Neuronal cascades shape whole-brain functional dynamics at rest

Giovanni Rabuffo*, Jan Fousek, Christophe Bernard, Viktor Jirsa

Institut de Neurosciences des Systèmes, Aix-Marseille University, Marseille, France

March 16, 2021

Abstract

At rest, mammalian brains display remarkable spatiotemporal complexity, switching between different brain states on a slow timescale of the order of tens of seconds. While the phenomenology of the resting state dynamics is valuable in distinguishing healthy and pathological brains, little is known about its underlying mechanisms. Here, using full-brain network modeling, we show that neuronal cascades spontaneously emerge on a slow timescale when neuronal populations are coupled via a detailed structural connectome. Neuronal cascades occur under conditions of near-criticality and are characterized by epochs of increased deviations from baseline firing rate activity. The ignition and subsequent propagation of cascades depend upon the brain state and connectivity of each region. The largest cascades produce bursts of Blood-Oxygen-Level-Dependent (BOLD) co-fluctuations at pairs of regions across the brain, which shape resting state network dynamics. We experimentally

*Corresponding authors: giovanni.rabuffo@univ-amu.fr, viktor.jirsa@univ-amu.fr

confirm these theoretical predictions. We demonstrate the existence and stability of intermittent epochs of functional connectivity comprising BOLD co-activation bursts in mice. We then provide evidence for the leading role of the neuronal cascades in humans with simultaneous EEG/fMRI recordings. These results show that neuronal cascades are the major determinant of spontaneous fluctuations in brain dynamics at rest.

1 Introduction

At rest, functional Magnetic Resonance Imaging (fMRI) reveals the existence of periods during which blood-oxygenation-level-dependent (BOLD) activity is highly correlated between specific brain regions, known as resting state networks (*RSNs*). *RSNs* are consistently observed across several mammalian species including human [1, 2] and non-human [3] primates, as well as in rats [4, 5] and mice [6, 7, 8, 9]. Functional Connectivity (FC) can be used to characterize such highly correlated periods of a measured signal (e.g. BOLD or electrophysiological) between nodes in a network. Dynamic FC (dFC) refers to changes in FC over time [10]. Remarkably, in all mammalian species studied so far, whole-brain dynamics is characterized by the switching between periods of stable FC patterns (hence *RSNs*) which give rise to large-scale infraslow fluctuations (< 0.1Hz) [11, 12, 13, 14, 15, 16, 17].

The switching behavior of *RSNs* can be quantified by dFC, providing reliable signatures of healthy, ageing and diseased brains [18, 19, 20, 21, 22, 23]. Our current understanding of these phenomena rests upon computational studies, performed at the whole brain level, demonstrating that the *RSN* dynamics is an emergent property of the network [24], which operates near criticality [25, 26, 27]. Nevertheless the range of possible mechanistic underpinnings remains vast and requires a further narrowing down to enable researchers to make use of the mechanistic understanding in future diagnostic and curative clinical applications.

Here we address this issue and investigate *in silico* the possible neuronal mechanisms that give rise to whole-brain functional dynamics at rest, generate predictions and test them *in vivo*. We use The Virtual Brain (TVB), a neuroinformatics and simulation platform, which allows connectome-based whole brain modeling of multiple species including human [28] and mice [29]. TVB includes a fixed number of network nodes (at least one per brain region) and the physical connections that link them, such as white matter tracts defined by diffusion tensor imaging or

axon projections obtained with virus injections [30]. A large number of neural mass models are readily available in TVB, which produce activity of neuronal populations and map on a range of brain imaging modalities including BOLD fMRI, EEG and MEG signals. We here adapt a novel neural mass model (NMM), which has been previously derived as the exact thermodynamic limit of an infinite number of coupled quadratic integrate-and-fire neurons [31] (Figure 1a). This analytic step allows to derive the average firing rate and membrane potential of a mesoscopic neuronal population, thus providing suitable neural mass variable while keeping track of the internal organization of the constituent neurons. Using this NMM, we demonstrate that slow neuronal cascades spontaneously emerge in the network. The largest neuronal cascades give rise to bursts of BOLD co-activation (BOLD-CA) events at pairs of brain regions across the brain, which account for stable long-lasting FC states. We verify experimentally the presence of neuronal cascades in brain network activations *in-silico* in two species, mouse and man, across BOLD-fMRI and simultaneous EEG-fMRI imaging modalities. These findings demonstrate that large neuronal cascade dynamics is the principal mechanism underlying the switching behavior of the intermittently stable FC states and can be robustly assessed in non-invasive brain imaging signals.

2 Results

2.1 Activity of single and connected nodes

We first consider the dynamics of a single isolated brain node (not connected to the network). If there is no input current ($I_{ext} = 0$, left panel in Figure 1b), solving the NMM equations (see Methods) unravels the dynamics of the node in terms of the mean firing rate $r(t)$ and mean membrane potential $V(t)$. The phase plane representation displays an 'up' stable focus and a 'down' fixed point (left panel in Figure 1b). If, at a given time, a node takes a given (r, V) value, the direction of the arrow at these coordinates indicates the direction of the flow, and where the node dynamics may end up: either the 'up' state in which the neurons populating the NMM exhibit an intense activity with high average firing rate and low average membrane potential; or the 'down' state in which the local neurons posses low firing rate and strong (negative) membrane potential.

The separatrix is the boundary line separating the up and down basins of attraction. In the absence of noise, neuronal activity of the node would rapidly freeze into a stable up or down level of activity. Introducing noise allows the node to explore the phase space. If the noise is sufficiently large and/or if the (r, V) coordinates of the node are not too far from the separatrix, noise driven movement in the phase plane can make the node cross the separatrix and switch from up→down or down→up states.

We now consider a non-null input current I_{ext}, which would correspond to synaptic inputs if the node is connected to other nodes. As I_{ext} increases (Figure 1b), the two branches of the V-nullcline move horizontally away from each other, while the r-nullcline remains fixed. Correspondingly, the location of the stable points in the phase plane shifts and their basins of attraction change in size. The separatrix is thus a dynamical element: its location depends on the amount of input that the node is receiving at a given time. If a node has a low firing

rate (down state) and if I_{ext} is small, the separatrix is close to the up state, making it nearly impossible for the node to reach the separatrix with random fluctuations. The node has a very low probability to switch to a high firing rate; it remains in the down state. Conversely, if the separatrix is close to the down state when I_{ext} is high, the node remains in the up state. This property of the NMM is key to understand the system's dynamic. At a given time t the inputs received by node n will change the phase plane and the separatrix, and make it easier (or not) for the node to change its activity as a function of noise and of its (r, V) coordinates at time t.

The inputs (I_{ext}) of a node are provided by the regions connected to it (Figure 1c). We use a brain avatar made of a large number of nodes corresponding to different brain areas (Figure 1d). The differential equation governing the evolution of V_n at node n contains the input current I_{ext}, which represents the synaptic drive from the firing rates r_m of all the m nodes that are connected to node n (Figure 1c; see Methods Section). The communication between regions is weighted by the structural links W_{nm} of the connectome (Figure 1e). As an effect of the coupling, the projected phase plane (r_n, V_n) can be thought of as a distorted version of the phase plane of a single isolated node, where the location of the down fixed point and the up stable focus depends on the structural role that each region plays in the network hierarchy. Poorly connected regions in the connectome receive a weak input current I_n so that the local phase plane is slightly distorted. Instead, regions with high centrality in the connectome, or that are part of strongly connected structural motifs, receive a strong input, which causes a greater distortion of the local phase plane.

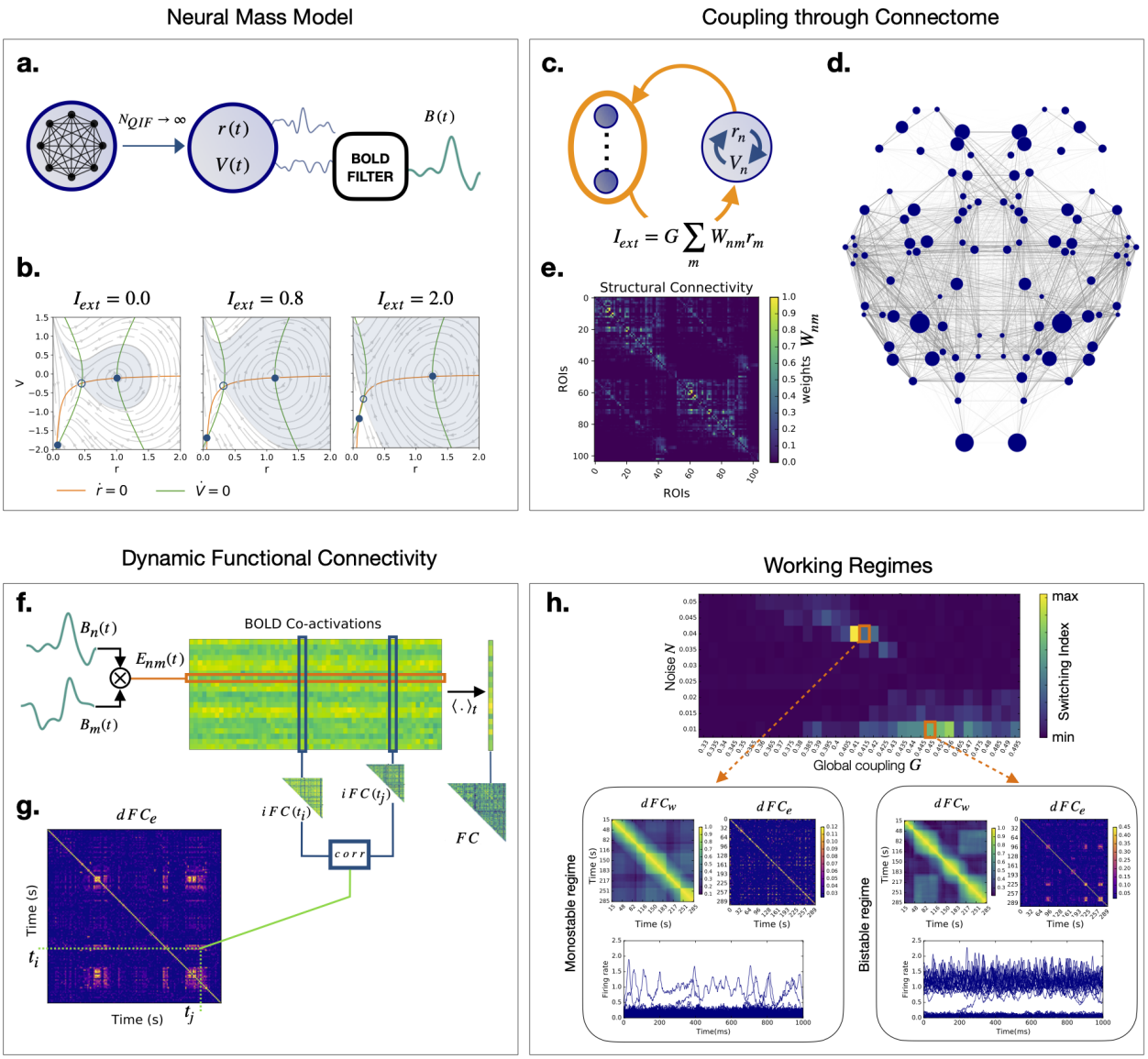

Neural Mass Model
Coupling through Connectome
a.
$N_{QIF} \to \infty$
$r(t)$
$V(t)$
BOLD FILTER
$B(t)$
b.
$I_{ext} = 0.0$
$I_{ext} = 0.8$
$I_{ext} = 2.0$
$\dot{r} = 0$
$\dot{V} = 0$
c.
r_n
V_n
$I_{ext} = G \sum_m W_{nm} r_m$
d.
e.
Structural Connectivity
ROIs
weights W_{nm}
Dynamic Functional Connectivity
Working Regimes
f.
$B_n(t)$
$B_m(t)$
$E_{nm}(t)$
BOLD Co-activations
$\langle \cdot \rangle_t$
$iFC(t_i)$
$iFC(t_j)$
corr
FC
g.
dFC_e
Time (s)
t_i
t_j
h.
Noise N
Global coupling G
max
Switching Index
min
Monostable regime
Bistable regime
dFC_w
dFC_e
Time (s)
Firing rate
Time(ms)

Figure 1:

Figure 1: Connectome based modeling. (a) The Neural Mass Model (NMM) used is derived as the thermodynamic limit of infinite quadratic integrate-and-fire (QIF) neurons. The outcome of the numerical integration of the NMM over the mouse connectome are neuroelectric time series for the average firing rate $r_n(t)$ and the average membrane potential $V_n(t)$ for each brain region. Applying the Balloon-Windkessel model to $V_n(t)$ we obtain the simulated BOLD signal $B_n(t)$. (b) The phase plane of each decoupled node ($I_{ext} = 0$) has a down fixed point and a up stable focus. As the external current I_{ext} is increased, the phase plane of the neural mass changes (see equations in Methods). In particular, the basin of attraction of the up state gradually becomes larger than that of the down state, while the fixed points move farther apart. (c) From the local node perspective, the network action can be thought of as an input current which provokes a distortion of the phase plane. In fact, the firing rate of the coupled regions act as an excitatory external current I_{ext} in the local neural mass equations. Therefore, strongly connected regions within the network will have a higher probability of being attracted into the up state with a large basin of attraction. (d-e) The mouse connectome and structural connectivity W_{nm} imported from the tracer experiments of the Allen Institute. The 104 cortical regions of interest (ROIs) are specified in the table in Figure S2b. (f) Given two nodes n and m the edge co-activation (CA) signal $E_{nm}(t)$ is defined as the product of the z-scored BOLD signal $B_n(t)$ and $B_m(t)$. (C) Each column of the BOLD-CA matrix represents and instantaneous realization of the Functional Connectivity (iFC). Averaging the $iFC(t)$ over time we obtain the static Functional Connectivity (FC). (g) The elements (i, j) of the edge-dynamic Functional Connectivity (dFC_e) matrix are defined as the Pearson Correlation between $iFC(t_i)$ and $iFC(t_j)$. (h) Two qualitatively distinct regimes of non-trivial functional dynamics. For every couple of global parameters (G, N) we calculated the dynamic functional connectivity (dFC) in a sliding window approach (dFC_w; as in Figure S1) and in an edge-centric approach (dFC_e; as shown in Figure 1g). The 'switching index' of each dFC_w matrix, was evaluated as the variance of the respective upper triangular part. We find two

regimes of activity, named *monostable* and *bistable*, where qualitatively distinct neuroelectric organizations give rise to a large-scale functional dynamics characterized by a non-vanishing switching index. In the monostable regime, the dFC_w and dFC_e display off-diagonal blocks, demonstrating a correlation between the functional activity at distinct times. The low global coupling G in the monostable regime does not guarantee a strong communication between the brain network regions, which most of the time populate the low firing rate ('down') state. A strong noise N pushes the brain regions in the high firing rate ("up") state for short transients. Also in the bistable regime, the dFC_w and dFC_e display off-diagonal blocks. A higher value of the global coupling promotes a subgroup of regions in the high firing rate ('up') state. Low levels of noise perturb the equilibrium of the system provoking localized switching in both up→down and down→up directions.

