $\bar{\Phi}$ is defined as follows.

$$\bar{\Phi} = \frac{1}{k} \sum_{i=1}^k \Phi_i(n) - \frac{1}{k} \sum_{i=1}^k \text{median}(\Phi_{\text{surr}(i)}(n)) \quad (6)$$

where n is the number of EEG channels for a sample unit and k is the number of sample units of the n EEG channels. $\Phi_i(n)$ measures the effective information of MIP for the n EEG channels, by definition, the integrated information of the sample unit. Here, we chose $n=8$ and $k=200$ following our previous study [49]. $\Phi_{\text{surr}}(n)$ is the spurious $\Phi_i(n)$ estimated from randomly shuffled EEG data sets. Subtracting the randomness, $\bar{\Phi}$ reflects the average integrated information of 200 sample units taken from the high-density EEG data that exceeds the spurious information integrated from the surrogate data. Since $\bar{\Phi}$ estimates the relative change of Φ , it is appropriate for our purposes to detect the maximum Φ to compare with the maximum criticality.

Functional connectivity: Phase Lag Index (*PLI*), a measure of phase locking between two signals, was used to define the functional connectivity in the EEG network [50]. We chose a Hilbert transform to extract the instantaneous phase of the EEG from each channel and calculate the phase difference $\Delta\theta_{ij}(t)$ between channels i and j , where $\Delta\theta_{ij}(t) = \theta_i(t) - \theta_j(t)$, $t = 1, 2, \dots, n$, and n is the number of samples within one epoch. *PLI*_{*ij*} between two nodes i and j is then calculated using equation (7):

$$PLI_{ij} = \left| \langle \text{sign}(\Delta\theta_{ij}(t)) \rangle \right|, \quad 0 \leq PLI_{ij} \leq 1. \quad (7)$$

Here, the sign() function yields: 1 if $\Delta\theta_{ij}(t) > 0$; 0 if $\Delta\theta_{ij}(t) = 0$; and -1 if $\Delta\theta_{ij}(t) < 0$. The mean $\langle \rangle$ is taken over all $t=1, 2, \dots, n$. If the instantaneous phase of one signal is consistently ahead of the other signal, the phases are considered locked and $PLI_{ij} \approx 1$. However, if the signals randomly alternate between a phase lead and phase lag relationship, there is no phase locking and $PLI_{ij} \approx 0$.

Surrogate data: To control for spurious connectivity of EEG, 20 surrogate data sets were generated with a random shuffling method, in which a time point is randomly chosen in each EEG channel; the EEG epochs are then shuffled before and after the time point. The shuffled

data have the same amplitude distribution and power spectrum of the original EEG, but there are disruptions of the original connectivity between two EEG signals.

Network construction: We expected that different EEG frequency bands and different states would have different levels of spurious connectivity [51]. Thus, after subtracting the median *PLI* of 20 surrogate data sets, if the remaining *PLI* was larger than 0.1 then the connectivity of two EEG signals was set as 1; otherwise, it was set as 0. The threshold (0.1) was chosen to avoid isolated nodes in the EEG network in the baseline states. The node degree of an EEG channel was defined as the number of functional links in the network.

EEG simulation: To test if the *PLI* network of EEG in conscious states is similar to the structural brain network in a critical state, we simulated 78 source signals in the structural brain network in a critical state and projected it into the 64 sensor signals on the scalp. We could then directly compare the *PLI* networks of EEG and the *PLI* network of the simulated EEG (; the 64 sensor signals on the scalp). The signals generated from the Kuramoto model and structural brain network represent source activities of the brain, which are under the surface of where EEG is measured. In reality, the electrical potentials generated by the neural activity in the brain conduct outwards through brain tissue and the skull and finally appear at the scalp surface where the EEG signal is measured. In order to compare experimental EEG and the model signals, we generated surface level signals from the simulated source signals. We used three concentric spherical head models; the three layers consist of the brain, skull, and scalp. The conductivity of the three layers was set to be 0.33, 0.0042, and 0.33 S/m, respectively [52]. The source activity was represented as a dipole moment. The coordinate of the dipole moment in the brain was determined by the region's standard coordinates and the orientation of the dipole moment was randomly assigned. The forward model simulation was conducted by using the Field Trip Toolbox [53].

Statistical Analysis: We performed one-way ANOVA ("anova1.m", MATLAB toolbox) with Tukey-Kramer correction ("multcompare.m" with alpha = 0.05 and ctype = "tukey-kramer" in MATLAB) for the comparison of the PCF and Φ among different states. The statistical tests were carried out for the modeling and empirical analysis separately. The adjusted P-values of 0.05 or lower (*P < 0.05, **P < 0.01, and ***P < 0.001) were considered to be statistically significant (S1 Table and S2 Table).

Results

Model Study: Correlation between Criticality and $\bar{\Phi}$

We simulated various brain network dynamics by adjusting the control parameter (i.e., coupling strength). First, we determined the critical state of the model by identifying the coupling strength that yielded a maximum criticality. The criticality of the system was defined by PCF, which reflects the susceptibility of the brain network dynamics to perturbation. As the coupling strength of the network was increased, the PCF reached a maximum at an intermediate coupling strength, while the order parameter increased monotonically (dotted line in Figure 2A). In Figure 2A, we illustrate the intermediate coupling strength (blue region) as the critical state of this brain network model and selected an incoherent (green region) and highly synchronized (red region) states of lower PCFs as a supercritical and subcritical state, respectively. The $\bar{\Phi}$ value was also maximized at a point between sub- and supercritical states, at a similar coupling strength that corresponds to the maximum PCF.

Here, we assumed that the maximum $\bar{\Phi}$ may arise due to the balance between the large variation of functional network configurations and the large constraints of structural network configurations, as a characteristic of a critical state. The disrupted balance in an incoherent or highly synchronized brain network results in a small $\bar{\Phi}$. Figure 2B presents the functional brain networks based on the phase lag index (*PLI*). The functional connectivity at the maximum PCF (in the critical state) resembles the structural brain network of DTI, whereas the functional connectivity at the lower PCFs (in the sub- and supercritical states) are relatively homogeneous and not similar with the structural brain network.

