

Multistability in large scale models of brain activity

Mathieu Golos¹, Viktor Jirsa¹, Emmanuel Daucé^{1,2}

1 Aix-Marseille Université, Inserm, INS UMR_S 1106, 13005 Marseille, France

2 Ecole Centrale Marseille, 13013 Marseille, France

* E-mail: mathieu.golos@univ-amu.fr

Abstract

Noise driven exploration of a brain network's dynamic repertoire has been hypothesized to be causally involved in cognitive function, aging and neurodegeneration. The dynamic repertoire crucially depends on the network's capacity to store patterns and their stability. Here we systematically explore the capacity of networks derived from human connectomes to store attractor states, as well as a range of network mechanisms to control the brain's dynamic repertoire. Using a deterministic "graded response" Hopfield model [1] with connectome-based interactions, we reconstruct the system attractor space through a uniform sampling of the initial conditions. Large fixed-point attractors sets are obtained in the low temperature condition, with a number of elements exceeding by several orders the number of attractors reported so far. Different variants of the initial model, including (i) a uniform activation threshold or (ii) a global negative feedback, report a similar robust multistability in a limited parameter range. A numerical analysis of the attractors empirical distributions identifies spatially-segregated components across the models, with a centro-medial core and several well-delineated regional patches. Those different modes share similarity with the fMRI independent components observed in the "resting state" condition. Temporally multistable behavior is reported in noise-driven generalizations of the models, with different meta-stable attractors visited along the same time course. Only the model with a global dynamic density control is found to display robust and long-lasting temporal multistability, with no tendency toward either overactivity or extinction. The best fit with the biological signals is observed at the edge of multistability, a parameter region that also corresponds to the highest entropy of the attractors distribution.

Author summary

Recent developments in non-invasive brain imaging allow to reconstruct brain-wide axonal tracts in human subjects. Physical models of the electrical transmission across brain areas, and corresponding computer implementations, allow to decipher how the different regions share signals and process information. Taking inspiration from the metastable dynamics of the spin glass model in statistical physics, we systematically explore the brain's capacity to process information and identify novel avenues for its enhancement. In particular, we consider the brains specific capability to activate different functional sub-networks across time, as observed in the resting state condition. We identify the conditions at which connectome-based brain networks robustly exhibit a switching behavior between characteristic patterns of activity. We characterize the principal patterns of coactive nodes sharing similarities with other modular structures as observed both in graph theoretical analysis and in resting-state functional Magnetic Resonance Imaging (rs-fMRI).

Introduction

Recent developments of tractography in diffusion tensor imaging (DTI) - or Diffusion Spectrum Imaging (DSI) - allow to build a global picture of the inter-areal connectivity with unprecedented precision [2]. Reconstructed axonal tracts link distant brain regions. The complete and countable set of axonal tracts is referred to as the connectome and can be expressed via the structural connectivity matrix, which includes matrix elements quantifying the connection strength [3]. The connectome contains thousands of links, hence requiring the use of analytical tools to unfold the underlying organization. In analogy with the structural connectivity, the *functional* connectivity refers to the total set of interaction patterns between brain regions [4, 5] as observed from brain imagery techniques. The functional connectivity quantifies the interactions in a statistical model-independent sense. The degree of interaction can be quantified by various metrics including cross-correlation, coherence, transfer entropy etc [6]. It can be applied in a large variety of situations, including temporally-windowed analysis of the resting state fMRI [7] (see also [8]).

A *functional network* is a particular set of distant neuronal assemblies displaying consistent co-activation through the realization of a particular task. The identification and interpretation of such functional networks from observation are the cornerstone of contemporary brain functional imagery. The richness and versatility of the brain activation patterns in a “task-free” condition was probably demonstrated first in the seminal paper by Biswal and colleagues [5]. The so-called “resting state activity” refers to the observed dynamics of the brain when a subject does not pursue a task: in presence of a task, consistent patterns of deactivation appear. They are referred to as the Default Mode Network [9]. The observation that there are relatively consistent distributed patterns of activity during rest has lead to the suggestion that it might be possible to characterize network dynamics through a low-dimensional set of Resting State Networks (RSN) patterns. Further studies reveal richer dynamics with different spatiotemporal patterns simultaneously active during the resting state [10–12], revealing a strong non-stationarity of the resting state.

One means of quantifying the degree of connectivity in a network is the application of graph theory tools, let it be structural or functional. Graph theory tools characterize the connectedness of network nodes, as well as topological features of the connectivity [13–15]. Graph theory, however, misses most of the temporal aspects of node interaction in its analysis. Thus, while RSNs may depend on the structural connectivity, they cannot be understood in those terms alone and require the inclusion of dynamics [16].

Simulation-based studies, as developed in [16–21], come in complement with the graph-theoretical studies. The

idea is to construct a dynamical system that should help unveil the dynamic aspects of the large-scale activity, like the build-up and collapse of transient large-scale assemblies as observed in the resting state condition. Dynamic models of resting state activity have essentially three ingredients: the underlying anatomical connectivity, the time delays due to signal transmission across different distances and the local brain area dynamics. Three types of models are proposed in the literature, with decreasing complexity: (i) spiking neural network simulations, as proposed in [22], account for the spontaneous self-organized criticality regimes as observed in brain slices. However, such models are both difficult to parametrize and computationally expensive; (ii) detailed mean-field models rely on a systematic reduction of the spiking activity to population dynamics. It allows a more computationally amenable exploration of the brain activity at larger scale, and relies on a lesser number of free parameters. Systematic exploration of the scaling parameter of the connectivity matrix, as proposed in [17, 23] reveals a bifurcation map toward different dynamical regimes. Ghosh et al [17] proposed that the mechanism underlying the emergence of RSN dynamics is actually the teetering at the edge of instability, in which the system is most susceptible to noise and thus expressing best its dynamic characteristics. Deco and Jirsa [23] demonstrated the existence of multistability with a maximum of 3-4 distinct attractors in this range; (iii) Spin-glass and Ising models of brain activity, as proposed in [19, 24, 25], gather even fewer free parameters, like the scaling factor and the temperature. Deco et al [19] report for instance an increased entropy and multistability in scale-free connectivity matrices. Marinazzo et al. [25] report an optimal information transfer at criticality in connectome-based spin-glass models. Convergent studies also report self-organized criticality signatures in physiological signals [26–29]. Two important messages seem to surface from this abundant literature, whatever the model resolution, namely the role of a scale-free structural organization and the role of (sub-)criticality in displaying a physiologically relevant activity.

Fewer studies, however, provide a detailed account of the shape of the underlying attractor landscape, and establish a connection between the simulations and the well-known resting-state networks observed in the physiological signals. Those few studies, [20] and [21], rely on a “best match” approach: large series of simulation are used to extract activity patterns that are compared with the best corresponding resting-state networks masks, where the masks are obtained from a literature review or from distinct datasets.

Here we provide a specific entry point to this approach, through the analysis of the fixed-point stability (and/or quasi-stability) of a connectome-based dynamical system. Our main hypothesis is that the brain capability to provide fluctuating patterns over time stems from the ability to stabilize a high variety of different attractors in a noise-free situation. It thus provides a comprehensive evaluation

of the multistable behavior in a deterministic neural network based on the connectome structure. All our analyses are performed on the basis of a fine-grained connectome (composed of about one thousand nodes). To this end, we choose a deterministic variant of the spin-glass dynamics, called the Hopfield “graded neuronal response” model [1]. This model shows interesting correspondences with neural mass models while being analytically tractable. We systematically investigate three variants of the original model and investigate their dynamical properties and the parametric range where multistability is possible.

A first section provides an illustrative example of our attractors sampling approach. Then, we provide in a second section a comprehensive examination of the multistability behavior, as observed in three different variants of the original model. To this end, we quantitatively estimate the number, density and variability of the attractor sets obtained in different ranges of the two principal parameters of our models, namely the gain and the scaling factor. We also compare with a null model, and with different connectomes. In a third section, we contrast the attractor sample-based functional connectivity matrices against a reference rs-fMRI-based functional connectivity matrix. Using a specific clustering method, we also provide a pattern-by-pattern comparison between the principal “modes” of our attractors sets and the typical resting state networks extracted from the rs-fMRI signal. In a fourth section, we generalize to the case of a stochastic dynamics, analyze its itinerant properties and compare with physiological activity in the specific case of the ultra-slow functional connectivity dynamics.

Results

Relaxation dynamics

A first study was made on the original graded-response Hopfield model. Our dynamical system is made of about one thousand nodes, each node representing a particular parcel of the brain according to the brain parcellation proposed in [2] (“SL” model – see the Material and Methods section). Figure 1 gives the relaxation dynamics obtained with initial conditions varying from very sparse to very dense. 100 trajectories were simulated for an initial activity varying from 2% to 98% average activation (with 3% steps), making a total of 3,300 simulations. The time course of the average activation provides a synthetic view of the relaxation dynamics toward the final attractors. We plot a sample of such average trajectories in figure 1a, with each gray line representing a different average time course from a different initial condition. The convergence toward a fixed point is rather fast, with most of the final attractors attained after approximately 100 ms. The final average activity seems to be widespread across the [0,1] interval. When close to 0, the corresponding pattern of activity is expected to have few nodes active. Conversely, when close

to 1, the corresponding pattern of activity is expected to have most of its nodes active. From a finer visual inspection, the final distribution of the average activities is not uniform, with a majority of trajectories converging either to the highest level of activity, or to the lowest one. These two final average states correspond to the two dominant attractors of the dynamics, i.e. the high activity state (“Up” state) and the low activity state (“Down” state).

For each initial density, a soft clustering was applied in order to provide a rough estimate of the number of final attractors (less than 0.9 similarity between two attractors in the final attractors set - see the Material and Methods section). The result is presented on fig. 1b. The peak of pattern variability is obtained for an initial density of 0.5 (50% nodes activation), with 56 different final attractors obtained for 100 different random initializations. For initial densities lower than 0.5, the dynamics almost surely converges to the “Down” state. Conversely, with initial densities higher than 0.5, it converges to the “Up” state.

The spatial spread of the final attractors can be visualized on the cortical surface. Figure 1c represents the “Down” state (inactive state). Figure 1g represents the “Up” state. Figures 1d-f display intermediary states, having contrasted activities with a clear separation between excited nodes and depressed nodes. The non-trivial attractors appear very varied in space, with densities varying from very sparse to very dense. General trends can however be observed, i.e. (i) a shared activity on neighbour nodes, reflecting the dominant local connections, (ii) a partial left-right symmetry, with contra-lateral regions often co-active and (iii) a minority of nodes disconnected from the rest of the network, displaying a central 0.5 activity level, being not influenced and having no influence on the other nodes (around the entorhinal cortex, parahippocampal cortex and fusiform gyrus). The final patterns obtained are rather “patchy”, with full regions being arbitrarily on or off from attractor to attractor. This co-activation (or co-deactivation) of brain regions can be seen as dynamically emerging from the strong local wiring and more sparse distant wiring.

Parametric exploration

Our approach to simulating brain function is based on a systematic exploration of the space of the model parameters. Few variants of the graded-Hopfield model are proposed in order to test different interplays between the excitatory and inhibitory influences (see fig. 2). The gain G , representing the node excitability, is disentangled from the scaling factor P , representing the ratio of excitation over inhibition. The different models are described in eqs. (1a-1b) and eqs. (2a-2c), and referred as the “Static and Local” (SL) threshold model, the “Static and Global” (SG) threshold model, and the “Dynamic and Global” (DG) threshold model.

The general objective is to identify parameter regions providing network activities having a spatial or temporal

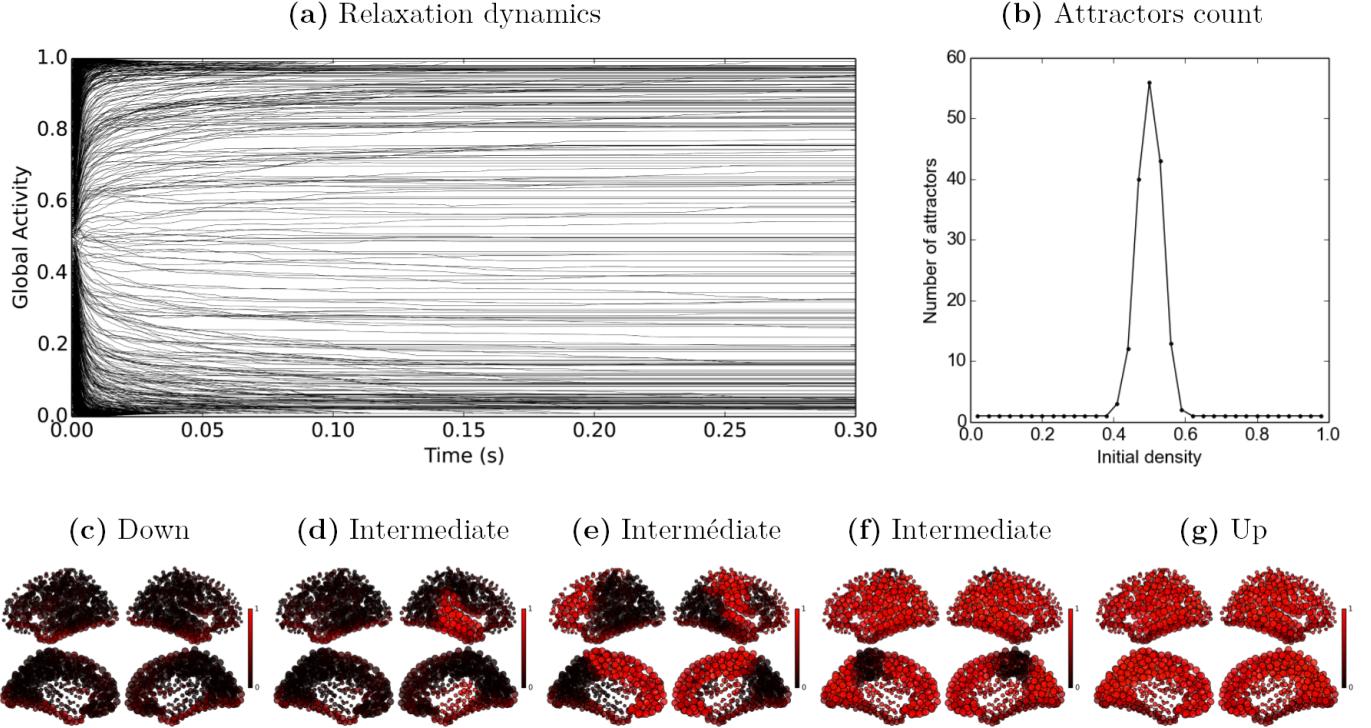


Figure 1. Hopfield graded-response connectome dynamics. **a:** Time course of the average network activity after 1,000 different random initializations. **b:** Distribution of the number of attractors depending on the initial density activity. **c–g:** Example final activation patterns: **c** is obtained for an initial density $f_0 = 0.02$, **d–f** for $f_0 = 0.5$ and **g** for $f_0 = 0.98$. $P = 1$, $G = 900$.