2.2 From node electrical activity to whole brain BOLD dynamics

We simulate whole-brain dynamics in a mouse brain avatar, which detailed structure (Figure 1d) is imported from the tracer experiments of the Allen Institute and is parcelated into anatomic regions of interest, associated with the network nodes [32]. Once a working regime is chosen by the selection of global and local model parameters, the raw outcome of the simulation consists of a high time-resolution neuroelectric signal: the average firing rate r and membrane potential V for each node. We refer to the average firing rate activity as the simulated EEG data. With a further processing step, we obtain a low time-resolution simulated BOLD activity by filtering the membrane potentials through the Balloon-Windkessel model [33] (Figure 1a).

In order to compare simulated and empirical whole brain imaging data, we use different metrics. There exists several dFC measures, which describe the dynamic evolution of functional brain patterns [10, 34, 35, 36, 37, 38]. The most diffuse, is based on a sliding window approach [12]. In brief, inside each time window, a static FC is computed as the correlation matrix of the BOLD activities. Then, the entries of the windowed-dFC (dFC_w) matrix are defined as the correlation between the FCs at different windows (see Methods and Figure S1a, top panel). For our analysis, we mainly exploit dynamic functional measures derived in an edge-centric approach, where the edge dynamics is defined by the product of the z-scored BOLD activity at each possible pair of nodes [39]. The resulting co-activation (CA) time series Figure 1f are associated to each network edge and have the same time resolution of the BOLD signals. A high co-activation amplitude corresponds to the simultaneous fluctuation of the nodes away from their respective baselines. Each column in Figure 1f represents an instantaneous realization of the Functional Connectivity (iFC) of all possible network edges at a time t. The static FC is obtained by averaging over all iFC in the time domain. The edge dynamic functional connectivity (dFC_e, Figure 1g) is defined (without the use of sliding windows) by the Pearson Correlation of the

iFCs at each couple of times t_i and t_j.

In the Methods section, we compare the edge-based and the windowed dFC measures. The window approach allows a more reliable measure of the correlations to the detriment of the resolution over the temporal structure. The edge approach maintains the full-time resolution, but it is more sensitive to spurious correlations. Here, we mainly use dFC_e because we are interested in the neuronal activity that gives rise to resting state dynamics at high temporal resolution.

2.3 Transient co-activations account for whole brain dynamics at rest

We analyze the situation in which the local model parameters are fixed everywhere such that any decoupled region is in a bi-stable configuration (see Methods and Figure 1b for $I_{ext} = 0$). We then explore the parameter space, varying the global parameters G (the global synaptic coupling) and N (noise), representing the impact of the structure over the local dynamics, and the stochastic currents simulating a generic environmental noise (thermodynamic, chemical, synaptic...), respectively. Then, we look for regions of the parameter space where the dFC matrices show off-diagonal block structures, as typically seen in empirical datasets. The block structures in the dFC_w matrix represent time periods during which a set of regions show stable and highly correlated BOLD values, which correspond to discrete FC states. We define the 'switching index' i.e., the variance of the upper triangular part of the dFC_w matrix (Figure 1h) to quantify whole brain dynamics, which depends on the number of different block structures of stable activity and how often they switch between one another. We identify two regimes giving rise to *RSN* dynamics, where the respective neuroelectric correlates are qualitatively different.

- In the first case, a weak structural coupling (low G) is insufficient to promote any region into a high firing rate state, because the excitatory drive received by any node is relatively small. However, all regions can transiently reach a high firing rate state because the stochastic

drive is strong. We refer to this as the *monostable* regime (Figure 1h, bottom left).

- In the second case, the stronger structural coupling pushes a subset of regions into a strongly active state leaving the remaining regions in the down state. A weak stochastic drive ensures stable dynamics with few regions jumping between the up and down states. We call this the *bistable* regime (Figure 1h, bottom right).

We conclude that the novel NMM introduced here reproduces the main feature of whole brain dynamics seen in experimental data in mice and humans: the off-diagonal block structure of the dFC matrices. In the following, we focus on the (G, N) values corresponding to the orange boxes in Figure 1h, i.e. the monostable and bistable regimes.

Summarizing, the balance between the local dynamics and the global connectivity, together with appropriate levels of perturbation of the system, allow the same local model to generate large-scale organizations similar to those measured empirically. The action of a node at the network level can be understood with the following intuitive scheme. Let us consider that a node n undergoes a sudden transition, from a low to a high firing rate state (i.e., from down to up state). Accordingly, this region will become 'visible' to the other nodes across the network, as it will start communicating with increased intensity. The regions more connected to n will be greatly influenced while the less connected ones will continue their normal activity almost undisturbed. In particular, such local transition will increase the up-state basin of attraction in the phase planes of the receiving nodes, thus favoring a transition towards (or consolidating) a high firing rate state. Conversely, a node n switching from up to down will produce a large decrease in the current driving its targets, hence favoring a switch to a down state.

The block structures in the dFC_w gives the impression that whole brain dynamics is organized by a sequence of transient but stable periods of correlated activity. However, the analysis of the corresponding dFC_e (Figure 1h, bottom panels) reveals that whole brain dynamics is in fact

characterized by much smaller periods of co-activations. When dFC_w is used, as the window slides in time, it will capture the transient co-activation as long as the latter is present in the window, giving rise to a block structure (Figure S1). Short duration bouts of BOLD co-activations seem to account for whole brain dynamics at rest, in keeping with results from empirical data analyzed in terms of static BOLD-CA *events* [39], and not dynamics BOLD-CA as here. We then analyze the underlying mechanisms at the neuronal scale.

2.4 Spontaneous emergence of neuronal cascades

Here we introduce a framework to interpret the simulated firing rate dynamics, its dependence on the structural connectome and the basic mechanism which allows the emergence of infra-slow fluctuations on the neuronal activity; the neuronal cascades. We focus on the bistable regime where the low noise level allows a clearer visualization of the cascades' mechanism. The hypotheses developed here will be later tested on the monostable regime as well as in empiric data. We identify five main categories of nodes based on their firing rate activity:

- (D) *down* regions (in light blue in Figure 2a) display a low firing rate state throughout the duration of the simulation, as the up state is practically unreachable by noise-driven fluctuations (the separatrix is close to the up state). Their low firing rate regime makes them poor communicators in the network hierarchy.

- (U) *up* regions (in light red in Figure 2a) always show a high firing rate, constantly providing inputs to their targets. The separatrix lays close to the down state and is never reached.

- (J) *jumping* regions (in green in Figure 2a) undergo regular local transitions, dwelling for a relatively long time in both low- and high- firing rate states. In the phase space of these regions, the separatrix lies midway between the up and down states. Therefore, they have the same probability to jump from up to down and vice versa. Note that the timing of the

(J) jumps defines a new *slow* time scale for the system. The fact that these jumps occur on a regular basis during the simulation ensures that the slow time scale is ever-present in the large scale network dynamics.

- (D*) *down-up* regions (in dark blue in Figure 2a) have a stable activity around the down fixed point but–in some occasions–manage to reach the up stable focus for a certain transient time.

- (U*) *up-down* regions (in dark red in Figure 2a) have a stable activity around the up state but are occasionally driven into a short-lived change of state in the down state.

In the shown simulation, the (D) regions include the default network regions (anterior cingulate, retrosplenial, and hippocampus) but also the cerebellum and some visual areas. The (U) regions are temporal/lateral regions. The (D*) regions mainly include motor regions, but also the subiculum and the piriform cortex. The (U*) regions are mainly somatosensory. The (J) regions are the insula and entorhinal cortex. The distribution of nodes into the five categories depends upon the choice of parameters (see supplementary Figure S2a for the list of region labels).

Since the nodes of class (J) can jump regularly, we expected them to have a strong effect on whole brain dynamics. A Principal Component Analysis (PCA) shows that nodes (J) explain most of the system variance (Figure S2b). However, given their uniform unfold throughout the simulation, the activity of regions (J) rarely grows above three standard deviations from baseline (green in Figure 2b).

In contrast, we find that a substantial contribution to the large deviations (over three SD) from baseline is made by the regions in classes (D*) and (U*), which have a less regular activity. During their occasional excursions in the other stable state, the activity of these regions substantially deviates from the mean (dark red and blue in Figure 2b). As we will describe in detail in the next sections, these events generate a neuronal cascade in the brain i.e., they subtend epochs of

increased deviations from baseline activity which last on the order of tens of seconds.

The topographic organization of the average firing rate in the connectome remarks the influence of the structure in organizing the whole-brain dynamics (Figure 2a). Depending on the location on connectome, a node is assigned a different role in the network dynamics. Let us consider the case of a node with jumping potential, which is also a hub for the network. Its local reshaping will have a wide influence throughout the network. However, the hubness of a jumping node does not guarantee the crossing of the separatrix by its target regions. If the hub is connected to a very stable subset of regions, following a jump of the hub, most of the separatrix lines in the targets will be slightly shifted, but not enough to allow any crossing. In contrast, if there exists a subset of strongly connected nodes with jumping potential, the jump of one will have strong effects within such subnetwork. In order to illustrate this point, let us consider that a group of nodes from class (U*), have very strong links binding them. Then, the occasional jump down of one of these nodes will strongly shift the separatrix of the other (U*) nodes towards the up stable focus, increasing their probability of jumping down. Then, many (U*) nodes will be dragged down one after the other, facilitating the next jumps, producing a cascade effect (see for example Figure 2b, bottom). During this event, the (J) nodes get also involved, interrupting for some time their regular control over the network. A windowed analysis of the leading eigenvector shows that the nodes in the $\star$-classes are driving the system during these rare events (Figure S2c).

Summarizing, we have shown that jumps of the firing rate at local sites are extremely relevant for the unfolding of the microscopic activity throughout the brain, as they provoke the largest perturbations in the system. If jumps happen with a certain regularity, they will generate a slow time scale in the system. The slow rhythm in our simulations can be traced back to a small subset of regions (J) which transit between the up and the down states. These regions lead the baseline evolution of the entire system, acting as homeostatic agents. In some rare,

but important occasions, this control fails during large cascades of activity. Such cascade effect takes place when a sudden local transition brings a typically stable region ((U*) or (D*)) away from baseline activity and induces other regions to reorganize. The cascade goes on until the noise and the structural pressure take the system back to its normal evolution e.g., through the normalizing action of the jumping regions (J). We then explore whether such rare deviations from baseline activity account for resting state brain dynamics.

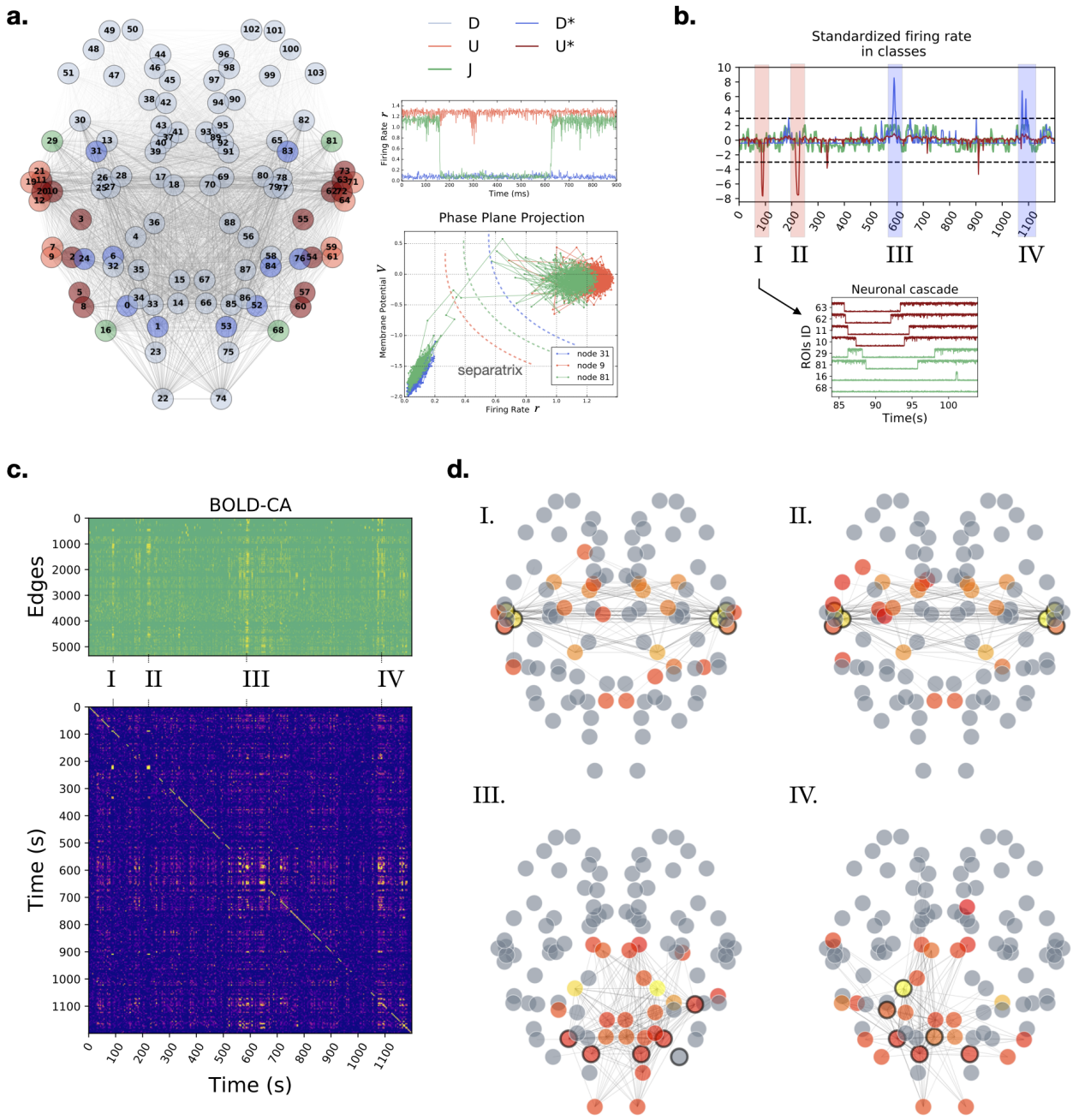

a.
D
U
J
D*
U*
Firing Rate r
Time (ms)
Phase Plane Projection
Membrane Potential V
separatrix
node 31
node 9
node 81
Firing Rate r
b.
Standardized firing rate
in classes
I
II
III
IV
Neuronal cascade
ROIs ID
Time(s)
c.
BOLD-CA
Edges
I
II
III
IV
Time (s)
Time (s)
d.
I.
II.
III.
IV.

