

Non-equilibrium brain dynamics as a signature of consciousness

Yonatan Sanz Perl,^{1,2} Hernán Bocaccio,² Ignacio Pérez-Ipiña,² Steven Laureys,³
Helmut Laufs,⁴ Morten Kringelbach,⁵ Gustavo Deco,⁶ and Enzo Tagliazucchi²

¹ Universidad de San Andrés, Buenos Aires, Argentina

² Physics Department, University of Buenos Aires, and Buenos Aires Physics Institute, Argentina

³ Coma Science Group, GIGA Consciousness, University of Liège, Liège, Belgium

⁴ Department of Neurology, Christian-Albrechts-University Kiel, Germany

⁵ Department of Psychiatry, University of Oxford, United Kingdom

⁶ Center for Brain and Cognition, Computational Neuroscience Group, Universitat Pompeu Fabra, Spain

(Dated: December 22, 2020)

The cognitive functions of human and non-human primates rely on the dynamic interplay of distributed neural assemblies. As such, it seems unlikely that cognition can be supported by macroscopic brain dynamics at the proximity of thermodynamic equilibrium. We confirmed this hypothesis by investigating electrocorticography data from non-human primates undergoing different states of unconsciousness (sleep, and anesthesia with propofol, ketamine, and ketamine plus medetomidine), and functional magnetic resonance imaging data from humans, both during deep sleep and under propofol anesthesia. Systematically, all states of reduced consciousness unfolded at higher proximity to equilibrium dynamics than conscious wakefulness, as demonstrated by entropy production and the curl of probability flux in phase space. Our results establish non-equilibrium macroscopic brain dynamics as a robust signature of consciousness, opening the way for the characterization of cognition and awareness using tools from statistical mechanics.

Introduction - One of the defining features of living matter is the scale-dependent violation of thermodynamic equilibrium [1]. While inanimate matter frequently displays non-equilibrium dynamics, these result from externally applied fields. In contrast, departures from equilibrium in living matter can also arise due to endogenous causes, such as complex chains of enzymatic reactions[2] driving mesoscopic mechanical processes[3]. Another characteristic of non-equilibrium dynamics in living matter is their dependence with the spatial and temporal grain; for instance, cells and larger biological structures might appear to be at equilibrium, while being sustained by non-equilibrium processes at the molecular scale [4].

Systems at thermodynamic equilibrium obey detailed balance, meaning that there are no net probability fluxes in the configuration space of the system. Detailed balance leads toward reversible dynamics with null entropy production rate. The vector field representing the probability flux in configuration space becomes purely rotational for steady non-equilibrium systems, the deviation from equilibrium can be quantified using the curl of this field [5], and entropy production can be estimated from the probability of transitioning between states [6]. This notion of equilibrium exists only in reference to a certain configuration space, which might reflect, in turn, a particular choice of spatio-temporal grain.

Models of non-equilibrium brain dynamics have been extensively studied in the literature (including the development of a field theory of stochastic neural activity [7]) in the contexts of associative learning (i.e. Hopfield networks) and decision making, among others [8]. The phase space of excitable systems includes the important particular case of scale-free critical behaviour [9]. In general, it is agreed that neural dynamics (as many other bio-

logical phenomena) are an intrinsically non-equilibrium phenomenon; however, this is less clear for macroscopic brain dynamics [10]. A recently proposed framework to quantify entropy production from neural data acquired at a macroscopic scale (functional magnetic resonance imaging [fMRI] recordings) showed that dynamics within a reduced 2D configuration space does not generally obey detailed balance, and that the extent of its violation (and thus the degree of entropy production) is task-dependent [11].

In spite of these advances, the relationship between conscious states and non-equilibrium dynamics remains to be elucidated. It seems unlikely that the dynamic nature of consciousness and cognition can be sustained by a system without strong deviations from thermodynamic equilibrium at some spatial and temporal scales. Thus, brain states associated with unconsciousness could obey detailed balance at the large scale, while simultaneously being supported by non-equilibrium homeostatic processes at the cellular scale. We investigated electrocorticography (ECoG) recordings acquired during different states of consciousness. Compared to fMRI, this modality is better suited for the estimation of transition probabilities in configuration space, given that higher temporal resolution results in better sampling of transition probabilities. We also investigated fMRI recordings by fitting semi-empirical whole-brain models to data from states of reduced consciousness, using them to produce synthetic time series which arbitrarily high temporal resolution.

