Complexity of resting brain dynamics shaped by multiscale structural constraints

Mengsen Zhang* and Manish Saggar

Department of Psychiatry and Behavior Sciences,
Stanford University, Stanford, CA 94304

Abstract

The brain is a complex, nonlinear system, exhibiting ever-evolving patterns of activities even without external inputs or tasks. Such intrinsic or resting neural dynamics has been found to play critical roles in the normal functioning of the brain and psychiatric disorders. It remains a challenge, however, to link the intrinsic dynamics to the underlying structure, in part, due to the nonlinearity. Here we use a nonlinear-dynamical model to examine how the complexity of intrinsic neural dynamics, in terms of multistability and temporal diversity, is sculpted by structural properties across scales. Our model combines a population-level model (Wilson-Cowan) with additional biophysical constraints (from the reduced Wong-Wang model). We show that multistability can emerge at the whole-brain level even when individual brain regions are by themselves monostable. The multi-functionality and memory capacity associated with multistability are thus synergistic properties of the whole-brain, irreducible to properties of its parts. The exact size of the functional repertoire and memory capacity is a joint product of the nonlinearity in the local dynamics and the topology of the large-scale network. Similarly, temporal diversity of the brain is determined by both local structural differences and the topology of the global network. Together, this work unravels an intertwined and circular relationship between local and global properties in defining the intrinsic dynamic organization of the brain. Looking forward, the model can be used to probe the multiscale mechanisms underlying psychiatric disorders and the effective scales for treatment.

1 Introduction

One of the fundamental goals of neuroscience is to understand how the structure of the brain constrains its function [1]. Two features of the brain make this goal

*Corresponding author: mengsenz@stanford.edu

challenging: the brain consists of functional units across multiple spatiotemporal scales, and these units interact with each other nonlinearly (see e.g. [2–4]). Here, nonlinearity means that the changes in the functional units are *not* proportional to the input it receives. Nonlinearity notoriously obscures the link between structural constraints imposed on the interaction and its dynamical consequences. Nonlinear dynamical models serve as essential tools for bridging structural and functional understanding of the brain [5, 6]. In the present work, we use a unified mean-field model to examine how structural properties across scales—from neurons to large-scale networks—jointly determine the *intrinsic* dynamic organization of the whole brain. We demonstrate that focusing solely on local or global properties is insufficient for understanding and predictably altering the complexity of intrinsic neural dynamics. Below we give a brief overview of why it is important to characterize the intrinsic dynamic organization of the brain, what are the key features, and how we construct a model to examine these features across scales.

Nonlinear dynamical models of the brain have provided mechanistic insights into specific functions such as memory [7, 8], visual processing [9, 10], decision making [11, 12], movement coordination [13, 14], just to name a few. They correspond to an empirical focus on task-positive activities of the brain—how the brain reacts to external stimuli or task demands. On the other hand, intrinsic dynamics of the brain has long been observed (e.g. [15, 16]), but commonly treated as a baseline subtracted from task-positive activities. This baseline, however, is more active than meets the eye: it consumes the largest fraction of the brain's energy resources, while the task-imposed additional cost is comparatively trivial [17]. Neural imaging studies further reveal the richness of such intrinsic brain dynamics—it is organized into distinct functional networks, which are visited intermittently (e.g. [18–24]). The intrinsic dynamics of the brain constrains behavior and task-related neural activities across multiple time scales (e.g. [25–27]) and sustains alteration in neurological and psychiatric disorders [28]. These empirical findings drive the development of large-scale models of whole-brain dynamics (see [29] for a review). They serve to improve our understanding of the nature and functional purpose of intrinsic brain dynamics [5, 30] and to provide tools for mechanistic classifications and diagnostics of psychiatric disorders [31].

Nonlinear dynamical systems are often chosen over linear ones to account for the ubiquitous multistability and coordinated rhythmic activities in biological systems—the brain is one of the best examples [32–36]. To say that a system is multistable is to say that multiple stable patterns of activities (attractors) are all achievable by the system. Which pattern is retrieved depends on the task (boundary conditions) or intrinsic noise. Multistability of the brain signifies its multi-functionality and its ability to form memory as persistent patterns of activity [2, 32, 37, 38]. The switching dynamics between functional networks in the resting brain is thought to reflect noise-driven exploration of the underlying multistable landscape, i.e. the brain's intrinsic functional repertoire [39]. On the other hand, brain dynamics is parsed in time by a hierarchy of diverse rhythmic activities [36, 40, 41]. The coordination across rhythmic activities at diverse

frequencies gives rise to the complex functional organization of the brain that involves dynamic integration and segregation across multiple scales [36, 42]. Disruption of such temporal coordination is associated with neuropsychiatric disorders [43]. Together, multistability and temporal diversity are key features of intrinsic neural dynamics that theorists seek to capture using large-scale nonlinear dynamical models [44–46].

In the present work, we use a nonlinear dynamical model to systematically demonstrate how multistability and temporal diversity are shaped by structural properties of the brain across multiple scales—from neurons, to local populations, and to large-scale networks. The model is a formal unification of the Wilson-Cowan model [9, 47] and the reduced Wong-Wang model [12, 48, 49]. The Wilson-Cowan model [47] is a well-known population-level model of brain dynamics. It describes the average firing activities of two interacting populations of neurons—one excitatory population and one inhibitory population (see Section 4.2 for details). It has been widely used for modeling multistability and rhythmic activity in large-scale brain networks, where each node (brain region) contains two such local populations (e.g. [44, 50–55]). While the Wilson-Cowan model exhibits the key dynamic features of interest, it is unconstrained by the biophysical properties of underlying neurons. These biophysical underpinnings, however, can be important for predicting the outcome of electrical or pharmacological stimulation in experimental or clinical settings (e.g. [56, 57]). The reduced Wong-Wang model, on the other hand, is a biophysical network model constrained by biologically plausible parameters at the neuronal level [11, 38, 58–61]. Its noise-driven dynamics near certain equilibria has been used to capture resting-state functional connectivity (e.g. [49, 62]). As we will show, however, its reduced nonlinearity limits the extent of multistability and makes it less viable for studying oscillations. Here, we combine the two models to retain both the nonlinearity of the Wilson-Cowan model and the biophysical constraints of the reduced Wong-Wang model. The unified model is thus suitable for both studying key nonlinear features of intrinsic brain dynamics and predicting the effect of biophysical perturbations in empirical settings.

Complementing existing studies where biophysical network models were used to fit specific data (e.g. [48, 49, 62]), here we go forward from a plausible model and ask what are the critical parameters controlling the complexity of intrinsic neural dynamics. Below we show that the intrinsic dynamics of the model brain is shaped jointly by structural properties across scales and cannot be predicted based on monoscale information. In particular, multistability and temporal diversity absent from individual brain regions can be created synergistically at the whole-brain level due to global connectivity. Conversely, local structural properties greatly constrain the size of the functional repertoire and the extent of temporal diversity.

2 Results

Our unified model describes the evolution of whole-brain dynamics as the mean-field activities of neuronal populations in each brain region. Each brain region is modeled as a pair of excitatory and inhibitory populations, i.e. the *local model* (Figure 1, left; equation 1-3). The mean-field activity of the excitatory and inhibitory populations is described by a pair of state variables S_E and S_I respectively, i.e. the gating variables. The excitatory population excites itself (via local connection w_{EE}) and the inhibitory population (via w_{EI}). The inhibitory population inhibits itself (via w_{II}) and the excitatory population (via w_{IE}). These local populations further connect to other brain regions into a global network, i.e. the *global model* (Figure 1, right; equation 4-6). Mathematical details of how the present model combines the Wilson-Cowan model and the reduced Wong-Wang model are provided in Materials and Methods.

In the following sections, we first examine the dynamic repertoire of isolated brain regions (the local model; Section 2.1) and how nonlinearity in the present model enhances multistability and produces realistic oscillation (Section 2.2). In particular, we show how local structural connectivity controls the emergence of multistability and oscillation in isolated brain regions. Building on this picture, we further show how long-range connectivity between these brain regions interacts with local properties in shaping the global multistable landscape (Section 2.3) and temporal diversity across regions (Section 2.4). The numerical results are illustrated in the main text while the corresponding analytical supports are provided in the Supplementary Materials.