Figure 2:

Figure 2: Mechanisms of cascade generation in the bistable synthetic model. (a) Different regions have a different fate depending on their structural role. If the overall input is too low a region is confined into the down fixed point (in the projected local phase plane) explored by noisy fluctuations (class D, light blue). If the input is too strong, a region remains stuck in the up stable focus plus noisy fluctuations (class U, light red). Only a subset of regions with the right level of input are capable of jumping over the separatrix between the up and down states. Few regions jump all the time throughout the simulation (class J, green). The dark red (class U*) and dark blue (class D*) regions dwell most of the time in the up and down states, respectively. Only in important rare occasions they cross the threshold to jump on the other side. The spatial distribution of the (D, U, J, D*, U*) classes is extremely symmetric and shows a clear role of the structure in defining the local dynamics. (a, right) Example of firing rates activity and exploration of the projected 2D phase spaces by the 'up', 'down' and 'jumping' regions. (b, top) Standardized average firing rate in the (J), (U*) and (D*) classes. The dark red (U*) and dark blue (D*) classes exceed the threshold (dashed lines) and carry other regions of the network away from baseline in a cascade effect. (b, bottom) Example of cascade: when the node 63 jumps into the down state, it first drags down the node 62 (with which it shares the strongest structural link in the network). After them, other strongly connected nodes follow the trend. The large neuronal cascades elicited by the jumps in the (U*) and (D*) classes (examples I, II and III, IV, respectively) have important consequences on the whole-brain dynamics. (c, top) During a neuronal cascade we observe a peak of BOLD co-activation at couples of regions across the whole brain. The blocks in the dFC_e matrix (c,bottom) appear in correspondence of CA events, showing that these bursts generate stable epochs of functional connectivity correlated in time. (d) In each selected epoch, the large neuronal cascades drive other nodes away from baseline activity (circled in black) and promote specific functional hubs at the BOLD level, represented by colored nodes in the network plots (strongest representatives in yellow. Gray regions have

been excluded as they do not contribute substantially.). A functional hub is defined by the components of the first leading eigenvector (eigenvalue $\lambda > 0.41$) associated to the istantaneous iFCs at times t_I, t_{II}, t_{III} and t_{IV}, respectively. Only the edges with the highest co-activations are displayed. Importantly, CA events generated from neuronal cascades at specific sites support distinct functional networks which are not correlated among themselves (e.g., no off-diagonal dFC_e block between I and III).

2.5 Neuronal cascades activate specific functional networks

The previous analysis focused on neuronal activity as defined by the firing rate r. We now look at the BOLD signal derived from the membrane potential V. Figure 2c (top) shows that whole brain dynamics at the BOLD level is characterized by the emergence of collective BOLD co-activations (CA), which are aligned to the large deviations from baseline of the firing rate activity of the (U*) class (Figure 2b, segments I and II) or the (D*) regions (Figure 2b, segments III and IV). When looking at the corresponding dFC_e matrix (Figure 2c, bottom), we also note that the most stable dFC_e blocks correspond to the largest BOLD-CA events. This observation is not a triviality, since a higher co-activation does not imply a stronger correlation. The presence of off-diagonal blocks in the dFC_e shows that CA patterns are correlated when generated either by the (U*) or in the (D*) regions. In this example, the CA patterns generated by the jumps of distinct classes are not correlated. Each CA pattern defines the upper triangular part of an instantaneous functional connectivity iFC (as in Figure 1f). Figure 2d displays, for each selected epoch, the edges with the highest CA values and the functional hubs, a subset of regions with a central role in the functional dynamics [29]. A functional hub is defined by the regions with the largest components of the iFC's leading eigenvector, represented as different colors in the brain network plots. We highlight with a black rim the nodes whose firing rate deviated from baseline activity. The same perturbation can re-occur spontaneously eliciting similar functional networks correlated at distinct times, as it is marked by the off-diagonal blocks of the dFC_e (see the network plots I, II or III, IV). Summarizing, neuronal perturbations starting at different locations in the connectome elicit different BOLD-CA events, which result in uncorrelated functional networks.

2.6 Neuronal cascades and BOLD co-activations in empiric data

In the previous sections we described how the large spontaneous deviations from baseline activity play a key role in the system dynamics by activating specific large-scale functional networks. Here we provide a working definition of *neuronal cascades* as a global measure of such long-lasting perturbations of the neuroelectric activity (neuronal cascades should not be confused with the concept of *neuronal avalanche* [40]. See supplementary Figure S3). First, we binarize the firing rate activity in every brain region by assigning a unitary value when it exceeds 3 standard deviations (Figure 3a). Then, we extract the magnitude of the deviations from baseline by summing over the regions (gray line in Figure 3b), we convolve the signal with a Gaussian kernel ($\sigma = 1.94s$ = BOLD Time Repetition), and we finally downsample it to the same time resolution of the BOLD signals. In the previous sections we presented a detailed functional analysis focusing on the bistable regime. In the following, we also analyze neuronal cascades and BOLD-CAs in the monostable to demonstrate the generality of our results. The resulting neuronal cascades signal (blue line in Figure 3b) is characterized by bumps which describe a long-lasting increase in the magnitude of non-standard perturbations. Our simulation shows localized increases in cascade activity, which are shorter in the noisy monostable regime and more prolonged in the bistable regime. Based on the results of the previous section (Figure 2d), neuronal cascades should give rise to BOLD-CAs. As anticipated, both monostable and bistable regimes display BOLD-CA patterns spontaneously emerging, simultaneously with neuronal cascades, in association with a non-trivial dFC_e (Figure 3c). In this section, the simultaneous nature of neuronal cascades and BOLD-CA is considered visually. The quantification will be provided in the following section.

These theoretical results predict that electrical activity in the brain is organized around transient neuronal cascades, which give rise to BOLD-CA. Using a unique dataset of simultaneous EEG-fMRI recordings in humans, we verify experimentally the presence of neuronal cascades and

BOLD-CA (Figure 3b).The results are remarkably similar (in particular for the bistable regime), all the more since the simulations are done using a mouse brain avatar. Although simultaneous EEG-fMRI recordings are not yet available in mice, we nevertheless verify the presence of BOLD-CA events in empiric mouse fMRI datasets (Figure 3d).

Another strong *in-silico* prediction is that BOLD-CA events and off-diagonal dFC_e blocks co-occur. We first extract from the BOLD-CA signals the root sum squared (RSS) time series, which quantifies the amplitude of the edge co-activations at each time point (Figure 3c-d, green lines). Then, we partition the RSS signal in CA *events* and nCA *non-events* (respectively above and below the 98th percentile of the RSS values). We show that the correlation between any two events is statistically higher than the correlation between events and non-events as well as between couples of non-events (Figure 4a). The empiric results are pooled over several experimental trials. In supplementary Figure S4a we show examples of the dFC_e, where times are sorted accordingly to increasing CA amplitude, and we explore the effect of a different thresholding on Figure 4a in the Human data. Since the correlation between co-activation patterns at different times correspond to the off-diagonal elements of the dFC_e matrix (as shown in Figure 1f-g), we conclude that bursts of BOLD-CA events account for the highest off-diagonal values in the dFC_e in both theoretical and experimental datasets, as predicted.

Since visual inspection suggests the co-occurrence of neuronal cascades and BOLD-CA, we quantified the phenomenon.

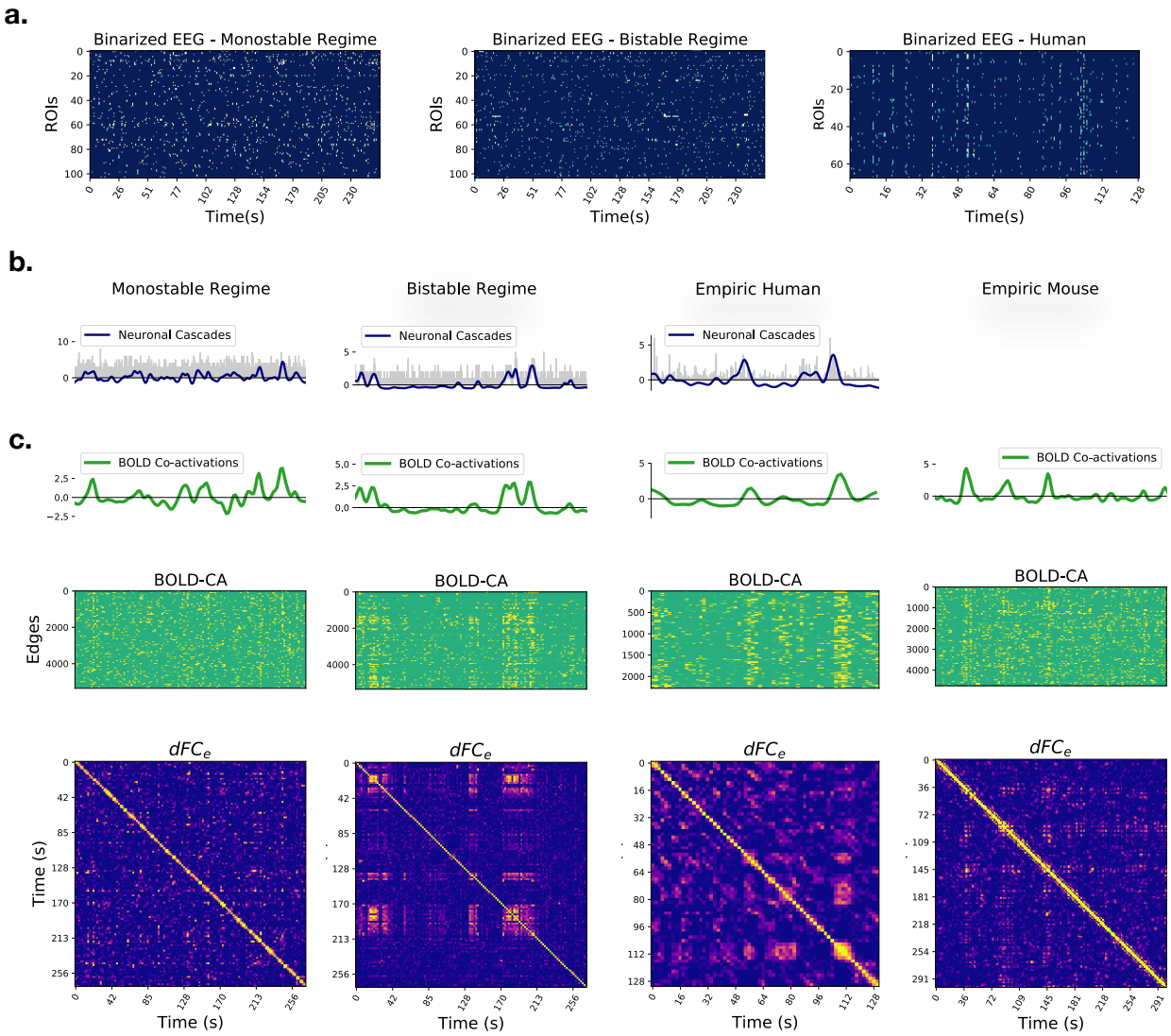

a.
Binarized EEG - Monostable Regime
Binarized EEG - Bistable Regime
Binarized EEG - Human
ROIs
Time(s)
b.
Monostable Regime
Bistable Regime
Empiric Human
Empiric Mouse
Neuronal Cascades
c.
BOLD Co-activations
BOLD-CA
Edges
dFC_e
Time (s)

Figure 3:

Figure 3: Neuronal cascades drive the functional dynamics. (a) Example raster plots of the firing rate deviations from baseline activity in the monostable and bistable synthetic regime and for a representative subject of the empiric EEG/fMRI Human dataset. In order to define the raster plot, a unitary value is assigned when a region goes above 3 standard deviations from its baseline activity, zero otherwise. (b) The binary sequence is summed over the brain regions to extract the global magnitude of the deviations from baseline (gray). This signal is convoluted with a Gaussian kernel (σ = BOLD Time Repetition (TR)) and downsampled to obtain the same resolution of the BOLD activity. This defines a sequence of *neuronal cascades* (blue). Notably, also in the Human data, the neuronal cascade signals are characterized by bumps unfolding over several seconds. (c, top and central panels) The BOLD-CA patterns in simulated and empiric mouse and human datasets are also characterized by sudden collective events involving large network parts. Its root sum square (RSS, green line) defines the collective edge CA amplitude. (c, bottom panels) The dFC_e matrices display non-trivial off-diagonal blocks. At a visual inspection, the BOLD-CA events seem to happen in coincidence with dFC_e blocks and, most notably, the neuronal cascade and BOLD-CA signals (blue and green lines in panels (b-c), respectively) seem to co-occur in most instance.

2.7 Neuronal cascades account for functional connectivity dynamics

When looking at the simulated activity, we note that the BOLD-CAs (RSS; green lines in Figure 3b) the neuronal cascades (blue lines in Figure 3a). Remarkably, a similar correspondence is present in a EEG/fMRI resting-state human dataset (best trial shown in Figure 3).

To characterize this correspondence, we correlate the cascades signal with the BOLD-CAs amplitude profile. In both the monostable and the bistable regimes, the two measures are significantly correlated ($\rho \sim 0.54$ and $\rho \sim 0.91$, respectively). In Figure 4b (top) we report the correlations between the cascades and the BOLD-CA amplitudes across several trials of a human cohort (see Methods section for trial selection). The correlation is carried out by shifting the RSS time series of a given time-lag. A peak of correlation appears naturally between the two measures when the BOLD signal is shifted backwards by 2 time points ($3.88s$) Figure 4b (bottom; blue line). Therefore, as expected, the EEG neuronal cascade signal precedes the BOLD co-activations by a few seconds. Notice that, in general, a shift forward of the BOLD activity results in a rapid loss of correlation.