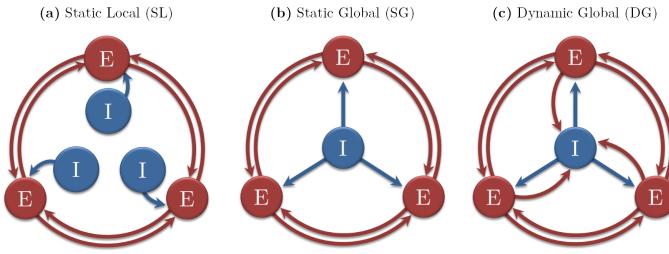


Figure 2. The different schemes used in simulations. Each excitatory node (E) corresponds to one anatomical region of interest (with 998 nodes simulated in most of our numerical investigations). The inhibition (I) takes the form of a linear threshold in our model. It can be local (L) or global (G), and static (S) or dynamic (D).

consistency with the physiological large-scale brain activity. A specific random sampling approach is applied over large regions of the parameter space. For each parameter set, multiple trajectories are generated from random initial conditions. The final attractors are retained, providing a *set of attractors*. Then a clustering algorithm is applied in order to effectively estimate the number of

distinct activation patterns (less than 0.9 similarity between two attractors in the attractors set - see Material and Methods). The density of the initial conditions (the proportion of initially active nodes), being found to play a major role in shaping the final attractors set, is carefully considered hereafter.

Attractors count

The first quantity of interest we consider is the number of distinct activity patterns attained by the dynamics from the space of initial conditions. This number reflects the degree of multistability (or scattering of the attractor basins). A *cluster of attractors* is a set of attractors having a greater than 0.9 similarity among each other. Each cluster is considered as a representative of a single final attractor hereafter. The number of clusters obtained for increasing values of G is represented on figure 3 for different models or different connectivity matrices. In the spin-glass literature, the gain G is sometimes referred to as the “inverse temperature” when a Glauber update scheme is used [19], where low values of G are interpreted as the “high temperature” case, and high values of G as the “low temperature” case. As expected, a key role of the gain G in enhancing multiple stability is found in every model, which is consistent with the spin-glass literature, where

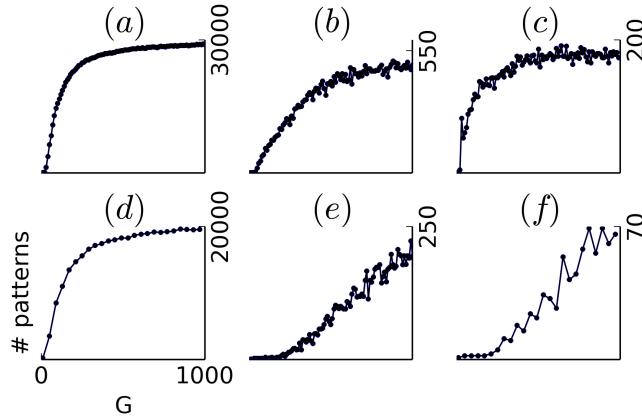


Figure 3. Estimated attractors count in function of G . from 33,000 random initializations and for different models and connectomes. **a,d-f:** SL model, with $P = 1$. **b:** SG model, with $P = 0.9$. **c:** DG model, with $P = 1$. Figures **a-c** and **e** use the the Hagmann composite connectome [2]. Figures **d** and **f** use a a DTI-based connectome extracted from an MRI image from the HCP project database [33]. Figures **e** and **f** use randomized versions of the initial connectomes (**a** and **d**).

an increased multistability is obtained for the decreasing “temperature” [30–32].

In the first row (figs 3a-c), the same connectivity matrix based on Hagmann [2] 998 ROI’s is used under the SL, SG, and DG models. A similar monotonic increase is observed on the three models. The SL case displays a very sharp increase of the number of attractors (approx. 500 attractors for $G = 20$, more than 1,000 for $G = 30$ etc...), with a rapid saturation (to approx. 29,000 attractors) possibly due to the limited size of our initial conditions set (33,000 different initial conditions). Despite the large number of initial conditions, every distinct initial condition seems to drive the dynamics toward a distinct final attractor, indicating a dramatic scattering of the attraction basins, and a probable underestimation of the total number of attractors. The SG case presents a more progressive increase, with a late plateau under 500 attractors for $G > 500$. An earlier plateau is observed in the DG case at much lower cardinalities, indicating a maximum of 150-175 distinct final attractors.

In the second row (figs 3d-f), different connectomes are used under the SL model only. The connectome used in fig. 3d has been extracted from the HCP project database [33], using the SCRIPTS pipeline [34]. It is composed of 1,120 ROI, including sub-cortical regions. The curve is quantitatively and qualitatively similar to fig. 3a, with a final number of distinct attractors estimated at approx 20,000. In figs 3e-f, the connectomes of fig. 3a (resp. fig. 3d) have been shuffled while maintaining the in and out-degrees of the different nodes (see Material and Methods). In this

case, a much smaller number of attractors is observed, with a maximum of 250 in the Hagmann case, and a maximum of 70 in the HCP case. This quantitative difference indicates a critical role of the graph structure in enhancing multistability, possibly due to an intrinsic small world and/or scale-free organization [14, 26], as suggested by [19] (see also [35]). Hagmann et al. [2] demonstrated that the cortical architecture of the brain contains hubs. We speculate that precisely because of the existence of hubs, cortical networks are able to reflect a large number of functional networks. In other words, the large number of functional networks reflects the richness of the dynamical repertoire [17] that the noise around the trivial stable spontaneous state can explore.

Cascade of bifurcations

The transition from monostability toward multistability on the SL model is presented in figure 4. This transition is a cascade of bifurcations, from monostability to bistability, quadrinstability etc. The first bifurcation, observed at $G = 13.2$, belongs to the supercritical Pitchfork type, with a first central attractor becoming unstable and two stable equilibria symmetrically appearing on each side of it. This bifurcation is rapidly followed by a second bifurcation for $G = 14.8$, giving rise to 4 different attractors. Numerous bifurcations follow for $G > 15.6$, with an exponential growth rate, but no symmetry anymore. We show on figure 4b the evolution of the Lyapunov function for $10 \leq G \leq 20$ (see eq. (11)). The value of G is on the y-axis. Each distinct attractor is represented by an unfilled circle, with its average density on the x-axis, and its Lyapunov value on the z-axis. The two main valleys, that go deeper for increasing G , correspond to the two extremal activity patterns (“up” and “down” states). The intermediary activity patterns, displaying a higher “energy” level, are less likely to be reached. The corresponding probability of appearance for each attractor is shown on fig. 4c using the Boltzmann-Gibbs distribution (see eqs. (9-10)), and compared on fig. 4d with their empirical proportion. A close to perfect match (0.95 correlation) is obtained with $\beta = 2$. As expected from the Lyapunov function, most of the attractors have a very low probability (close to 0) at the exception of the “up” and “down” attractors that show a similar symmetrical probability close to 0.5 (for $G > 13.2$).

Role of the scaling factor

The high gain case was thoroughly studied in order to identify the role of the scaling factor P and initial vector density f_0 on the different models. The figure 5 presents the maps obtained in the 3 models considered, namely SL (top), SG (middle) and DG (bottom). For $G = 900$, we systematically varied the scaling factor P and the density of the initial conditions f_0 . For numerical reasons, only 3,300 initial conditions were sampled at each value (P, f_0) ,

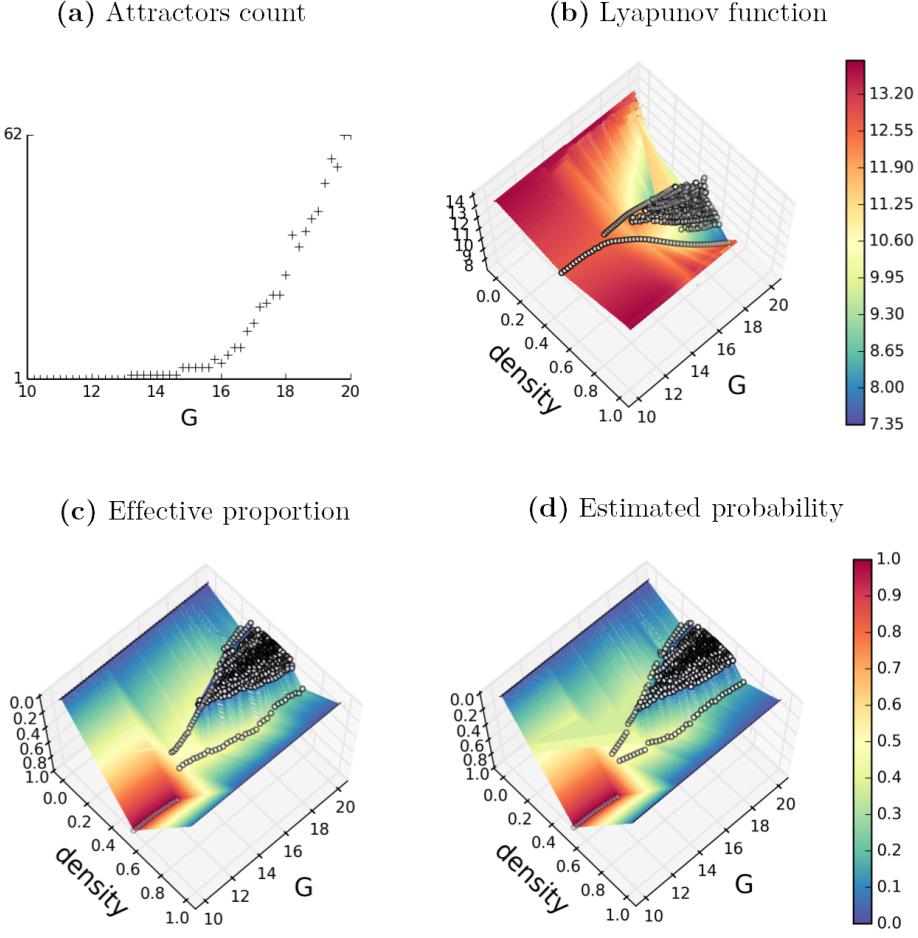


Figure 4. Bifurcation diagrams in the SL model. See text. $P = 1$, $\beta = 2$.

leading to an underestimation of the attractor count at the maxima (see fig. 3), but having no consequence on their location. Three observables are considered, namely the *number* of distinct attractors (i.e. the number of clusters), the average *density* (average proportion of active nodes) and the *cardinality* attached to each cluster in the final set. In figs 5a,c,e, the color code gives the number of distinct attractors at each value of (P, f_0) , and the black lines are isodensity lines obtained after calculating the average density of the final attractors. In figs 5b,d,f, the color code gives the empirical *entropy* of the final set, based of the cardinality of the final clusters (see Material and Methods).

The two maps at the top (fig. 5a-b) correspond to the SL model. For low values of P and/or low initial densities, the system is mono-stable and systematically converges toward the “Down” attractor where the activity on every node remains close to zero ($\text{density} < 0.05$). A comparable monostable behavior is observed for high values of P and/or high initial densities, with the trivial “Up” attractor having a maximal activity on almost every node

(pattern density > 0.9). The multistable behavior is obtained in the intermediate range (around $P = 1$, and with an initial density between 0.3 and 0.6), i.e. the system can relax to a large set of distinct attractors, with varied densities (see for instance Figures 1d-f). The peak of multistability is attained at $(P = 1, f_0 = 0.5)$, where almost every final attractor is different. The empirical entropy map (fig. 5b) provides an additional insight about the predictability of the final activity patterns, where 0 corresponds to a full predictability (single output, no information). The map shows here a similar shape than the one at the left, with the highest entropy at the multistability peak. The maximal value is close to the entropy upper bound (Here at 11.69, obtained with an alphabet of 3,300 symbols with equal probability). The high entropy values indicate here a varied set of equally probable final activity patterns, supporting a strong flexibility of the system.

Consider now the SG case (fig. 5c-d). The general shape of the left map is qualitatively similar to fig. 5a, with a much lower attractor count. The upper right corner corresponds to very dense patterns (“Up” state) and the lower

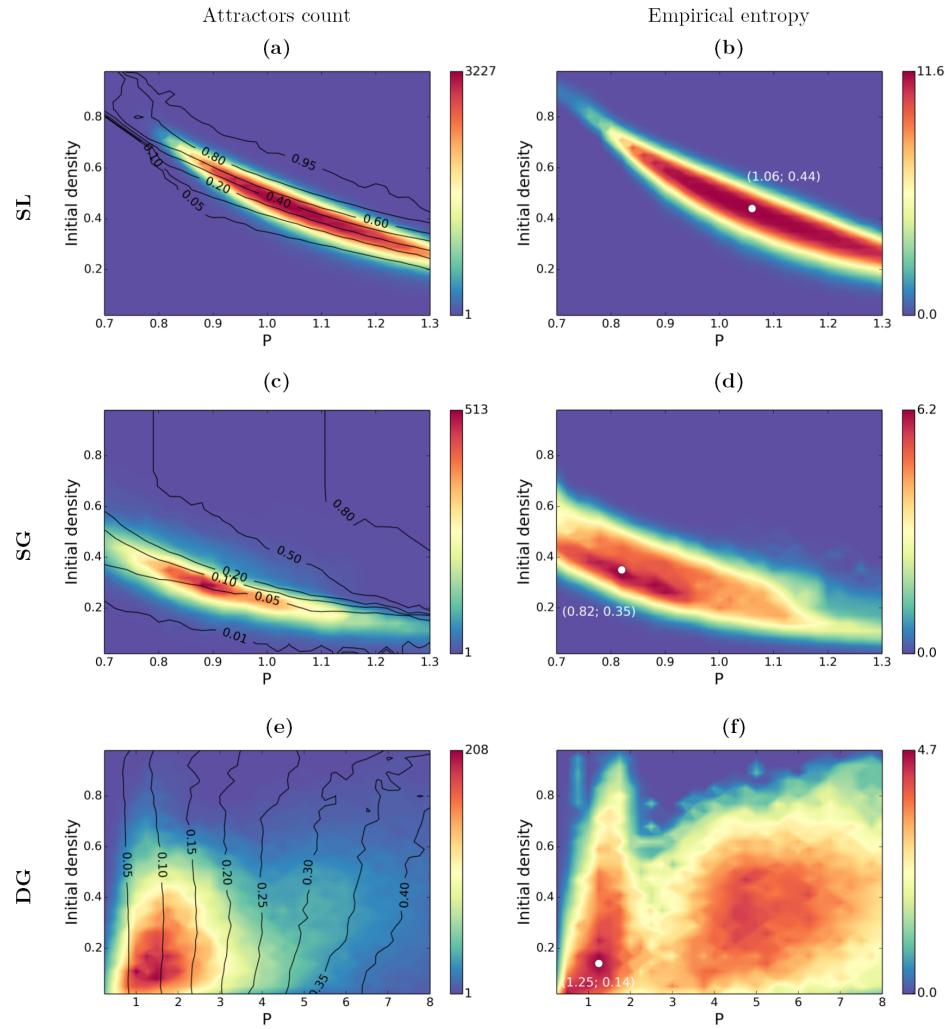


Figure 5. Attractors count (left) and empirical entropy (right) in function of P and initial density f_0 . 3,300 initial conditions are randomly sampled for each value of (P, f_0) . Black lines are isodensity lines corresponding to the attractors average densities. **a-b :** SL model. **c-d :** SG model. **e-f :** DG model. $G = 900$.

left to very sparse (“Down state”). The two trivial attractors thus similarly appear to exert a strong attractivity. The multistability peak is found on a slightly different range, close to $P = 0.9$. The maximal number of attractors is close to 500, which is significantly less than in the SL case. The main difference lies in the much smaller attractor density (of the order of 0.03 – 0.1), which indicates much sparser activities than in the previous case. In this network composed of purely excitatory nodes, each sparse pattern denotes a pool of neurons sharing enough excitatory links to locally keep a self-sustained activity, but not enough to excite the rest of the network. Interestingly, the entropy map shows a wider region of maximal entropy, covering regions where a relatively lower number of attractors is observed. This reveals a complexity that is not quantitatively different when the number of attractors is close to 500 or close to 30-50.