Consciousness and non-equilibrium dynamics in monkey ECoG data- We first considered ECoG data non-human primates (128 channels from electrodes located as shown in Fig. 1A) during wakefulness, deep sleep (2 animals, 21 sessions with 8 of wakefulness), and under

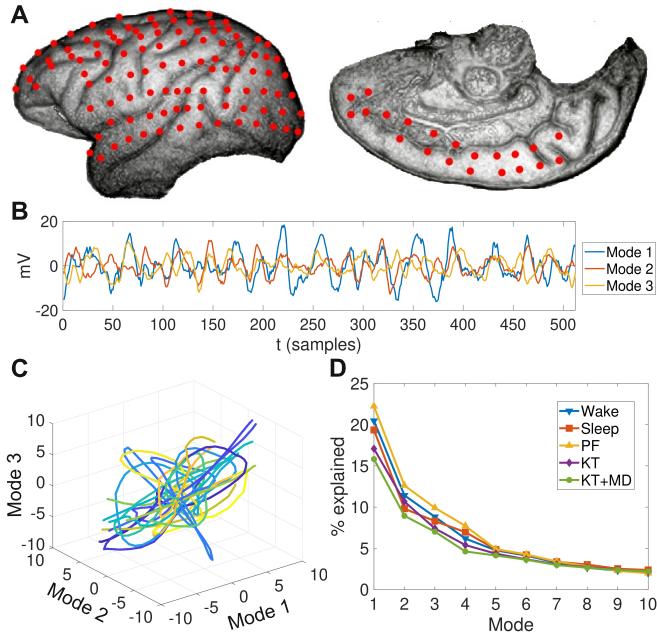


FIG. 1. Principal components of non-human primate ECoG signals. A) Sample ECoG array electrode locations. B) Temporal evolution of the three main modes of ECoG activity for an awake monkey obtained using PCA. C) Time series of panel B projected on the three-dimensional configuration space (the colour gradient represents the temporal evolution of along the trajectory). D) Variance explained as a function of retained principal components, plotted for all states of consciousness.

the effects of the following anaesthetic drugs: **propofol** (PF; 2 animals, 4 sessions of wake, 4 of anaesthesia, 4 of recovery), **ketamine** (KT; 2 animals, 4 sessions of wake, 4 of anaesthesia, 4 of recovery), and **ketamine plus medetomidine** (KF+MD; 4 animals, 11 sessions of wake, 11 sessions of anaesthesia, and 10 sessions of recovery). All sessions were recorded with eyes closed. Time series were notch filtered to remove the power line frequency and its first harmonics (50 Hz, 100 Hz, 150 Hz) and then bandpass filtered between 5 and 500 Hz. Filtered signals were downsampled from 1 kHz to 256 Hz using linear interpolation prior to application of PCA. For more information on this data set visit <http://neurotycho.org/>.

Sample results obtained from linear dimensional reduction are shown in Fig. 1B as the temporal evolution of the first three modes obtained applying PCA to ECoG of an awake animal. Fig. 1C shows the projection of these time series on the space defined by the three principal modes, with temporal evolution represented by a colour gradient from blue scales (earlier times) to yellow scales (later times). The percentage of the variance explained as a function of retained principal components is shown in Fig 1D separately for all states of consciousness.

We then assessed the violation of detailed balance by dividing the two-dimensional space spanned by the two main modes of ECoG activity into 10×10 grids from -2 to

2 standard deviations (s.d.) [3, 11]. Figure 2A presents the average probability density and flux in configuration space for wake, sleep and their subtraction.

We then estimated the entropy production as a measure of equilibrium/non-equilibrium behaviour. We associated with each brain state a probability transition P_{ij} representing the probability of transitioning from state x_i at time t to x_j at time $t + 1$, where the state of the system is defined by the grid in the configuration space defined by the two main modes. The arrows in Fig. 2A are scaled to the transition probabilities P_{ij} between adjacent states; rotational fluxes are typical of non-equilibrium dynamics and result in net entropy production. We calculated the entropy production as: $S = \sum_{ij} P_{ij} \log(\frac{P_{ij}}{P_{ji}})$ [6, 11]. Figure 2B (up) shows that entropy production is minimal during states of unconsciousness (sleep, KT, PF, KT+MD) compared to wakefulness before and after anaesthesia, i.e. recovery. Figure 2B (bottom) presents the curl of the probability flux, significantly larger for all states of unconsciousness vs. controls with the exception of PF.