Finally, to prove the statistical significance of our observations, we build surrogate BOLD signals and we compare the associated CA amplitudes to the neuronal cascade signal in each trial. In particular, we compare the observed cross-correlation profile (blue line Figure 4b) with the profiles extracted by correlation of neuronal cascades and surrogate CA amplitudes of three kinds:

- Time-shuffled: Example of cross-correlation profile in Figure 4b, orange line.
- Phase-randomized, cross-spectrum preserved: Example in Figure 4b, green line.
- Phase-randomized, cross-spectrum not preserved: Example in Figure 4b, purple line.

In all the surrogates, the functional dynamics is disrupted (see Methods). In the first two models, the static FC is preserved but its dynamic evolution is made trivial. Namely, the time-shuffled

model consists of random FC jumps around a fixed pattern, while in the phase surrogate (cross-spectrum preserved), stationarity is strictly imposed by destroying any coherent fluctuation around the static FC. These two surrogates also preserve the original fat-tail distribution of the BOLD CA amplitudes (RSS values), and therefore supports bursty CA events (Supplementary Figure S5a-b). When the cross-spectrum is not preserved, the static FC is corrupted and the CA amplitude distribution is normalized. For each surrogate type, we compute 1000 cross-correlation profiles. The distribution of their mean, maximum and variance (averaged across all the trials) is shown in Figure 4c. The observed values are always significant when compared with the Phase-randomized surrogate with cross-spectrum not preserved. The observed mean values lie between the random (time-shuffled) and the ordered (phase-randomized) scenario. Only the Phase-randomized surrogate, where the cross-spectrum is preserved, reaches the peaks of correlation of the observed data, and the significance of the observed maximum is $p = 0.108$. The variance of the correlation profile is always significant as compared to all the surrogate models. The result holds in several instances at the single subject level (Supplementary Figure S5c). We conclude that neuronal cascades drive BOLD-CAs and thus shape whole brain fMRI connectivity.

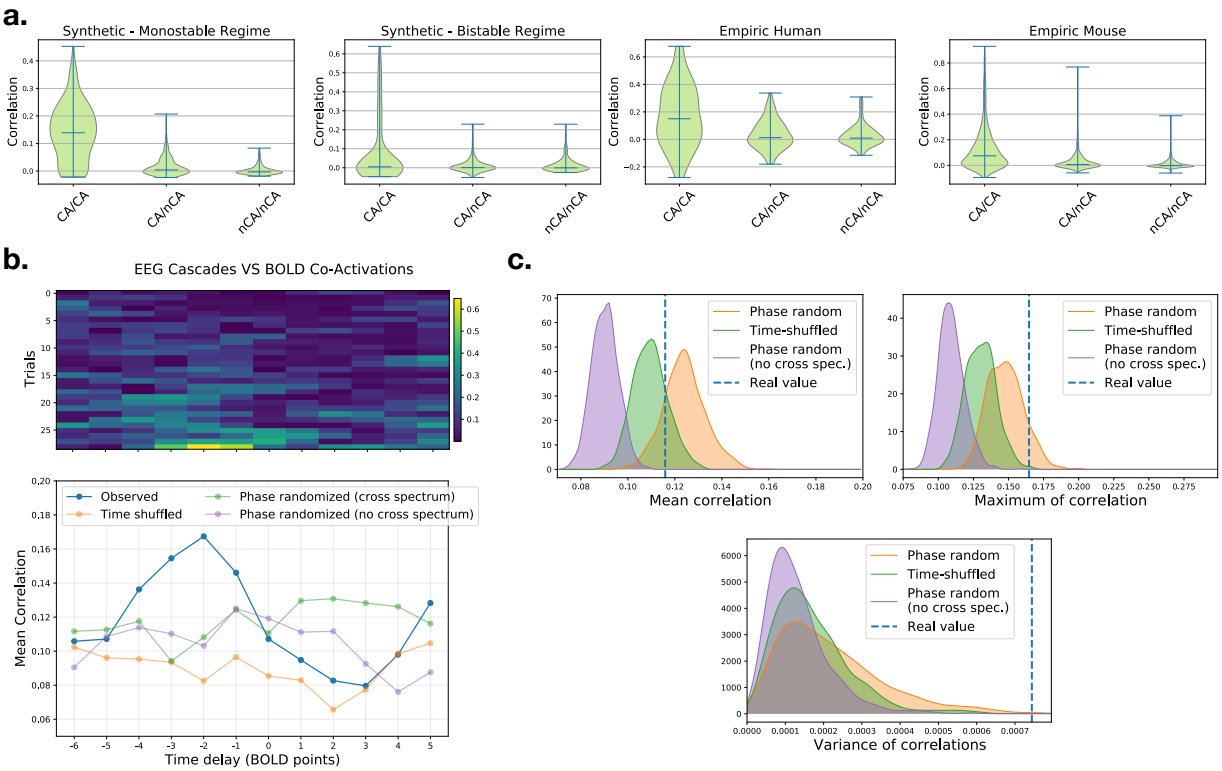

a.
Synthetic - Monostable Regime
Synthetic - Bistable Regime
Empiric Human
Empiric Mouse
Correlation
CA/CA
CA/nCA
nCA/nCA
b.
EEG Cascades VS BOLD Co-Activations
Trials
Observed
Time shuffled
Phase randomized (cross spectrum)
Phase randomized (no cross spectrum)
Mean Correlation
Time delay (BOLD points)
c.
Phase random
Time-shuffled
Phase random (no cross spec.)
Real value
Mean correlation
Maximum of correlation
Variance of correlations

Figure 4:

Figure 4: (a) The largest BOLD co-activations events (CA, above the 98th percentile of the RSS, green line in Figure 3c) are distinguished from *non-events* (nCA, below threshold). We report the synthetic and empiric correlations between iFCs at times within CA events, between CA events and non-events nCA, within non-events nCA. These correlations are by definition the off-diagonal values of the dFC_e matrix (see Figure 1f-g). The distribution of the correlations within events (CA/CA) is wider and explains the greatest off-diagonal correlation values of the dFC_e across all the synthetic and empiric datasets. In both empiric and simulated data sets, the neuronal cascades seem to precede in most cases the BOLD-CA events (see Figure 3). We quantify this correspondence in the Human EEG/fMRI datasets by evaluating the correlation between the cascade magnitude and the BOLD-CA amplitude (blue and green lines in Figure 3) for different time lags in several (EEG/fMRI) trials extracted from a Human cohort of resting subjects (b, top). Negative lag is associated with a shift backward of the BOLD signal. (b,bottom) The cross-correlation averaged across trials shows a clear trend (blue line). In particular, the peak of correlation at lag -2 sampling points ($1pt = 1.94s$), as well as the rapid fall for positive lags, confirms that the EEG precedes the BOLD activity by few seconds. The same profile is evaluated by comparing the largest cascades with the BOLD-CA signals extracted from 1000 time-shuffled (example in green), 1000 phase-randomized (cross-spectrum preserved, example in orange) and 1000 phase-randomized (cross-spectrum not preserved, example in red) BOLD surrogates for every subject. (c) The distribution of the mean, maximum and variance of the cross-correlations for each surrogate model is displayed and compared to the empiric values. In particular, the variance plot shows a clear significance of the results.

3 Discussion

Cognitive function requires the coordination of neural activity across many scales, from neurons and circuits to large-scale networks. As such, it is unlikely that an explanatory framework focused upon any single scale will yield a comprehensive theory of brain activity, cognitive function or brain disease. Fast time scales evolve within the quasi-stationary constraints imposed by slow time scales. As the latter change, the organization of the former must change consequently. We demonstrated this principle along the organization of the fast neuroelectric correlates of the slowly evolving large-scale dynamic Functional Connectivity (dFC) patterns for two mammalian species. Large neuronal cascades, defined by collective deviations of the simulated EEG signals away from baseline activity, can emerge as co-activation (CA) BOLD events, i.e. large simultaneous activation at pairs of regions across the brain network. In turn, the intermittent set of BOLD-CA events underlies specific stable FC epochs. Thus, large neuronal cascades have an organizational role in the resting state brain dynamics.

In this work, we have adopted a theory driven approach to reproduce qualitatively the resting state FC, while keeping track of the electric neuronal activity [31]. To address multiscale-phenomena, we maintain the co-existence of microscopic and macroscopic scales of organization. We have constructed a model that maps the activity of an infinite number of coupled neurons, each described by a simple phase-oscillator equation, into the mean field firing rate and membrane potential variables. Future applications allow the possibility of including synaptic dynamics [41, 42], adaptation [43], excitatory versus inhibitory populations [31, 44, 45], among others. Although not giving full access to biophysical processes, as when using detailed single neuron models, the model we used carries a low computational load, enabling the simulation of fMRI-like sessions in TVB while preserving important information about the microscopic neuronal organization (see Methods).

While the debate over the actual nature of functional connectivity dynamics is still open [14, 46, 47], evidence was provided in favor of the dFC as a legitimate measure of the functional evolution. In fact, the temporal unfolding of the FC displays transient epochs of slowdowns and accelerations. The distribution of the FC 'speeds' lies between a completely ordered scenario and a random one, a hallmark of complexity [23]. We found that a large part of the functional information is condensed in large BOLD-CA events, in keeping with previous results [48] (see Figure S4c). The occurrence of the BOLD-CA events cannot be ascribed to the spectral properties of the BOLD signals nor to motion artifacts [48], which supports a genuine phenomenon. Our results provide an *in silico* support to a bursts-based system dynamics, which might be a general property for a class of high-dimensional systems organized at large spatial and slow temporal scales. Indeed, the dynamics is driven by short activation peaks, rather than by continuous processes, in accordance with previous works showing that the relevant information about the major resting state networks is carried by the largest fluctuations of the fMRI blood oxygenation level-dependent (BOLD) signals [49, 50, 51, 52, 53].

Varying the global couplings in the model, we discovered two regions of the parameter space–defining the monostable and the bistable regimes–where the large-scale organization is qualitatively different e.g., in terms of the statistical properties of the BOLD-CA events (Figure 4a) or in the life-time of the stable functional states (i.e., the size of the dFC blocks in Figure 3c, bottom). The generation of a variety of dynamic regimes from changes in the global 'environmental' parameters is biologically realistic, and is in keeping with the degeneracy principle that characterizes the activity of neuronal networks [54]. For example, sleep deprivation is associated with changes in the speed of the FC evolution [55]. Accordingly, distinct features of the dFC should correspond to a different underlying neuroelectric organization (and perhaps to distinct mechanisms of communication across the brain regions [56]). The transition between wakefulness and sleep brain states offers an example of sudden change of the whole-brain dynamics associated

with the microscopic neuroelectric re-organization [57, 58, 59, 60]. In our model, the neuroelectric correlates of the monostable and the bistable regimes define distinct pathways towards a large-scale dynamic organization (Figure 1h). Further exploration of the parameter space may reveal new mechanisms of pattern formation.

During the BOLD-CA events the evolution of the FC slows down, as shown by the high correlation between functional states at consecutive times (Figure 2c and Figure 3c). A slowing down of the dynamics could be associated to longer absorption times of the largest BOLD fluctuations. An emergent stable FC elicited by a BOLD-CA event can persist for several seconds. The same state can re-appear after a few minutes in coincidence with another CA. Therefore, the spontaneous bursts of co-activation at pairs of nodes across the brain highlight special sets of network edges. How the system selects these preferential sets and what triggers the sudden co-fluctuations are open questions. We suggest that large neuronal cascades may play a central role in determining fluctuations in resting, perceptual and cognitive processes. These findings highlight the strong interdependence between the multi-scale neural processes.

In many dynamical systems, slowdowns and large-scale events are typically observed at the transition between different states (or phases) i.e., at the critical point. Critical dynamic systems are characterized by specific properties, including the presence of fluctuations at all spatiotemporal scales [61]. Historically, the hallmark of a critical organization of the brain activity came from the observation of neuronal *avalanches* whose size and duration fit a power-law [40] (for a 'critical' look at brain criticality see [62]). This pioneering work provided a proof-of-principle that the science of criticality could be used to inform our understanding of complex patterns of activity in the brain and, by extension, behaviour, even though, with few notable exceptions [63, 64], these two streams of research have largely evolved in parallel. Establishing the presence of critical instabilities in neurological disorders opens a number of novel therapeutic and diagnostic windows (for review, see [65]). Schizophrenia for instance is associated with less suppression

of activity in the default-mode network when subjects engage in external tasks [66, 67, 68, 69]. These examples remain sparse and often disconnected across domains, which renders the need to an overarching framework across scales important. Our model makes a constructive contribution in this sense. Avalanches-like activity is observed across orders of magnitude by analyzing the simulated EEG signals (Supplementary Figure S3a-b). Importantly, both regimes feature rare large-scale events. In order to compare the fast EEG and the slower BOLD signals, we focused on the slow unfolding of the EEG deviations from baseline activity. In other words, instead of focusing on the size and duration of the neuronal avalanches, we analyzed their collective occurrence at the resolution of seconds, which we refer to as neuronal cascades (Figure S3c). We showed that the cascade magnitude is correlated to the BOLD co-activation amplitude (Figure 4b). In particular, in the empiric human dataset we observe a rapid drop of this correlation when the BOLD activity is shifted backward w.r.t. the EEG signal. This fact makes our result more robust, since we expect a change of the BOLD activity to happen *after* a sustained neuroelectric activity.

Notice that in the empiric data set, not all the neuronal cascades give rise to BOLD CAs, and not all the BOLD CAs are preceded by large neuronal cascades. In general we should expect a non-trivial interaction between these phenomena and other infraslow processes. For example, a growing corpus of evidences relates the slow functional and neuroelectric dynamics with other physiological and cognitive infraslow processes such as neuromodulation [70], visceral (heart and gut) inputs [71], and cognitive performance [72].

In the cases in which neuronal cascades and BOLD CA events co-occur, one could hypothesize that the stable FC state evoked by a CA event is related to the specific structural channels in which the neuronal cascade spreads. In fact, we showed in the simulated data that the largest neuronal cascades underlie specific functional patterns (Figure 2), depending on the location of the avalanche onset. Also, the observation that the same avalanche appears at regular intervals

(see for example the three peaks at second ~ 100, ~ 200 and ~ 300 in Figure 2b) suggests that the re-occurrence of the same stable patterns in time could be traced back to the increased probability of the same cascade to occur following a first large event, in the same way an after-shock follows the main earthquake. More generally, aspects of the functional dynamics phenomenology might be be inherited by the (seemingly) critical properties of the neuronal activity. Traces of criticality were in fact observed also in the hemodynamic activity [49]. The above hypotheses require further exploration on larger empiric data sets. In the model, we have shown that the mechanistic origin of the largest avalanches can be understood ultimately by the interplay between the local dynamics, structural constraints and stochastic events.