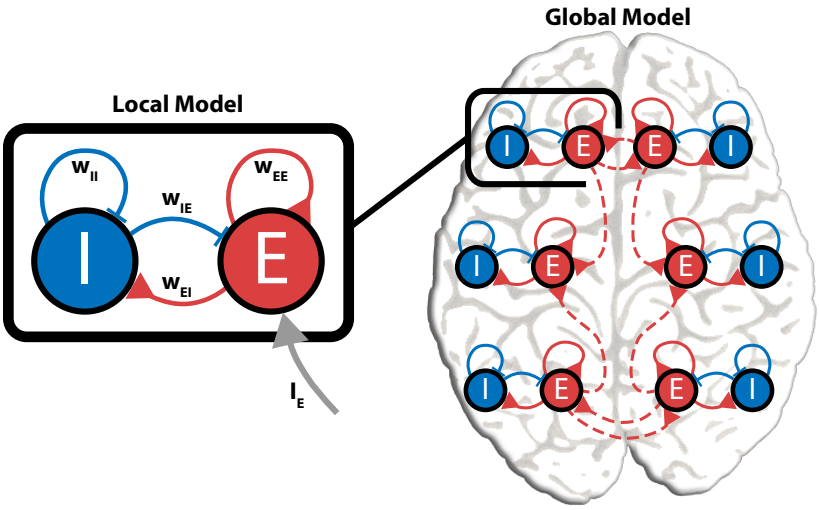

Figure 1: **A schematic of the structure of the present model.** For illustrative purposes, the cortex is modeled in this figure as a network 6 brain regions, 3 in each hemisphere (right panel, the global model). Each brain region is modeled as two interacting local neural populations — an excitatory population E and an inhibitory population I (partly enclosed by a black box). The excitatory population of each brain region can receive current input from the excitatory population of other brain regions, constrained by large-scale structural connectivity (red dashed connections). Zooming in to a single brain region, we have the local model (left box). Local populations interact with each other via two excitatory connections (red, w_{EE} and w_{EI}) and two inhibitory connections (blue, w_{II} and w_{IE}). The local excitatory population can receive an input I_E from outside of this brain region (gray arrow), which can be the total input from other brain regions.

2.1 Local structural connectivity controls dynamic repertoire of an isolated brain region

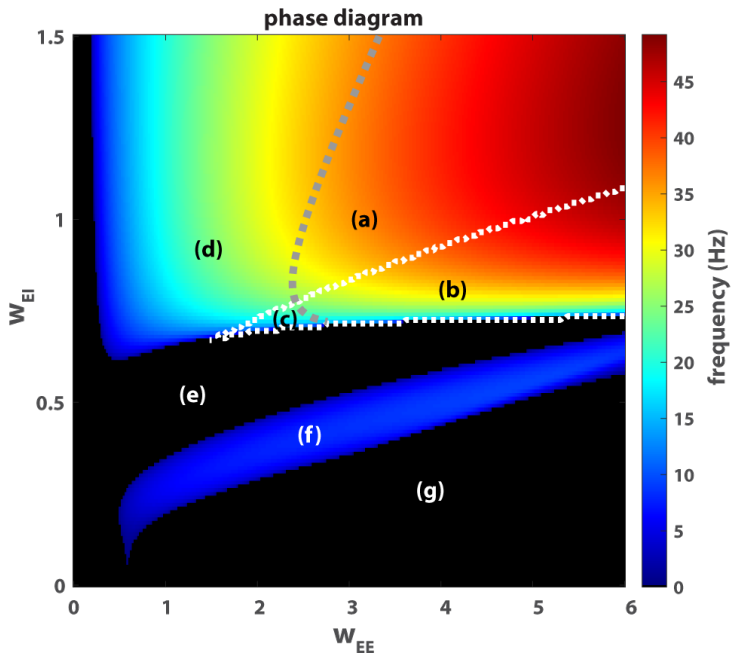

Figure 2: **Local dynamics critically controlled by the strength of excitatory-to-excitatory connection w_{EE} and excitatory-to-inhibitory connection w_{EI}.** (a)-(g) are seven different dynamic regimes of the local model (equation 1-2) in the 2-dimensional parameters space (w_{EE}, w_{EI}). Here the local inhibitory-to-excitatory connectivity w_{IE}—the inhibitory feedback—is matched to the excitatory-to-excitatory connectivity, i.e. $w_{IE} = w_{EE}$; and $I_E = 0.382$ as in [49]. Black areas (e, g) are the regimes of stable equilibrium. Colored areas are the oscillatory regimes: (a)-(b) for limit cycles and (c), (d), (f) for damped oscillations. The color reflects the frequency of oscillation. A gray dashed line indicates the Hopf bifurcation. The triangular area enclosed by white dashed lines (saddle-node bifurcation) is the bi-stable regime (b, c). An example from each regime is provided in Figure 3.

The local model exhibits a rich repertoire of dynamical features, including multistability (c, b in Figure 2, 3), damped oscillation (c, d, f), and limit cycles (sustained oscillation; a, b). Mathematical analysis of the local model (Section S4 in Supplementary Materials) shows that nonlinearity in the dynamics can essentially be controlled by two local structural properties: the strength of excitatory-to-excitatory connection w_{EE} and the strength of the excitatory-to-inhibitory connection w_{EI}. The two structural properties are responsible for

"twisting" the flow of dynamics, manifested in the curvature of the nullclines (dashed lines in Figure 3). Specifically, stronger w_{EE} introduces a deeper twist and fold of the red nullcline (compare Figure 3a and d), whereas stronger w_{EI} introduces a more vertical twist of the blue nullcline (compare Figure 3d and e). These twists are the key sources of dynamic complexity—multistability and oscillation. For example, when w_{EE} is sufficiently large (equation S19), multistability becomes possible: a deep fold of the red nullcline makes possible multiple intersections between the nullclines, thereby the number of attractors (see Section S4, Multistability). When w_{EI} is sufficiently large (equation S27), oscillatory activity becomes possible. Moreover, the combination of large w_{EE} and w_{EI} gives rise to sustained oscillation (equation S50). The characteristic frequency of such oscillation further depends on the specific values of w_{EE} and w_{EI}. It is important to note that the general qualitative effects of these two local structural properties are consistent with those of the Wilson-Cowan model, but the specific boundaries at which transitions occur are determined by the biophysical constraints inherited from the reduced Wong-Wang model (see equations S27, S50). Analytical results (Section S4) provide detailed quantification of how these boundaries are shifted by different local structural properties.

To maintain a sufficient twist in the red nullcline (red dashed line in Figure 3) and associated dynamic complexity, inhibitory-to-excitatory feedback w_{IE} needs to be proportional to self-excitation w_{EE} (c.f. equation S9). In the present study, we simply let $w_{IE} = w_{EE}$. This equality is a simpler alternative to the Feedback Inhibition Control adopted in [49] in both numerical and mathematical analyses.