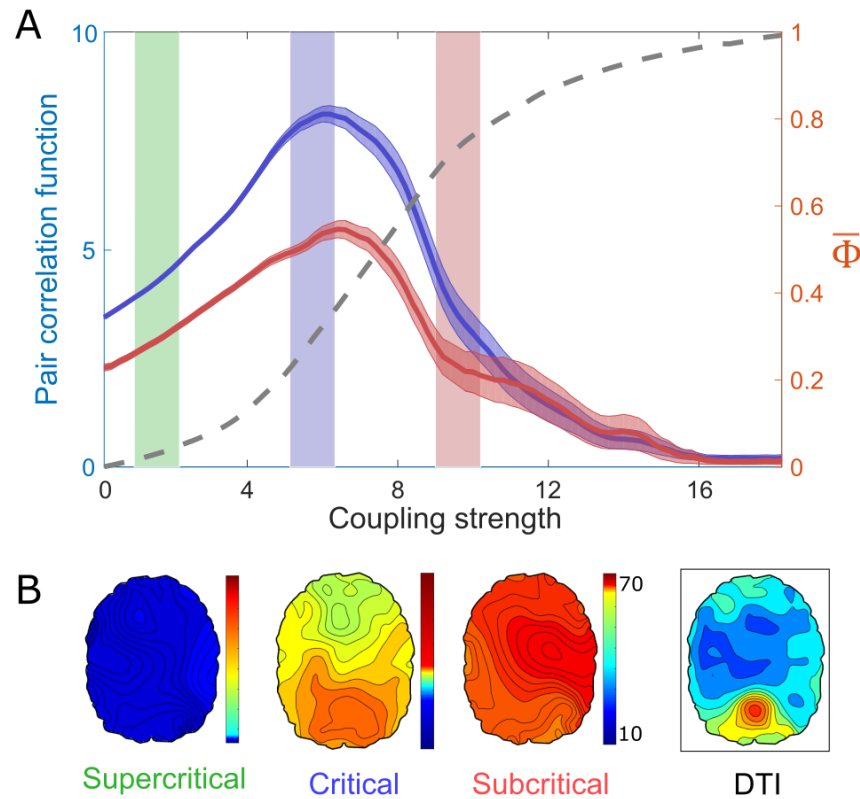


Figure 2. Criticality, $\bar{\Phi}$, and structured functional connectivity in the brain network. (A) When modulating the coupling strength (a control parameter), the maximum $\bar{\Phi}$ coincides with the maximum criticality, as measured by the pair correlation function (PCF), while the order parameter (dotted line) increases in a monotonic way. (B) Only in the critical state (blue region in Figure 2A), a salient structured functional connectivity, which resembles the structural brain network, emerges in the brain network. Color bar indicates the node degree.

A Network Mechanism of the Maximal $\bar{\Phi}$ in a Critical State

In a critical state, a network synchronization is balanced by incoherent and synchronous connections through partial phase locking. At a coarse-grained level, the distribution of incoherent and synchronous connections is shaped by the network topology; oscillations at the nodes with dense connections become slower and more synchronous, while oscillations at the nodes with sparse connections are faster and incoherent [25,26]. As a consequence, a coarse-grained functional network resembles a structural network in a critical state [14,54,55].

Figure 3A presents the Spearman correlation coefficients between the node degrees (in the structural brain network) and the *PLIs* of the 78 nodes (in the functional brain network) as coupling strength increases. The Spearman correlation coefficient is at a maximum in the critical state (blue region), while both the super- and subcritical states (green and red regions, respectively) are associated with smaller correlations. The results imply that the constraint of the structural network on a functional network is maximized in the critical state, with higher degree nodes having a larger *PLI*. Figure 3B presents the scatter plots of the *PLIs* versus the node degrees of the 78 nodes in the three states. A large positive correlation coefficient appears only in the critical state (Figure 3B blue, $r=0.57$, $p\text{-value} < 0.001$). However, the large correlation does not mean that the functional network is static. Figure 3C presents the temporal evolution of the Spearman correlation coefficients between the instantaneous phases of alpha oscillations in the functional brain network and the node degrees in the structural brain network for the three states. The varying correlation coefficients indicate the temporal change of the phase lead-lag relationships among 78 brain regions upon the structural brain network. The large variation of functional networks is one of the characteristics of a critical state and measured by a large PCF. As a result, the large variation of the functional brain network at a small-time scale (\sim seconds) (Figure 3C) under the large constraint from the structural brain network (Figure 3B) in a large time scale (\sim minutes) in the critical state may be the network condition for the maximal $\bar{\Phi}$ in the brain network.