The last map (fig.5e-f) corresponds to DG model. In this model, a dynamic threshold regulates the average activity (see Material and Methods), destabilizing the two extremal patterns, and giving more “space” to the intermediate patterns. In consequence, the final density is no more dependent on the initial density, showing a clear dependence on P : sparse activity patterns are obtained for low values of P ; dense activity patterns are obtained for high values of P . Contrary to the previous cases, multistability is robustly obtained in a large region of the parametric space, for $0.2 < P < 10$. The multistability region is about 10 times wider than in the static thresholds cases, with a peak of multistability attained between $P = 1$ and $P = 3$. The densities obtained at the peak of multistability, ranging from 0.05 to 0.15, are much lower than in the SL case, and similar to the ones obtained in the SG case. A smaller number of final distinct patterns (less than 200)

is obtained than in the SL and SG cases. When looking at entropy map, two regions of maximal variability are observed. A first maximum is found for $0.5 < P < 2$, corresponding to low densities (around 0.1), and a second maximum is observed for $4 < P < 6$, corresponding to higher densities (around 0.3). The maximum value is not quantitatively different at the two maxima (around 4), while the number of attractors is. This indicates, at the first maximum, a great proportion of attractors unlikely to happen, and only a small proportion showing up in practice. This also indicates, on contrary, the different attractors happening more uniformly over trials at the second maximum.

The scaling factor P , whose general effect is to regulate the excitation, is thus found to play distinct roles across the models. In the case of a fixed threshold (SL and SG models), it helps to settle a range of “viable” activity, where the excitation can spread over limited regions, without igniting the full brain. In contrast, the DG model is very tolerant to a variable scaling factor (a variable E/I ratio), and the role of P is here to regulate the density of the final activity patterns, allowing to unfold attractor sets of different densities. The density of the final activity pattern is more generally an important quantity to consider. In the models having a fixed activation threshold (SL and SG), the final density is very variable and depends on the initial conditions, while it is tightly controlled in the DG case. In addition, all densities can coexist in the SL model, while only “sparse” activity patterns (between 0.02 and approx. 0.4) can coexist the SG and DG model.

Attractor set analysis

One way to validate a model is to compare the simulations results with activity patterns obtained from fMRI signals in patients. The richness of the simulated dynamic repertoire can be expressed via different means, including multistability of attractors, coexistence of multiple time scales (so-called slow and fast manifolds) and diversity of attractor types. Much of the discussion on the mechanisms underlying the resting state dynamics of the brain has focused on the exploration of the brains dynamic repertoire. The latter has been demonstrated in a series of papers from various laboratories [10–12, 36–42] to be composed of intermittent activation patterns comprising well-known resting state networks. Here each element of a final attractors set provides a specific activation/deactivation pattern (a sub-network) reflecting a correlation structure that is potentially present in the connectome. Each pattern can be mapped on the cortical surface, and interpreted from the functional perspective : an attractor provides a set of nodes that are expected to “work together”, possibly reflecting an underlying brain function. In analogy with the Hopfield associative memory, typical patterns of activity (or prototypes) embedded in the weights matrix are expected to be revealed by sampling the initial conditions space.

To guide our comparison with experimental results, we consider a set of resting-state functional MRI (rs-fMRI) time courses that were observed on 5 healthy in [43]. A total of 35 minutes of rs-fMRI signals were recorded in two sessions per subject. The signals were mapped onto a set of 998 ROI’s on the cortical surface. Pearson correlation was applied on the full set of data, resulting in a cross-subject 998 x 998 correlation matrix considered here as the reference functional connectivity matrix (Empirical Functional Connectivity matrix – EFC).

Correlation matrices

Patterns of co-activation can be compared across the different matrices, whether functional, structural or “attractor based”. A simple, but quite effective metric, is the measure of correlation among matrix values. This comparison can be done globally, or restricted to specific regions of the matrices. For instance, Honey et al [43] extracted the nonzero structural links, and compared their values with the corresponding values in the EFC. A correlation of 0.53 was found in that case (using a Gaussian resampling of the initial DSI weights). When directly correlating the binary structure of a single hemisphere (containing zero when the link is absent and one when the link is present) with the corresponding EFC values, a correlation of 0.44 is found. A lesser value (0.39) is obtained when using the original weights. On contrary, a simple mathematical manipulation (a log-rescaling of the original weights) allows to reach a 0.48 correlation. This similarity however drops to 0.22 when considering only the interhemispheric links structure (0.19 when using the original weights, 0.24 when using a log-rescaling), which reflects a poorer detection of the callosal fibers by the DSI algorithm. Those quantities must be contrasted against the intra and inter-subject variability as observed in the rs-fMRI signal. When comparing the EFC calculated on two distinct sessions for the same subject, an average 0.58 correlation is observed. When comparing the EFC obtained on two different subjects, an average 0.32 correlation is observed. These comparisons indicate that 0.55-0.60 correlations reflect a close fit, and 0.3-0.35 correlations reflect a more scarce but still significant fit. In consequence, the value greater than 0.40 observed between the structural matrix and the EFC reflects a good (better than scarce) initial fit, that raises the question of the specific contribution of network simulations in improving the prediction of large-scale connectivity patterns from structural data.

When multiple attractors are obtained for a particular (P, f_0) value, an attractor-based functional connectivity matrix can be calculated. This matrix (Attractor-based Functional Connectivity matrix - AFC) is a correlation matrix that reflects the co-activations (and co-deactivations) across the different attractors of the set. For each model and each value of (P, f_0) , AFC matrices were calculated. When multiple attractors are presents, the AFC matrices display a checkerboard structure (not

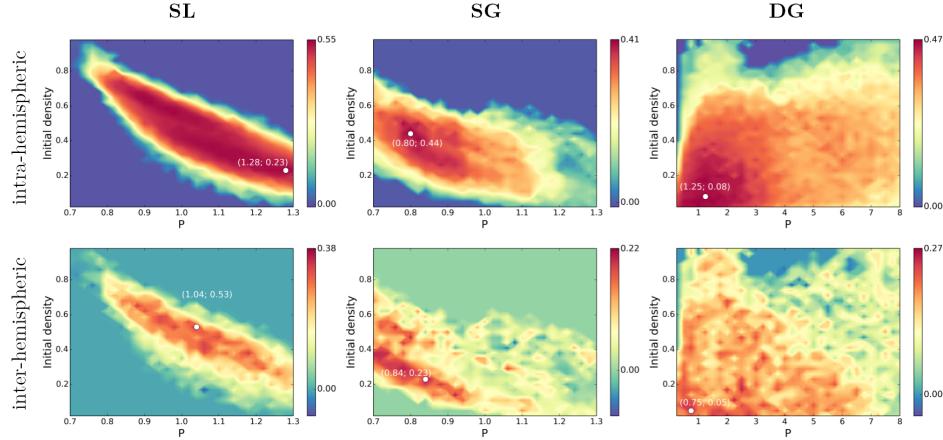


Figure 6. Attractor-based functional connectivity matrices are compared with the Empirical Functional Connectivity matrix. Intra- and inter-hemispheric AFC matrix values are correlated to the EFC values. The comparisons are done on the different models on the same range as in figure 5. $G = 900$.

shown) that reflects the modal spatial organization of the attractors, with well-localized anatomical groups having a tendency to share activity. This checkerboard covariance organization is a correlate of the “patchy” aspect of the attractors seen on figure 1.

The correlation maps between the intra- and inter-hemispheric AFC and the intra- and inter-hemispheric EFC are shown on figure 6. In the case of the SL model, a strong intra-hemispheric similarity (between 0.50 and 0.55) is obtained for decreasing values of f_0 in the $0.8 < P < 1.3$ range. Those maximal values are obtained along the two high and low-density lines (0.10 and 0.90 densities) and mostly reflect the local correlation structure. The best report of the inter-hemispheric correlation structure (0.32) is obtained for intermediate densities ($P = 1.04$, $f_0 = 0.53$, around 50% nodes active). The general trend is thus a better prediction of the local correlations at low/high attractor densities, and a better prediction of the large scale correlations at intermediate attractor densities. The SL model is thus found to display a significant improvement of the functional connectivity prediction, at both local and long range, when compared with the sole structural prediction.

The similarity map obtained on the SG model (figure 6b) indicates a maximal correlation (0.41) at a region that is above the maximal entropy line (see fig.5d), and corresponds to intermediate density attractors at the upper-edge of multistability (around 20-30% of the nodes active on average). The best correlation with the inter-hemispheric organization (0.22) is obtained in a region of low-density attractors, at the lower edge of multistability (around 2-5% of the nodes active on average). Whether reporting non-trivial correlation structures at intermediate densities, the model must however be considered as not improving the prediction (i.e. the non-connected nodes that co-activate in the SG simulated dynamics do not match

with the ones that co-activate in the physiological signal).

When considering the DG model (fig. 6c), the intra-hemispheric best match (0.47) is obtained for $P < 2$, on an interval that corresponds to the highest entropy (fig.5f), with average activities ranging from 5% to 15%. The inter-hemispheric best match (0.27) is obtained on a smaller range for $P < 1$, at the leftmost edge of multistability, for a set of attractors having on average 5% of their nodes active. While slightly better than the SG model, only the inter-hemispheric correlation can here be considered as a significantly improving the prediction regarding the structural information. In consequence, $0.5 < P < 1$ can be considered a relevant range in predicting the physiological activity in this model.

Clustering approach

Brain dynamics analysis identifies large scale networks in the rs-fMRI signal by extracting the most frequent co-activation/co-deactivation patterns. A popular approach is using independent components analysis (ICA) to extract the principal patterns. In contrast with the high number of attractors displayed by our model, the large-scale dynamics, as observed in the rs-fMRI signal, is known to display a relatively small set of modes, each mode reflecting known associations of brain regions in the realization of cognition, perception and action [10, 36–40, 42].

From the spin-glass literature, the number of attractors is known to combinatorially explode in the low temperature condition [30], leading to numerous “spurious” attractors. In that case, not every final attractor should represent an effective prototype, but rather a combination of them [32]. In the absence of an explicit set of prototypes, this renders the final attractor landscape quite puzzling and difficult to read. Empirical statistics must be used in order to approach the underlying prototypic organization

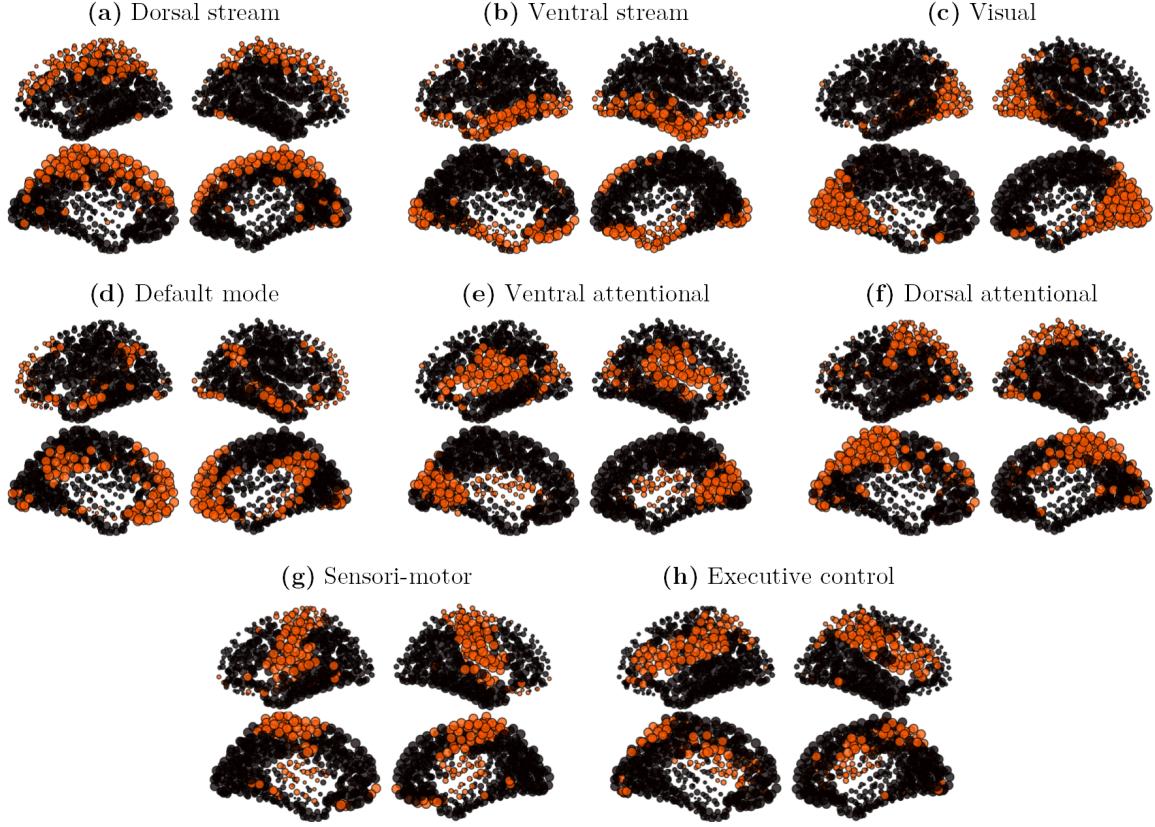


Figure 7. 8 clusters cores, cross-subject BOLD signal. The initial set contains 5,200 binary observation vectors. Thresholded rs-fMRI signal ($s > 2$), 5 subjects \times 35 minutes. Double-pass clustering, with inclusion match similarity ($k = 0.5$).

from the attractors distribution. The idea is to identify the principal "modes" of the distribution, where a mode is a region of the sample space where the data points have a stronger density. Dense regions of the attractor space are expected to represent larger (and thus more stable) attraction basins, and reflect the putative prototypic organization of the connectome structure. The combination of attractor sampling and clustering we develop in this section thus provides an insight into the statistics of the functional modules (or sub-networks) embedded in a connectome, encompassing strong nonlinear effects that are not considered in the graph-theoretical approaches.