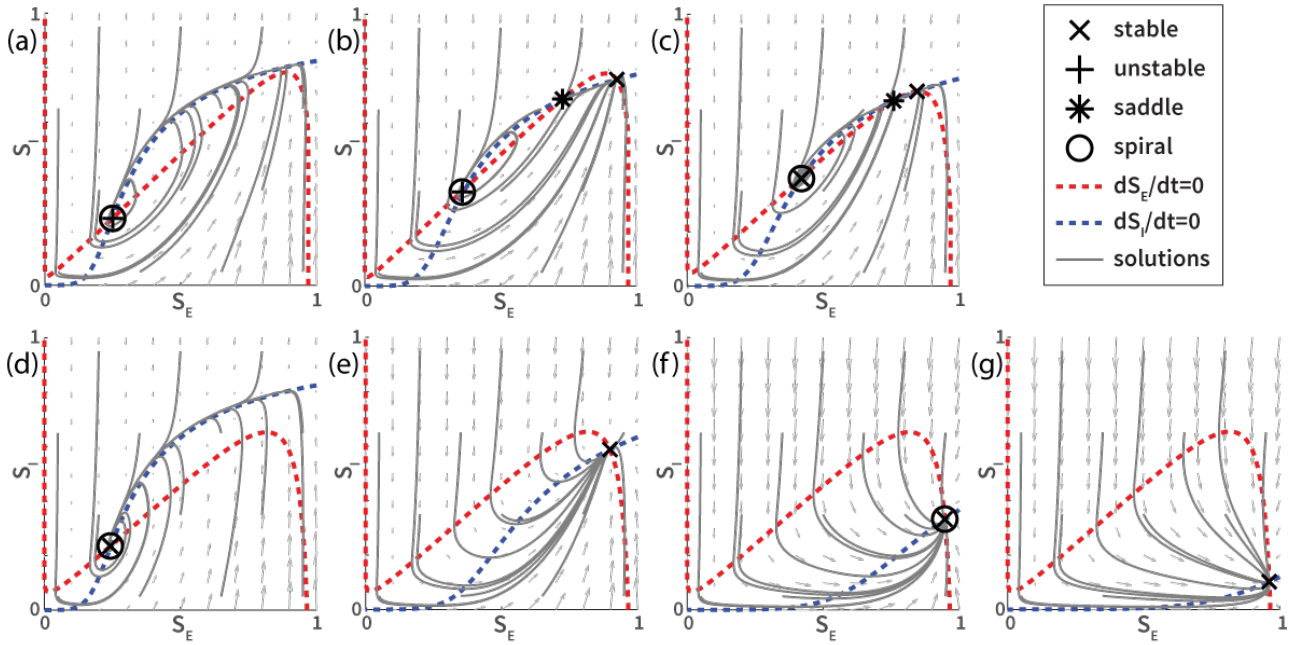

Figure 3: **Example phase portraits from different regimes of the local model**. Phase portraits (a) to (g) are examples chosen from the corresponding regimes in Figure 2. The specific parameters defining local structural connectivity are (a) $w_{EE} = 4$, $w_{EI} = 1$; (b) $w_{EE} = 4$, $w_{EI} = 0.8$; (c) $w_{EE} = 2.3$, $w_{EI} = 0.75$; (d) $w_{EE} = 1.5$, $w_{EI} = 1$; (e) $w_{EE} = 1.5$, $w_{EI} = 0.5$; (f) $w_{EE} = 1.5$, $w_{EI} = 0.3$; (g) $w_{EE} = 1.5$, $w_{EI} = 0.2$. The vector fields (arrows) reflect the underlying dynamics at different points in the state space. Gray trajectories following the vector fields are solutions of the local model (equation 1, 2) given a fixed sets of ten different initial conditions. Nullclines (dashed lines) indicate where the flow of the dynamics is either purely vertical (red) or purely horizontal (blue). The intersections between the nullclines are the fixed points. Different types of fixed points are labeled with different markers (see legend). A fixed point is stable ($\times$) if nearby trajectories converge to it over time, unstable ($+$) if nearby trajectories diverge from it, or a saddle ($*$) if nearby trajectories approach it in some direction(s) but diverge from it in some other direction(s). A fixed point is said to be a spiral ($\circ$) if trajectories near the fixed point rotate either towards the fixed point (damped oscillation) or away from the fixed point (sustained oscillation or limit cycle in the present case). Strong oscillation mainly appears on the ascending branch of the red nullcline. Overall, we see that local connectivity defines the dynamics in each regime essentially by controlling the geometry of the nullclines.

2.2 The effects of nonlinearity in the local model

Before getting into the global model, we briefly show how the added nonlinearity from the Wilson-Cowan model [47] extends the reduced Wong-Wang model [12, 49] to more complex scenarios. As discussed in detail in Section 4.2-4.3, this unified model is both more in line with Wong-Wang's initial derivation [12] and the geometric form of the Wilson-Cowan model that defines its qualitative behavior. Here we demonstrate the result numerically.

Formally, the present model matches the reduced Wong-Wang model for low levels of excitation (c.f. equation 18 and Figure 7). Consequently, their dynamics also match under weak local excitatory connectivity (Figure 4a-b). In a regime of stronger local excitatory connectivity, as explored in [62], the two models diverge (Figure 4c-d). In the present model (Figure 4c), all trajectories are well-confined within a physiologically plausible range—state variables S_E and S_I denote the fraction of open channels, which by definition are between 0 and 1. In contrast, certain trajectories of the reduced Wong-Wang model (Figure 4d) overshoot beyond the physiologically plausible range. The effect of added nonlinearity in the present model manifests through the curvature of the blue nullclines, which confines the flow of oscillatory activities and creates extended multistability (see e.g. Figure 3b). Thus, the present model is more suitable for studying key nonlinear dynamical features in the resting brain.

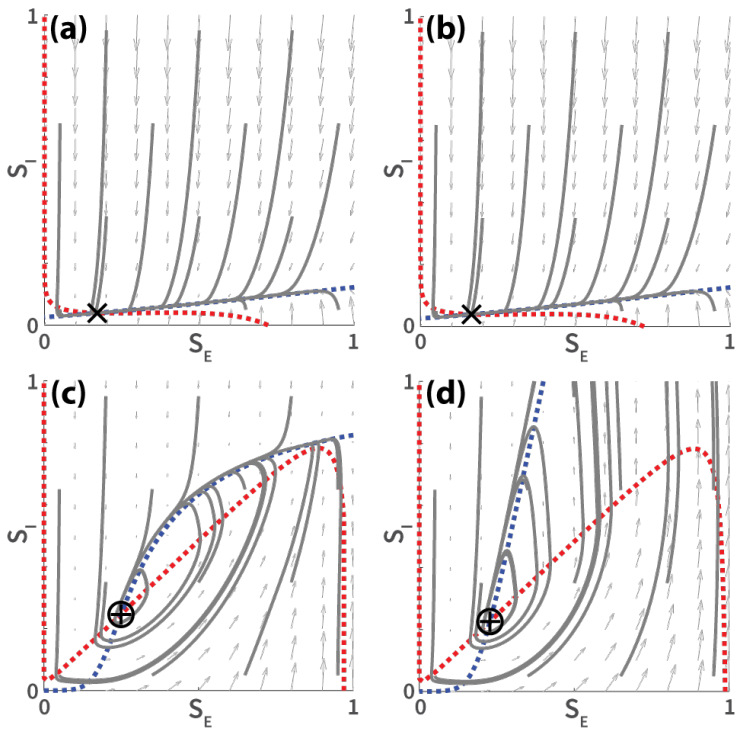

Figure 4: **Comparisons between the present model (a, c) and the reduced Wong-Wang model (b, d) in two dynamic regimes.** (a) and (b) show the phase portraits of the present model (equations 1-2) and the reduced Wong-Wang model (equations 11-12) respectively in a regime of weak local excitatory connectivity. Parameter values are obtained from [49] and identical across the two models: $w_{EE} = 0.21$, $w_{EI} = 0.15$, $w_{IE} = 1$, $w_{II} = 1$, $I_E = 0.382$ and $I_I = 0.267$ (unspecified parameters follow Table 1). The resulted dynamics are virtually identical. (c) and (d) show a similar comparison between the two models in an oscillatory regime, where the local excitatory connectivity is stronger ($w_{EE} = 4$, $w_{EI} = 1$). While the dynamics of the present model (c) is well confined within a realistic range (S_E, $S_I \in [0, 1]$), it is not the case for the reduced Wong-Wang model (d).

2.3 Global dynamic landscape shaped by structural properties across scales

Now we show the added effects of global connectivity on the dynamic landscape by comparing the bifurcation diagrams of the global model (equation 4-6; Figure 5d-i) with those of the local model (equation 1-3; Figure 5a-c). A bifurcation diagram here describes how potential dynamic patterns (attractors) of different types emerge and disappear as a certain control parameter varies. The control parameter for the local model is the input level I_E and that of the global model

is the global coupling G (equation 6; see Section S1 for computational details).

For the local model (equation 1-2), Figure 5a-c show three bifurcation diagrams for succeedingly greater local excitatory connectivity (i.e. w_{EE} and w_{EI}). As expected, there is a corresponding increase in multistability and oscillation. The increase in multistability can be seen as the addition of new folds in the bifurcation diagrams (b,c) compared to (a). Figure S3 gives an example of a monostable bifurcation diagram (no fold) for very low local excitatory connectivity. The new folds in Figure 5b-c create a new middle branch of attractors with either damped (green trace in b) or sustained oscillation (blue trace in c). With this additional branch, the response of a single brain region to different inputs (I_E) is no longer binary and static as in (a), but with an in-between option that evolves in time. In the present work, we only focus on this middle branch when we speak of oscillations; spirals outside the middle branch, e.g. Figure 3f, is highly damped, not observable in the time series. This middle branch can encode a wide range of input levels as the oscillation frequency of the local population (insets of Figure 5 b,c). The specific range of encoding frequency can differ from brain region to brain region contingent upon the local structural connectivity (see, e.g., the insets of Figure 5 b,c and regions a-d in Figure 2). Overall, the local model exhibits more complex behavior, i.e. multistability and oscillations, with greater local excitatory connectivity.