Changes in grid size- We repeated the procedure for grids with different numbers of cells, from $N = 6^2$ to $N = 20^2$ cells, and computed entropy production for each size. Figure 3 shows the results obtained for sleep, KT, PF and KT+MD, as well as for the corresponding control conditions. Entropy production presented a minimum value when phase space for $N = 10 - 12$ cells by and then increased monotonically. Significant differences were observed for most grid sizes.

Second order transition probabilities- The calculation of entropy production by means of probability flux in configuration space assume Markovian dynamics [6]. This represents an important limitation in our analyses, since higher order temporal dependencies have been suggested by prior reports. To assess the potential impact of this limitation, we included a second order term in the computation of the transition probabilities, i.e. $P_{i,j,k}$ [11]. The results preserved all significant differences shown in Fig. 2.

Entropy production vs. spectral content- We compared the assessment of non-equilibrium dynamics by means of entropy production with the spectral power in different bands (1-4 Hz, 4-12 Hz, 12-30 Hz, 30-60 Hz), an empirical metric that is frequently used to assess (un)consciousness in electrophysiological signals [12]. Figure 4 presents scatter plots of spectral content in different bands (indicated by symbol type and colour) vs. entropy production for all conditions. We did not observe significant correlations that could be interpreted as redundant information between spectral content and entropy production.

fMRI data enhancement using whole-brain semi-empirical models- To investigate non-equilibrium dynamics as a signature of consciousness in human fMRI recordings, we first developed a data enhancement technique allowing us to improve the sampling of configuration space.

For this purpose, we implemented computational models linking together multiple sources of empirical data by

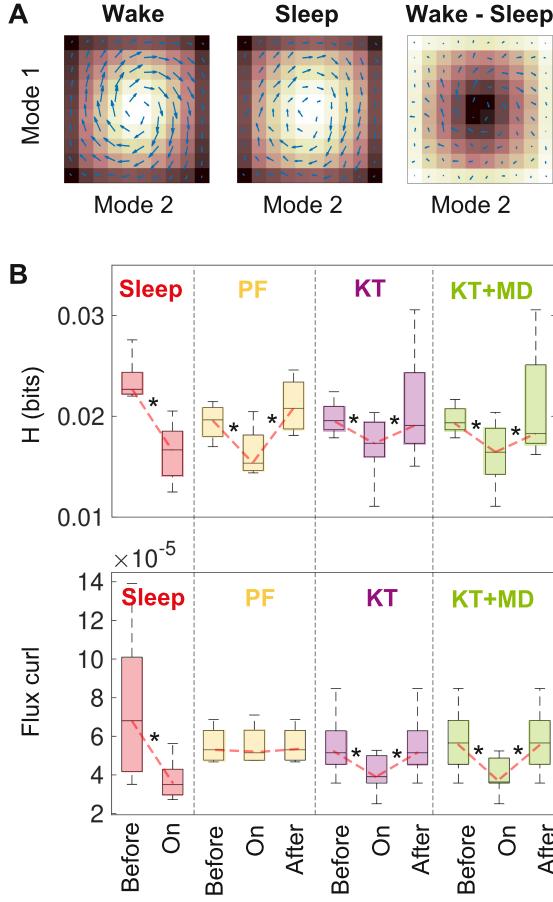


FIG. 2. Non-equilibrium entropy production is a signature of conscious states. A) 2D configuration space defined by the two main ECoG modes for sleep, wake and their subtraction. B) Up: entropy production per state and condition. Bottom: curl of the probability flux per state and condition.* $p<0.05$, Wilcoxon's test, Bonferroni corrected for multiple comparisons.

means of coupled local dynamics with different qualitative behaviours [13]. Our model incorporates empirical estimates of the large-scale structural connectivity of the brain obtained using diffusion tensor imaging (DTI), a non-invasive methodology to estimate axon bundles from water diffusion anisotropies. We computed observable based on the large-scale functional connectivity of the brain, i.e. the linear correlation between mean all pairs of fMRI time series extracted from 90 pre-defined regions of interest spanning the whole cortical and sub-cortical grey matter [14]. Finally, we constructed models with heterogeneous local parameters whose variation is constrained by anatomical priors that represent boundaries between brain regions associated with specific functional systems (resting state networks [RSN]) [15].