In order to allow comparison with our attractor sets, a specific clustering algorithm was applied. The general idea is to consider each observation vector from the rs-fMRI signal as representing a particular attractor of the underlying brain dynamics. 5,200 observations vectors were extracted from the five subjects time courses of the Hagmann rs-fMRI database. A double-pass clustering, using inclusion match similarity (see Material and methods), was applied on the observation vectors set, with a similarity threshold $k = 0.5$. 601 clusters were extracted, with cluster size strongly varying in size, the biggest cluster gathering 928 elements (17.8%), and many small ones containing only

1 element. Not every cluster being considered significant, only the characteristic vectors of the 8 principal clusters, ranked by importance, and gathering 61% of the signal, are presented on figure 7, where the characteristic vectors are here defined as the 25% most active nodes of the cluster. All patterns of the set spread on a large scale, displaying strongly bilateral and symmetrical components. Our set of patterns contains (a) a superior and medial pre-frontal mode, including superior parietal elements (Dorsal stream, 928 observations), (b) an orbito-frontal mode including inferior temporal and para-hippocampal components (Ventral stream, 617 observations), (c) a medial and lateral occipital mode corresponding to the visual network (528 observations), (d) a centro-medial mode with inferior parietal and temporal components corresponding to the default mode network (DMN, 382 observations), (e) a superior temporal and ventral prefrontal mode including a primary visual component (Ventral attentional, 236 observations), (f) a lateral occipital mode including superior and medial parietal components (Dorsal attentional, 191 observations) (g) a pre and post-central mode corresponding to the sensori-motor network (163 observations), and (h) a superior parietal and dorso-lateral prefrontal mode corresponding to the fronto-parietal executive control net-

work (143 observations). Those components are clearly delineated and consistent with the literature [10,36–40,42].

The question is now how much of the large functional networks are predicted by the activation patterns obtained in our simulation-based attractor sets. To allow comparison, a tighter clustering was applied on two attractor sets obtained on two different models. On the basis of previous comparisons (see fig. 6), only the SL and DG models are considered here, for only them display a significant improvement in predicting the empirical functional connectivity.

A first set of patterns is built from the graded-response Hopfield model (SL model) using 33,000 random initial conditions. The attractor set, composed of 28,849 attractors, displays dense activity patterns with on average 50% of the nodes active. When looking at the different attractors, each attractor seems to be composed of several anatomical “patches”, and the same patches seem arbitrarily active (or inactive) across the different attractors. Significant co-activations do not show up at first sight. A double-pass clustering, using inclusion match similarity (see Material and methods), was applied over the first set of 28,849 attractors, with a similarity threshold $k = 0.8$. A total of 725 clusters was found, with cluster size strongly varying in size, with the biggest cluster gathering 4,089 similar elements (14.2% of the total), and many small ones containing only 1 element. Half of the nodes being on average active in the SL model, repeatedly active nodes in a cluster identify the cluster core: characteristic vectors are here defined as the 10% most active nodes in a cluster. The characteristic vectors of the 8 principal clusters, ranked by importance, and gathering 50% of the total attractors of the set, are presented on figure 8. From this set, 2 patterns (a and e) present a bilateral symmetry, 3 patterns (b, g and h) are left-lateralized and 3 patterns (c, d and f) are right-lateralized.

Despite the strong core selection, a widespread organization of the patterns is observed. The different patterns are constituted of several well-delineated, but not necessarily contiguous, anatomical regions. Each cluster core should thus be seen as predicting a typical large-scale interaction pathway, either within or across the hemispheres. The logic of the comparison is thus identifying which *parts* of the functional organization are correctly reported by the simulation, and which parts of it differ. For this purpose, an “inclusion match” similarity is used which indicates which proportion of a cluster core (10% most active nodes) correctly match with the 25% most active nodes of a rs-fMRI-based cluster. This match is considered significant when the proportion of correct predictions is greater than 37% (t-value > 3), good when the proportion is greater than 51% (t-value > 6), and strong when the proportion is greater than 65% (t-value > 9). All patterns show a significant correspondence with at least one functional pattern of figure 7.

The first cluster core (fig. 8a, 4,089 attractors) links bilaterally the frontal pole, the orbito-frontal cortex and the

anterior prefrontal cortex, displaying a strong match with the anterior part of the Default-mode network. The second cluster core (fig. 8b, 3,048 attractors) is a left-lateralized pattern linking the medial and lateral occipital cortex with the posterior temporal cortex (left ventral visual pathway) displaying a good match with the visual and the ventral functional networks. The third cluster core (fig. 8c, 2,045 attractors) is a right-lateralized pattern linking superior temporal and inferior parietal regions with the ventral prefrontal cortex. A fair match with the ventral attentional network is observed. The fourth cluster core (fig. 8d, 1,573 attractors) is the right contralateral homologous of fig. 8b (right ventral visual pathway), displaying a good match with the visual and the ventral functional networks. The fifth cluster core (fig. 8e, 907 attractors) is a centro-medial pattern, linking bilaterally the mesio-parietal and cingulate cortex. The sixth cluster core (fig. 8f, 892 attractors) is a right-lateralized pattern, linking the superior temporal and the inferior parietal regions, showing a fair match with the ventral attentional network (similar to 8c except for the ventral prefrontal regions). The seventh cluster core (fig. 8g, 881 attractors) is the left contralateral homologous of pattern 8f, showing a fair match with the ventral stream. At last, the eighth cluster core (fig. 8h, 844 attractors) is a left-lateralized pattern linking the lateral parietal regions with the ventral prefrontal cortex, displaying a consistent match with the executive control network. Two main differences with the rs-fMRI functional networks can be highlighted here. A first difference is the ventral visual pathway here separated in a left and right components (b and d), while separated in bilateral visual and ventral components in the rs-fMRI signal analysis (b and c). Both interpretations can however be considered as physiologically relevant. On contrary, the patterns (c), (f) and (g) seem relatively specific to simulation, for they report a “vertical” network, from the inferior parietal to a large portion of the temporal cortex (except the temporal pole), that has no obvious counterpart in physiology. The non-contiguous link between the inferior parietal regions and the inferior prefrontal regions, reported in pattern (c) is however consistent with the non-contiguity of the executive control network.

A second set of attractors is built from the DG model on a series of simulations obtained with P varying from 0.5 to 7 (with 0.25 steps), providing a set of 36,459 different attractors. This second set of attractors is more regular, the attractors are much sparser with many of them looking closely similar and fewer ones having a different spatial distribution. The same clustering algorithm was applied over this second set of attractors. A total of 80 clusters was found, with an even stronger cluster size disbalance, with the biggest cluster grouping 11,141 similar elements (30.6% of the total), and the many tiny ones having only one element like previously. A small number of configurations thus tends to dominate the dynamics here.

The characteristic vectors of the 8 principal clusters, ranked by importance, and gathering 89% of the total at-

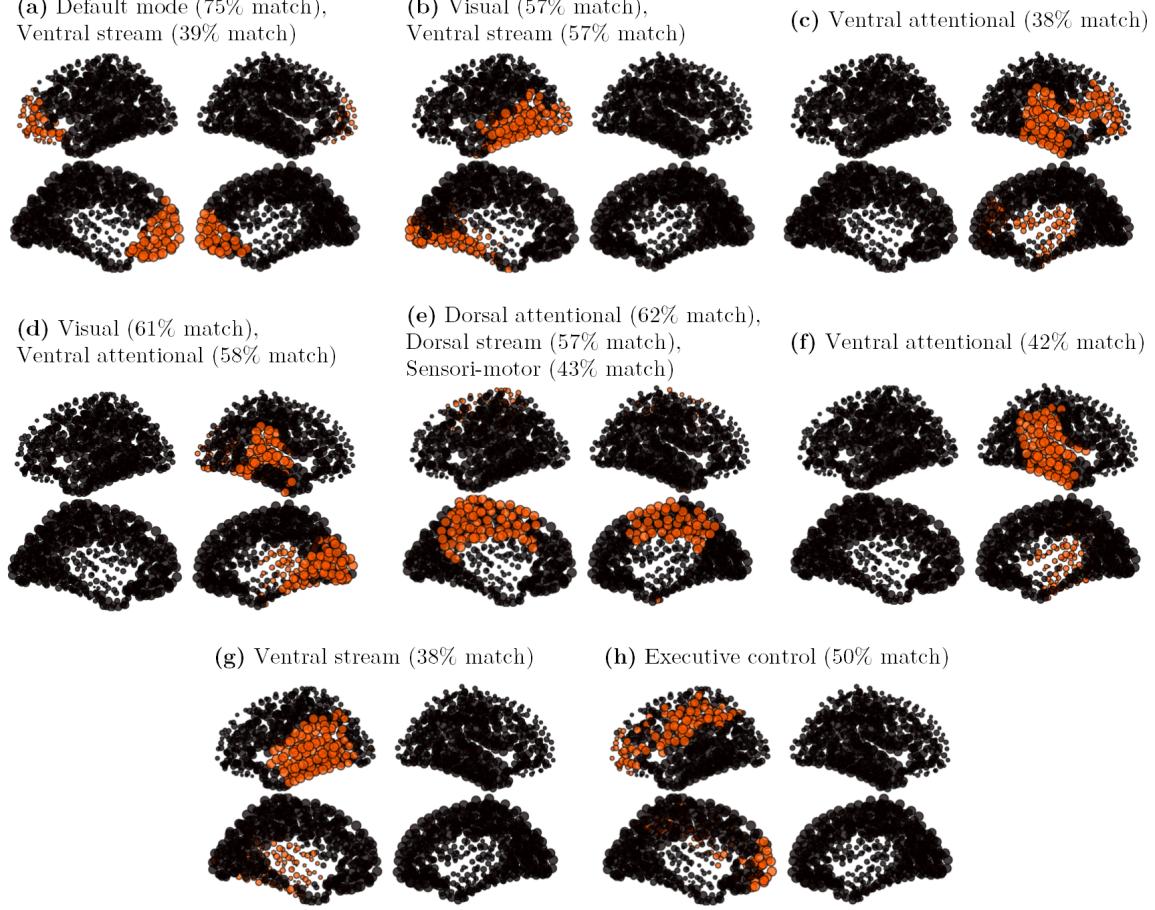


Figure 8. 8 clusters cores, on a SL attractor set. The initial attractors set contains 28,849 attractors, with $P = 1$ and $G = 900$. Double-pass clustering, with inclusion match similarity ($k = 0.8$).

tractors of the set, are presented on figure 9. The characteristic vectors are here again defined as the 10% most active nodes in a cluster, and an “inclusion match” comparison with the eight most characteristic rs-fMRI-based patterns (fig.7) is done. From this set, 3 patterns (a, b and h) present a full or partial bilateral symmetry, 3 patterns (c, f and g) are left-lateralized and 2 patterns (d and e) are right-lateralized. The first cluster core (fig. 9a, 11,141 attractors) links bilaterally the cuneus, the paracentral lobe and the posterior cingulate. This centro-medial pattern is sometimes considered as the functional core of the Default mode network [44], while also matching with the dorsal attentional network in our analysis. The second cluster core (fig. 9b, 6,376 attractors) links bilaterally the cuneus, lingual and the inferior part of the precuneus, displaying a strong match with the visual network. The third cluster core (fig. 9c, 5,494 attractors) links the left inferior parietal region with the cuneus through the superior parietal cortex, displaying a fairly good match with the Dorsal attentional network. The fourth cluster core (fig. 9d, 2,833 attractors) links the left inferior parietal region with the superior temporal region, showing an inciden-

tal match with the ventral attentional network. The fifth cluster core (fig. 9e, 2,832 attractors) links the left superior parietal region with the post-central sulcus, showing a good match with the executive control and dorsal attentional networks. The sixth cluster core (fig. 9f, 1,347 attractors) is the contralateral homologous to fig. 9e. The seventh cluster core (fig. 9g, 1,275 attractors) is the contralateral homologous to fig. 9d. At last, the eighth cluster core (fig. 9h, 1,236 attractors) is a leftmost homologous of fig. 9a, showing here again a good match with the dorsal attentional network. In that model, the frontal and orbito-frontal regions are only present in smaller clusters not represented on the figure. The most prominent regions are here again the precuneus, the inferior parietal region, the bank of the superior temporal sulcus, and the post-central sulcus.

A general observation from this detailed pattern analysis is a clear difference in the spatial dominance of the attractors across the two models. The SL model, with a more widespread distribution of activity, provides a greater variety of patterns. In particular, the ventral visual stream and the orbitofrontal and medial prefrontal

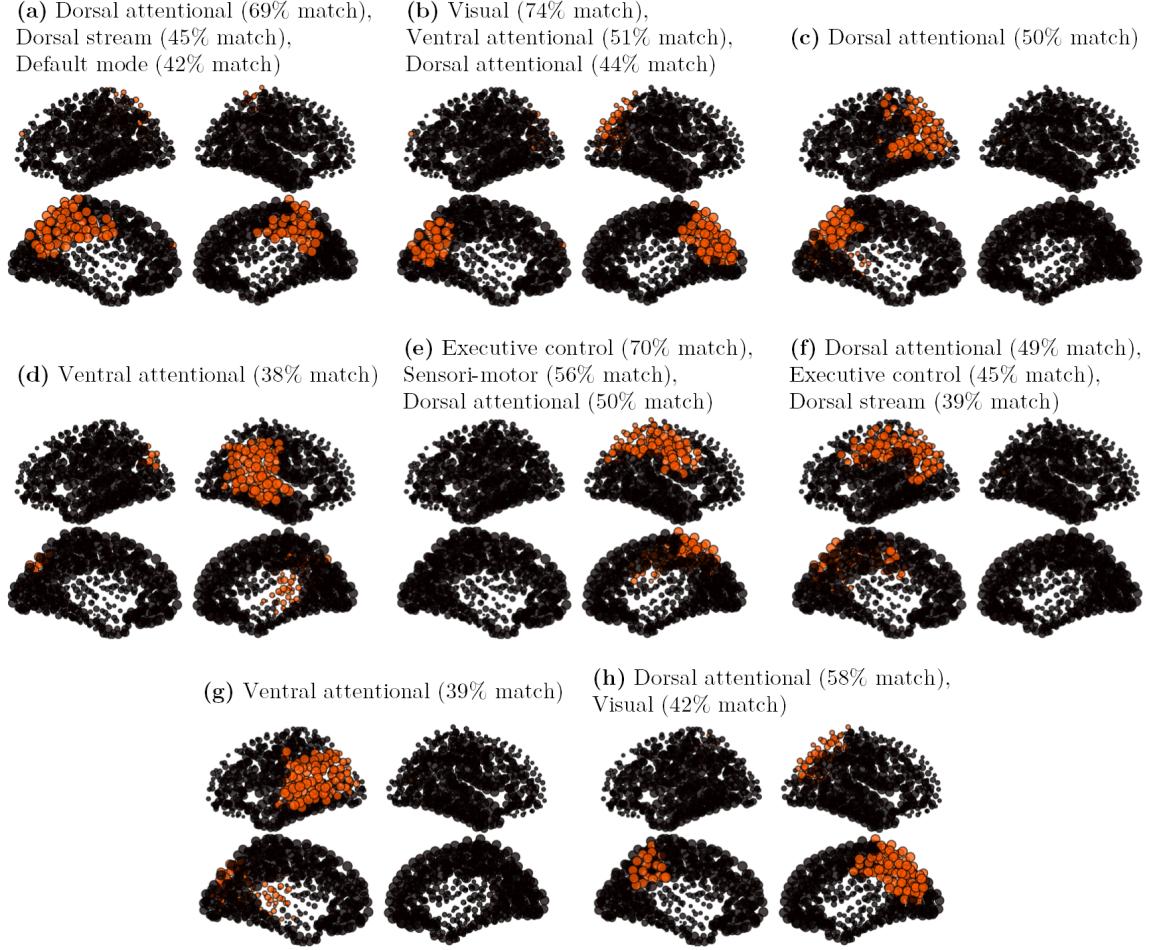


Figure 9. 8 clusters cores, on a DG attractor set. The initial attractors set contains 36,459 attractors, with $P \in [0.5; 7]$ and $G = 900$. Double-pass clustering, with inclusion match similarity ($k = 0.8$).