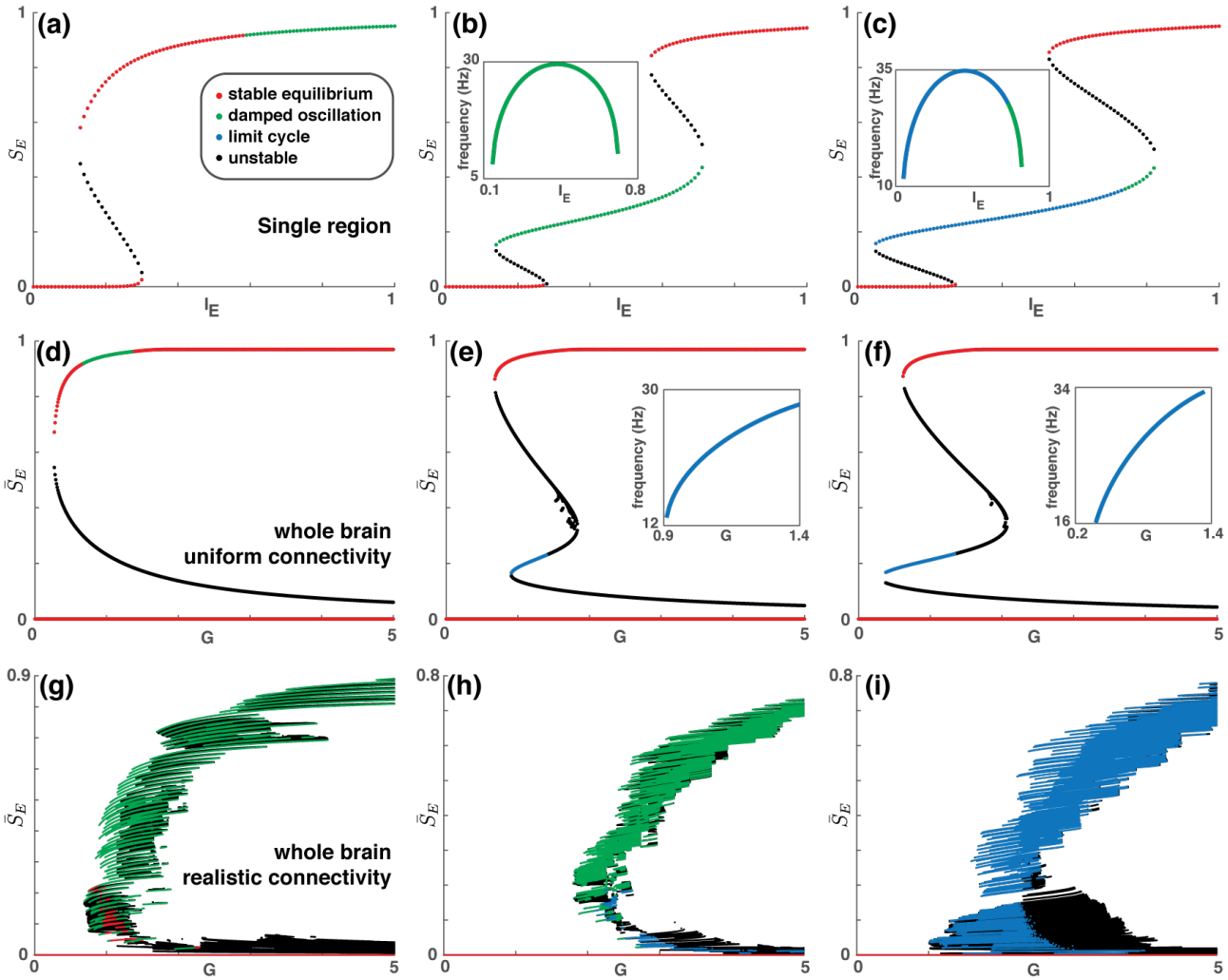

Figure 5: **Local and global structural properties jointly determine the complexity of whole-brain dynamics.** (a-c) show the bifurcation diagrams of the local model for three different types of local excitatory connectivity: (a) $w_{EE} = 0.7$ and $w_{EI} = 0.35$; (b) $w_{EE} = 2$ and $w_{EI} = 1$; (c) $w_{EE} = 2.8$ and $w_{EI} = 1$. Overall, local connectivity increases from (a) to (c). The activity of the excitatory population S_E is used as an order parameter, indicating the location of each attractor. The external input I_E is used as a control parameter. Each point in the diagram indicates the location of a particular fixed point. The color denotes the type of each fixed point: non-black points represent attractors, black points unstable fixed points that are not associated with a limit cycle. Some of the attractors connect to each other and form an attractor branch (roughly horizontal curves). All (a)-(c) have a upper branch and a lower branch. (b)-(c) have an additional branch in the middle, where the brain region oscillates. Insets of (b) and (c) show the oscillation frequency of the brain region as a function of the input current. (d)-(f) show the corresponding bifurcation diagrams for three uniform global networks, i.e. the large-scale structural connectivity C_{ij}'s are identical between any two brain regions $i \neq j$ (equation 6). The average activity of all excitatory populations ($\bar{S}_E$) is used as an order parameter and the global coupling G (equation 6) as an control parameter. Similarly, (g)-(i) show the corresponding bifurcation diagrams for three realistic global networks, i.e. C_{ij}'s reflect the human connectome (see text for details).

Next, we look at the dynamic landscape at the whole-brain level. Figure 5d-f show the bifurcation diagrams of the global model (equations 4-6) with 66 *uniformly* connected brain regions, i.e. the structural connectivity is the same between any two regions. Here the 66 regions correspond to the 66 anatomical regions of the cortex as in [48, 63]. The local structural properties within these global networks correspond exactly to those of Figure 5a-c respectively. The global bifurcation diagrams share general qualitative features with their local counterparts—the number of major branches of attractors and the presence of oscillatory activity. One exception is that the damped oscillation in isolated brain regions (green branch in Figure 5b) becomes sustained oscillation when connected to other regions of the same characteristic frequency (blue branch in Figure 5e). One remarkable difference between the uniformly connected global models and the corresponding local models is that, with the absence of persistent input, the global model can retain memories of prior input while the local model cannot. That is, when $I_E = 0$, the local model is monostable (lower red branches in Figure 5a-c), i.e. returning to the same state regardless of prior input. The global model (equation 4-6) by definition does not receive external input; yet the model is multistable for a sufficient amount of global coupling, e.g. $G > 1$ (Figure 5d-f). We further substantiate and generalize this result analytically and numerically (Section S5 Multistability) to the case where each isolated brain region is monostable for any I_E. These findings suggest that the coupling between brain regions can synergistically create a functional repertoire or memory capacity that isolated brain regions do not possess.

Uniform structural connectivity is, of course, a rather naive assumption for a whole-brain model. Therefore, we next substitute this uniform global structural connectivity with a realistic one (Figure 5g-i). To obtain a realistic human connectome, we average over the connectome of 11 unrelated subjects from the Human Connectome Project [64, 65] based on a 66-region parcellation of the cortex [63]. Given a realistic global structural connectivity, the global memory capacity is greatly amplified (Figure 5g-i). Comparing to Figure 5d-f, the number of attractor branches in Figure 5g-i is much higher: there are 173 branches of attractors in (g), 609 in (h), and 554 in (i) (branches are computed using single-linkage clustering). The heterogeneous nature of the human connectome leads to much greater whole-brain memory capacity by breaking the spatial symmetry. The greatest memory capacity (Figure 5h,i) results from the combination of local dynamic complexity (multistability) and realistic global connectivity.

In short, we have shown that the complexity of the global dynamical landscape depends crucially on both local structural properties and the synergistic effects of global connectivity. Additional analytical supports are provided in Section S5.

2.4 Local and global causes of temporal diversity

Having examined the multistability of the global dynamic landscape, now we turn to the structural constraints on temporal diversity. In particular, we show how spectral properties of the simulated neural activities and corresponding hemodynamic responses are affected by both the diversity of local structural

properties and the structure of the large-scale connectome.

Given a uniform global network, temporal diversity across the whole brain can be induced by the diversity of local excitatory-to-excitatory connection (w_{EE}), as shown in Figure 6a. Brain regions with relatively weak w_{EE} (blue) have low characteristic frequencies around 10 Hz (alpha range), while brain regions with strong w_{EE} (red) have higher characteristic frequencies around 30 Hz (beta/gamma range). In other words, the characteristic frequency of the oscillation increases monotonically with w_{EE} (see also Figure S1a). This is expected from the behavior of isolated brain regions (Figure 2d). In addition to the expected diversity, signs of coordination between regions can be seen as the wide-spread alpha peaks (Figure 6a). In contrast, regions with a higher characteristic frequency (beta/gamma range) are not as influential to other regions. That is, low-frequency oscillations, rather than high-frequency ones, are responsible for global coordination.

The above observations concern high-frequency dynamics typically measured using, e.g. electroencephalography (EEG) and magnetoencephalography (MEG). For low-frequency dynamics typical for functional magnetic resonance imaging (fMRI), we examine the low-frequency content (0.01-0.1 Hz) of the normalized power spectra of BOLD activities, derived from the same simulated neural dynamics (see Section S2 in Supplementary Materials for details). The result is shown in Figure 6b: there is no significant dependency of low-frequency power on w_{EE} (Spearman correlation $\rho = -0.029$, $p = 0.81$). In short, we find differential effects of local structural diversity on neural dynamics at the time scales typical for different neural imaging modalities.