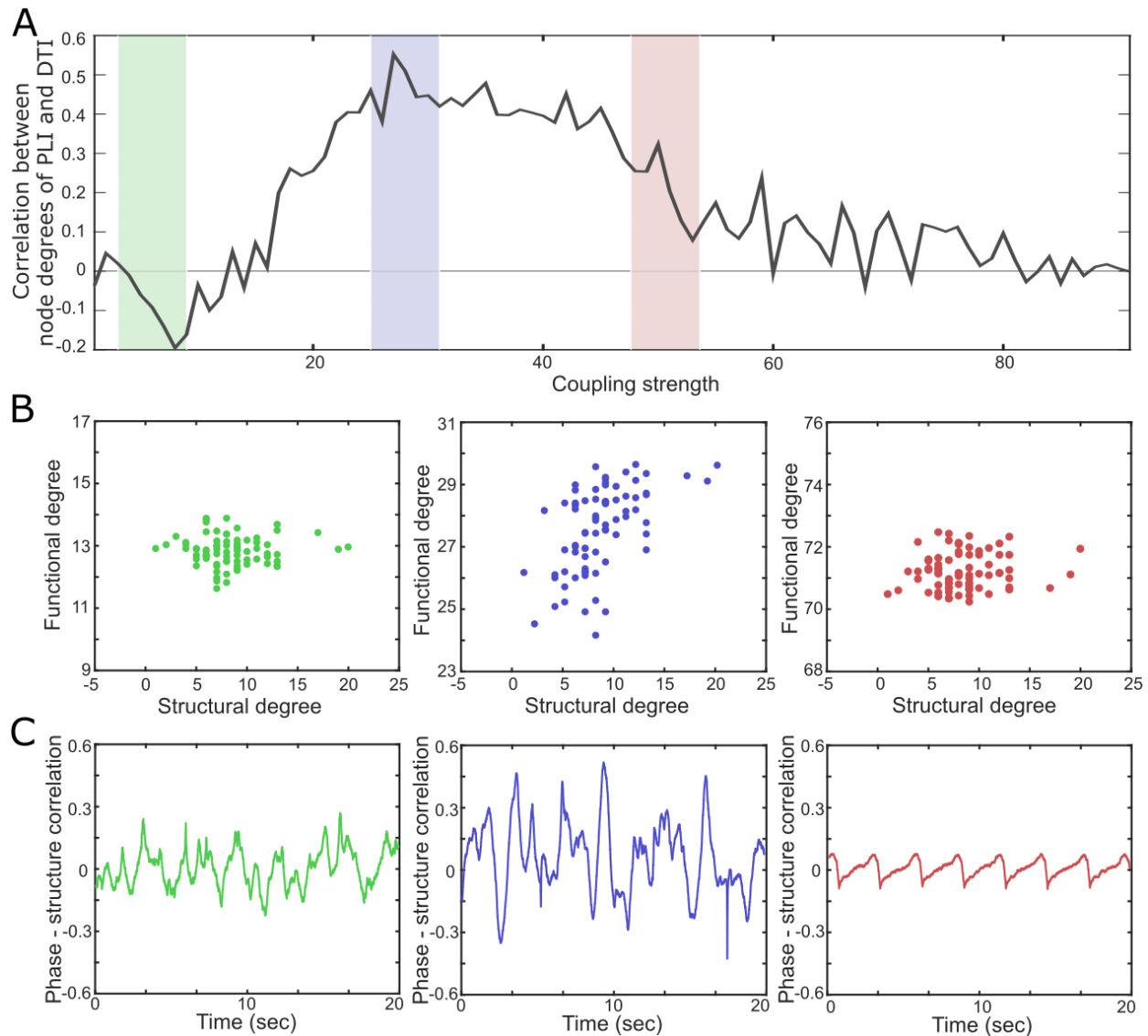


Figure 3. Correlations between the functional and structural brain networks near to and far from a critical state. (A) The Spearman correlation coefficient between the 78 phase lag index (*PLI*) values in the functional brain network and the 78 node degrees in the structural brain network is maximal in the critical state (black line, blue shaded region). (B) The scatter plots (the 78 *PLI* values versus the 78 node degrees) for the supercritical (green), critical (blue), and subcritical (red) states. The large correlation in the critical state implies a large constraint of the structural network on the functional network. (C) The Spearman correlation coefficients between the instantaneous phases of alpha oscillations and the node degrees of the 78 nodes in the structural brain network. The large temporal variation indicates a large repertoire of functional connectivity, which is a characteristic of a critical state.

EEG study: Correlation between Criticality, $\bar{\Phi}$, and Human Consciousness

To investigate the relationship between criticality, integrated information, and level of human consciousness, we compared the PCF and $\bar{\Phi}$ of high-density EEG and the response rate during general anesthesia. The behavioral response rate, which is inversely proportional to the drug concentration, was used as a surrogate for the level of consciousness. In figure 4A, the conscious resting state and the conscious recovery state have a higher PCF and $\bar{\Phi}$ than in the unconscious states. For the continuous EEG data, the change of PCF correlates with the change of $\bar{\Phi}$. Moreover, both the measures reflect the response rate during the significant state transition. These results are consistent with the model prediction and empirically demonstrate for the first time a direct relationship between PCF, $\bar{\Phi}$, and the level of human consciousness.

The model study also predicted a large correlation of the structural brain network with the functional brain network in consciousness, which corresponds to the functional brain network in a critical state. However, since we recorded only the scalp EEG, we were not able to test the correlation between functional (*PLI*) and structural (node degree) brain networks. Instead, we compared the *PLI* networks of the EEG and a simulated EEG. For the simulation of EEG in the conscious state, we first simulated the source signals in the structural brain network in a critical state (the same simulation in Figure 2 and 3) and then projected the source signals onto the scalp (see the Method in details). If the model prediction is correct, both the *PLI* networks should be largely correlated.

Indeed, we found high Spearman correlation coefficients between the *PLI* network of the EEG in the conscious states and the *PLI* network of the simulated EEG in a critical state. This correlation decreases along with the reduced response rate in the unconscious states (Figure 4B). The high correlation coefficients in conscious states indicate a strong constraint of the structural brain network on the EEG.