network, that regularly show up in the SL model, are barely present in the DG model. On the contrary, centromedial as well as superior and inferior parietal activations, characteristic of the dorsal stream and dorsal attentional networks, seem to dominate the DG attractor sets. Like in the SL model, a non-physiological “vertical” parieto-temporal network, barely corresponding with the ventral attentional network, is found. The reason of this artifactual network appearing may be an overestimation of the parieto-temporal links in the initial tractography, or an activation disbalance due to an underestimation of other links of the connectome (like the callosal fibers for instance, see also [45]). Bilateral large-scale and non-contiguous functional networks, like the Default mode network, are not reported in full by the simulation. This must not be considered a drawback for the subcortical activity and many bilateral callosal interactions are missing, while probably crucial in shaping the large-scale brain activity. Simulation results thus only provide substantial clues about large-scale affinities. Some of them, like the visual ventral stream (figures 8b and 8d), or the non-

contiguous executive control network (figures 8h), were not obviously seen from the graph-theoretical analyses.

Noisy dynamics

If only traditional functional connectivity is considered across the entire time series, then the major determinant of the resting state is the structural connectivity [45]. However, recently the non-stationary behavior at rest has been linked to a fast switching dynamics between epochs of temporarily invariant functional connectivity. Hansen et al [21] demonstrated that nonlinearity and multistability of network states are the necessary condition for the emergence of the switching dynamics of the functional connectivity, whereas the actual details of the intrinsic network node dynamics are less relevant. It is this property of the non-stationary Functional Connectivity Dynamics (FCD) that motivates our choice of our extended noise-driven graded-response Hopfield model.

We consider the case of noise-driven dynamics (see eqs 3–4), which are the stochastic generalization of the de-

terministic activity studied in previous sections. Noise-driven dynamics are not expected to relax on an attractor, but rather to implement infinite wandering across the total state space, where the convergence toward attractive states is counterbalanced by centrifugal random excursion. The existence of multiple attractor basins makes it possible however to estimate the stability of the different attractors, either theoretically when a Lyapunov function exists (see Material and Methods), or empirically by measuring the duration of stay in the vicinity of the different attractor basins.

Temporal multistability

We study in figure 10 the temporal behavior of the noise-driven extension of our 3 dynamical models (eqs 3–4), in the parametric range where a prominent multistability is expected (namely the regions of the parameter space displaying the highest entropy – see fig. 5). We separate the time scale of the activity and the thresholds, with $\tau_x = 10$ ms and $\tau_\theta = 80$ ms. The thresholds dynamics is thus slower than the activation dynamics, making possible the emergence of more complex dynamics. We choose a moderate level of noise ($\sigma_x = \sigma_\theta = 0.2$) in order to have the attractor basins close enough to the ones obtained in the deterministic case. Each simulation lasts 10 seconds: in the first 2 seconds, the dynamics is deterministic; then the noise is turned on for the remaining time (from 2 s to 10 s). For each dynamical system, 1000 initial conditions are chosen among the set of final attractors obtained after random sampling search (see previous sections). Each gray line corresponds to the time course of the average activity for a different initial condition. The red line corresponds to a particular time course we picked up for its large temporal variability. The corresponding 998-nodes spatio-temporal time course of the nodes potentials is shown below in color code.

Consider first the SL case (Hopfield model) (fig. 10a and 10c). The overlapping noisy trajectories form a fuzzy cloud that covers almost the totality of the [0,1] interval, reflecting the strong variability of the average activities. The darker the gray, the stronger the density of trajectories. The two dark lines at the high and low levels of activity indicate a significant attraction toward the “Up” and “Down” states. Intermediate trajectories are however maintained, with random excursions around the varied intermediate patterns. The red time course represents a typical multistable trajectory. Several steps are observed that systematically push the trajectory toward stronger levels of activity, finally reaching the “Up” state and remaining on it. The spatio-temporal time course (fig. 10c) displays the progressive recruitment of more and more nodes, issuing a pattern having most of its nodes active. Other observations of temporally multistable trajectories (not shown) confirm the systematic drift toward one of the two trivial attractors. Either by recruiting new nodes, the dynamics converges toward the high activity state, or by

progressive extinction, the dynamics converges to the low-activity regime with nodes displaying a noisy activity at low rates. Long-lasting temporal multistability appears thus difficult to implement in this model. Despite their number and variety, the intermediary attractors appear not stable enough to maintain a realistic switching activity in the long run.

In the SG case (fig. 10e and 10g), the noisy dynamics simulations offer a slightly different picture. Most of the initial conditions correspond to very sparse (i.e. very local) activation patterns. The introduction of noise at 2 seconds leads to a strong remapping of the activity, with strong contrast against the set of attractors obtained by sampling. In some cases, the dynamics converges toward the “Down” state and remains stuck on it (lower line). In most of the cases, the average activity gradually increases toward an average activity between 0.2 and 0.5. No tendency toward higher activities is observed. By inspection of a spatio-temporal trajectory (fig. 10g), a clear temporal multistability is obtained, with different spatial patterns assembling and disassembling over time. A strong core activity in the centro-medial cortex is observed (cuneus, precuneus, paracentral lobule, posterior cingulate...), with a possible occipital and inferior parietal component, and a variable number of nodes activate in rest of the cortex (with an apparent tendency toward bilateral activation).

In the DG case (fig. 10f and 10h), a consistent multi-stable regime is obtained for various values of P . The most temporally varied trajectories are obtained for $P = 0.6$, which corresponds to a very sparse level of activation (around 0.05) and to the leftmost limit of the multistability region (see fig. 5). Every initial condition results in an apparent low activity regime. When considering a particular trajectory (red trajectory), small excursions to the “Down” activity are observed. When spatially displayed (fig. 10h), distinct spatial patterns, with similar density, are visited during a single trajectory. Because of the low amplification, the activity is prone to fall to the “Down” state, which plays the role of a reset facilitating excursion toward a new attractor basin.

Those exploratory simulations have shown that spatial multistability, as observed in the previous sections, can convert to temporal multistability when a moderate noise is introduced in the dynamical system. In the static threshold models, intermediate activity patterns remain stable enough to provide effective temporal multistability. The attraction of the “Up” and “Down” is enhanced by noise and a long term tendency toward either overactivation or extinction is observed. The DG model, having unstable extremal patterns, displays the most consistent multistable behavior : varied patterns of similar density are explored during the same time course, with temporary excursions toward the “Down” state (“reset” state), and no long-term tendency toward overactivation nor extinction.

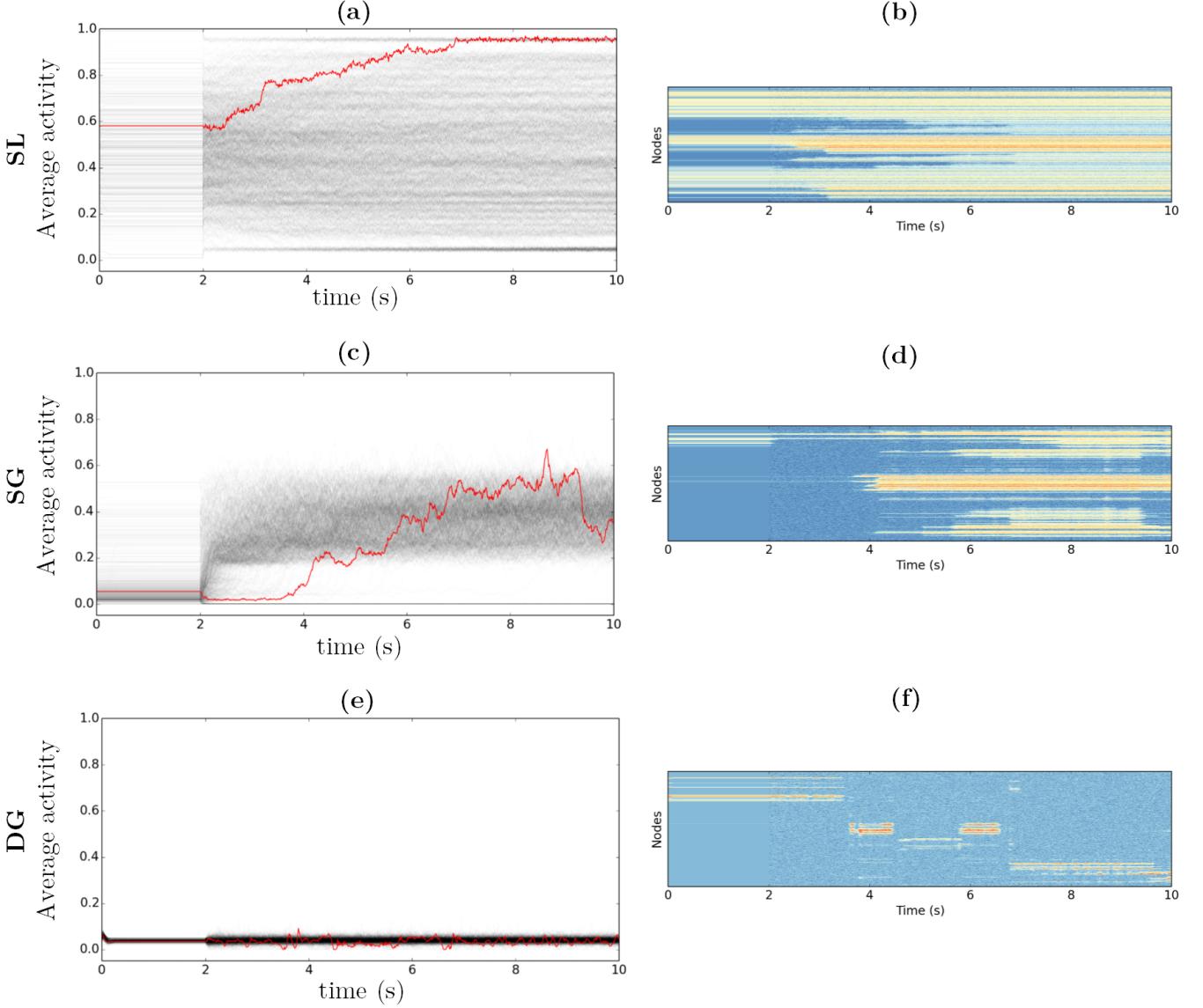


Figure 10. Noisy dynamics time course. The dynamics is noiseless for 2 s, and noisy for the rest of the simulations. $G = 900$, $\tau_x = 10$ ms, $\tau_\theta = 80$ ms, $\sigma_x = \sigma_\theta = 0.2$. **a, b:** SL model, $P = 1$. **c, d:** SG model, $P = 0.86$. **e, f:** DG model, $P = 0.6$. **a, c, e:** gray lines : 1,000 average time courses – red line : sample average time course. **b, d, f:** sample spatio-temporal time course. Color code : from blue (low potential) to red (high potential).

Functional Connectivity Dynamics

Simulation-based approaches allow to create synthetic time series that can be treated in the same manner as empirical data. Real and synthetic signals are not expected to fit in time, and need to be compared through first and second order statistics reflecting the consistent co-activations of the different nodes across time. The Functional Connectivity Dynamics (FCD) is a novel metric capturing the non-stationary aspects of the functional connectivity [7,21] (see Material and Methods). The FCD provides indications about the ultra-slow changes in the global organization of the dynamics in the range of minutes (10^{-2} Hz

variations), consistently with [7,29] expected to reflect behavioral changes.

A FCD calculated on a single subject rs-fMRI signal is compared on figure 11 with two synthetic FCD's issued from the SL and the DG models. In the physiological FCD (fig. 11a), each row shows a typical fluctuation of correlation across the time axis with regards to the temporal reference on the diagonal, with alternations of high ($\simeq 0.5$) and low ($\simeq 0.3$) correlations on large temporal intervals. Those alternations were also obtained in simulation in very specific parametric ranges on the SL and DG models. In the SL model, it is observed at the critical gain $G_c = 12.6$.

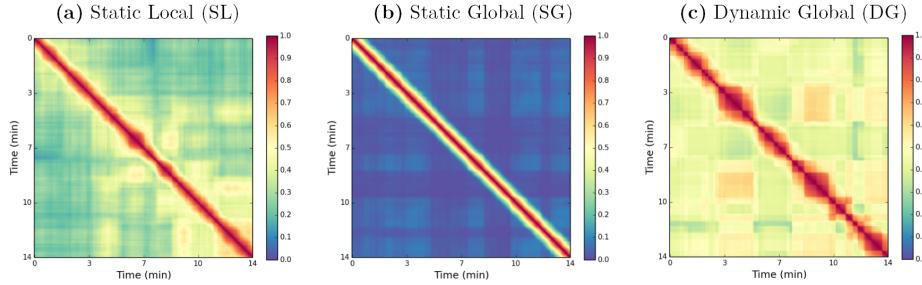


Figure 11. Functional Connectivity Dynamics matrices. **a:** from the empirical BOLD data of one healthy subject (15 min session). **b:** from a 15 min simulated BOLD signal on the SL model, with $P = 1$, $G = 12.6$ and $\sigma_x = 0.1$. **c:** from a 15 min simulated BOLD signal on the DG model, with $P = 0.75$, $G = 900$, $\sigma_x = \sigma_\theta = 0.04$.