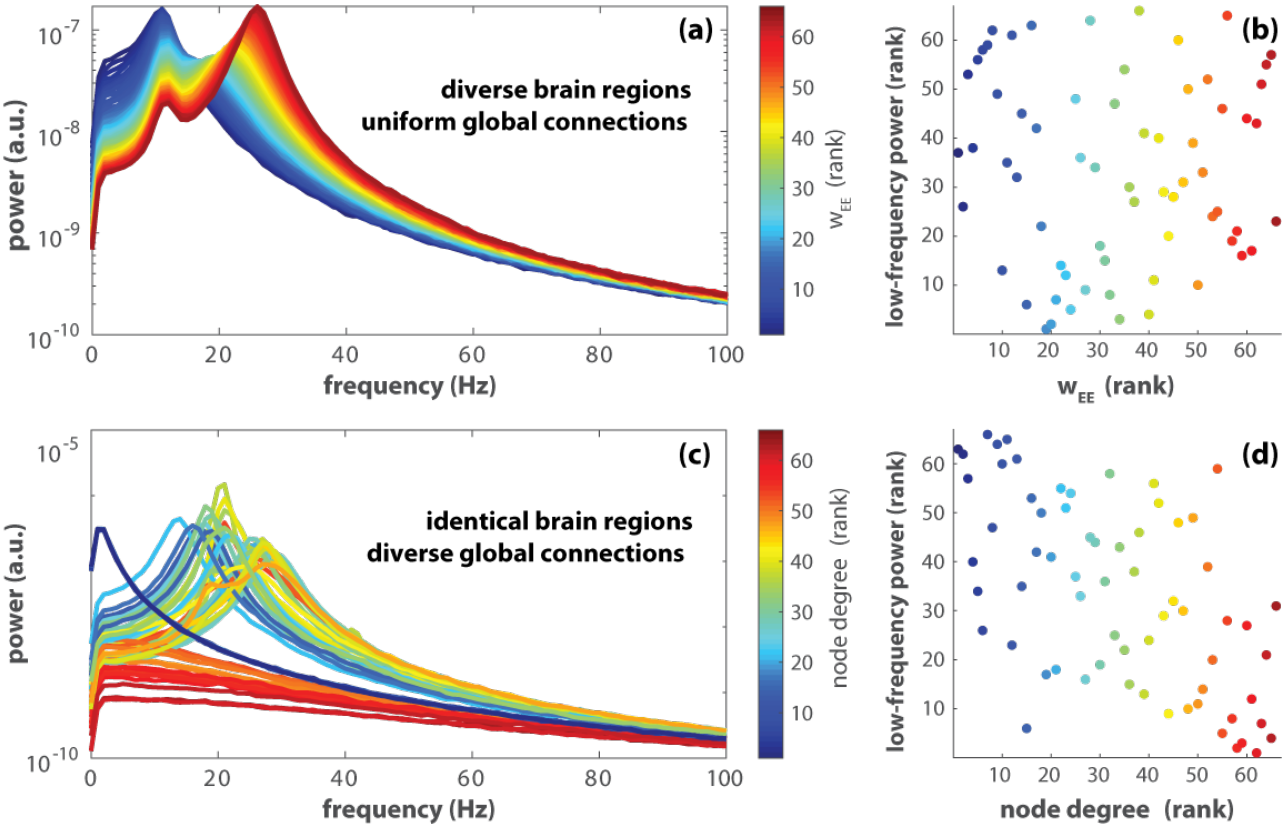

Figure 6: **Temporal diversity induced by diversity in local (a, b) and global structural connectivity (c, d)**. Spectral analyses are based on two simulated trials of the global model (equation 4-6 with $N = 66$) each with identical initial conditions $S_E^{(i)}(0) = S_I^{(i)}(0) = 0.2$, a 1200s duration, and a moderate level of noise $\sigma = 0.01$. For the first simulated trial (a, b), different brain areas are endowed with different local connectivity, $w_{EE}^{(i)}$, evenly spread in the interval $[1, 2]$; the large-scale structural connectivity is set to be uniform, i.e. $C_{ij} = 1/(N-1)$, for $i \neq j$. In addition, the global coupling $G = 1.35$. (a) shows the power spectra of the excitatory gating variables $S_E^{(i)}$ for $i = 1, \cdots, N$. The spectrum for each brain region is color coded by the rank of w_{EE}—blue to red indicate the smallest to the largest w_{EE}. The peak frequency of these spectra clearly increases with w_{EE}. (b) shows the rank of the low-frequency power of the corresponding BOLD signal, integrated over the frequency range $[0.01, 0.1]$ Hz (see Section S2 for details), which depends little on the rank of w_{EE}. (c) and (d) show results of similar analyses but for the second simulated trial, where the individual brain regions are identical ($w_{EE}^{(i)} = 2$ for all i) but the global structural connectivity is realistic, i.e. C_{ij} here reflects the human connectome [64, 65] with the global coupling $G = 2.5$. Both low-frequency (d) and high-frequency (c) activities are highly affected by the degree of the brain region in the global network (rank color-coded).

On the other hand, temporal differentiation does not mandate the brain regions themselves to be structurally different. As shown in Figure 6c-d, locally identical brain regions can behave very differently due to the topology of the large-scale network (human connectome as in Section 2.3). The influence of large-scale structural connectivity on temporal diversity is manifested in both the high-frequency neural dynamics (Figure 6c; Figure S2a) and the low-frequency power of the BOLD signals (Figure 6d; Figure S2b). Specifically, the low-frequency power is inversely related to the degree of each node (brain region) in the large-scale network (Spearman correlation $\rho = -0.584$, $p < 10^{-6}$).

In Section S3, we demonstrate that the above effects are robust over 200 simulated trials of the same parameter settings. Overall, both local (Figure 6a,b; Figure S1) and large-scale structural connectivity (Figure 6c,d; Figure S2) contribute to the diversification of local dynamics. The contribution of local structural differences is stonger in a higher-frequency range (Figure 6d; Figure S1a), while the contribution of global structural connectivity is stronger in a very-low frequency range (Figure 6d; Figure S2b). Modeling real neural dynamics requires considering both sides of the spectrum.

3 Discussion

In the present work, we use a unified mean-field model to show how the complexity of intrinsic brain dynamics can be shaped by structural constraints across scales. The model combines the nonlinearity of the Wilson-Cowan model [47] and the biophysical constraints of the reduced Wong-Wang model [12, 49]. The nonlinearity enhances the model's multistability and oscillatory dynamics, essential for understanding the nonlinear dynamics of the resting brain (Section 2.2). The biophysical constraints enhance its utility for generating theoretical predictions regarding the effect of electrical or pharmacological stimulation [31]. Moreover, the unification connects neuronal, population-level, and network-level properties. This allows for a multiscale interrogation of the structure-function relation in the brain, essential for understanding its orders and disorders [4, 66–69]. Using this model, we show that local structural properties fundamentally constrain multistability, thereby, the functional repertoire and memory capacity of the whole brain (e.g., compare Figure 5i with Figure S3c). Nevertheless, synergistic effects at the global level play an important role in creating and amplifying multistability (e.g. compare Figure 5e with Figure 5h). Similarly, the temporal diversity of the whole brain depends on both the diversity of local structural connectivity and the topology of the large-scale network (Figure 6). These findings suggest that the intrinsic dynamic organization of the brain cannot be determined by either the local structural differences or the large-scale network topology alone.

In isolated brain regions, nonlinearity in the dynamics can be effectively controlled by two key local structural properties: local excitatory-to-excitatory connectivity (self-excitation, w_{EE}) and local excitatory-to-inhibitory connectivity (w_{EI}). In the brain, local excitatory-to-excitatory connections are abundant,

especially relative to excitatory-to-inhibitory connections [70]. In our model, they contribute indispensably to multistability. The contribution of multistability to biological complexity is not restricted to the brain but permeates living systems from molecular to social levels [34, 35]. Multistability in living systems is often linked to self-excitation, or in more general terms, positive feedback loops [71–73]. Mathematically, we show that self-excitation creates multistability by twisting and folding the flow of dynamics (Section S4 in Supplementary Materials). While the qualitative effect of self-excitation applies to Wilson-Cowan type models in general, the quantitative threshold for multistability to emerge depends on the specific transfer function and thereby the biophysical properties at the neuronal level (equation S19; c.f. [47, eq.17] for the counterpart in the Wilson-Cowan model). The biophysical underpinnings of the present model help connect the mathematical prediction of increasing nonlinearity to physical manipulations. For example, varying the level of local connectivity in the model can be realized by manipulating the conductance of N-methyl-D-aspartate (NMDA) receptors in local neuronal populations. Empirically, the conductance can be modulated by various pharmacological and endogenous antagonists and agonists, such as ketamine [74] and dopamine [75]. Such manipulations have been theoretically predicted and shown to affect memory capacity [56, 57, 76].