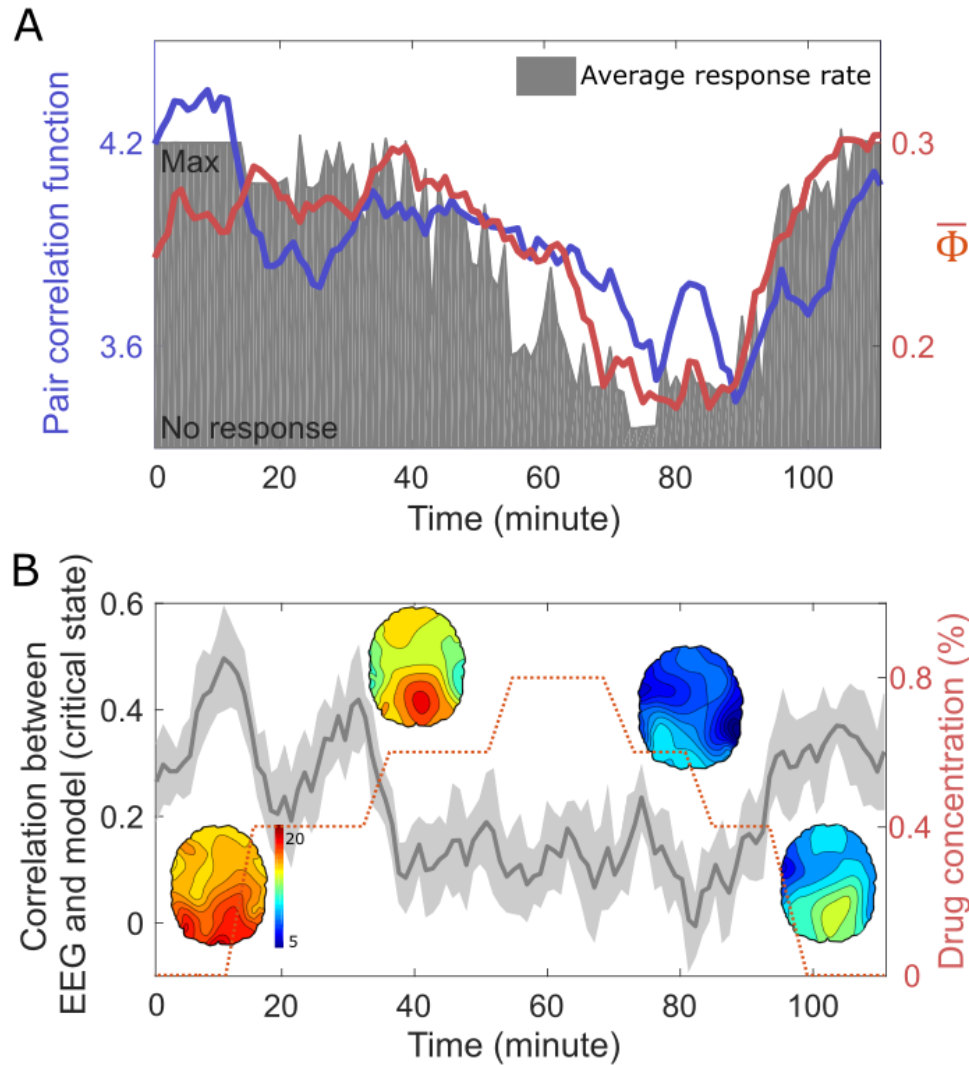


Figure 4. Criticality, integrated information, and level of human consciousness during general anesthesia. (A) The PCF and $\bar{\Phi}$ of 64-channel EEG correlate with the response rate (grey area), which was modulated with increasing anesthetic concentrations. (B) The Spearman correlation coefficients between the *PLI* networks of the EEG and the simulated EEG based on the anatomical brain network and critical state. The conscious states (baseline, induction, and recovery) show larger correlations, which implies a stronger constraint of the structural brain network on the EEG in conscious states. As a result, the balance between the large constraint of the structural brain network and the large repertoire (i.e., large PCF) of the functional brain network may be the network basis of the large $\bar{\Phi}$ of EEG in the conscious brains.

Discussion

Summary of the findings: In this paper, we performed a computational modeling and EEG data analysis to study the relationship between integrated information, criticality, and level of human consciousness. In the modeling study, we simulated the brain network activities near and far from a critical state and quantified the criticality and integrated information with PCF and $\bar{\Phi}$. We found that when the brain network reaches a maximal PCF, the $\bar{\Phi}$ is also maximized under the largest constraint of the structural brain network (that is, with the largest correlation between the functional and structural brain networks). To verify the modeling results, we gradually modulated the level of consciousness in humans with a general anesthetic and compared the PCF and $\bar{\Phi}$ of high-density EEG with the behavioral response rate. We found that, in the conscious resting states, the subjects have the highest PCF and $\bar{\Phi}$ s when the behavioral response rates are highest. We also showed that the functional brain network of EEGs in conscious states largely correlates with the functional brain network of the simulated EEGs, which were modeled to reflect the underlying structural brain network in a critical state. From both the modeling study and EEG analysis, we propose that the balance between the large variation of functional brain network (as measured as the large PCF) and the large constraint from the structural brain network in a critical state is a necessary condition to generate the high $\bar{\Phi}$ of conscious states.

Criticality and maximal $\bar{\Phi}$: IIT's composition principle states that any subset of elements within a system can be a mechanism of the system if the intrinsic cause-effect power is irreducible [1,3]. Furthermore, the set of all mechanisms and their constraints within the system comprises a system's cause-effect structure. Algorithmically, the intrinsic cause-effect structure can be identified by the effective information among partitions (φ), with the integrated information of a system (Φ) defined by the information transmission of the minimal information bipartitions. However, in the application to our brain network model, since we did not give any cause-effect structure initially, a non-zero $\bar{\Phi}$ may emerge through only the interaction of 78 nodes in the brain network. This raises the following questions. How does a non-zero $\bar{\Phi}$ emerge spontaneously in the brain network? What is the network basis that facilitates the emergence of a cause-effect structure? In this study, we propose that the phase lead-lag relationship that is shaped by the structural brain network may create conditions that facilitate the emergence of a cause-effect structure.