For generating the synthetic BOLD signal, we first binarize the activity and then apply the convolution with the Balloon/Windkessel kernel. In that case, long-lasting alternations between the "up" and "down" patterns are observed, visible on the resulting FCD. However, the low correlations observed outside the diagonal ($\simeq 0.15$), reflect stronger than realistic changes in the FC organization across time. Similar FC variations are observed in DG model for a large span of P values ($P \in [0.5; 2.5]$). The most robust alternating behavior is obtained around the lowest bifurcation ($P = 0.75$, see fig. 11c). For moderate levels of noise, a more constant FC organization is maintained across time, with correlation values alternating between $\simeq 0.3$ and $\simeq 0.6$. In the two models, the most realistic ultra-slow alternations are observed near the critical values of the control parameters. The low recurrence of FC states observed in the SL dynamic can however be considered less realistic than the one observed in the DG dynamics.

Discussion

Here we provide a systematic analysis of the connectome attractor landscape in a noise-free condition and establish a pattern-to-pattern comparison between a series of BOLD signals and a series of attractors generated from a corresponding connectome. The variety of attractor landscapes is enhanced in our case by considering several variants of the initial "graded-response" Hopfield model [1], either by relaxing the local threshold constraint, or by enriching the network dynamics through adaptive thresholds imposing constraints on the attractor density (average number of active nodes). All three models we consider display a robust multistability in a clearly defined parametric range, while providing contrasted results regarding the distribution, density and temporal behavior in noise-free and noisy conditions.

First, our attractors sampling approach identifies the parametric range under which a robust multistable behavior is obtained. Large attractors sets are obtained at

the high gain (low spin-glass temperature) condition, every node having its activity either close to 0 or 1. This large number of static attractors exceeds by several orders the size of attractor sets generally reported in connectome based simulations [19, 23]. The number of attractors in a given network model may subserve the network capability to attain various functional configurations [46], which has been termed the dynamic repertoire [8, 17]. Previous works on the resting state dynamics that have explicitly extracted and counted network states have been typically concerned with equilibrium states of the network [21, 23]. Here we have explored the network mechanisms influencing the network's capacity to store patterns and demonstrated that the attractor number can be expanded by four orders of magnitude beyond so far reported number of patterns and that this attractor state can indeed be explored when driven through noise. The difference with the other models lies, on one side, in the local thresholds setting, introducing a symmetry that helps the system to visit every possible stable configuration. It also relies on a formal separation of the model free parameters in a gain (G) controlling the node excitability and a scaling factor (P) controlling the excitatory/inhibitory ratio. Those two parameters are found to independently control the multistability, with a cascade of bifurcations of the Pitchfork type observed when controlling G , and a non-monotonic behavior, reminiscent of the more detailed mean-field models behavior [23], when controlling P . We moreover show this strong multistability to be consistently obtained on different connectome datasets, while considerably weakened when using randomized versions of the initial connectomes, this finding attesting the idea of an intrinsic small-world/scale-free structure of the connectome [2, 26, 47], providing a support to the multistability, as already suggested by [19].

Second, clear spatially-segregated components are identified from the attractors distributions. Sets of principal modes are obtained, in different proportions, across the different models, using a specific "inclusion match" clustering analysis. The strongest complexity in number, variability and spatial extent is obtained on the orig-

inal Hopfield graded-response (SL) model. Our clustering analysis report specific large-scale networks, displaying a strong similarity with the independent components [10, 36] or communities [41] obtained from fMRI resting state time course analysis. Some of them (frontal pole core, centro-medial core, primary visual core) were already identified in graph-theoretical approaches [2], while others, like left and right ventral visual streams, or a left-lateralized parieto-frontal network, were not observed in the graph-theoretical approaches. In contrast, adaptive-threshold based variants of the Hopfield model (SG and DG), although displaying sparser activity patterns, provide a lesser variability, with the centro-medial component dominating the samples, representing between 50 and 75 % of the total attractor sets. This central activity is reminiscent of the default mode network [9], although having the orbitofrontal component absent in that case. The other modes include a primary visual modality, a superior parietal element, a pre/post-central medial element, and a superior-temporal element. Although stemming from different principles, those different modes share similarity with the “regional modules” identified in [2]. Attractor cross-correlation matrices are compared with the rs-fMRI-based functional connectivity matrices, also providing contrasted results when comparing the different models, with an advantage to the SL model in providing predictions on both intra-and inter-hemispheric functional connectivities. A first conclusion from these comparisons is the more flexible and more widespread-distributed attractor sets in the SL model, providing a better account of the large-scale structural organization than models having a tighter control of the nodes average activity.

Third, when noise is introduced in the dynamics, a temporally multistable behavior is obtained (with alternating meta-stable attractors visited along the same time course) in the parametric ranges corresponding to the maximal spatial multistability (and/or maximal entropy) in the noise-free condition. This observation is confirmed on every model under consideration. Noise-driven exploration of the brain’s dynamic repertoire has been hypothesized to play an important role in the execution of cognitive functions and serve as a biomarker of aging [48]. It is for this reason that the relation between network capacity and noise strength is significant. Noise causes a large proportion of attractors to vanish and become invisible, leaving space to a much smaller attractor sets, including trivial attractors like the “Up” (full brain activation) and “Down” (full brain deactivation) sets. In the absence of density control, a “centripetal” tendency toward either the “Up” or “Down” state is observed. Only the density control case (DG model), imposing stable density across time, provides a condition where no tendency toward overactivation or extinction is observed. Only this central control of the average activity seems capable of maintaining a temporal multistability in the long run, this highly flexible dynamics being moreover obtained, like in [23], in a parametric range close to the bifurcation point. The existence

of a central control in the brain is of course highly conjectural, while hypothesized in some studies to be present in the claustrum, which is a small region near the insula that is expected to coordinate distant synchronized activities [49, 50]. Nevertheless, our findings demonstrate that a dynamic feedback mechanism (here through the dynamic adaptation of the thresholds) has the capacity to significantly enhance the complexity of the time courses and qualitatively capture the non-stationary behavior of brain networks at rest.

In conclusion, we propose a novel approach to large-scale brain simulation that encompasses a capability to generate large sets of spatially-distributed attractors, reminiscent of well-known resting state networks, and theoretically interpret the parameters that control the dynamics. Sensible differences in the final distribution of activities are induced in case of local or global control of the nodes activity. More precisely, physiologically plausible temporally alternating dynamics are obtained under a global control of the average activity (DG model), at the detriment of more physiological-relevant distributions of activities in the original Hopfield graded-response model (SL model). The multistable behavior is obtained on a large parameter range, but the best fit with the ultra-slow functional connectivity dynamics, as observed in the BOLD time courses, is obtained at the edge of multistability, a parameter region that also corresponds to the highest entropy of the attractors distribution. The general conclusion is the importance of the noise-free dynamics in analyzing the attractors landscape, for identifying high-multistability/high entropy parameter regions that both fit with the most physiological distributions of activity, and the most relevant time courses in the noisy condition.

Material and Methods

Structural connectivity

In the case of large scale (brain-wide) simulations, the data set to analyze is a connection matrix, obtained by tractography. The nonzero weights represent the diameter of the axonal tracts, expected to reflect the strength of the bilateral interactions between several mm³-wide regions of the cortex. Due to the technique utilized, the matrix is symmetric and the weights are positive (since large-scale connections are generally considered excitatory).

The structural matrix we consider is extracted from DSI (Diffusion Spectrum Imaging), realized on five healthy subjects, then treated with tractography [2]. Diffusion MRI gives informations on fibers orientations in vivo with anisotropic diffusion of water in the brain (the anisotropy being mainly caused by the barrier created by myelin sheath insulating neurons axon). In parallel, 998 regions of interest (ROI) are extracted from a parcellation applied on a standard MRI. The resulting connectome is averaged from the most significant 10,000 connections of 5 individ-

ual connectomes after a k-Core decomposition (details are in [2]).

Null models are generated using a randomization algorithm (Maslov and Sneppen [51]) that preserve the degree of the node (random switch of the non-null edges).

Functional connectivity dynamics (FCD)

FCD computes the functional connectivity for windowed portions of the time series and then calculates the Pearson cross-correlation for time-shifted consecutive FC matrices. The FCD is visualized as a correlation matrix over the time shifted values. Time series of 15 minutes were computed on the SL and DG models at 998 nodes resolution. The signals were convolved with a Balloon/Windkessel kernel [52–54], resulting 0.5 Hz synthetic BOLD time course. First, a 1 minute sliding window is used to generate on each link a series of temporally indexed Pearson correlation values, resulting a $N \times N \times T$ matrix, where T is the total duration minus the time window. Second, a subset of n nodes is selected (for instance the upper triangle of the matrix excluding the diagonal on figure 11). Third, a $T \times T$ matrix is calculated where each (t_1, t_2) couple reports the correlation among the two n -values vectors indexed at time t_1 and t_2 . This $T \times T$ matrix is referred as the Functional Connectivity Dynamics matrix (FCD).

Spatial vs. temporal multistability

Fixed-point multistability (or “spatial multistability”) is the property of a dynamical system to reach different final static attractors stemming from different initial conditions (otherwise it is said monostable). The connectome being structurally closed, it can connect a recurrent network composed of mutually-feeding excitable nodes, and offers a natural support for fixed-point multistability in a fashion similar to the spin-glass model [30] and corresponding auto-associative memories extensively studied in the 80’s [55]. Hopfield’s seminal paper popularized this concept as a key element of associative memory, introducing the tools and ideas of statistical physics to neural networks modeling. Besides, the stability of the attractors has been shown to decrease with the number of items stored in the network [31], and the maximal number of possible memories has been shown to linearly grow with the actual number of neurons [56], or with the number of synapses in the large dilution limit [57]. The extension of the initial model to graded neuronal responses [1] qualitatively provides similar properties [58].

In contrast to spatial multistability, “temporal multistability” refers to the case when a single temporal trajectory displays transient dynamics across several “pseudo”-attractor basins. Temporal multistability can be obtained by various means in neural networks, either by implementing bistability at the local nodes level [21, 59, 60], or introducing a noisy drive [61–64] to a system initially dis-

playing fixed-point multistability. This behavior is sometime called “attractor itinerancy” [61, 65] or “winnerless activity” [60, 62], and the corresponding set of unstable orbits are called “Milnor attractors” [66]. Metastable systems have been implemented in various control architectures (e.g [67–69]), allowing biologically-plausible sequential switching between elementary behaviors, and bringing more flexibility to unexpected environmental changes. Itinerant dynamics is thus an appealing concept in neurosciences [8], accounting for the versatility of the brain activity observed both in the presence or in the absence of stimulation, where each transiently stable attractor may be interpreted as a different functional network, representative of an underlying large-scale functional connectivity.

Models

The original Hopfield model [55] is a discrete time dynamical system composed of multiple nodes, each node owning a distinct binary state variable. The interactions between the nodes rely on a connectivity matrix built from a pre-existing set of prototypes. The update, inspired from the spin-glass [30] model (used to describe magnetic properties of dilute alloys), is based on a random scanning of every node. The existence of a fixed point dynamics is guaranteed by a Lyapunov function. In his seminal paper [55], Hopfield gave the conditions under which every prototype is expected to be an attractor of the dynamics (multistable dynamical system): when the initial conditions are close enough to one of the prototypes, the dynamical system is expected to relax on a corresponding attractor. A possible output of the system is the final attractor (interpreted as a recollection from a set of stored memories).

Various extensions of the initial model have been proposed, among which a “graded-response” version [1] having a similar Lyapunov function when the diagonal is null, showing the same multistability properties, with a more straightforward physiological interpretation. In its graded-response form, the dynamical system is described by its N state variables x_1, \dots, x_N . The connections between the nodes are described by a matrix W , this matrix being symmetric by construction. In the following, we consider the weights matrix to be normal, i.e. $\|W\| = 1$. The weights matrix we use in particular is a normalized connectome, i.e. $W = C/\|C\|$ where C is the original connectome.

The state update is based on a weighted sum across the activity of every input nodes, with a linear decay (see eq. 1a), where τ_x defines the temporal scale of the dynamics. The output of a node is calculated with a monotonic sigmoid transfer function with activation threshold θ_i , so that the output of a node is bounded between 0 (no output) and 1 (strong output) (see eq. 1b).

$$\tau_x \frac{dx_i}{dt} = -x_i + \sum_{j=1}^N W_{i,j} A_j \quad (1a)$$

$$A_i = \frac{1}{2}(1 + \tanh(G(Px_i - \theta_i))) \quad (1b)$$

where x_i is the node potential, A_i the node output, W the connectivity matrix, θ_i the threshold, P the scaling factor, G the gain and τ_x the time constant.

The initial Hopfield model was not aimed at modeling brain architectures but rather studying the storage and retrieval capacities of recurrent neural networks. The continuous version has however strong similarities with the Wilson-Cowan model [70], and can be encompassed in the family of neural mass models where the activity of a node is interpreted as the (normalized) average firing rate of a population of neurons, assuming that every single neuron behaves on average like the field. The continuous state variable x is then interpreted as an average membrane potential in a population, the time constant τ_x as a membrane time constant (of the order of few milliseconds) and the connections weights between populations as the average synaptic weights.

The activation threshold is an essential component of the model. Classical Hopfield models use a different threshold on every node (Static and Local) (eq. 2a). The threshold values are calculated in order to compensate for the average activation. Note that only the $P = 1$ case exactly corresponds to the original Hopfield model, in which case the dynamical system is symmetrical around $A_i = 0.5$, with each node having an equal probability to be active or inactive (provided its local weights sum is different than 0). As soon as $P \neq 1$, the activity is biased, either toward a higher proportion of nodes active for $P > 1$, or a lesser proportion for $P < 1$.

We consider here several variants of the model, namely the case of global thresholds (Static and Global – SG) (eq. 2b), and the case of both a global and dynamic threshold (Dynamic and Global – DG) (eq. 2c).

$$\theta_i = \frac{1}{2} \sum_{j=1}^N W_{i,j} \quad SL \quad (2a)$$

$$\theta = \frac{1}{2N} \sum_{i=1}^N \sum_{j=1}^N W_{i,j} \quad SG \quad (2b)$$

$$\tau_\theta \frac{d\theta}{dt} = -\theta + \sum_{i=1}^N A_i / N \quad DG \quad (2c)$$

The SG case (eq. 2b) is a natural simplification of the original Hopfield model. Replacing the local thresholds by a single one has strong implications, discarding the symmetry around the central 0.5 state. Like previously, each node is expected to be active when a large enough proportion of its inputs are active. In that setting however, as the

weight averages are not balanced between the nodes, some nodes with greater (resp. lower)-than-average weights input have a greater (resp. lower) probability to be active than the others. This should result in nodes being more (resp. less) frequently active in the final attractor patterns. In order to keep similar parametric ranges as in the first model, the global threshold is calculated as the average over the local thresholds of (2a).