Building on the local dynamics, important synergistic effects are present at the global network level that cannot be attributed to the properties of isolated brain regions. Such synergistic effects include creating new *types* of dynamics, e.g. sustained oscillation in global network not present in isolated nodes (Figure 5e-h vs. Figure 5b); and enlarging the functional repertoire or memory capacity of the whole brain through enhanced multistability. Non-trivial memory capacity emerges at the whole-brain level even when the local nodes themselves are memory-less (Section S4 Multistability). Moreover, global memory capacity is determined by how well the global network structure amplifies local complexity (compare Figure 5h). In particular, a realistic global network (Figure 5 g-i) better amplifies local complexity than a uniform one (Figure 5 d-f). The non-uniform nature of the human connectome breaks the spatial symmetry of the global model, whereas symmetry breaking is often a key contributor to complex dynamics [2, 42, 77–80]. Of course, the human connectome is not only non-uniform, but also endowed with more specific features such as modularity, small-worldness, and multiscale characteristics—embracing both structural integration and segregation [1, 69, 81, 82]. A detailed study of how such specific structural features alter the geometry of the global dynamic landscape is worthy of further theoretical investigation (see Section S5).

The temporal diversity of the model brain is also affected by both local and global structural constraints. In the local model, oscillatory activity requires a sufficiently strong excitatory-to-inhibitory connection. The oscillation may be damped or sustained at various characteristic frequencies, contingent upon the strength of excitatory-to-excitatory connection (see Figure 2a-d and Section S4 Oscillation). The importance of inhibitory neurons, and their interaction with pyramidal cells, for generating rhythmic activity has been well demonstrated in both theoretical and empirical studies, e.g. [83–87]. For multiple

oscillatory processes to form complex spatiotemporal patterns, it often requires the coexistence of diverse time scales [35, 42, 88]. In the present model, temporal differentiation can be caused by local structural differences between brain regions, i.e. the strength of local excitatory-to-excitatory connection (Figure 6ab). It has been shown that incorporating such local structural diversity in the reduced Wong-Wang model better describes real neural dynamics [62], demonstrating its empirical relevance. On the other hand, temporal differentiation can also be induced solely by the structure of the global network—the whole defining the parts (Figure 6cd). The diversity of node degree is a key contributor to the spectral diversity in the low-frequency range (Figure 6d), which has been observed empirically (e.g. [89]). It resonates with earlier findings that slow dynamics is more reflective of the large-scale network structure (e.g. [90]). These multiscale structural sources of temporal diversity may influence each other through their joint-influence on the dynamics due to plasticity. Further theoretical investigation incorporating plasticity may shed light on how connectivity at both the local and the global level changes with development (see [91]).

To summarize, we have provided a systematic demonstration of how intrinsic neural dynamics can be shaped by structural constraints across scales. The complexity of the intrinsic dynamic organization is characterized by the extent of multistability and temporal diversity. These complex dynamic features can emerge from the nonlinearity in the dynamics controlled by local structural connectivity. They can also be created or amplified synergistically at the large-scale network level. The work highlights the importance of multiscale modeling of intrinsic neural dynamics. This theoretical framework, constrained by the biophysics, may be used in stimulation-based hypothesis generation and model fitting. Instead of fitting models to unperturbed resting brain dynamics, electrical or magnetic stimulation of selected brain regions may be used for a more extensive and controlled exploration of the underlying multistable landscape. This stimulation-based approach may provide additional empirical constraints for individualized and generalizable model-fitting. Thus, the work provides basic theoretical and computational elements for experiment-modeling integrated approaches to understanding nonlinear brain dynamics.

4 Materials and Methods

4.1 The present model

The present model describes the dynamics of multiple interconnected brain regions (see Figure 1 right panel for a schematic). Each brain region consists of two local neuronal populations (Figure 1 left box), an excitatory and an inhibitory population, denoted by 'E' and 'I' respectively. The average activities within these two populations are described by a pair of average synaptic gating variables S_E and S_I (i.e. the fraction of open channels in each population). The local dynamics of a single brain region is described by the following equations,

which we refer to as *the local model*.

$$\frac{\mathrm{d}\,S_E}{\mathrm{d}\,t} = -\frac{S_E}{\tau_E} + (1 - S_E)\gamma_E\, H_E(w_{EE}S_E - w_{IE}S_I + I_E) \tag{1}$$

$$\frac{\mathrm{d}\,S_I}{\mathrm{d}\,t} = -\frac{S_I}{\tau_I} + (1 - S_I)\gamma_I\, H_I(w_{EI}S_E - w_{II}S_I + I_I) \tag{2}$$

The activity of each population has a natural decay at the time scale of τ_E and τ_I respectively, as described by the first term on the right-hand-side of equation 1-2. At the same time, each population's activity tends to increase with the fraction of the currently closed channels $(1 - S_p)$ and the population firing rate (H_p), scaled by a factor γ_p for $p \in \{E, I\}$. This is described by the second term on the right-hand-side of equation 1-2. H_E and H_I are transfer functions that map synaptic current input to population firing rate of the excitatory and the inhibitory population respectively (for example, H_E shown in Figure 7 as a black curve). In particular, they are sigmoidal functions of the form

$$H_p(x) = \frac{r_{max} + \dfrac{a_p x - b_p - r_{max}}{1 - e^{d_p(a_p x - b_p - r_{max})}}}{1 - e^{-d_p(a_p x - b_p)}}, \tag{3}$$

whose output increases with input monotonically and saturates at r_{max}—the maximal firing rate limited by the absolute refractory period of neurons (around 2 ms in certain cell types [92, 93]). The specific shape of each transfer function is determined by three additional parameters a_p, b_p and d_p (a_p and b_p determine the location and slope of the near-linear segment in the middle; d_p determines the smoothness of the corners bordering the said near-linear segment). This transfer function is converted from Wong and Wang's original formulation [12, 94] (a soft rectifier function, equation 13, orange dashed line in Figure 7) into a sigmoidal form (black solid line in Figure 7), while retaining the original value of parameters a_p, b_p, and d_p (shown in Table 1). The parameters were chosen to approximate the average response of a population of spiking pyramidal cells ($p = E$) and interneurons ($p = I$) respectively, incorporating physiologically plausible parameters [11, 12]. The justification and technicality of the said conversion is discussed in Section 4.3.

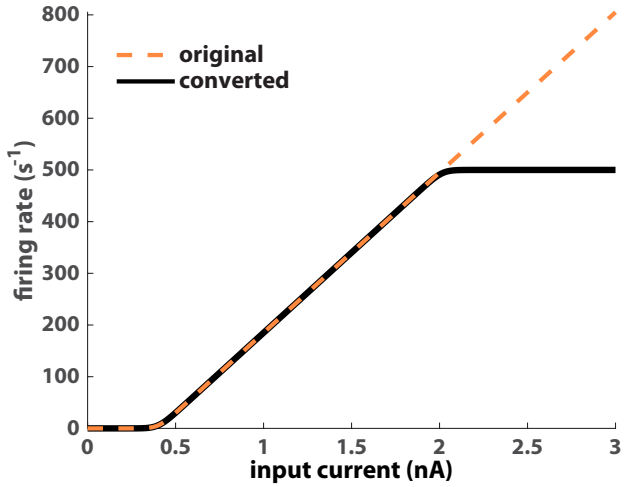

Figure 7: **Converting a rectifier transfer function to a sigmoidal form.** Black solid line is the transfer function used in the present model H_E (equation 3). It matches the transfer function used by Wong and Wang in their original formulation [12, 94] (orange dashed line; equation 13) for a low level of input. For a high level of input, $H_E(x)$ saturates at $r_{max} = 500$ (Hz), while the Wong-Wang's original version continues to scale (almost) linearly with input.

Local interaction between the populations is modulated by four coupling parameters $w_{pq} \geqslant 0$ in equation 1-2, indicating the influence from the local population p to q, where $p, q \in \{E, I\}$ (Figure 1 left box). These coupling parameters reflect the local structural connectivity. The excitatory population excites both the excitatory and the inhibitory population (via w_{EE} and w_{EI} respectively), while the inhibitory population inhibits both (negative signs precede w_{IE} and w_{II}).

In addition to the interaction within a brain region, the local populations are capable of responding to external current inputs denoted as I_E and I_I in equation 1-2, respectively. Importantly, such input can come from other brain regions in a globally connected network (Figure 1 right panel, dashed lines).