Previous studies have examined how brain network topology modulates the frequencies and phases of local node dynamics, subsequently shaping a pattern of global information flow and

functional connectivity. A mathematical relationship between the node degree and the phase of node dynamics was identified analytically [25,26], which enables us to estimate analytically the phase of a node (in a functional network) with only its node degree and local connectivity (in a structural network). This was tested with diverse brain networks (human, monkey, and mouse), which demonstrated that when considering long-term and spatially coarse-grained brain activities (> minutes) the global phase lead-lag relationship and information flow pattern (measured by transfer entropy and Granger causality) were predictable based only on the structural brain network topologies of the three species. In particular, the difference between the "hub node" and "peripheral node" becomes most salient at a critical state. In other words, the structural complexity of overall coupled node dynamics is maximized at a critical state. The differences between frequency and phase among coupled node dynamics naturally give rise to information flow in the brain network, consequently dividing the nodes in the brain network into 'senders' and 'receivers' of information flow. However, in sub- and supercritical states, there is no information flow between node dynamics because of a highly-synchronized state (i.e., no difference that would allow for information flow) and incoherent state (i.e., no interaction that would allow for information flow). Therefore, a cause-effect structure may emerge between these two extreme states and be maximized at a balanced state. Extrapolating from results based on the phase lead-lag relationship, emerging a cause-effect structure without external stimuli is likely not random but rather is organized by a mathematical relationship between network structure and dynamics in a critical state [25,26,56,57].

A large variation under a large constraint: In our modeling study and EEG analysis, we found that a large variation of the functional brain network occurs in spite of a strong constraint from the structural brain network. Notably, the large variation and the strong constraint were observed at different time scales. When we investigated the correlation coefficient between the instantaneous phases and the node degrees of the 78 nodes, the correlation coefficient at each time point varied widely at a short time scale (~ seconds) (Figure 3C). By contrast, the averaged instantaneous phases and the node degrees at a longer time scale (~ minutes) have a large positive correlation coefficient ($r=0.57$, Figure 3B), which implied a bias toward positive values in the short time scale variation. Our previous EEG study demonstrated that the correlation coefficient between the EEG network and the structural brain network was pronounced when calculated with large temporal windows (> 60 seconds) but diminished when calculated with small windows (< 5 seconds). Interestingly, the scale-dependency appeared only in the baseline

conscious states, not in the altered states of consciousness such as anesthetized state, minimally conscious state, and unresponsive wakefulness syndrome [23]. The conscious brain can be characterized by diverse repertoires of global functional connectivity on a short time scale. However, when combining all repertoires of small windows in a large window, the different functional connectivity patterns across small windows may be averaged out and only the common functional connectivity patterns that are constrained by the structural brain network remain. Such an averaged functional connectivity in a large window can be simulated by the mean-field method upon a structural brain network. Nevertheless, how the large variance and the large constraint contribute to the large integrated information in conscious states remains elusive. Further studies are required to understand the mechanistic association between the functional variance, structural constraint, and Φ in a network.

Deviations from criticality and $\bar{\Phi}$: When brain dynamics deviate from a critical state, the capacity to shape global node dynamics into a structure that resembles the network structure is lost. In a supercritical state, there is no functional interaction between nodes. In a subcritical state, node dynamics cannot be structured due to strong functional interactions between nodes that eliminate functional heterogeneity. As a result, the $\bar{\Phi}$ of a brain not in a critical state—regardless of whether the deviation is toward subcriticality or supercriticality—will decrease in proportion to the distance from the critical state.

Many empirical studies support the criticality hypothesis in consciousness by comparing the dynamics of various conscious states with unconscious states (such as sleep, anesthesia, seizure, and coma). They have commonly shown that unconsciousness is associated with a deviation from criticality, which is quantified with various methods (power law, susceptibility, pair correlation function, and correlation between the functional and structural brain networks) [15,22,23,40,58–68]. However, no one has yet associated criticality, human consciousness, and Φ because of the explosive computational demands required for Φ calculations. In a previous study, we introduced a method that statistically estimates the increase/decrease of Φ for a continuous EEG, termed $\bar{\Phi}$, and showed the applicability to characterizing levels of consciousness with high-density EEG and also enabled us to study a relationship with network criticality. Our current results show that the empirical and computational model studies support the hypothesis that the characteristic network properties in a critical state naturally give rise to a structured, asymmetric phase lead-lag relationship among oscillators and, in turn, maximize Φ in the conscious brain.

442

443 **Limitations:** This study has several limitations. First, $\bar{\Phi}$ does not measure the precise Φ .
 444 Instead, we estimated the relative change of Φ , focusing on the increase/decrease along with
 445 the change in criticality in the network model and based on empirical EEG during anesthetic-
 446 induced unconsciousness. Even so, this relative measure is sufficient for the purpose of this
 447 study, which is to examine the relationship between Φ , criticality, and levels of consciousness.
 448 And considering the significant differences among the six versions of Φ in the recent
 449 comparison [69], testing our conclusion with other versions of Φ is required. Second, in this
 450 study, we derived a relationship between asymmetric phase lead-lag relationships and $\bar{\Phi}$, but
 451 this relationship does not explain how the phase lead-lag relationship generates a cause-effect
 452 relationship defined by information partition for Φ . It may require an analytic study to identify
 453 conditions to link phase lead-lag and Φ . Third, to find an association between Φ , criticality, and
 454 consciousness, we used the subjects' response rate during exposure to a general anesthetic as
 455 a surrogate for the level of consciousness. However, it is well known that unresponsiveness
 456 does not necessarily correlate with lack of consciousness. Although previously studied in the
 457 context of vegetative and minimally conscious states, our research team has recently identified
 458 covert consciousness during propofol sedation in which overt motoric responses were absent
 459 but brain network responses suggestive of volition were present [70]. Future studies might
 460 incorporate EEG data, behavioral responsiveness, and neuroimaging protocols to determine
 461 covert consciousness in order to more precisely identify the relationship between critical
 462 dynamics, consciousness, and Φ . Finally, the structural brain network used in the modeling
 463 study includes only the cortex. Including subcortical networks such as thalamocortical and
 464 hippocampocortical connections, etc. could improve the modeling performance for complex
 465 state transitions during general anesthesia.

466

467 **Conclusion:** We demonstrated for the first time an explicit relationship between criticality,
 468 integrated information, and human consciousness with computational modeling and EEG
 469 analysis. We propose that network criticality, that is, a balanced state between large variation of
 470 functional network configurations and strong constraint of structural network configurations, is a
 471 network condition for integrated information. Understanding this relationship may open a new
 472 way to study diverse states of consciousness situated near to and far from a critical state in
 473 terms of integrated information. It may also provide a theoretical foundation for controlling the
 474 level of consciousness and integrated information by modulating criticality at a network level.