We outline this model in Fig. 5A. Local dynamics present dynamical criticality when a region undergoes a bifurcation from stable fixed point dynamics ($a<0$) to oscillatory behaviour ($a>0$) according to a Hopf bifurca-

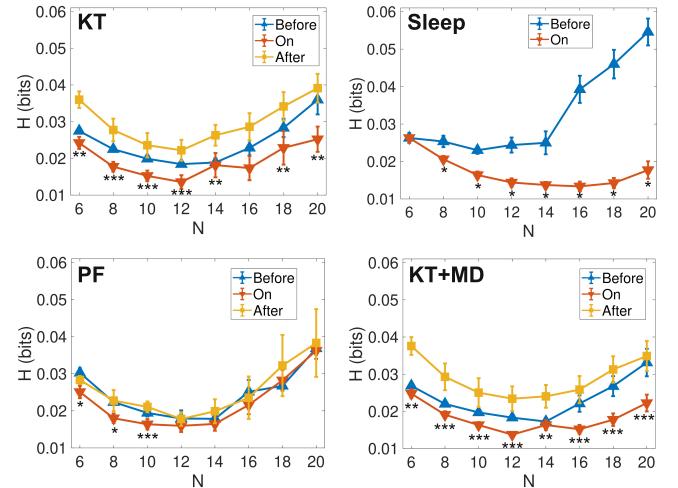


FIG. 3. Entropy production for different grid sizes. Each panel presents entropy production as a function of the square root of the number of cells for a given condition (sleep, PF, KT, KT+MD). * $p<0.05$ before vs. on, ** $p<0.05$ after vs. on, *** $p<0.05$ after and before vs. on, Wilcoxon's test, Bonferroni corrected for multiple comparisons.

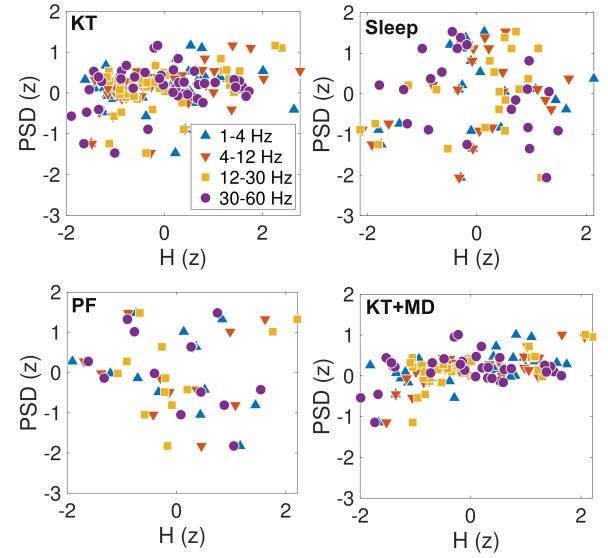


FIG. 4. Comparison between entropy production and spectral content for frequency bands 1-4 Hz, 4-12 Hz, 12-30 Hz, 30-60 Hz, indicated with different symbol shapes and colours. No significant correlation between these variables were found.

tion; these dynamics are coupled by the structural connectivity network inferred from DTI data, and scaled by a global coupling parameter (G). As shown in [15] this scaling factor G and the local bifurcation parameters corresponding to each anatomical prior (i.e. a value for each RSN) can be optimized to reproduce the functional connectivity measured during different states of consciousness.