This leads us to *the global model*. Formally, we substitute I_E in the local model (equation 1) with a global input I_G (equation 4),

$$\frac{\mathrm{d}\, S_E^{(i)}}{\mathrm{d}\, t} = -\frac{S_E^{(i)}}{\tau_E} + (1 - S_E^{(i)})\gamma_E H_E \left(w_{EE}^{(i)} S_E^{(i)} - w_{IE}^{(i)} S_I^{(i)} + I_G^{(i)}(\vec{S}_E) \right) + \sigma \xi_E^{(i)}(t) \tag{4}$$

$$\frac{\mathrm{d}\, S_I^{(i)}}{\mathrm{d}\, t} = -\frac{S_I^{(i)}}{\tau_I} + (1 - S_I^{(i)})\gamma_I H_I \left(w_{EI}^{(i)} S_E^{(i)} - w_{II}^{(i)} S_I^{(i)} + I_I \right) + \sigma \xi_I^{(i)}(t) \tag{5}$$

where $S_E^{(i)}$ and $S_I^{(i)}$ are the synaptic gating variable of the excitatory and the inhibitory population of the ith brain region respectively, and $\xi_\bullet^{(i)}$ is a noise term

scaled to an amplitude σ. The state of all excitatory populations is denoted as a vector $\vec{S}_E$, the ith element of which is $S_E^{(i)}$. The global input to the ith brain region depends on both its connectivity with, and the ongoing state of, other brain regions,

$$I_G^{(i)}(\vec{S}_E) = G \sum_{\substack{j=1 \\ j \neq i}}^{N} C_{ij} S_E^{(j)} \tag{6}$$

where N denotes the total number of brain areas, $C_{ij} \geqslant 0$ the long-range structural connectivity from the jth to the ith brain region and G is a global coupling parameter that controls the overall level of interaction across brain regions. Since C_{ij} is only intended to represent long-range connectivity, we let $C_{ij} = 0$ for any $i = j$ to preclude recurrent connections. For the effects of G and C_{ij} to be independently comparable, here we impose a normalization condition on the matrix norm,

$$\|\mathbf{C}\|_\infty = \max_i \left(\sum_{j=1}^{N} |C_{ij}| \right) \equiv 1. \tag{7}$$

Since the global coupling parameter G modulates the level of input to each brain region, one would expect it to have comparable influence on the local dynamics as I_E in the local model (equation 1).

Next, we discuss its formal connection to two well-studied mean-field models of brain dynamics, namely, the Wilson-Cowan model (Section 4.2) [9, 47] and the reduced Wong-Wang model (Section 4.3) [12, 49].

parameter	interpretation	value
τ_E	decay time of NMDA receptor	0.1 (s)
τ_I	decay time of GABA receptor	0.01 (s)
γ_E	kinetic parameter of excitatory population	0.641
γ_I	kinetic parameter of inhibitory population	1
a_E	parameter of H_E	310 (nC^{-1})
b_E	parameter of H_E	125 (Hz)
d_E	parameter of H_E	0.16 (s)
a_I	parameter of H_I	615 (nC^{-1})
b_I	parameter of H_I	177 (Hz)
d_I	parameter of H_I	0.087 (s)
r_{max}	maximal firing rate	500 (Hz)
w_{EE}	excitatory-to-excitatory coupling	$\sim$ (nA)
w_{EI}	excitatory-to-inhibitory coupling	$\sim$ (nA)
w_{IE}	inhibitory-to-excitatory coupling	$\sim$ (nA)
w_{II}	inhibitory-to-inhibitory coupling	0.05 (nA)
I_E	external input to excitatory population	$\sim$ (nA)
I_I	external input to inhibitory population	0.1 (nA)
G	global coupling	$\sim$ (nA)
C_{ij}	structural connectivity between brain regions	$\sim$
σ	noise amplitude	$\sim$

Table 1: **The interpretation and value of model parameters.** Here we summarize the parameters used in equation 1-5. Most parameters assume a fixed value, which was introduced by [12]. A "$\sim$" indicates that this parameter is manipulated in the present study to explore the behavior of the model.

4.2 Relation to the Wilson-Cowan model

Formally, the above model can be considered a special variant of the Wilson-Cowan model [9, 47]. Though the specific interpretation of certain parameters differ, the two models describe similar dynamic relations between neuronal populations. Here we briefly review the form of the Wilson-Cowan model in comparison to the present model.

The Wilson-Cowan model, in its initial form [47], concerns the dynamics of a pair of interacting excitatory and inhibitory neuronal populations. The activities of the two populations are denoted as $E(t)$ and $I(t)$—the proportion of firing excitatory/inhibitory cells averaged over a period of time (the refectory period). The model takes the form

$$\tau_E \frac{\mathrm{d}\,E}{\mathrm{d}\,t} = -E + (k_E - r_E E)\mathscr{S}_E(c_1 E - c_2 I + P) \tag{8}$$

$$\tau_I \frac{\mathrm{d}\,I}{\mathrm{d}\,t} = -I + (k_I - r_I I)\mathscr{S}_I(c_3 E - c_4 I + Q). \tag{9}$$

τ_E and τ_I are time constants of the dynamics of the excitatory and inhibitory

population respectively. $c_\bullet$'s are the coupling parameters between the two population. Coefficients $k_\bullet$ and $r_\bullet$ result from a temporal coarse-graining procedure in the initial derivation (see [47] for detail). $\mathscr{S}_\bullet$ is a sigmoidal transfer function, rising monotonically from 0 to 1 with non-negative input. P and Q are external inputs to their respective populations. If we divide both sides of equation 8-9 by the time constants, we are looking at the same general form as equation 1-2.

The main difference is between the respective transfer functions. Wilson and Cowan [47] chose a particular form of $\mathscr{S}$ for mathematical analysis:

$$\mathscr{S}(x) = \frac{1}{1+\exp\left[-a(x-\theta)\right]} - \frac{1}{1+\exp(a\theta)}, \tag{10}$$

where parameter a determines the maximal slope of the function $\mathscr{S}$ and parameter θ the location of the maximal slope. Technically, Wilson and Cowan [47] only requires $\mathscr{S}$ to be of a general sigmoidal form. It may reflect the average response of a population of neurons with heterogeneous firing thresholds or heterogeneous afferent connections. The distribution of the said thresholds or connections is reflected in the parameters a and θ.

In other words, the choice of the transfer function and the parameters is *non-specific* to a predefined microscopic model. Moreover, Wilson and Cowan [47] took a function-oriented approach to analyzing the model. The key was whether the model was able to produce fundamental behaviors expected from a neural model—multistability, hysteresis, and oscillation—for *some* specific choice of parameters and transfer function. Qualitative conclusions from their analysis depend on the general geometric properties of the transfer function rather than the specific form of equation 10.

The transfer function of the present model (equation 3) follows the general geometric properties assumed by Wilson and Cowan [47]. The difference is that the parameters in equation 3 are associated specifically with a microscopic model [11], a network of leaky integrate-and-fire neurons with biologically plausible parameters, as inherited from the reduced Wong-Wang model [12, 49]. This choice provides a channel of correspondence between parameters of the models at different scales of description. To expand on this point, we next elaborate on the connection between the present model and the reduced Wong-Wang model.

4.3 Relation to the reduced Wong-Wang model

The present model can also be considered as a variant of the Wong-Wang model [12] and its high-dimensional generalizations, here referred to as the reduced Wong-Wang model [48, 49, 62]. In particular, we consider the model of whole-brain dynamics [49, 62],

$$\frac{\mathrm{d}\, S_E^{(i)}(t)}{\mathrm{d}\, t} = -\frac{S_E^{(i)}}{\tau_E} + \left(1 - S_E^{(i)}\right)\gamma_E \tilde{H}_E \left(w_{EE}^{(i)} S_E^{(i)} - w_{IE}^{(i)} S_I^{(i)} + I_G^{(i)}(\vec{S}_E)\right) + \sigma\xi^{(i)}(t) \tag{11}$$

$$\frac{\mathrm{d}\, S_I^{(i)}(t)}{\mathrm{d}\, t} = -\frac{S_I^{(i)}}{\tau_I} + \tilde{H}_I \left(w_{EI}^{(i)} S_E^{(i)} - w_{II}^{(i)} S_I^{(i)} + I_I\right) + \sigma\xi^{(i)}(t), \tag{12}$$

following the same notations as in equation 4-6, where

$$\tilde{H}_p(x) = \frac{a_p x - b_p}{1 - e^{-d_p(a_p x - b_p)}} \tag{13}$$

with $p \in \{E, I\}$ denoting the excitatory and the inhibitory population respectively (see Figure 7 dashed line for $\tilde{H}_E$). The parameters a_p, b_p and d_p were chosen such that $\tilde{H}_p$ approximates the average firing rate of an ensemble of leaky integrate-and-fire neurons receiving uncorrelated noisy inputs.