475

476 **Acknowledgments:** This work is supported by grant No. R01 GM098578 (PIs: GM and UL)
477 from the National Institutes Health.

478

479 **References**

- 480 1. Oizumi M, Albantakis L, Tononi G. From the phenomenology to the mechanisms of
481 consciousness: Integrated Information Theory 3.0. PLoS Comput Biol. 2014;10:
482 e1003588. doi:10.1371/journal.pcbi.1003588
- 483 2. Tononi G. An information integration theory of consciousness. BMC Neurosci. 2004;5: 42.
484 doi:10.1186/1471-2202-5-42
- 485 3. Tononi G, Boly M, Massimini M, Koch C. Integrated information theory: from
486 consciousness to its physical substrate. Nat Rev Neurosci. 2016;17: 450–461.
487 doi:10.1038/nrn.2016.44
- 488 4. Tononi G, Koch C. The neural correlates of consciousness: an update. Ann N Y Acad
489 Sci. 2008;1124: 239–261. doi:10.1196/annals.1440.004
- 490 5. Tononi G. Integrated information theory of consciousness: an updated account. Arch Ital
491 Biol. 2012;150: 56–90. doi:10.4449/aib.v149i5.1388
- 492 6. Tononi G. Information integration: its relevance to brain function and consciousness. Arch
493 Ital Biol. 2010;148: 299–322. doi:10.4449/AIB.V148I3.1224
- 494 7. Bak P, Tang C, Wiesenfeld K. Self-organized criticality. Phys Rev A. 1988;38: 364–374.
495 doi:10.1103/PhysRevA.38.364
- 496 8. Bak P, Paczuski M. Complexity, contingency, and criticality. Proc Natl Acad Sci U S A.
497 1995;92: 6689–96. doi:10.1073/pnas.92.15.6689
- 498 9. Kitzbichler MG, Smith ML, Christensen SR, Bullmore E, Bak P, Paczuski M, et al.
499 Broadband criticality of human brain network synchronization. PLoS Comput Biol. 2009;5:
500 e1000314. doi:10.1371/journal.pcbi.1000314
- 501 10. Werner G. Metastability, criticality and phase transitions in brain and its models.
502 Biosystems. 2007;90: 496–508. doi:10.1016/J.BIOSYSTEMS.2006.12.001

11. Hesse J, Gross T. Self-organized criticality as a fundamental property of neural systems. *Front Syst Neurosci.* 2014;8: 166. doi:10.3389/fnsys.2014.00166
12. Tagliazucchi E, Balenzuela P, Fraiman D, Chialvo DR. Criticality in large-scale brain FMRI dynamics unveiled by a novel point process analysis. *Front Physiol.* 2012;3: 15. doi:10.3389/fphys.2012.00015
13. Hidalgo J, Grilli J, Suweis S, Muñoz MA, Banavar JR, Maritan A. Information-based fitness and the emergence of criticality in living systems. *Proc Natl Acad Sci.* 2014;111: 10095–10100. doi:10.1073/PNAS.1319166111
14. Haimovici A, Tagliazucchi E, Balenzuela P, Chialvo DR. Brain organization into resting state networks emerges at criticality on a model of the human connectome. *Phys Rev Lett.* 2013;110: 178101. doi:10.1103/PhysRevLett.110.178101
15. Marinazzo D, Pellicoro M, Wu G, Angelini L, Cortés JM, Stramaglia S. Information transfer and criticality in the Ising model on the human connectome. *PLoS One.* 2014;9: e93616. doi:10.1371/journal.pone.0093616
16. Moon J-Y, Lee U, Blain-Moraes S, Mashour GA. General relationship of global topology, local dynamics, and directionality in large-scale brain networks. *PLoS Comput Biol.* 2015;11: e1004225. doi:10.1371/journal.pcbi.1004225
17. Cocchi L, Gollo LL, Zalesky A, Breakspear M. Criticality in the brain: a synthesis of neurobiology, models and cognition. *Prog Neurobiol.* 2017;158: 132–152. doi:10.1016/J.PNEUROBIO.2017.07.002
18. Beggs JM. The criticality hypothesis: how local cortical networks might optimize information processing. *Philos Trans R Soc A Math Phys Eng Sci.* 2008;366: 329–343. doi:10.1098/rsta.2007.2092
19. Khajehabdollahi S, Abeyasinghe PM, Owen AM, Soddu A. The emergence of integrated information, complexity, and consciousness at criticality. *bioRxiv.* 2019; 521567. doi:10.1101/521567
20. Kim M, Mashour GA, Moraes S-B, Vanini G, Tarnal V, Janke E, et al. Functional and topological conditions for explosive synchronization develop in human brain networks with the onset of anesthetic-induced unconsciousness. *Front Comput Neurosci.* 2016;10: 1. doi:10.3389/fncom.2016.00001