In the “dynamic threshold” setup (eq. 2c), the threshold changes over time as a function of the node activities. The general idea is to decrease the threshold when the nodes activity is too low, and, on the contrary, to increase the threshold when the activity is too high. As such, the threshold participates, as new state variable, to the dynamics of the system, having a regulatory influence on the dynamics, i.e. controlling the average level of activity (in the spirit of e.g. [56]). The threshold is thus interpreted as a local population of inhibitory neurons (with linear response). In that case, the scaling parameter P represents the Excitatory/Inhibitory (E/I) ratio. $P < 1$ means that the inhibition dominates the excitation, and $P > 1$ means the contrary. Consistently with eq. (2a), the dynamic threshold is calculated as the average over every node activity, representing a spatio-temporal average, i.e. the average proportion of active nodes in the system. From a biological standpoint, it is interpreted as a global inhibitory node (with linear response) feeding back the average activity toward every excitatory node of the system (see fig. 2).

In order to study the temporal multistability behavior, we consider a stochastic generalization of the deterministic systems described previously. Namely, two noise terms are considered. First, on each node, additional independent continuous white noises η_i 's are added to the state variable x_i (eq. 3) with diffusion parameter σ_x . Second, independent additive threshold noises ξ_i 's are considered in the different threshold modalities addressed previously, with σ_θ representing the threshold diffusion parameter.

$$\tau_x \frac{dx_i}{dt} = -x_i + \sum_{j=1}^N W_{i,j} A_j + \sigma_x \eta_i(t) \quad (3)$$

$$\tau_\theta \frac{d\theta_i}{dt} = -\theta_i + \theta_i^0 + \sigma_\theta \xi_i(t) \quad SL \quad (4a)$$

$$\tau_\theta \frac{d\theta}{dt} = -\theta + \theta^0 + \sigma_\theta \xi(t) \quad SG \quad (4b)$$

$$\tau_\theta \frac{d\theta}{dt} = -\theta + \sum_{i=1}^N A_i / N + \sigma_\theta \xi(t) \quad DG \quad (4c)$$

The additive noise on state variables and thresholds results in two supplementary hyper-parameters, namely the node potential noise drift σ_x and the threshold noise drift σ_θ . In the static case, the thresholds are now expected to display a leaky excursion around their fixed point θ_i^0 (resp. θ^0), calculated from (2a) (resp. (2b)). In consequence, the static threshold setups also rely on a dynamical update in this case.

In contrast with deterministic systems, noisy systems are not expected to relax on a particular attractor. Relying on non-deterministic components, they are also expected to develop richer temporal behaviors than mere fixed-point dynamics. In the case of contracting (or dissipative) dynamical systems, the noise drift is expected to counterbalance the contraction of the dynamics. When the level of noise is low, the trajectories remain generally trapped in a particular attractor basin and resemble those of the corresponding deterministic systems. With higher noise levels however, the noisy trajectories can occasionally escape from attractor basins, displaying a capability to switch from one attractor basin to the other.

Numerical simulations

The computer simulations have been done with continuous time dynamics reduced to discrete steps of 0.1 ms, using Euler discretization scheme. The weights matrix is a normalized connectome as described in the "Structural connectivity" section. When not pointed out otherwise, the initial state is a random binary vector \mathbf{A}_0 , whose proportion of zeros and ones is set according to a density factor f_0 . The initial conditions are set randomly according to a binomial draw with an expectation varying from 0 to 1 according to f_0 . After initialization, the activity of the nodes is recurrently transmitted to the other nodes through the structural connectivity matrix. In the noiseless case, the dynamics is expected to relax on a stable attractor in short time. Let $\mathbf{A}^k(0)$ be the k^{th} vector of initial conditions and $\mathbf{A}^k(t)$ the corresponding vector of activity at time t . The average activity at time t (see figures 1a, 10a, 10c, and 10e) is defined as $\bar{A}^k(t) = \frac{1}{N} \sum_{i=1}^N A_i^k(t)$.

The free parameters of the system need to be set prior to simulation. We have 4 free parameters in the deterministic case, namely τ_x , τ_θ , P and G , and two additional parameters in the stochastic case, namely σ_x and σ_θ .

Two of those parameters were specifically disentangled for our study, namely P , which is a scaling factor applied to the node potential, and G , the gain of the transfer function. $Px_i - \theta_i$ is the effective potential and $G/2$ is the effective slope of the transfer function. In our setting, the scaling parameter P is expected to tune the average level of activation needed to activate the node. For instance, assuming $\|W\| = 1$ and with $P < 1$, the node is expected to be active when more than half of its inputs are active. Conversely, with $P > 1$, the node is expected to be active when less than half of its inputs are active. More generally, the lower P , the lower the chance for the node to be active, the lower the expected density of the final attractor.

The network node activity is interpreted as a mean-field representing the underlying physiological activity. The network dynamics thus operates with physiological time constants taking place in the milliseconds range. The state variable time constant is set to a fixed value $\tau_x = 10$ ms, by analogy with the membrane time constant of individual

neurons. The threshold time constant τ_θ , mostly representing inhibitory influences, is considered slower and varied from 10 ms (identical time scales) to 80 ms (time scales separation). The scaling factor P is a pivotal parameter interpreted as the system excitatory/inhibitory balance. Depending on the model, P varies from 0 (no excitation) to 20 (strong excitation). Last, the gain G shapes the node outputs from soft nonlinearity 1 to strong nonlinearity 900 (the final activity being close to binary in the last case).

In the noisy dynamics case, the two noise drift parameters σ_x and σ_θ are set in the $[0, 1]$ interval, making a total of 6 free parameters.

In the dynamical systems we consider, a node is said active when its activity is greater than 0.5, and inactive the other way. The self-sustained activity is the capability of the dynamical system to keep a certain proportion of its nodes active in the long run. In a self-sustained regime, the internal activity is strong enough to recurrently excite the same nodes over and over without additional stimulation. The stronger the strength of the self-excitatory weights, the higher the chance of self-sustained activity to take place. With normalized weights, the excitatory strength mostly depends on the scaling factor P , with a higher P corresponding to denser self-sustained patterns of activity.

Random sampling-based attractor search

Following the analogy with the Hopfield neural network, we interpret the connectome as a dynamical system, having several attractors "stored" in its structure. We use the following reverse approach: given a particular connectivity matrix, we try to infer the set of prototypes embedded in it. In the absence of prior knowledge, we sample the initial conditions space.

We randomly initialize the system with binary activation patterns (namely active or inactive node). The number of initial configurations growing exponentially with the size (i.e. 2^{998} configuration available for discrete initial states), we only take a sample from a subset of possible initial states. Importantly, we consider different initial densities in order to better sample the state space, varying from 0.02 to 0.98 with 0.03 steps (here 33 densities distributed linearly on the $[0, 1]$ interval – excluding 0 corresponding to a trivial solution for the continuous Hopfield model). In the DG model, the threshold time constant is set to the same value than the potential time constant, for it has no influence on a final static attractor. Then, for each density, we let n different initial conditions relax on their attractor. For each value of the parameter space to study, $33 \times n$ random initialization are thus processed. In the particular case where the initial density is a parameter to study, $n = 3,300$ initializations are done on a given density parameters. The dynamics is stopped when the system reaches its equilibrium $\frac{\langle \bar{x}(t) \rangle_T - \bar{x}(t)}{\bar{x}(t)} < \varepsilon$ or after 1 second if the equilibrium condition is not reached, where $\bar{x}(t)$ is the average potential over the nodes (at time

$t)$, $\langle \bar{x}(t) \rangle_T$ is the temporal average of $\bar{x}(t)$ on the $[t - T, t]$ temporal interval ($T = 100$ ms), and ε a (small) numerical constant (10^{-6}).

Any final pattern of activity satisfying a double dissimilarity condition with any previous attractor is saved, otherwise the cardinality of the better corresponding attractor set is incremented. The double dissimilarity condition considers both a Pearson correlation *and* a Euclidean similarity lower than 0.9. The goal of this double dissimilarity condition is to discard the low-density patterns, possibly having a low correlation but a high Euclidian similarity, and the high density patterns possibly having a lower Euclidean similarity but nevertheless being strongly correlated. A final set of $m \leq n$ attractors is obtained, each attractor being associated with its cardinality, representing the weight of its attraction basin.

$$\begin{aligned} \text{sim}_{\text{cor}}(\mathbf{x}, \mathbf{y}) &= \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{(\text{var}(\mathbf{x})\text{var}(\mathbf{y}))^{1/2}} \\ \text{sim}_{\text{eucl}}(\mathbf{x}, \mathbf{y}) &= \frac{1}{1 + (\sum_i (x_i - y_i)^2)^{1/2}} \end{aligned}$$

Clustering algorithm

With very large attractor sets, we additionally use a clustering algorithm (see Algorithm 1) in order to identify smaller groups of attractors sharing strong spatial similarity. Considering a set of n final attractors $\mathbf{x}_1, \dots, \mathbf{x}_n$, the objective of a clustering algorithm is to assign a cluster (a set) to each of them, where k is a similarity threshold such that the similarity between separate clusters should be less than k . The clustering algorithm operates on binary patterns sets. A binary pattern is composed of active nodes (having their output > 0.5) and inactive nodes. Each binary pattern thus defines a set of active nodes A (and a complementary set of inactive nodes \bar{A}). For the purpose of extracting large-scale structural invariants, we use a specific “inclusion match” metric that indicates which proportion of a binary pattern A is included in a binary pattern B :

$$\text{incl}(A, B) = \begin{cases} \frac{|A \cap B|}{|A|} & \text{if } |A| > 0 \\ 0 & \text{if } |A| = 0 \end{cases}$$

Using this metric, patterns of variable size, but sharing elements, can be considered similar. For instance, if all the elements of the pattern A are included in the pattern B , the inclusion match is equal to 1. The metric being non-symmetric, we use the following symmetrical similarity

$$\text{sim}_{\text{incl}}(A, B) = \max\{\text{incl}(A, B), \text{incl}(B, A)\}$$

indicating how much of one set is included in the other.

The similarity between two clusters E and F is the similarity between the reference patterns μ_E and μ_F of the two sets, i.e. $\text{sim}_{\text{ref}}(E, F) = \text{sim}_{\text{incl}}(\mu_E, \mu_F)$. The reference patterns can be calculated different ways (the average

pattern of the cluster, the one with the highest average similarity with all the others, etc.).

Algorithm 1 Clustering algorithm

```

1: parameter :  $k$ 
2: initialize  $n$  sets:
    $S \leftarrow \{\{\mathbf{x}_1\}, \dots, \{\mathbf{x}_n\}\}$ 
3:  $\text{sim}_{\text{test}} \leftarrow 1$ 
4: while  $\text{sim}_{\text{test}} > k$  do
5:    $E, F \leftarrow \underset{(E', F') \in S \times S}{\text{argmax}} \text{sim}_{\text{ref}}(E', F')$ 
6:    $\text{sim}_{\text{test}} \leftarrow \text{sim}_{\text{ref}}(E, F)$ 
7:   remove  $E$  and  $F$  from  $S$ 
8:    $G \leftarrow E \cup F$ 
9:   Add  $G$  to  $S$ 
10: end while

```

In the “double pass” clustering case, the algorithm 1 is applied twice to the same set of patterns. In the first pass, the reference patterns of the clusters are the patterns having the highest “inclusion score”, where the inclusion score is the sum of the inclusion match with all the other patterns of the cluster. In the second pass, the reference patterns are the average binary patterns, gathering the nodes that are active in more than 50% of the patterns. The first pass results in a set of clusters whose constituents have a variable density, but share an essential “core” activity that is present in all the patterns of a cluster. The second pass is a smoothing pass that gathers together the clusters being similar on average (where the final average patterns are then possibly composed of several “cores”).

BOLD signal reconstruction

When simulating the rs-fMRI dynamics, a strong low-pass filtering is applied in order to identify low-frequency components operating in the range of seconds. A first 100 Hz downsampling is applied (average values over 10 ms windows on the temporal signals). Then a synthetic BOLD signal is generated by convolving this signal with a Balloon/Windkessel kernel [52–54]:

$$H(t) = \exp\left(\frac{-0.5 t}{\tau_s}\right) \frac{\sin\left(t \sqrt{\frac{1}{\tau_f} - \frac{1}{4 \tau_s T}}\right)}{\sqrt{\frac{1}{\tau_f} - \frac{1}{4 \tau_s T}}} \quad (5)$$

with $T = 10$, $\tau_s = 0.8$ s, $\tau_f = 0.4$ s. The resulting signal is finally downsampled it at 0.5 Hz and detrended (by suppressing the time-averaged value).

Potential function

Our modeling approach allows determining analytically a potential function, from which many dynamic and stochastic properties of the network can be derived in a simple manner. We begin our discussion with the formulation of the Fokker-Planck equation and its solutions. As

the resting state dynamics evolves, it traces out a trajectory in a M -dimensional state space. The M -dimensional state vector $\mathbf{q}(t) = (\dots q_i(t) \dots)$ obeys the Langevin equation $\dot{\mathbf{q}}(t) = \mathbf{K}(\mathbf{q}(t)) + \mathbf{F}(t)$ where the rate of change of the state vector (time derivative on the left) depends on its deterministic influences $\mathbf{K} = (\dots K_i(t) \dots)$ that are non-linearly dependent on its current state and stochastic forces $\mathbf{F} = (\dots F_i(t) \dots)$. Here we consider only δ -correlated fluctuating forces \mathbf{F} , i.e. $\langle F_i(t)F_j(t') \rangle = Q_{i,j}\delta(t - t')$. The probability density function $f(\mathbf{q}, t)$ defines the distribution realizations of trajectories in \mathbf{q} -space and its dynamics is determined by the Fokker-Planck equation.

$$\dot{f} = -\nabla_{\mathbf{q}} \cdot \{\mathbf{K}f\} - \frac{1}{2} \sum_{k,l} Q_{k,l} \frac{\partial^2 f}{\partial q_k \partial q_l} \quad (6)$$

We can rewrite the Fokker-Planck equation as a continuity equation by means of the abbreviation $j_k = (K_k f - \frac{1}{2} \sum_l Q_{k,l} \frac{\partial f}{\partial q_l})$ and obtain $\dot{f} = -\nabla_{\mathbf{q}} \cdot \mathbf{j}$, where the temporal change of the probability density $f(\mathbf{q})$ is equal to the negative divergence of the probability current $\mathbf{j} = (\dots j_k \dots)$. For $\dot{f} = 0$, the stationary solution of the Fokker-Planck equation is time-independent. The Fokker-Planck equation does not imply zero current, $\mathbf{j} = \mathbf{0}$, because conditions of detailed balance could be realized, in which closed probability flows may persist. Nevertheless, when the deterministic components may be expressed via a gradient dynamics as in the case of the SL and SG models ($\mathbf{q} = \mathbf{x}$, $M = N$), then this allows us to determine an explicit stationary solution of the Fokker-Planck equation. In particular, when $K_k = -\frac{\partial V(x)}{\partial x_k}$ and the diffusion coefficients obey the condition $Q_{k,l} = \delta_{k,l}Q$, then the time independent probability density function reads:

$$f(\mathbf{x}) = n \exp(-2V(\mathbf{x})/Q) \quad (7)$$

where n is the normalization coefficient, satisfying the natural boundary conditions that $f(\mathbf{x})$ vanishes for $|\mathbf{x}| \rightarrow \infty$. For our large-scale network, the potential function $V(\mathbf{x})$ reads:

$$\begin{aligned} V(\mathbf{x}) = & \frac{1}{2} \sum_i \left[x_i^2 - x_i \sum_j W_{i,j} \right. \\ & - x_i \sum_{j \neq i} W_{i,j} \tanh(G(Px_j - \theta_j)) \\ & \left. - W_{i,i} \frac{1}{GP} \ln \cosh(G(Px_i - \theta_i)) \right] \end{aligned} \quad (8)$$

Stationary distribution

The minima of the potential function $V(\mathbf{x})$ determine the set of potential states that the large-scale brain network can occupy and thus defines its dynamical repertoire. The stationary time independent solution of the probability density allows characterizing the most likely paths to be

taken in state space by identifying regions of high probability. This analytical approach turns out to be a major advantage as opposed to computational approaches, which would require time-consuming simulations of many realizations of trajectories. The following crucial distinction needs to be made here: what is commonly referred to as non-stationary brain dynamics at rest does generally not refer to a non-stationary dynamics of the probability density, but rather to the evolution of the trajectory in state space occupying certain subspaces for a finite time, followed by a rapid switch to occupy another subspace and dwell there for another characteristic time. The realized probability density is stationary and described by equation (7). Time-dependent solutions of the Fokker-Planck equation cannot be expressed analytically in general terms, but formulated at least locally around a given brain state (minimum of $V(\mathbf{x})$), which allows the explicit computation of the time dependent moments such as the time-dependent mean around a brain state or its two-time correlation function [71].