The present work provides a new understanding of whole-brain functional dynamics, which is shaped by neuronal cascades giving rise to large BOLD co-activation. Since most neurological disorders are characterized by complex reorganizations at the neuronal scale, it will be interesting to determine whether specific alterations in neuronal cascades sign specific neurological disorders, and explain the already identified alterations in whole brain dynamics.

4 Materials and Methods

4.1 Empirical Mice and Human data sets

The connectome used for the simulations was extracted using The Virtual Mouse Brain pipeline described in [29], which processes the tracer experiments performed at the Allen Institute [32]. There, adult male C57Bl/6J mice are injected in specific brain regions with recombinant adeno-associated virus, which expresses the EGFP anterograde tracer. In each experiment, the tracer migration signal is detected by a serial two-photon tomography system. The anterograde tracing provides information about the axonal projections starting at the injected site. We define the injection density of the source region as the number of infected pixels normalized by the total number of pixels in the region. Similarly, the projection density is defined in every region as the number of pixels detected in the target region following an injection at the source, normalized by the total number of pixels belonging to the target region. The tracer-based connectome is built by averaging over injection experiments performed in the right brain areas and targeting regions in ipsilateral and contralateral hemispheres. Through the Allen Connectivity Builder interface in TVB we parceled the brain in 104 anatomical regions of interest (ROIs, table in Figure S2a). Then, we defined the connection strength between source region i and target region j i.e., the link ij, as the ratio between the projection density at j and the injection density at i. The tracer connectome, with links normalized between 0 and 1 is shown in Figure 1e.

The empirical fMRI mice dataset was imported from the publicly available collection [73]. For our analysis we used the cohort of 20 control wild-type (resting) animals registered in the sub-dataset [74]. The data was previously preprocessed according to a common pipeline, and registered on the Allen volumetric atlas. In our analysis we considered the activity in those voxels corresponding to the imported Allen connectome (Table in Figure S2a). The empiric dataset did not distinguish between specific parts (e.g.,ventral and dorsal) of certain brain regions (blue and

red pairs in Table in Figure S2a). A unique time series was associated to each of these pairs of regions.

The empirical EEG/fMRI dataset was acquired and preprocessed at Charité - Universitätsmedizin Berlin [75] and made available in online repository[1] [76]. In summary, from a larger cohort (49 subjects, 18-80 years), 15 youngest subjects (18-31 years) were selected based on the quality of the EEG recording after correction of the MR artifacts. For each subject, diffusion-weighted MRI, T1-weighted MRI, and EEG-fMRI in resting state paradigm were obtained. The T1-weighted image was used to parcellate the cortical gray matter into 68 regions according to the Desikan-Killiany atlas [77]. This definition of regions was then used to estimate the structural connectivity from the dw-MRI data [78], and to extract the regional average time series from the fMRI data. The EEG data (Easy-cap; 64 channels, MR compatible) was treated with a high pass filter at 1.0 Hz followed with MR imaging acquisition artefact removal using Analyser 2.0 (v2.0.2.5859, Brain Products, Gilching, Germany). Resulting sensor-level time series was downsampled to 200 Hz and low pass filtered to 30 Hz before correction for physiological artifacts (ballistocardiogram, muscle activity). Next, EEG source imaging was performed to obtain projected activity on the cortical surface, and averaged within the 68 regions of the Desikan-Killiany parcellation. See [75] for detailed description of both fMRI and EEG processing. After the preprocessing, we made a quality check of every subject data and excluded BOLD points and associated EEG time windows (BOLD TR = 1.94 s = 388 EEG samples) which presented residual artifacts (at either EEG or fMRI level) based on the following criteria:

- if any of the the 6 time series of the motion degrees of freedom (resulting from the FSL MCFLIRT head movement correction step) presented a peak.
- if the EEG window contained fluctuations simultaneously affecting most of the frequencies (time-frequency analysis).

[1] https://osf.io/mndt8/

- if the EEG presented fluctuations above 7 standard deviations.

We finally selected 30 artifact-free trials of consecutive EEG/fMRI acquisition (minimum duration 2 minutes) across the cohort.

4.2 Neural Mass Model

The dynamics of a brain region of interest (i.e., a node in the network) is described by a Neural Mass Model (NMM) derived analytically as the limit of infinitely all-to-all coupled θ-neuron phase oscillators, whose dynamics is equivalent to that of quadratic integrate-and-fire (QIF) neurons (Figure 1a) [31]. The i-th neuron is described by the equation

$$\dot{V}_i(t) = V_i^2(t) + \eta_i + J\bar{s}(t) + I(t) \quad \text{if} \quad V_i \geq V_p \quad \text{then} \quad V_r \leftarrow V_i. \tag{1}$$

As the membrane potential reaches a peak value V_p, it resets to V_r. The limit $V_p = -V_r \to \infty$ is considered. The parameter η_i is the neuron excitability, and it enters as an heterogeneous current in the membrane potential. A time-dependent current $I(t)$ homogeneously affects all the neurons in the neural mass. The synaptic activation $s(t)$ of a single neuron respects the equation $Q \cdot s(t) = \sum_k \delta(t - t_k)$, being $Q = (1 - \tau \frac{\mathrm{d}}{\mathrm{d}t})$ and t_k the arrival time of the presynaptic action potential. The neuronal coupling among N_{QIF} neurons inside a neural mass is given by the average synaptic activation $\bar{s}(t) = \frac{1}{N_{QIF}} \sum_{j>1}^{N_{QIF}} s(t)$ scaled by the synaptic weight J. Assuming a Lorentzian distribution of the membrane potentials across the coupled neuronal population it is possible to perform the $N_{QIF} \to \infty$ limit exactly according to the Ott-Antonsen ansatz [79]. Also the heterogeneous currents are distributed according to a Lorentzian distribution with width at mid height Δ and peak location η. After the limit we can describe the activity in a neural mass n in terms of the average membrane potential $V_n(t)$ and the average firing rate $r_n(t)$ (which corresponds to $\bar{s}(t)$ in the limit $\tau \to 0$). The dynamics associated to a local network node

is then described by the following NMM equations

$$\dot{r}_n(t) = \frac{\Delta}{\pi} + 2r_n(t)V_n(t), \tag{2}$$

$$\dot{V}_n(t) = V_n^2(t) + \eta + Jr_n(t) - \pi^2 r_n^2(t) + I(t). \tag{3}$$

The synaptic weight $J = 14.5$, the average neuronal excitability $\eta = -4.6$, and the spread of the heterogeneous noise distribution $\Delta = 0.7$, are homogeneously tuned so that each decoupled node is in a bistable regime with a 'down' fixed point and an 'up' stable focus in the 2D phase space (Figure 1b, $I_{ext} = 0$). Differently from standard NMMs, the chosen model keeps track of the internal amount of synchronization of the constituent neurons by preserving the analytic form of the Kuramoto parameter [80]

$$Z_n(t) = |\sum_i \exp(i\theta_i)| = K_n(t)\exp(i\theta_n(t)), \tag{4}$$

during the thermodynamic limit [2]. Here, K_n and θ_n are the average amount of synchronization and the average phase of the neurons in the node n. The real and imaginary components of Z_n are connected to the average membrane potential V_n and firing rate r_n of the neural population n via the relation

$$Z_n = \frac{1 - W_n^\star}{1 + W_n^\star}, \qquad W_n \equiv \pi r_n + iV_n. \tag{5}$$

4.3 Connectome based large-scale Network Modeling

The local model is then coupled over an empirically extracted connectome (Figure 1d). The coupling term enters as an additive current in the average membrane potential equations (see also Figure 1c)

$$I(t) = G \sum_{m \neq n} W_{nm} r_m(t - \tau_{nm}) \tag{6}$$

The global coupling parameter G sets the impact of the connectivity matrix W_{nm} over the local dynamics. The time delay $\tau_{nm} = l_{nm}/v$ is given by the tract length between nodes n and

[2] The θ_i-representation of the QIF equations can be achieved from (1) by a change of variables $V_i = \tan(\theta_i/2)$

m divided by the average conduction speed which we set to a realistic value $v = 4\ m/s$. Since the firing rate is by definition greater than or equal to zero, all the interactions are excitatory.

The numeric integration of the nodal equations over the network is performed using The Virtual Brain (TVB) open access platform [30]. The solution of the coupled system consist of a neuroelectric raw dataset describing the evolution of the variables $(r_n(t), V_n(t))$ in each region n of the connectome. These variables are our measure of the microscopic activity. The sampling rate of these neuroelectric variables is set to 1000Hz. The surrogate BOLD activity $B_n(t)$ in each region is derived by filtering the membrane potential with the Balloon-Windkessel model [33] (default values implemented in TVB). We use a repetition time of 2 seconds so that the BOLD rate is 0.5 Hz.

4.4 Time dependent functional connectivity

To quantify the temporal evolution of the brain connectivity, we have employed two approaches: windowed dynamical functional connectivity (dFC_w), and edge dynamical functional connectivity (dFC_e). Let us denote $B_n(t)$ as the regional BOLD time-series for each node $n = 1 \dots N$. To compute the dFC_w [12, 23], we first obtain the series of functional connectivity matrices $FC(w)$ at each sliding window $w = 1 \dots W$, defined as the correlation matrices for the segments $B_n(t)|_{t_w}^{t_w+\tau}$ (in the example Figure S1a we fix the window size $\tau = 60s$ and window step $2s$). Next we correlate the vectorized upper triangular parts of the $FC(w)$ matrices at different time windows to obtain the dFC_w matrix

$$dFC_w(w_i, w_j) = corr\left[triu(FC(w_i), triu(FC(w_j))\right]. \tag{7}$$

On the other hand, the computation of the dFC_e starts with z-scoring the BOLD time series $z_n = \frac{B_n - \mu_n}{\sigma_n}$. The edge time series is then computed as element wise multiplication along time for each pair of regions $E_{nm}(t) = z_n(t) z_m(t)$ for $n, m = 1 \dots N$ (see Figure 1C). Next, for each

pair of time points t_1, t_2 we compute the correlation of the edge vectors as

$$dFC_e(t_i, t_j) = corr(E_{nm}(t_i), E_{nm}(t_j)). \tag{8}$$

While both dFC variants are similar, the main difference lies in the scope of z-scoring of the time series B_n. In the case of dFC_e the z-score z_n is computed from the whole time series, whereas in the dFC_w the z-score is performed within the Pearson correlation in each time window $[t_w, t_w + \tau]$ separately. To check the correspondence of the dFC_e and dFC_w matrices, we have also computed a hybrid dFC from the edge time series E_{nm}. Here we select the segments $w = 1 \ldots W$ of the edge time series $E_{nm}(t)$ for $t_w \leq t \leq t_w + \tau$. Next, to keep the global scope of the z-scoring, we compute the $FC'(w)$ as a sum $FC'(w) = \sum_{t_w}^{t_w+\tau} E_{nm}(t)$. Lastly, we compute the hybrid dFC matrix from the $FC'(w)$ vectors as

$$dFC_h(w_i, w_j) = corr\left[FC'(w_i), FC'(w_j)\right]. \tag{9}$$

See Figure S1 for a schema illustrating the computation of all three variants of the dFC, and for comparison of the dFC variants in empirical and simulated data.

4.5 Surrogate BOLD models

In order to compare our results in Figure 4b with the null-hypotheses of a random evolution and of inter-regional stationarity of the FC, we build time-shuffled and phase-randomized surrogates of the FC dynamics, respectively. The time-shuffled surrogate is obtained by randomizing the order of the instantaneous functional connectivities $iFC(t)$ i.e., the columns of the edge CA time series $E_{ij}(t)$ (see Figure 1C). According to [14] the phase randomized surrogate is obtained by adding a uniformly distributed random phase to each frequency of the Fourier transformed signal, and subsequently retrieving the phase randomized signals by anti-Fourier transform. Importantly, the phase shift is different at every time point but can be applied uniformly to all the brain regions (cross-spectrum preserved) or separately in every region (cross-spectrum not preserved).

Only in the first case the bursts of co-activation are not destroyed but shifted (Figure S4c) and the static FC is preserved. The coherent fluctuations around the stationary FC, however, are destroyed. For more details see also [81].

References

[1] Jonathan D Power, Alexander L Cohen, Steven M Nelson, Gagan S Wig, Kelly Anne Barnes, Jessica A Church, Alecia C Vogel, Timothy O Laumann, Fran M Miezin, Bradley L Schlaggar, et al. Functional network organization of the human brain. *Neuron*, 72(4):665–678, 2011.

[2] Michael D Fox and Marcus E Raichle. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature reviews neuroscience*, 8(9):700–711, 2007.

[3] Justin L Vincent, Gaurav H Patel, Michael D Fox, Abraham Z Snyder, Justin T Baker, David C Van Essen, John M Zempel, Lawrence H Snyder, Maurizio Corbetta, and Marcus E Raichle. Intrinsic functional architecture in the anaesthetized monkey brain. *Nature*, 447(7140):83–86, 2007.

[4] Hanbing Lu, Qihong Zou, Hong Gu, Marcus E Raichle, Elliot A Stein, and Yihong Yang. Rat brains also have a default mode network. *Proceedings of the National Academy of Sciences*, 109(10):3979–3984, 2012.

[5] Jaymin Upadhyay, Scott J Baker, Prasant Chandran, Loan Miller, Younglim Lee, Gerard J Marek, Unal Sakoglu, Chih-Liang Chin, Feng Luo, Gerard B Fox, et al. Default-mode-like network activation in awake rodents. *PloS one*, 6(11):e27839, 2011.

[6] Francesco Sforazzini, Adam J Schwarz, Alberto Galbusera, Angelo Bifone, and Alessandro Gozzi. Distributed bold and cbv-weighted resting-state networks in the mouse brain. *Neuroimage*, 87:403–415, 2014.

[7] James M Stafford, Benjamin R Jarrett, Oscar Miranda-Dominguez, Brian D Mills, Nicholas Cain, Stefan Mihalas, Garet P Lahvis, K Matthew Lattal, Suzanne H Mitchell, Stephen V

David, et al. Large-scale topology and the default mode network in the mouse connectome. *Proceedings of the National Academy of Sciences*, 111(52):18745–18750, 2014.

[8] Alessandro Gozzi and Adam J Schwarz. Large-scale functional connectivity networks in the rodent brain. *Neuroimage*, 127:496–509, 2016.