21. Kim M, Kim S, Mashour GA, Lee U. Relationship of topology, multiscale phase synchronization, and state transitions in human brain networks. *Front Comput Neurosci.* 2017;11: 55. doi:10.3389/fncom.2017.00055
22. Beggs JM, Timme N. Being critical of criticality in the brain. *Front Physiol.* 2012;3: 163. doi:10.3389/fphys.2012.00163
23. Lee H, Golkowski D, Jordan D, Berger S, Ilg R, Lee J, et al. Relationship of critical dynamics, functional connectivity, and states of consciousness in large-scale human brain networks. *Neuroimage.* 2019;188: 228–238. doi:10.1016/J.NEUROIMAGE.2018.12.011
24. Yoon S, Sorbaro Sindaci M, Goltsev A V, Mendes JFF. Critical behavior of the relaxation rate, the susceptibility, and a pair correlation function in the Kuramoto model on scale-free networks. *Phys Rev E.* 2015;91: 32814. doi:10.1103/PhysRevE.91.032814
25. Moon J-Y, Lee U, Blain-Moraes S, Mashour GA. General relationship of global topology, local dynamics, and directionality in large-scale brain networks. *PLoS Comput Biol.* 2015;11: e1004225. doi:10.1371/journal.pcbi.1004225
26. Moon J-Y, Kim J, Ko T-W, Kim M, Iturria-Medina Y, Choi J-H, et al. Structure shapes dynamics and directionality in diverse brain networks: mathematical principles and empirical confirmation in three species. *Sci Rep.* 2017;7: 46606. doi:10.1038/srep46606
27. Breakspear M. Dynamic models of large-scale brain activity. *Nat Neurosci.* 2017;20: 340–352. doi:10.1038/nn.4497
28. Gong G, He Y, Concha L, Lebel C, Gross DW, Evans AC, et al. Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography. *Cereb Cortex.* 2009;19: 524–536. doi:10.1093/cercor/bhn102
29. Naruse Y, Matani A, Miyawaki Y, Okada M. Influence of coherence between multiple cortical columns on alpha rhythm: a computational modeling study. *Hum Brain Mapp.* 2009;31: 703–715. doi:10.1002/hbm.20899
30. Breakspear M, Heitmann S, Daffertshofer A. Generative models of cortical oscillations: neurobiological implications of the kuramoto model. *Front Hum Neurosci.* 2010;4: 190. doi:10.3389/fnhum.2010.00190
31. Nakagawa TT, Woolrich M, Luckhoo H, Joensson M, Mohseni H, Kringelbach ML, et al.

- How delays matter in an oscillatory whole-brain spiking-neuron network model for MEG alpha-rhythms at rest. *Neuroimage*. 2014;87: 383–394.
doi:10.1016/J.NEUROIMAGE.2013.11.009
32. Flores FJ, Hartnack KE, Fath AB, Kim S-E, Wilson MA, Brown EN, et al. Thalamocortical synchronization during induction and emergence from propofol-induced unconsciousness. *Proc Natl Acad Sci U S A*. 2017;114: E6660–E6668.
doi:10.1073/pnas.1700148114
33. Deco G, Cabral J, Saenger VM, Boly M, Tagliazucchi E, Laufs H, et al. Perturbation of whole-brain dynamics in silico reveals mechanistic differences between brain states. *Neuroimage*. 2018;169: 46–56. doi:10.1016/j.neuroimage.2017.12.009
34. Deco G, Cabral J, Woolrich MW, Stevner ABA, van Hartevelt TJ, Kringelbach ML. Single or multiple frequency generators in on-going brain activity: A mechanistic whole-brain model of empirical MEG data. *Neuroimage*. 2017;152: 538–550.
doi:10.1016/J.NEUROIMAGE.2017.03.023
35. Jobst BM, Hindriks R, Laufs H, Tagliazucchi E, Hahn G, Ponce-Alvarez A, et al. Increased stability and breakdown of brain effective connectivity during slow-wave sleep: mechanistic insights from whole-brain computational modelling. *Sci Rep*. 2017;7: 4634.
doi:10.1038/s41598-017-04522-x
36. Kim M, Kim S, Mashour GA, Lee U. Relationship of topology, multiscale phase synchronization, and state transitions in human brain networks. *Front Comput Neurosci*. 2017;11: 55. doi:10.3389/fncom.2017.00055
37. Zhang H, Watrous AJ, Patel A, Jacobs J. Theta and alpha oscillations are traveling waves in the human neocortex. *Neuron*. 2018;98: 1269–1281.e4.
doi:10.1016/j.neuron.2018.05.019
38. Caminiti R, Ghaziri H, Galuske R, Hof PR, Innocenti GM. Evolution amplified processing with temporally dispersed slow neuronal connectivity in primates. *Proc Natl Acad Sci U S A*. 2009;106: 19551–19556. doi:10.1073/pnas.0907655106
39. Blain-Moraes S, Tarnal V, Vanini G, Alexander A, Rosen D, Shortal B, et al. Neurophysiological correlates of sevoflurane-induced unconsciousness. *Anesthesiology*. 2015;122: 307–316. doi:10.1097/ALN.0000000000000482

- 593 40. Beggs JM, Plenz D. Neuronal avalanches in neocortical circuits. *J Neurosci.* 2003;23:
594 11167–77. doi:10.1523/JNEUROSCI.23-35-11167.2003
- 595 41. Brochini L, de Andrade Costa A, Abadi M, Roque AC, Stolfi J, Kinouchi O. Phase
596 transitions and self-organized criticality in networks of stochastic spiking neurons. *Sci*
597 *Rep.* 2016;6: 35831. doi:10.1038/srep35831
- 598 42. Kim H, Moon J-Y, Mashour GA, Lee U. Mechanisms of hysteresis in human brain
599 networks during transitions of consciousness and unconsciousness: Theoretical
600 principles and empirical evidence. *PLOS Comput Biol.* 2018;14: e1006424.
601 doi:10.1371/journal.pcbi.1006424
- 602 43. Lv P, Guo L, Hu X, Li X, Jin C, Han J, et al. Modeling dynamic functional information
603 flows on large-scale brain networks. *Med Image Comput Comput Assist Interv.* 2013;16:
604 698–705.
- 605 44. Schmidt R, LaFleur KJR, de Reus MA, van den Berg LH, van den Heuvel MP. Kuramoto
606 model simulation of neural hubs and dynamic synchrony in the human cerebral
607 connectome. *BMC Neurosci.* 2015;16. doi:10.1186/s12868-015-0193-z
- 608 45. Deco G, Kringelbach ML, Jirsa VK, Ritter P. The dynamics of resting fluctuations in the
609 brain: metastability and its dynamical cortical core. *Sci Rep.* 2017;7: 3095.
610 doi:10.1038/s41598-017-03073-5
- 611 46. Stanley HE, Wong VK. Introduction to phase transitions and critical phenomena. *Am J*
612 *Phys.* 1972;40: 927–928. doi:10.1119/1.1986710
- 613 47. Oizumi M, Tsuchiya N, Amari S. Unified framework for information integration based on
614 information geometry. *Proc Natl Acad Sci U S A.* 2016;113: 14817–14822.
615 doi:10.1073/pnas.1603583113
- 616 48. Barrett AB, Seth AK. Practical Measures of Integrated Information for Time-Series Data.
617 *PLoS Comput Biol.* Public Library of Science; 2011;7: e1001052.
618 doi:10.1371/journal.pcbi.1001052
- 619 49. Kim H, Hudetz AG, Lee J, Mashour GA, Lee U, Group the RS, et al. Estimating the
620 integrated information measure Phi from high-density electroencephalography during
621 states of consciousness in humans. *Front Hum Neurosci.* 2018;12: 42.
622 doi:10.3389/FNHUM.2018.00042