For a large excitability G , the sigmoidal saturation function in our network model approximates a Heavyside function which approximates a reduced Ising-spin attractor model, which allows further analytical investigation as performed by Deco et al [19]. The model is then a network of stochastic binary units (spins), where each unit s_i takes output value $A_i = 1$ with probability f_i and value $A_i = 0$ with probability $1 - f_i$. For symmetric connectivity, the Boltzmann-Gibbs distribution giving the probability of finding the network in a specific state \mathbf{A}^α can be expressed analytically by

$$p^\alpha = \frac{e^{-\beta E^\alpha}}{Z}. \quad (9)$$

where Z is the partition function defined by

$$Z = \sum_\alpha e^{-\beta E^\alpha} \quad (10)$$

and E^α is the energy function

$$E^\alpha = -\frac{1}{2} \sum_{i,j} W_{i,j} A_i^\alpha A_j^\alpha - \sum_i \theta_i A_i^\alpha \quad (11)$$

The probability p^α gives the probability of finding the configuration \mathbf{A}^α . Therefore, in order to describe the attractor landscape of the spin network, we can characterize the existence and probability of each possible attractor (here corresponding to a specific configuration \mathbf{A}^α) by the entropy of the system, which can be derived analytically, yielding:

$$H = -\sum_\alpha p^\alpha \log p^\alpha = \frac{\sum_\alpha \beta E^\alpha e^{-\beta E^\alpha}}{Z} + \log Z \quad (12)$$

Deco et al. [19] computed explicitly different types of structural networks and investigated how the entropy of

the Ising-spin network evolves as a function of connectivity. The higher is the entropy, the higher is the number of ghost attractors that efficiently structure the fluctuations of the system at the edge of the bifurcation.

The empirical entropies (shown in figures 4 and 5) rely on the cardinality of the different sets of attractors obtained after sampling the initial condition space. Namely, for each attractor i , $\tilde{p}_i = \frac{n_i}{n}$ where n_i is the cardinal of the i^{th} attractor set and n is the number of samples. Then, $\forall i$:

$$\tilde{H} = - \sum_i \tilde{p}_i \log \tilde{p}_i \quad (13)$$

Acknowledgements

Thanks to O. Sporns (Indiana University) for sharing the connectome data and T. Proix (INS, Marseille) for support in tractographic analysis and connectome generation [34].

References

1. Hopfield JJ (1984) Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences of the United States of America* 81: 3088–3092.
2. Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, et al. (2008) Mapping the structural core of human cerebral cortex. *PLoS biology* 6: e159.
3. Sporns O, Tononi G, Kötter R (2005) The human connectome: a structural description of the human brain. *PLoS computational biology* 1: e42.
4. Friston K, Frith C, Liddle P, Frackowiak R (1993) Functional connectivity: the principal-component analysis of large (pet) data sets. *Journal of cerebral blood flow and metabolism* 13: 5–5.
5. Biswal B, Zerrin Yetkin F, Haughton VM, Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic resonance in medicine* 34: 537–541.
6. Wibral M, Vicente R, Lizier JT Directed information measures in neuroscience .
7. Allen EA, Damaraju E, Plis SM, Erhardt EB, Eichele T, et al. (2014) Tracking whole-brain connectivity dynamics in the resting state. *Cerebral cortex (New York, NY : 1991)* 24: 663–76.
8. Deco G, Jirsa VK, McIntosh AR (2013) Resting brains never rest: computational insights into potential cognitive architectures. *Trends in Neurosciences* : 1–7.
9. Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, et al. (2001) A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America* 98: 676–82.
10. Damoiseaux J, Rombouts S, Barkhof F, Scheltens P, Stam C, et al. (2006) Consistent resting-state networks across healthy subjects. *Proceedings of the national academy of sciences* 103: 13848–13853.
11. Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, et al. (2011) Functional network organization of the human brain. *Neuron* 72: 665–678.
12. Wu GR, Stramaglia S, Chen H, Liao W, Marinazzo D (2013) Mapping the voxel-wise effective connectome in resting state fmri. *PloS one* 8: e73670.
13. Sporns O, Honey CJ, Kötter R (2007) Identification and classification of hubs in brain networks. *PloS one* 2: e1049.
14. Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience* 10: 186–198.
15. Rubinov M, Sporns O (2010) Complex network measures of brain connectivity: uses and interpretations. *NeuroImage* 52: 1059–1069.
16. Deco G, Jirsa VK, McIntosh AR (2011) Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nature Reviews Neuroscience* 12: 43–56.
17. Ghosh A, Rho Y, McIntosh AR, Kötter R, Jirsa VK (2008) Noise during rest enables the exploration of the brain's dynamic repertoire. *PLoS computational biology* 4: e1000196.
18. Deco G, Rolls ET, Romo R (2009) Stochastic dynamics as a principle of brain function. *Progress in neurobiology* 88: 1–16.
19. Deco G, Senden M, Jirsa V (2012) How anatomy shapes dynamics: a semi-analytical study of the brain at rest by a simple spin model. *Frontiers in computational neuroscience* 6.
20. Haimovici A, Tagliazucchi E, Balenzuela P, Chialvo DR (2013) Brain organization into resting state networks emerges at criticality on a model of the human connectome. *Physical review letters* 110: 178101.
21. Hansen EC, Battaglia D, Spiegler A, Deco G, Jirsa VK (2014) Functional connectivity dynamics: Modeling the switching behavior of the resting state. *NeuroImage* .
22. Rubinov M, Sporns O, Thivierge JP, Breakspear M (2011) Neurobiologically realistic determinants of self-organized criticality in networks of spiking neurons. *PLoS computational biology* 7: e1002038.
23. Deco G, Jirsa VK (2012) Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 32: 3366–75.
24. Fraiman D, Balenzuela P, Foss J, Chialvo DR (2009) Ising-like dynamics in large-scale functional brain networks. *Physical Review E* 79: 061922.
25. Marinazzo D, Pellicoro M, Wu G, Angelini L, Cortés JM, et al. (2014) Information transfer and criticality in the ising model on the human connectome. *PloS one* 9: e93616.
26. van den Heuvel MP, Stam CJ, Boersma M, Hulshoff Pol H (2008) Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain. *NeuroImage* 43: 528–539.
27. Musso F, Brinkmeyer J, Mobsacher A, Warbrick T, Winterer G (2010) Spontaneous brain activity and eeg microstates. a novel eeg/fmri analysis approach to explore resting-state networks. *NeuroImage* 52: 1149–1161.
28. Van de Ville D, Britz J, Michel CM (2010) Eeg microstate sequences in healthy humans at rest reveal scale-free dynamics. *Proceedings of the National Academy of Sciences* 107: 18179–18184.
29. Palva JM, Zhigalov A, Hirvonen J, Korhonen O, Linkenkaer-Hansen K, et al. (2013) Neuronal long-range temporal correlations and avalanche dynamics are correlated with behavioral scaling laws. *Proceedings of the National Academy of Sciences* 110: 3585–3590.
30. Sherrington D, Kirkpatrick S (1975) Solvable model of a spin-glass. *Physical review letters* 35: 1792.
31. Amit DJ, Gutfreund H, Sompolinsky H (1985) Spin-glass models of neural networks. *Physical Review A* 32: 1007.

32. Amit DJ, Gutfreund H, Sompolinsky H (1987) Statistical mechanics of neural networks near saturation. *Annals of Physics* 173: 30–67.
33. Van Essen D, Smith S, Barch D, Behrens T, Yacoub E, et al. (2013) The wu-minn human connectome project: An overview. *Neuroimage* 80: 62–79.
34. Proix T (2014). Surface and Connectivity Reconstruction: Imaging Pipeline & ToolS - SCRIPTS. <https://github.com/timpx/scripts>. [Online; 25-Mars-2014].
35. Sollich P, Tantari D, Annibale A, Barra A (2014) Extensive parallel processing on scale-free networks. *Physical review letters* 113: 238106.
36. Beckmann CF, DeLuca M, Devlin JT, Smith SM (2005) Investigations into resting-state connectivity using independent component analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360: 1001–1013.
37. Fox MD, Corbetta M, Snyder AZ, Vincent JL, Raichle ME (2006) Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences* 103: 10046–10051.
38. Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, et al. (2007) Dissociable intrinsic connectivity networks for salience processing and executive control. *The Journal of neuroscience* 27: 2349–2356.
39. Mantini D, Perrucci MG, Del Gratta C, Romani GL, Corbetta M (2007) Electrophysiological signatures of resting state networks in the human brain. *Proceedings of the National Academy of Sciences* 104: 13170–13175.
40. Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, et al. (2009) Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences* 106: 13040–13045.
41. Meunier D, Lambiotte R, Fornito A, Ersche KD, Bullmore ET (2009) Hierarchical modularity in human brain functional networks. *Hierarchy and dynamics in neural networks* 1: 2.
42. Raichle ME (2011) The restless brain. *Brain Connectivity* 1: 3–12.
43. Honey C, Sporns O, Cammoun L, Gigandet X, Thiran JP, et al. (2009) Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences* 106: 2035–2040.
44. Utevsky AV, Smith DV, Huettel SA (2014) Precuneus is a functional core of the default-mode network. *The Journal of Neuroscience* 34: 932–940.
45. Messé A, Rudrauf D, Benali H, Marrelec G (2014) Relating structure and function in the human brain: relative contributions of anatomy, stationary dynamics, and non-stationarities. *PLoS computational biology* 10: e1003530.
46. Jirsa V, Sporns O, Breakspear M, Deco G, McIntosh AR (2010) Towards the virtual brain: network modeling of the intact and the damaged brain. *Archives italiennes de biologie* 148: 189–205.
47. Zalesky A, Fornito A, Harding IH, Cocchi L, Yücel M, et al. (2010) Whole-brain anatomical networks: does the choice of nodes matter? *NeuroImage* 50: 970–83.
48. McIntosh AR, Kovacevic N, Lippe S, Garrett D, Grady C, et al. (2010) The development of a noisy brain. *Archives italiennes de biologie* 148: 323–337.
49. Crick FC, Koch C (2005) What is the function of the claustrum? *Philosophical Transactions of the Royal Society B: Biological Sciences* 360: 1271–1279.
50. Koubeissi MZ, Bartolomei F, Beltagy A, Picard F (2014) Electrical stimulation of a small brain area reversibly disrupts consciousness. *Epilepsy & Behavior* 37: 32–35.
51. Maslov S, Sneppen K (2002) Specificity and stability in topology of protein networks. *Science (New York, NY)* 296: 910–913.
52. Buxton RB, Frank LR (1997) A model for the coupling between cerebral blood flow and oxygen metabolism during neural stimulation. *Journal of Cerebral Blood flow and Metabolism* 17: 64–72.
53. Buxton RB, Wong EC, Frank LR (1998) Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Magnetic Resonance in Medicine* 39: 855–64.
54. Friston KJ, Mechelli a, Turner R, Price CJ (2000) Nonlinear responses in fMRI: the Balloon model, Volterra kernels, and other hemodynamics. *NeuroImage* 12: 466–77.
55. Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. *Biophysics : Hopfield I T ., V . Proceedings of the National Academy of Sciences of the United States of America* 79: 2554–2558.
56. Amit DJ, Gutfreund H, Sompolinsky H (1987) Information storage in neural networks with low levels of activity. *Physical Review A* 35: 2293.
57. Tsodyks M, Feigel'man M (1988) The enhanced storage capacity in neural networks with low activity level. *Europhysics Letters* 6: 101.
58. Kühn R, Bös S, van Hemmen JL (1991) Statistical mechanics for networks of graded-response neurons. *Physical Review A* 43: 2084.
59. Adachi M, Kazuyuki A (1997) Associative dynamics in a chaotic neural network. *Neural Networks* : 83–98.
60. Rabinovich M, Volkovskii A, Lecanda P, Huerta R, Abarbanel H, et al. (2001) Dynamical encoding by networks of competing neuron groups: winnerless competition. *Physical Review Letters* 87: 068102.
61. Tsuda I (1992) Dynamic link of memory – chaotic memory map in nonequilibrium neural networks. *Neural Networks* 5: 313–326.
62. Rabinovich MI, Huerta R, Varona P, Afraimovich VS (2008) Transient cognitive dynamics, metastability, and decision making. *PLoS computational biology* 4: e1000072.
63. Deco G, Jirsa V, McIntosh A, Sporns O, Kötter R (2009) Key role of coupling, delay, and noise in resting brain fluctuations. *Proceedings of the National Academy of Sciences* 106: 10302–10307.
64. Freyer F, Roberts JA, Becker R, Robinson PA, Ritter P, et al. (2011) Biophysical mechanisms of multistability in resting-state cortical rhythms. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 31: 6353–6361.
65. Kaneko K, Tsuda I (2003) Chaotic itinerary. *Chaos (Woodbury, NY)* 13: 926–936.
66. Ashwin P, Timme M (2005) Unstable attractors: existence and robustness in networks of oscillators with delayed pulse coupling. *Nonlinearity* 18: 2035.
67. Tani J (1996) Model-based learning for mobile robot navigation from the dynamical systems perspective. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* 26: 421–436.
68. Yamashita Y, Tani J (2008) Emergence of functional hierarchy in a multiple timescale neural network model: a humanoid robot experiment. *PLoS computational biology* 4.
69. Perdikis D, Huys R, Jirsa VK (2011) Time scale hierarchies in the functional organization of complex behaviors. *PLoS computational biology* 7.
70. Wilson HR, Cowan JD (1973) A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 13: 55–80.
71. Haken H (1983) Advanced synergetics. Springer.