[9] Joanes Grandjean, Carola Canella, Cynthia Anckaerts, Gülebru Ayrancı, Salma Bougacha, Thomas Bienert, David Buehlmann, Ludovico Coletta, Daniel Gallino, Natalia Gass, et al. Common functional networks in the mouse brain revealed by multi-centre resting-state fmri analysis. *Neuroimage*, 205:116278, 2020.

[10] R Matthew Hutchison, Thilo Womelsdorf, Elena A Allen, Peter A Bandettini, Vince D Calhoun, Maurizio Corbetta, Stefania Della Penna, Jeff H Duyn, Gary H Glover, Javier Gonzalez-Castillo, et al. Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage*, 80:360–378, 2013.

[11] Enzo Tagliazucchi, Frederic Von Wegner, Astrid Morzelewski, Verena Brodbeck, and Helmut Laufs. Dynamic bold functional connectivity in humans and its electrophysiological correlates. *Frontiers in human neuroscience*, 6:339, 2012.

[12] Elena A Allen, Eswar Damaraju, Sergey M Plis, Erik B Erhardt, Tom Eichele, and Vince D Calhoun. Tracking whole-brain connectivity dynamics in the resting state. *Cerebral cortex*, 24(3):663–676, 2014.

[13] Jian Qin, Shan-Guang Chen, Dewen Hu, Ling-Li Zeng, Yi-Ming Fan, Xiao-Ping Chen, and Hui Shen. Predicting individual brain maturity using dynamic functional connectivity. *Frontiers in human neuroscience*, 9:418, 2015.

[14] Rikkert Hindriks, Mohit H Adhikari, Yusuke Murayama, Marco Ganzetti, Dante Mantini,

Nikos K Logothetis, and Gustavo Deco. Can sliding-window correlations reveal dynamic functional connectivity in resting-state fmri? *Neuroimage*, 127:242–256, 2016.

[15] Maria Giulia Preti and Dimitri Van De Ville. Dynamics of functional connectivity at high spatial resolution reveal long-range interactions and fine-scale organization. *Scientific reports*, 7(1):1–12, 2017.

[16] Joanes Grandjean, Maria Giulia Preti, Thomas AW Bolton, Michaela Buerge, Erich Seifritz, Christopher R Pryce, Dimitri Van De Ville, and Markus Rudin. Dynamic reorganization of intrinsic functional networks in the mouse brain. *Neuroimage*, 152:497–508, 2017.

[17] Javier Gonzalez-Castillo and Peter A Bandettini. Task-based dynamic functional connectivity: Recent findings and open questions. *Neuroimage*, 180:526–533, 2018.

[18] Eswar Damaraju, Elena A Allen, Aysenil Belger, Judith M Ford, S McEwen, DH Mathalon, BA Mueller, GD Pearlson, SG Potkin, A Preda, et al. Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia. *NeuroImage: Clinical*, 5:298–308, 2014.

[19] Barnaly Rashid, Eswar Damaraju, Godfrey D Pearlson, and Vince D Calhoun. Dynamic connectivity states estimated from resting fmri identify differences among schizophrenia, bipolar disorder, and healthy control subjects. *Frontiers in human neuroscience*, 8:897, 2014.

[20] Jianpo Su, Hui Shen, Ling-Li Zeng, Jian Qin, Zhening Liu, and Dewen Hu. Heredity characteristics of schizophrenia shown by dynamic functional connectivity analysis of resting-state functional mri scans of unaffected siblings. *Neuroreport*, 27(11):843–848, 2016.

[21] Yuhui Du, Godfrey D Pearlson, Dongdong Lin, Jing Sui, Jiayu Chen, Mustafa Salman, Carol A Tamminga, Elena I Ivleva, John A Sweeney, Matcheri S Keshavan, et al. Identifying

dynamic functional connectivity biomarkers using gig-ica: Application to schizophrenia, schizoaffective disorder, and psychotic bipolar disorder. *Human brain mapping*, 38(5):2683–2708, 2017.

[22] Maria Giulia Preti, Thomas AW Bolton, and Dimitri Van De Ville. The dynamic functional connectome: State-of-the-art and perspectives. *Neuroimage*, 160:41–54, 2017.

[23] Demian Battaglia, Thomas Boudou, Enrique C.A. Hansen, Diego Lombardo, Sabrina Chettouf, Andreas Daffertshofer, Anthony R. McIntosh, Joelle Zimmermann, Petra Ritter, and Viktor Jirsa. Dynamic functional connectivity between order and randomness and its evolution across the human adult lifespan. *NeuroImage*, 222:117156, 2020.

[24] Adrián Ponce-Alvarez, Gustavo Deco, Patric Hagmann, Gian Luca Romani, Dante Mantini, and Maurizio Corbetta. Resting-state temporal synchronization networks emerge from connectivity topology and heterogeneity. *PLoS Comput Biol*, 11(2):e1004100, 2015.

[25] Anandamohan Ghosh, Y Rho, Anthony Randal McIntosh, Rolf Kötter, and Viktor K Jirsa. Noise during rest enables the exploration of the brain's dynamic repertoire. *PLoS Comput Biol*, 4(10):e1000196, 2008.

[26] Gustavo Deco and Viktor K Jirsa. Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. *Journal of Neuroscience*, 32(10):3366–3375, 2012.

[27] Viktor K Jirsa, T Proix, Dionysios Perdikis, Michael Marmaduke Woodman, H Wang, J Gonzalez-Martinez, Christophe Bernard, Christian Bénar, Maxime Guye, Patrick Chauvel, et al. The virtual epileptic patient: individualized whole-brain models of epilepsy spread. *Neuroimage*, 145:377–388, 2017.

[28] Paula Sanz-Leon, Stuart A Knock, Andreas Spiegler, and Viktor K Jirsa. Mathematical

framework for large-scale brain network modeling in the virtual brain. *Neuroimage*, 111:385–430, 2015.

[29] Francesca Melozzi, Marmaduke M Woodman, Viktor K Jirsa, and Christophe Bernard. The virtual mouse brain: a computational neuroinformatics platform to study whole mouse brain dynamics. *Eneuro*, 4(3), 2017.

[30] Paula Sanz Leon, Stuart A Knock, M Marmaduke Woodman, Lia Domide, Jochen Mersmann, Anthony R McIntosh, and Viktor Jirsa. The virtual brain: a simulator of primate brain network dynamics. *Frontiers in neuroinformatics*, 7:10, 2013.

[31] Ernest Montbrió, Diego Pazó, and Alex Roxin. Macroscopic description for networks of spiking neurons. *Physical Review X*, 5(2):021028, 2015.

[32] Seung Wook Oh, Julie A Harris, Lydia Ng, Brent Winslow, Nicholas Cain, Stefan Mihalas, Quanxin Wang, Chris Lau, Leonard Kuan, Alex M Henry, et al. A mesoscale connectome of the mouse brain. *Nature*, 508(7495):207–214, 2014.

[33] Karl J Friston, Andrea Mechelli, Robert Turner, and Cathy J Price. Nonlinear responses in fmri: the balloon model, volterra kernels, and other hemodynamics. *NeuroImage*, 12(4):466–477, 2000.

[34] Waqas Majeed, Matthew Magnuson, Wendy Hasenkamp, Hillary Schwarb, Eric H Schumacher, Lawrence Barsalou, and Shella D Keilholz. Spatiotemporal dynamics of low frequency bold fluctuations in rats and humans. *Neuroimage*, 54(2):1140–1150, 2011.

[35] Stephen M Smith, Karla L Miller, Steen Moeller, Junqian Xu, Edward J Auerbach, Mark W Woolrich, Christian F Beckmann, Mark Jenkinson, Jesper Andersson, Matthew F Glasser, et al. Temporally-independent functional modes of spontaneous brain activity. *Proceedings of the National Academy of Sciences*, 109(8):3131–3136, 2012.

[36] Martin A Lindquist, Yuting Xu, Mary Beth Nebel, and Brain S Caffo. Evaluating dynamic bivariate correlations in resting-state fmri: a comparison study and a new approach. *NeuroImage*, 101:531–546, 2014.

[37] Fikret Işik Karahanoğlu and Dimitri Van De Ville. Transient brain activity disentangles fmri resting-state dynamics in terms of spatially and temporally overlapping networks. *Nature communications*, 6(1):1–10, 2015.

[38] James M Shine, Oluwasanmi Koyejo, Peter T Bell, Krzysztof J Gorgolewski, Moran Gilat, and Russell A Poldrack. Estimation of dynamic functional connectivity using multiplication of temporal derivatives. *NeuroImage*, 122:399–407, 2015.

[39] Joshua Faskowitz, Farnaz Zamani Esfahlani, Youngheun Jo, Olaf Sporns, and Richard F. Betzel. Edge-centric functional network representations of human cerebral cortex reveal overlapping system-level architecture. *Nature Neuroscience*, October 2020.

[40] John M Beggs and Dietmar Plenz. Neuronal avalanches in neocortical circuits. *Journal of neuroscience*, 23(35):11167–11177, 2003.

[41] Stephen Coombes and Áine Byrne. Next generation neural mass models. In *Nonlinear Dynamics in Computational Neuroscience*, pages 1–16. Springer, 2019.

[42] Halgurd Taher, Alessandro Torcini, and Simona Olmi. Exact neural mass model for synaptic-based working memory. *arXiv preprint arXiv:2010.07071*, 2020.

[43] Richard Gast, Helmut Schmidt, and Thomas R Knösche. A mean-field description of bursting dynamics in spiking neural networks with short-term adaptation. *Neural Computation*, 32(9):1615–1634, 2020.

[44] Carlo R Laing. Phase oscillator network models of brain dynamics. *Computational models of brain and behavior*, pages 505–517, 2017.

[45] Grégory Dumont and Boris Gutkin. Macroscopic phase resetting-curves determine oscillatory coherence and signal transfer in inter-coupled neural circuits. *PLOS Computational Biology*, 15(5):e1007019, May 2019.

[46] Timothy O Laumann, Abraham Z Snyder, Anish Mitra, Evan M Gordon, Caterina Gratton, Babatunde Adeyemo, Adrian W Gilmore, Steven M Nelson, Jeff J Berg, Deanna J Greene, et al. On the stability of bold fmri correlations. *Cerebral cortex*, 27(10):4719–4732, 2017.

[47] Raphael Liegeois, Timothy O Laumann, Abraham Z Snyder, Juan Zhou, and BT Thomas Yeo. Interpreting temporal fluctuations in resting-state functional connectivity mri. *Neuroimage*, 163:437–455, 2017.

[48] Farnaz Zamani Esfahlani, Youngheun Jo, Joshua Faskowitz, Lisa Byrge, Daniel P. Kennedy, Olaf Sporns, and Richard F. Betzel. High-amplitude cofluctuations in cortical activity drive functional connectivity. *Proceedings of the National Academy of Sciences*, 2020.

[49] Enzo Tagliazucchi, Pablo Balenzuela, Daniel Fraiman, and Dante R Chialvo. Criticality in large-scale brain fmri dynamics unveiled by a novel point process analysis. *Frontiers in physiology*, 3:15, 2012.

[50] Xiao Liu and Jeff H Duyn. Time-varying functional network information extracted from brief instances of spontaneous brain activity. *Proceedings of the National Academy of Sciences*, 110(11):4392–4397, 2013.

[51] Thomas W Allan, Susan T Francis, Cesar Caballero-Gaudes, Peter G Morris, Elizabeth B Liddle, Peter F Liddle, Matthew J Brookes, and Penny A Gowland. Functional connectivity in mri is driven by spontaneous bold events. *PloS one*, 10(4):e0124577, 2015.

[52] Ignacio Cifre, Mahdi Zarepour, Silvina G Horovitz, Sergio Cannas, and Dante R Chialvo.

On why a few points suffice to describe spatiotemporal large-scale brain dynamics. *arXiv preprint arXiv:1707.00759*, 2017.

[53] Daniel Gutierrez-Barragan, M Albert Basson, Stefano Panzeri, and Alessandro Gozzi. Infraslow state fluctuations govern spontaneous fmri network dynamics. *Current Biology*, 29(14):2295–2306, 2019.

[54] Eve Marder and Adam L Taylor. Multiple models to capture the variability in biological neurons and networks. *Nature neuroscience*, 14(2):133–138, 2011.

[55] Diego Lombardo, Catherine Cassé-Perrot, Jean-Philippe Ranjeva, Arnaud Le Troter, Maxime Guye, Jonathan Wirsich, Pierre Payoux, David Bartrés-Faz, Régis Bordet, Jill C Richardson, et al. Modular slowing of resting-state dynamic functional connectivity as a marker of cognitive dysfunction induced by sleep deprivation. *BioRxiv*, 2020.

[56] Demian Battaglia and Andrea Brovelli. Functional connectivity and neuronal dynamics: insights from computational methods. In *The Cognitive Neurosciences*. 2020.

[57] Mélanie Boly, Vincent Perlbarg, Guillaume Marrelec, Manuel Schabus, Steven Laureys, Julien Doyon, Mélanie Pélégrini-Issac, Pierre Maquet, and Habib Benali. Hierarchical clustering of brain activity during human nonrapid eye movement sleep. *Proceedings of the National Academy of Sciences*, 109(15):5856–5861, 2012.

[58] Linda J Larson-Prior, John M Zempel, Tracy S Nolan, Fred W Prior, Abraham Z Snyder, and Marcus E Raichle. Cortical network functional connectivity in the descent to sleep. *Proceedings of the National Academy of Sciences*, 106(11):4489–4494, 2009.

[59] Enzo Tagliazucchi, Nicolas Crossley, Edward T Bullmore, and Helmut Laufs. Deep sleep divides the cortex into opposite modes of anatomical–functional coupling. *Brain Structure and Function*, 221(8):4221–4234, 2016.

[60] Mazen El-Baba, Daniel J Lewis, Zhuo Fang, Adrian M Owen, Stuart M Fogel, and J Bruce Morton. Functional connectivity dynamics slow with descent from wakefulness to sleep. *PloS one*, 14(12):e0224669, 2019.

[61] Luca Cocchi, Leonardo L Gollo, Andrew Zalesky, and Michael Breakspear. Criticality in the brain: A synthesis of neurobiology, models and cognition. *Progress in neurobiology*, 158:132–152, 2017.

[62] John M Beggs and Nicholas Timme. Being critical of criticality in the brain. *Frontiers in physiology*, 3:163, 2012.

[63] Viktor K Jirsa, R Friedrich, Hermann Haken, and JA Scott Kelso. A theoretical model of phase transitions in the human brain. *Biological cybernetics*, 71(1):27–35, 1994.