We applied this optimization procedure to reproduce

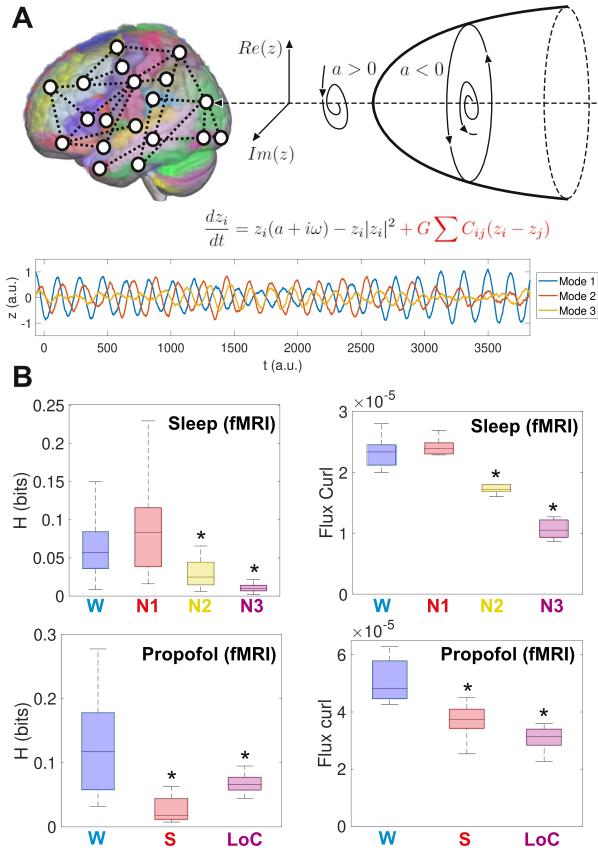


FIG. 5. Model-based fMRI data enhancement used to assess the non-equilibrium dynamics in the human brain. A) Up: a whole-brain semi-empirical model presenting dynamical criticality (local Hopf bifurcation) fitted to the empirical functional connectivity of different sleep stages, using structural connectivity (DTI) to couple the local dynamics, and different RSN as priors to constraint the independent variation of local bifurcation parameters. Bottom: the temporal evolution of the three principal modes obtained using PCA over the simulated time series. B) Entropy production (left) and flux curl (right) computed from the model fitted to empirical fMRI data obtained during wakefulness and three progressively deeper stages of human sleep (up), as well as or wakefulness and two levels of propofol-induced loss of consciousness (bottom). * $p < 0.05$, Wilcoxon's test, Bonferroni corrected for multiple comparisons.

the empirical observables for wakefulness, for all human non-rapid eye movement (NREM) sleep stages (N1, N2, N3), and for propofol-induced sedation and loss of consciousness. We then used the optimal parameters to enhance the BOLD signal lengths of each subject ($n=15$) up to 30.000 samples for each region of interest [16, 17], allowing us to estimate entropy production as done with the ECoG data (Fig. 2). Figure 5A (bottom) presents the three main modes extracted from the simulated BOLD time series using PCA for parameters corresponding to wakefulness. Figure 5B presents entropy production and curl of probability flux estimated

from the time series corresponding to all sleep stages and depths of propofol sedation. We found that N2 and N3 sleep (the deepest stages of NREM sleep) presented dynamics significantly closer to detailed balance than conscious wakefulness; we found the same for propofol sedation and loss of consciousness, but not for early (N1) sleep.

Conclusions- The combined assessment of ECoG and fMRI data (assisted by model-based temporal enhancement) demonstrated that macroscopic non-equilibrium dynamics are a signature of consciousness spanning two model organisms measured with two different neural activity recording techniques, both during deep sleep (N2 and N3 sleep, but not N1), and under the effects of several anaesthetic drugs (propofol, ketamine, and ketamine plus medetomidine for monkeys; two doses of propofol for humans).

Most of the leading theories agree that consciousness and cognition emerge from the coordinated and dynamic interplay of distributed brain activity. For example, the global workspace theory posits that consciousness is equivalent to the non-linear percolation of sensory information throughout an anatomical connectivity backbone, allowing flexible access by different modular cognitive processors [18]. The dynamic core hypothesis by Edelman and Tononi identifies consciousness with an integrated dynamic process capable of achieving an enormous number of configurations [19]. Even if these explanations are considered satisfactory from a mechanistic perspective, the problem stands still: how does the neural collective self-organizes to fulfil the dynamic requirements of these theories? Our work shows that non-equilibrium brain dynamics is a general feature of conscious states. In turn, this type of dynamics can be achieved through different mechanisms, such as statistical [20] and dynamic criticality [21].

Importantly, we established the independence between non-equilibrium entropy production and spectral features that are frequently used to characterize unconscious states. This observation is relevant since some of the general anaesthetic we investigated (e.g. ketamine [22]) fail to induce the type of low frequency and complexity oscillations that considered indicative of unconsciousness, suggesting that our metrics are more general.

In summary, we demonstrated a link between a very general property of the brain as a macroscopic physical system, and the emergence of consciousness and cognition. Future studies should refine our conclusions, attempting to converge towards the relationship between dynamics and computation in neural tissue, one of the most challenging and long-standing problems in the field.

Authors acknowledge funding from Agencia Nacional De Promocion Cientifica Y Tecnologica (Argentina), grant PICT-2018-03103. The authors acknowledge the Toyoko 2020 program for granting cloud computing services.