50. Stam CJ, Nolte G, Daffertshofer A. Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. *Hum Brain Mapp.* 2007;28: 1178–1193. doi:10.1002/hbm.20346
51. Lee U, Lee H, Müller M, Noh G-J, Mashour GA. Genuine and spurious phase synchronization strengths during consciousness and general anesthesia. *PLoS One.* 2012;7: e46313. doi:10.1371/journal.pone.0046313
52. Wen P, Li Y. EEG human head modelling based on heterogeneous tissue conductivity. *Australas Phys Eng Sci Med.* 2006;29: 235–240. doi:10.1007/BF03178571
53. Oostenveld R, Fries P, Maris E, Schoffelen J-M. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci.* 2011;2011: 156869. doi:10.1155/2011/156869
54. Deco G, Jirsa VK. Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. *J Neurosci.* 2012;32: 3366–75. doi:10.1523/JNEUROSCI.2523-11.2012
55. Tagliazucchi E, Chialvo DR, Siniatchkin M, Amico E, Brichant J-F, Bonhomme V, et al. Large-scale signatures of unconsciousness are consistent with a departure from critical dynamics. *J R Soc Interface.* 2016;13: 20151027. doi:10.1098/rsif.2015.1027
56. Hu X, Boccaletti S, Huang W, Zhang X, Liu Z, Guan S, et al. Exact solution for first-order synchronization transition in a generalized Kuramoto model. *Sci Rep.* 2014;4: 7262. doi:10.1038/srep07262
57. Ko T-W, Ermentrout GB. Partially locked states in coupled oscillators due to inhomogeneous coupling. *Phys Rev E Stat Nonlin Soft Matter Phys.* 2008;78: 16203. doi:10.1103/PhysRevE.78.016203
58. Bédard C, Destexhe A. Macroscopic models of local field potentials and the apparent 1/f noise in brain activity. *Biophys J.* 2009;96: 2589–2603. doi:10.1016/J.BPJ.2008.12.3951
59. Hudetz AG, Humphries CJ, Binder JR. Spin-glass model predicts metastable brain states that diminish in anesthesia. *Front Syst Neurosci.* 2014;8: 234. doi:10.3389/fnsys.2014.00234
60. Liang Z, King J, Zhang N. Intrinsic organization of the anesthetized brain. *J Neurosci.* 2012;32: 10183–10191. doi:10.1523/JNEUROSCI.1020-12.2012

61. Liu X, Ward BD, Binder JR, Li S-J, Hudetz AG. Scale-free functional connectivity of the brain is maintained in anesthetized healthy participants but not in patients with unresponsive wakefulness syndrome. PLoS One. 2014;9: e92182. doi:10.1371/journal.pone.0092182
62. Lee U, Mashour GA. Role of network science in the study of anesthetic state transitions. Anesthesiology. 2018; 1. doi:10.1097/ALN.0000000000002228
63. Moretti P, Muñoz MA. Griffiths phases and the stretching of criticality in brain networks. Nat Commun. 2013;4: 2521. doi:10.1038/ncomms3521
64. Priesemann V, Valderrama M, Wibral M, Le Van Quyen M. Neuronal avalanches differ from wakefulness to deep sleep – evidence from intracranial depth recordings in humans. PLoS Comput Biol. 2013;9: e1002985. doi:10.1371/journal.pcbi.1002985
65. Lee U, Oh G, Kim S, Noh G, Choi B, Mashour GA. Brain networks maintain a scale-free organization across consciousness, anesthesia, and recovery: evidence for adaptive reconfiguration. Anesthesiology. 2010;113: 1081–1091. doi:10.1097/ALN.0b013e3181f229b5
66. Krzemiński D, Kamiński M, Marchewka A, Bola M. Breakdown of long-range temporal correlations in brain oscillations during general anesthesia. Neuroimage. 2017;159: 146–158. doi:10.1016/J.NEUROIMAGE.2017.07.047
67. Meisel C, Storch A, Hallmeyer-Elgner S, Bullmore E, Gross T. Failure of adaptive self-organized criticality during epileptic seizure attacks. PLoS Comput Biol. 2012;8: e1002312. doi:10.1371/journal.pcbi.1002312
68. Bonachela JA, de Franciscis S, Torres JJ, Muñoz MA. Self-organization without conservation: are neuronal avalanches generically critical? J Stat Mech Theory Exp. 2010;2010: P02015. doi:10.1088/1742-5468/2010/02/P02015
69. Mediano P, Seth A, Barrett A, Mediano PAM, Seth AK, Barrett AB. Measuring integrated information: comparison of candidate measures in theory and simulation. Entropy. 2018;21: 17. doi:10.3390/e21010017
70. Huang Z, Vlisides PE, Tarnal VC, Janke EL, Keefe KM, Collins MM, et al. Brain imaging reveals covert consciousness during behavioral unresponsiveness induced by propofol. Sci Rep. 2018;8: 13195. doi:10.1038/s41598-018-31436-z