[64] JAS Kelso, SL Bressler, S Buchanan, GC DeGuzman, MAFAH Ding, A Fuchs, and T Holroyd. A phase transition in human brain and behavior. *Physics Letters A*, 169(3):134–144, 1992.

[65] Stephen Coombes and John R Terry. The dynamics of neurological disease: integrating computational, experimental and clinical neuroscience. *European Journal of Neuroscience*, 36(2):2118–2120, 2012.

[66] Ben J Harrison, Murat Yücel, Jesus Pujol, and Christos Pantelis. Task-induced deactivation of midline cortical regions in schizophrenia assessed with fmri. *Schizophrenia research*, 91(1-3):82–86, 2007.

[67] Andreas Meyer-Lindenberg, Jean-Baptiste Poline, Philip D Kohn, John L Holt, Michael F Egan, Daniel R Weinberger, and Karen Faith Berman. Evidence for abnormal cortical functional connectivity during working memory in schizophrenia. *American Journal of Psychiatry*, 158(11):1809–1817, 2001.

[68] Ayna B Nejad, Bjørn H Ebdrup, Hartwig R Siebner, Hans Rasmussen, Bodil Aggernæs, Birte Y Glenthøj, and William FC Baaré. Impaired temporoparietal deactivation with working memory load in antipsychotic-naive patients with first-episode schizophrenia. *The World Journal of Biological Psychiatry*, 12(4):271–281, 2011.

[69] Susan Whitfield-Gabrieli, Heidi W Thermenos, Snezana Milanovic, Ming T Tsuang, Stephen V Faraone, Robert W McCarley, Martha E Shenton, Alan I Green, Alfonso Nieto-Castanon, Peter LaViolette, et al. Hyperactivity and hyperconnectivity of the default network in schizophrenia and in first-degree relatives of persons with schizophrenia. *Proceedings of the National Academy of Sciences*, 106(4):1279–1284, 2009.

[70] James M Shine, Patrick G Bissett, Peter T Bell, Oluwasanmi Koyejo, Joshua H Balsters, Krzysztof J Gorgolewski, Craig A Moodie, and Russell A Poldrack. The dynamics of functional brain networks: integrated network states during cognitive task performance. *Neuron*, 92(2):544–554, 2016.

[71] Damiano Azzalini, Ignacio Rebollo, and Catherine Tallon-Baudry. Visceral signals shape brain dynamics and cognition. *Trends in Cognitive Sciences*, 23(6):488–509, 2019.

[72] Simo Monto, Satu Palva, Juha Voipio, and J Matias Palva. Very slow eeg fluctuations predict the dynamics of stimulus detection and oscillation amplitudes in humans. *Journal of Neuroscience*, 28(33):8268–8272, 2008.

[73] Joanes Grandjean. A common mouse fmri resource through unified preprocessing, 2020.

[74] Francesca Mandino, Yeow Ling Yun, John Gigg, Malini C. Olivo, and Joanes Grandjean. Mouse rest 3 x TG, 2019.

[75] Michael Schirner, Anthony Randal McIntosh, Viktor Jirsa, Gustavo Deco, and Petra Ritter. Inferring multi-scale neural mechanisms with brain network modelling. *Elife*, 7:e28927, 2018.

[76] Michael Schirner. Hybrid brain model data, Sep 2018.

[77] Rahul S. Desikan, Florent Ségonne, Bruce Fischl, Brian T. Quinn, Bradford C. Dickerson, Deborah Blacker, Randy L. Buckner, Anders M. Dale, R. Paul Maguire, Bradley T. Hyman, Marilyn S. Albert, and Ronald J. Killiany. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, 31(3):968–980, July 2006.

[78] Michael Schirner, Simon Rothmeier, Viktor K Jirsa, Anthony Randal McIntosh, and Petra Ritter. An automated pipeline for constructing personalized virtual brains from multimodal neuroimaging data. *NeuroImage*, 117:343–357, 2015.

[79] Edward Ott and Thomas M Antonsen. Low dimensional behavior of large systems of globally coupled oscillators. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 18(3):037113, 2008.

[80] Yoshiki Kuramoto. *Chemical oscillations, waves, and turbulence.* Courier Corporation, 2003.

[81] Dean Prichard and James Theiler. Generating surrogate data for time series with several simultaneously measured variables. *Physical review letters*, 73(7):951, 1994.

5 Supplementary Figures

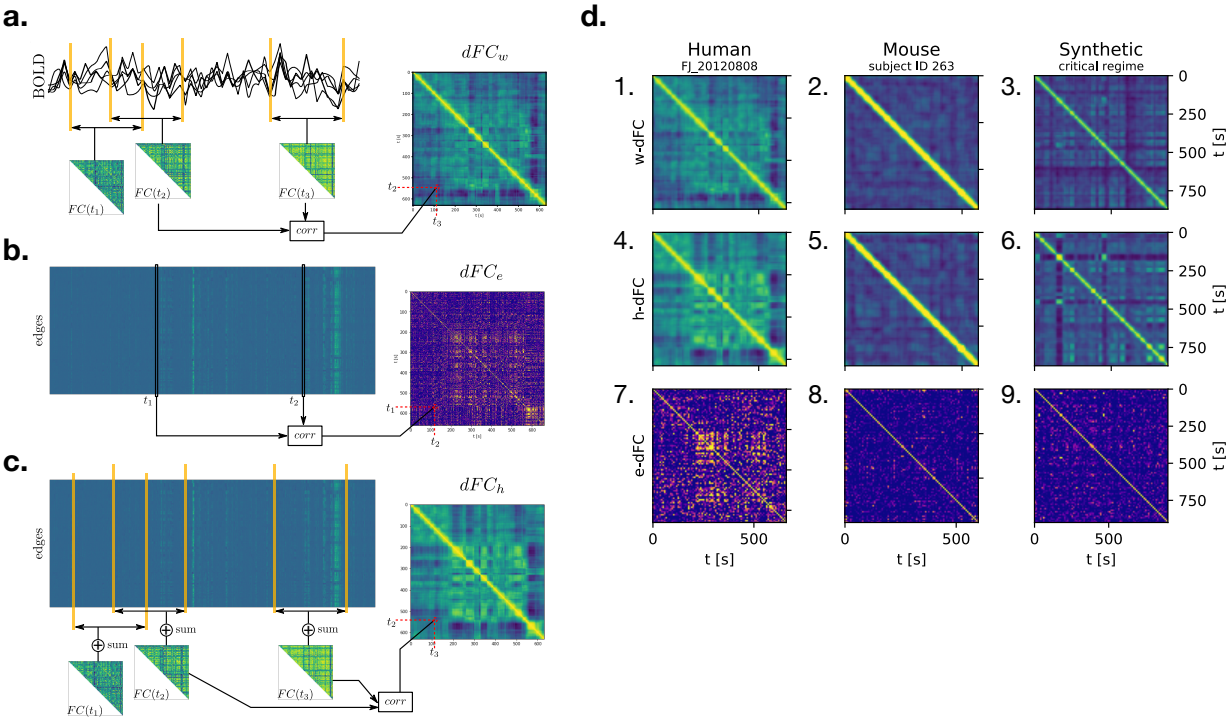

a.
BOLD
dFC_w
$FC(t_1)$
$FC(t_2)$
$FC(t_3)$
corr
b.
edges
dFC_e
c.
dFC_h
sum
d.
Human
FJ_20120808
Mouse
subject ID 263
Synthetic
critical regime
w-dFC
h-dFC
e-dFC
t [s]

Figure S1:

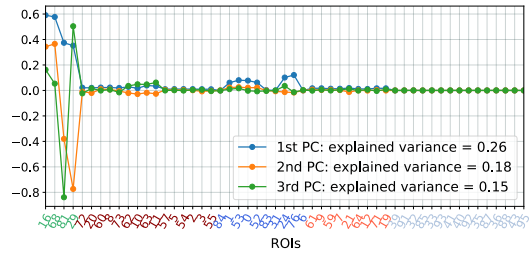
Figure S1: The workflows for three variants of computing dFC. (A) Windowed dFC_w is computed from BOLD time series by correlating FC matrices extracted from the overlapping windows. (B) Edge dFC_e is computed by correlating time points from the BOLD-CA. (C) the hybrid dFC_h retains the global z-score of the BOLD-CA (as in dFC_e), but computes the dFC_h matrix by correlating the sum of overlapping windows applied to BOLD-CA (similar to dFC_w). (D) Comparison between windowed (1, 2, 3), hybrid (4, 5, 6), and hybrid (7, 8, 9) dFC shown on example data of empirical human (1, 4, 7), mouse (2, 5, 8), and simulated data in the bistable regime (3, 6, 9).

a.

Allen Mouse Atlas- regions of interest

0	Right Primary motor area	52	Left Primary motor area
1	Right Secondary motor area	53	Left Secondary motor area
2	Right Primary somatosensory area, nose	54	Left Primary somatosensory area, nose
3	Right Primary somatosensory area, barrel field	55	Left Primary somatosensory area, barrel field
4	Right Primary somatosensory area, lower limb	56	Left Primary somatosensory area, lower limb
5	Right Primary somatosensory area, mouth	57	Left Primary somatosensory area, mouth
6	Right Primary somatosensory area, upper limb	58	Left Primary somatosensory area, upper limb
7	Right Supplemental somatosensory area	59	Left Supplemental somatosensory area
8	Right Gustatory areas	60	Left Gustatory areas
9	Right Visceral area	61	Left Visceral area
10	Right Dorsal auditory area	62	Left Dorsal auditory area
11	Right Primary auditory area	63	Left Primary auditory area
12	Right Ventral auditory area	64	Left Ventral auditory area
13	Right Primary visual area	65	Left Primary visual area
14	Right Anterior cingulate area, dorsal part	66	Left Anterior cingulate area, dorsal part
15	Right Anterior cingulate area, ventral part	67	Left Anterior cingulate area, ventral part
16	Right Agranular insular area, dorsal part	68	Left Agranular insular area, dorsal part
17	Right Retrosplenial area, dorsal part	69	Left Retrosplenial area, dorsal part
18	Right Retrosplenial area, ventral part	70	Left Retrosplenial area, ventral part
19	Right Temporal association areas	71	Left Temporal association areas
20	Right Perirhinal area	72	Left Perirhinal area
21	Right Ectorhinal area	73	Left Ectorhinal area
22	Right Main olfactory bulb	74	Left Main olfactory bulb
23	Right Anterior olfactory nucleus	75	Left Anterior olfactory nucleus
24	Right Piriform area	76	Left Piriform area
25	Right Cortical amygdalar area, posterior part	77	Left Cortical amygdalar area, posterior part
26	Right Field CA1	78	Left Field CA1
27	Right Field CA3	79	Left Field CA3
28	Right Dentate gyrus	80	Left Dentate gyrus
29	Right Entorhinal area, lateral part	81	Left Entorhinal area, lateral part
30	Right Entorhinal area, medial part, dorsal zone	82	Left Entorhinal area, medial part, dorsal zone
31	Right Subiculum	83	Left Subiculum
32	Right Caudoputamen	84	Left Caudoputamen
33	Right Nucleus accumbens	85	Left Nucleus accumbens
34	Right Olfactory tubercle	86	Left Olfactory tubercle
35	Right Substantia innominata	87	Left Substantia innominata
36	Right Lateral hypothalamic area	88	Left Lateral hypothalamic area
37	Right Superior colliculus, sensory related	89	Left Superior colliculus, sensory related
38	Right Inferior colliculus	90	Left Inferior colliculus
39	Right Midbrain reticular nucleus	91	Left Midbrain reticular nucleus
40	Right Superior colliculus, motor related	92	Left Superior colliculus, motor related
41	Right Periaqueductal gray	93	Left Periaqueductal gray
42	Right Pontine reticular nucleus, caudal part	94	Left Pontine reticular nucleus, caudal part
43	Right Pontine reticular nucleus	95	Left Pontine reticular nucleus
44	Right Intermediate reticular nucleus	96	Left Intermediate reticular nucleus
45	Right Central lobule	97	Left Central lobule
46	Right Culmen	98	Left Culmen
47	Right Simple lobule	99	Left Simple lobule
48	Right Ansiform lobule	100	Left Ansiform lobule
49	Right Paramedian lobule	101	Left Paramedian lobule
50	Right Copula pyramidis	102	Left Copula pyramidis
51	Right Paraflocculus	103	Left Paraflocculus

b.



c.

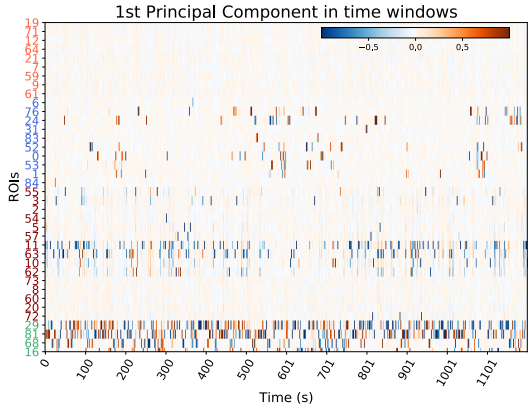


Figure S2:

Figure S2: (a) List of brain regions of interest of the Allen Mouse Atlas considered in the simulations. (b) The first three Principal Components extracted from the simulated firing rate of the bistable regime reveal a dominant contribution of the jumping regions (class J in Figure 2a). (c) Analysis of the first leading eigenvector in non-overlapping time windows reveals the intermittent contribution of regions from classes U* and D* which become active carriers of the system variance during times associated to the neuronal cascades (the time labels are the same as in Figure 2 panel b and d).