-
- [1] F. Gnesotto, F. Mura, J. Gladrow, and C. P. Broedersz, Broken detailed balance and non-equilibrium dynamics in living systems: a review, *Reports on Progress in Physics* **81**, 066601 (2018).
- [2] X. Fang, K. Kruse, T. Lu, and J. Wang, Nonequilibrium physics in biology, *Reviews of Modern Physics* **91**, 045004 (2019).
- [3] C. Battle, C. P. Broedersz, N. Fakhri, V. F. Geyer, J. Howard, C. F. Schmidt, and F. C. MacKintosh, Broken detailed balance at mesoscopic scales in active biological systems, *Science* **352**, 604 (2016).
- [4] D. A. Egolf, Equilibrium regained: from nonequilibrium chaos to statistical mechanics, *Science* **287**, 101 (2000).
- [5] J. Wang, Landscape and flux theory of non-equilibrium dynamical systems with application to biology, *Advances in Physics* **64**, 1 (2015).
- [6] É. Roldán and J. M. Parrondo, Estimating dissipation from single stationary trajectories, *Physical review letters* **105**, 150607 (2010).
- [7] M. A. Buice and J. D. Cowan, Field-theoretic approach to fluctuation effects in neural networks, *Physical Review E* **75**, 051919 (2007).
- [8] H. Yan, L. Zhao, L. Hu, X. Wang, E. Wang, and J. Wang, Nonequilibrium landscape theory of neural networks, *Proceedings of the National Academy of Sciences* **110**, E4185 (2013).
- [9] S. De Franciscis, J. Torres, and J. Marro, Unstable dynamics, nonequilibrium phases, and criticality in networked excitable media, *Physical Review E* **82**, 041105 (2010).
- [10] M. Esposito, Stochastic thermodynamics under coarse graining, *Physical Review E* **85**, 041125 (2012).
- [11] C. W. Lynn, E. J. Cornblath, L. Papadopoulos, M. A. Bertolero, and D. S. Bassett, Non-equilibrium dynamics and entropy production in the human brain, arXiv preprint arXiv:2005.02526 (2020).
- [12] C. Koch, M. Massimini, M. Boly, and G. Tononi, Neural correlates of consciousness: progress and problems, *Nature Reviews Neuroscience* **17**, 307 (2016).
- [13] R. Cofré, R. Herzog, P. A. Mediano, J. Piccinini, F. E. Rosas, Y. Sanz Perl, and E. Tagliazucchi, Whole-brain models to explore altered states of consciousness from the bottom up, *Brain Sciences* **10**, 626 (2020).
- [14] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain, *Neuroimage* **15**, 273 (2002).
- [15] I. P. Ipiña, P. D. Kehoe, M. Kringelbach, H. Laufs, A. Ibañez, G. Deco, Y. S. Perl, and E. Tagliazucchi, Modeling regional changes in dynamic stability during sleep and wakefulness, *NeuroImage* , 116833 (2020).
- [16] Y. S. Perl, C. Pallacivini, I. P. Ipina, M. L. Kringelbach, G. Deco, H. Laufs, and E. Tagliazucchi, Data augmentation based on dynamical systems for the classification of brain states, *bioRxiv* (2020).
- [17] Y. S. Perl, H. Bocaccio, I. Pérez-Ipiña, F. Zamberlán, J. Piccinini, H. Laufs, M. Kringelbach, G. Deco, and E. Tagliazucchi, Generative embeddings of brain collective dynamics using variational autoencoders, *Physical Review Letters* **125**, 238101 (2020).
- [18] G. A. Mashour, P. Roelfsema, J.-P. Changeux, and S. Dehaene, Conscious processing and the global neuronal workspace hypothesis, *Neuron* **105**, 776 (2020).
- [19] G. Tononi and G. M. Edelman, Consciousness and complexity, *science* **282**, 1846 (1998).
- [20] E. Tagliazucchi, D. R. Chialvo, M. Siniatchkin, E. Amico, J.-F. Brichant, V. Bonhomme, Q. Noirhomme, H. Laufs, and S. Laureys, Large-scale signatures of unconsciousness are consistent with a departure from critical dynamics, *Journal of The Royal Society Interface* **13**, 20151027 (2016).
- [21] G. Solovey, L. M. Alonso, T. Yanagawa, N. Fujii, M. O. Magnasco, G. A. Cecchi, and A. Proekt, Loss of consciousness is associated with stabilization of cortical activity, *Journal of Neuroscience* **35**, 10866 (2015).
- [22] A. Maksimow, M. Särkelä, J. Långsjö, E. Salmi, K. Kaisti, A. Yli-Hankala, S. Hinkka-Yli-Salomäki, H. Scheinin, and S. Jääskeläinen, Increase in high frequency eeg activity explains the poor performance of eeg spectral entropy monitor during s-ketamine anesthesia, *Clinical neurophysiology* **117**, 1660 (2006).