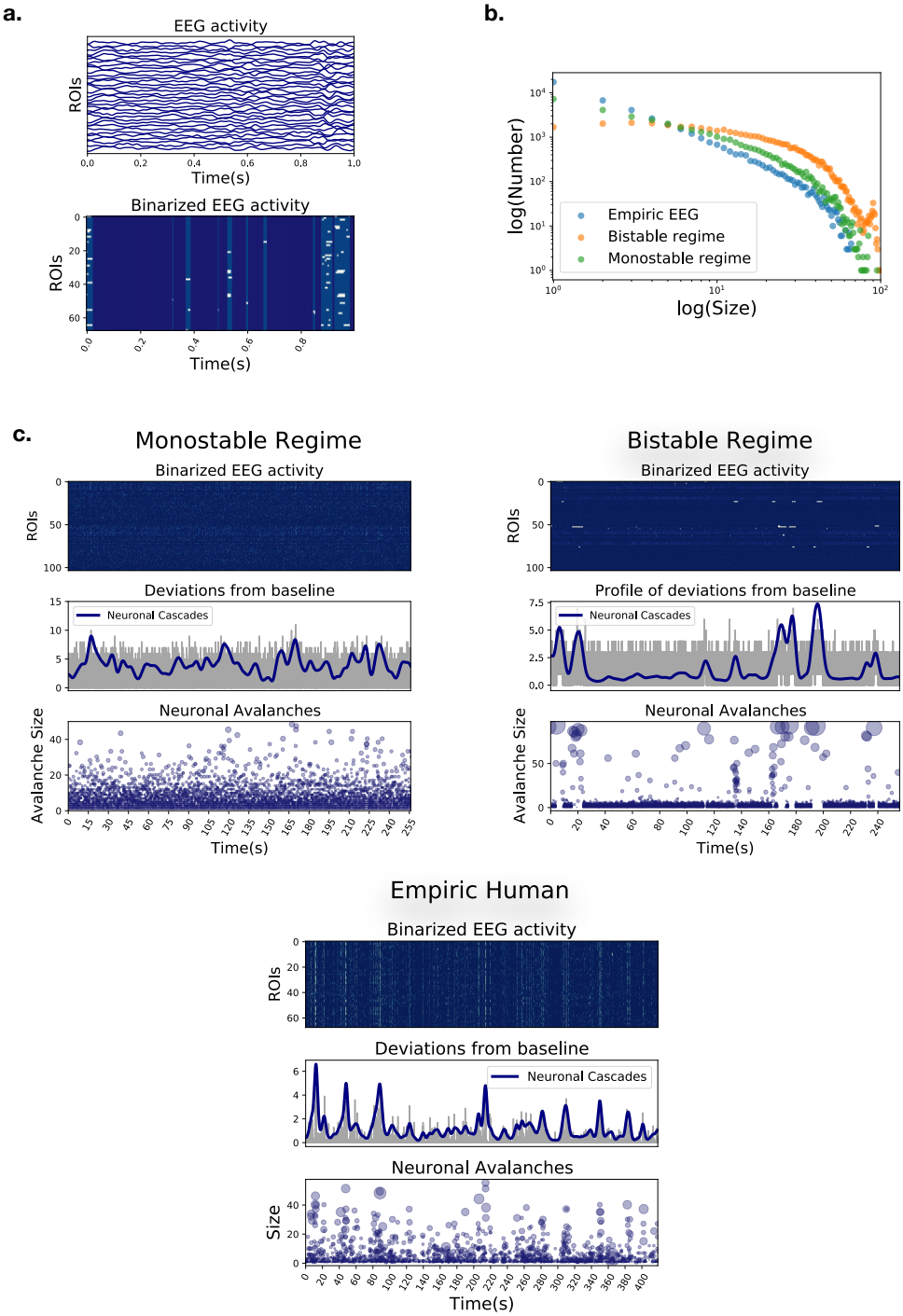

a.
EEG activity
ROIs
Time(s)
Binarized EEG activity
ROIs
Time(s)
b.
log(Number)
log(Size)
Empiric EEG
Bistable regime
Monostable regime
c.
Monostable Regime
Binarized EEG activity
ROIs
Deviations from baseline
Neuronal Cascades
Neuronal Avalanches
Avalanche Size
Time(s)
Bistable Regime
Binarized EEG activity
ROIs
Profile of deviations from baseline
Neuronal Cascades
Neuronal Avalanches
Avalanche Size
Time(s)
Empiric Human
Binarized EEG activity
ROIs
Deviations from baseline
Neuronal Cascades
Neuronal Avalanches
Size
Time(s)

Figure S3:

Figure S3: Neuronal cascades versus neuronal avalanches: (a) We binarize the EEG signals assigning a unitary/null value every time the activity in a region is above/below a certain threshold (e.g., above 3 standard deviations). Roughly speaking, a neuronal avalanche consists of a *consecutive* deviation from baseline activity of a subset of brain regions (or localized neuronal groups). The number of regions in the subset defines the size of an avalanche, while the duration of the consecutive activations defines its lifetime. (b) Power law distribution of avalanche sizes have been considered as an evidence of brain criticality [40], and their occurrence has been associated to optimal information processing in the network, among others (see for example [61] for a review). Interestingly, avalanche-like activity is observed in both simulated regimes, with sizes distributed across several orders of magnitude (not necessarily following a power law distribution). (c) Typically, neuronal avalanches have a lifetime of the order of tens of milliseconds. The neuronal cascades observed here describe a slower phenomenon at the scale of seconds. Looking at the occurrence of avalanches in time (bottom panels; each dot marks the occurrence of an avalanche of a certain size, the dot size represents the duration of the avalanche) we notice an increased magnitude in correspondence of the largest cascades (central panels). In fact, we can think of the neuronal cascades as collections of neuronal avalanches. Let us notice that, in general, the fat tail distribution of the neuronal avalanches does not explain the presence of large neuronal cascades and the clustering of avalanches in time as observed in empirical data. As a counter example we can look at the occurrence in time of the avalanches in the monostable regime which is rather sparse (despite having a distribution similar to the empirical data) and associated with small fluctuations of the neuronal cascade signal.

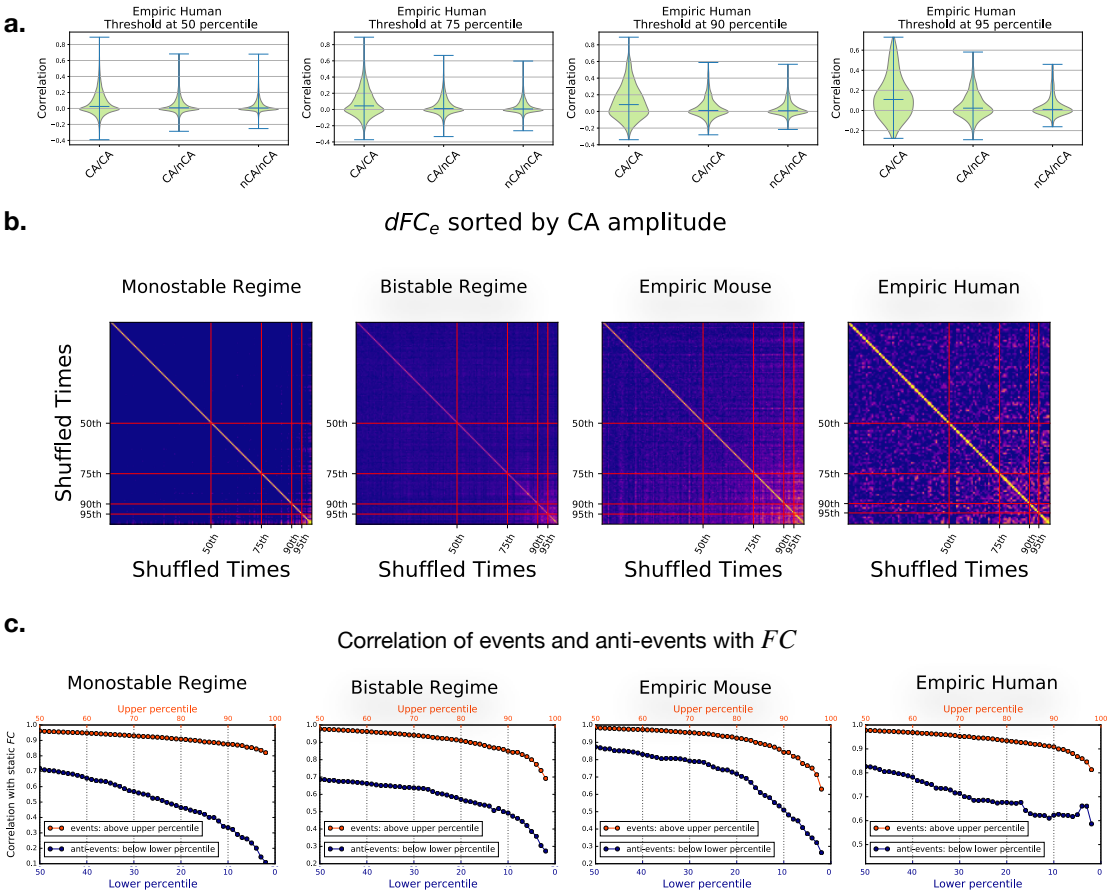

a.
Empiric Human
Threshold at 50 percentile
Empiric Human
Threshold at 75 percentile
Empiric Human
Threshold at 90 percentile
Empiric Human
Threshold at 95 percentile
Correlation
CA/CA
CA/nCA
nCA/nCA
b.
dFC_e sorted by CA amplitude
Monostable Regime
Bistable Regime
Empiric Mouse
Empiric Human
Shuffled Times
50th
75th
90th
95th
c.
Correlation of events and anti-events with FC
Monostable Regime
Bistable Regime
Empiric Mouse
Empiric Human
Upper percentile
Lower percentile
Correlation with static FC
events: above upper percentile
anti-events: below lower percentile

Figure S4:

Figure S4: (a) Analysis of Figure 3B, human subject, for different values of the threshold. The inter-event correlations distribution is always wider than between nonevents or between events and nonevents. (b)Examples of dFC_e in synthetic and empiric datasets. The times are sorted depending on the amplitude of the associated BOLD co-activations. The red lines mark the 50th, 75th, 90th and 95th percentile of the CA amplitudes distribution. (c) In all the synthetic and empiric examples listed, the full resting-state FC is highly correlated with the FC evaluated considering only the times of the CA events i.e., averaging the $iFC(t)$ across the times when the BOLD-CA amplitude goes above the upper percentile. Correlation for different values of the upper percentile threshold is shown in red. In particular, these correlations are always greater than the correlations between the full FC and the FC evaluated considering only the times of the anti-events (defined at times where the BOLD-CA amplitude goes below the lower percentile).

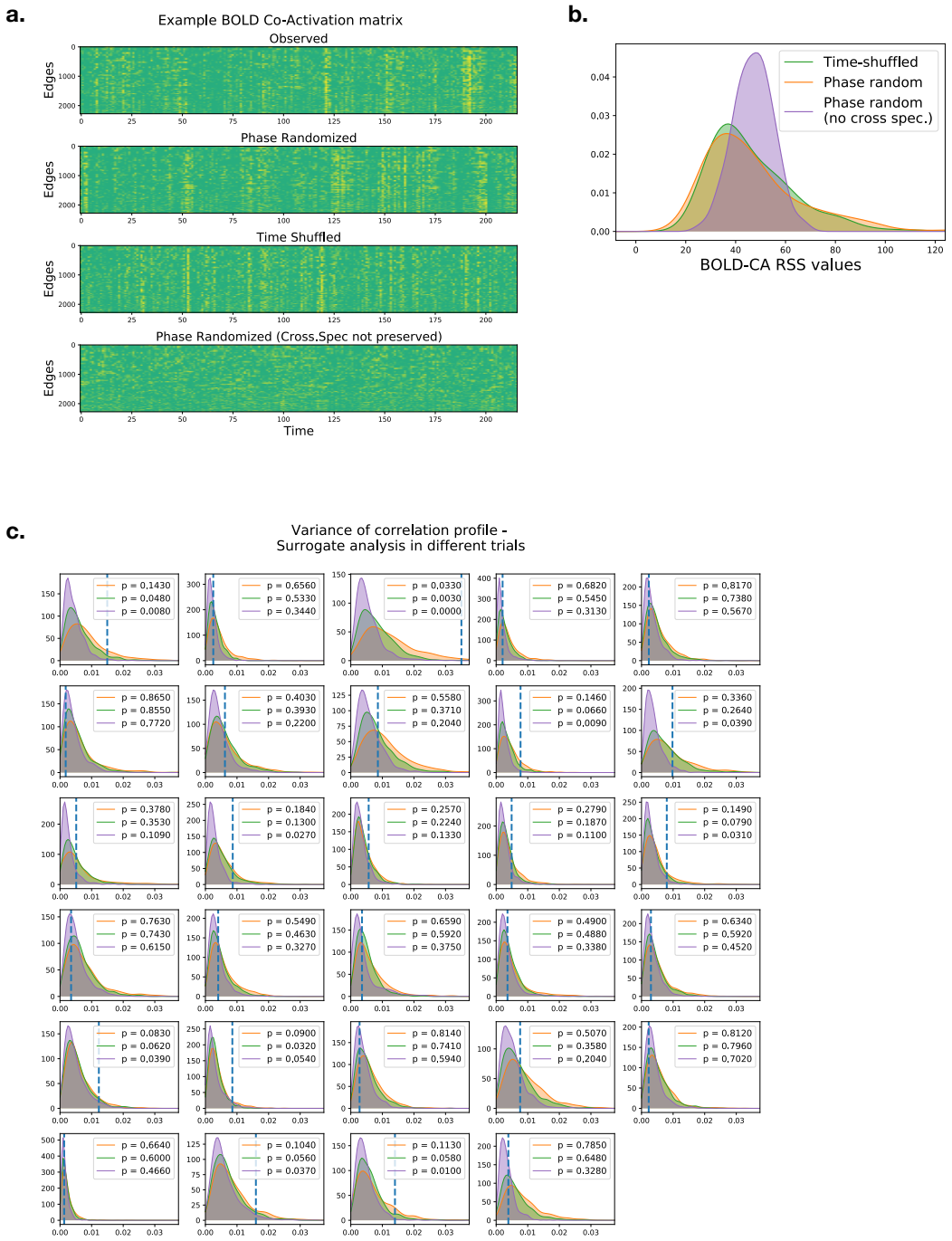

a.
Example BOLD Co-Activation matrix
Observed
Phase Randomized
Time Shuffled
Phase Randomized (Cross.Spec not preserved)
Edges
Time
b.
Time-shuffled
Phase random
Phase random
(no cross spec.)
BOLD-CA RSS values
c.
Variance of correlation profile -
Surrogate analysis in different trials

Figure S5:

Figure S5: (a) Example of the edge CA time series in the original human dataset, compared to the respective phase-randomized and time-shuffled surrogates. Notice that the large co-activation events are preserved only in the first two surrogate models, but they are shifted in time. (b) Distribution of the BOLD-CA amplitudes (the root sum squared of the edge co-activations) in the time-shuffled (same distribution as in the original dataset) and in the phase randomized surrogates where the cross-spectrum is preserved (orange) or not (purple). (c) Single trial analysis of variance distribution (similar to Figure 4c). The legend reports the p-values of the real observed variance versus the three surrogate distributions.