More specifically, the sub-threshold dynamics of the membrane potential $V(t)$ of each neuron can be described as

$$C_m \frac{\mathrm{d}\, V(t)}{\mathrm{d}\, t} = -g_L \left(V(t) - V_L\right) + I_{syn}(t) \tag{14}$$

where C_m is the membrane capacitance, g_L the leak conductance, and V_L the resting potential of the membrane. The total synaptic input current $I_{syn}(t)$ is a random process with an average μ_C and standard deviation σ_C. When $V(t)$ reaches a threshold V_{th}, the neuron emits a spike after which the membrane potential returns to a reset voltage V_{reset} and stays there for a duration τ_{ref}, i.e. the refractory period.

The average firing rate ν of an ensemble of such neurons can be derived from the Fokker-Planck approximation that describes the evolution of the membrane voltage distribution of an ensemble of neurons (see e.g. [95, Section 1], [61] for descriptions of the Fokker-Planck approach). This eventually leads to the first-passage time equation (average time for crossing the threshold),

$$\nu = \left(\tau_{\mathrm{ref}} + \tau_m \sqrt{\pi} \int_{\frac{V_{\mathrm{reset}} - V_{\mathrm{ss}}}{\sigma_V}}^{\frac{V_{\mathrm{th}} - V_{\mathrm{ss}}}{\sigma_V}} e^{x^2} (1 + \mathrm{erf}(x)) dx \right)^{-1} \tag{15}$$

where $\tau_m = C_m / g_L$ is the membrane time constant, $\sigma_V = \sqrt{\tau_m} \sigma_C / C_m$ the standard deviation of the depolarization, $\mathrm{erf}(x)$ the error function

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du, \tag{16}$$

and V_{ss} the steady state voltage

$$V_{\mathrm{ss}} = V_L + \frac{I_{syn}}{g_L}. \tag{17}$$

The transfer function employed by Wong and Wang [12, 94], i.e. equation 13 with appropriate choice of parameters, is a good approximation of equation 15 when the input level is low.

Thus, the first passage equation 15 provides a bridge between the transfer function (equation 13) and the single-cell level model (equation 14) incorporating realistic biophysical parameters (Table 2). In other words, it allows one to use empirically measurable quantities at the neuronal level to directly constrain the

the transfer function and the entire model. This is a major difference with the Wilson-Cowan model [9, 47] in its initial development.

parameter	interpretation	value
C_m	membrane capacitance	0.5, 0.2 (nF)
g_L	leak conductance	25, 20 (nS)
τ_m	membrane time constant	20, 10 (ms)
τ_{ref}	refractory period	2 (ms)
V_L	resting membrane potential	-70 (mV)
V_{th}	threshold for firing	-50 (mV)
V_{reset}	reset potential	-55 (mV)

Table 2: **Biophysical parameters of a single leaky-integrate-and-fire neuron.** If two parameter values are provided in right-most column, the first value is for a generic pyramidal cell and the second is for a generic interneuron. Differences between the biophysical parameters of different cell types lead to differences in the transfer functions (equation 13).

According to the first passage equation 15, the firing rate ν is a sigmoidal function of the input, which saturates at $r_{max} \equiv 1/\tau_{ref}$. This is not the case, however, for the transfer function $\tilde{H}$ (equation 13). It can be seen from equation 13 that for a large input x, the exponential term in the denominator becomes negligible (a_p, b_p, and d_p are all positive), and as a result, $\tilde{H}_p \approx a_p x - b_p$ is approximately a linear function. To make the transfer function a better approximation of the first passage equation and at the same time retain the mapping between their parameters, we can simply convert the transfer function by substituting the numerator of $\tilde{H}$ as below

$$H_p(x) = \frac{r_{max} - \tilde{H}_p(r_{max} - x)}{1 - e^{-d_p(a_p x - b_p)}}. \tag{18}$$

Thus, we obtain the transfer function used in the present model (equation 3). As shown in Figure 7, H_p matches $\tilde{H}_p$ for low levels of input but flattens out eventually at r_{max} as one would expect from equation 15, instead of increasing indefinitely (e.g. black curve H_E overlaps with orange curve $\tilde{H}_E$ for input lower than 1.5 nA, and then saturates for input greater than 2 nA). H_p is therefore a better approximation of the first passage equation 15.

In short, the present model is endowed with the geometric properties of the Wilson-Cowan model [9, 47] and at the same time consistent with the neuronal level-to-population level mapping of the reduced Wong-Wang model [12, 49].

5 Acknowledgements

This work is supported by a NIH Director's New Innovator Award to M.S. (MH119735).

References

1. Sporns, O. & Tononi, G. Classes of network connectivity and dynamics. *Complexity* **7,** 28–38 (2001).
2. Kelso, J. A. S. *Dynamic Patterns: The Self-Organization of Brain and Behavior* (The MIT Press, Cambridge, Massachusetts, 1995).
3. Freeman, W. J. & Vitiello, G. Nonlinear brain dynamics as macroscopic manifestation of underlying many-body field dynamics. *Physics of Life Reviews* **3,** 93–118 (2006).
4. Bassett, D. S. & Siebenhühner, F. in *Multiscale Analysis and Nonlinear Dynamics* 179–204 (John Wiley & Sons, Ltd, 2013).
5. Deco, G., Jirsa, V. K. & McIntosh, A. R. Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nature Reviews Neuroscience* **12,** 43 (2011).
6. Park, H.-J. & Friston, K. Structural and functional brain networks: from Connections to cognition. *Science* **342,** 1238411 (2013).
7. Yamaguchi, Y. A theory of hippocampal memory based on theta phase precession. *Biological Cybernetics* **89,** 1–9 (2003).
8. Jahnke, S., Timme, M. & Memmesheimer, R.-M. A unified dynamic model for learning, replay, and sharp-wave/ripples. *Journal of Neuroscience* **35,** 16236–16258 (2015).
9. Wilson, H. R. & Cowan, J. D. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* **13,** 55–80 (1973).
10. Johnson, J. S., Spencer, J. P., Luck, S. J. & Schöner, G. A dynamic neural field model of visual working memory and change detection. *Psychological Science* **20,** 568–577 (2009).
11. Wang, X.-J. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* **36,** 955–968 (2002).
12. Wong, K.-F. & Wang, X.-J. A recurrent network mechanism of time integration in perceptual decisions. *Journal of Neuroscience* **26,** 1314–1328 (2006).
13. Schöner, G. & Kelso, J. A. S. Dynamic pattern generation in behavioral and neural systems. *Science* **239,** 1513–1520 (1988).
14. Jirsa, V. K., Fuchs, A. & Kelso, J. A. S. Connecting cortical and behavioral dynamics: Bimanual coordination. *Neural Computation* **10,** 2019–2045 (1998).
15. Berger, H. Über das elektrenkephalogramm des menschen. *Archiv für Psychiatrie und Nervenkrankheiten* **87,** 527–570 (1929).
16. Bishop, G. H. Cyclic changes in excitability of the optic pathway of the rabbit. *American Journal of Physiology-Legacy Content* **103,** 213–224 (1932).

17. Raichle, M. E. & Mintun, M. A. Brain work and brain imaging. *Annual Review of Neuroscience* **29,** 449–476 (2006).

18. Raichle, M. E. *et al.* A default mode of brain function. *Proceedings of the National Academy of Sciences* **98,** 676–682 (2001).

19. Greicius, M. D., Krasnow, B., Reiss, A. L. & Menon, V. Functional connectivity in the resting brain: a Network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences* **100,** 253–258 (2002).

20. Fox, M. D. *et al.* The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences* **102,** 9673–9678 (2005).

21. Damoiseaux, J. S. *et al.* Consistent resting-state networks across healthy subjects. *Proceedings of the National Academy of Sciences* **103,** 13848–13853 (2006).

22. Raichle, M. E. The restless brain. *Brain Connectivity* **1,** 3–12 (2011).

23. Hutchison, R. M., Womelsdorf, T., Gati, J. S., Everling, S. & Menon, R. S. Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques. *Human Brain Mapping* **34,** 2154–77 (2012).

24. Hipp, J. F., Hawellek, D. J., Corbetta, M., Siegel, M. & Engel, A. K. Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nature Neuroscience* **15,** 884–890 (2012).

25. Fox, M. D., Snyder, A. Z., Zacks, J. M. & Raichle, M. E. Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. *Nature Neuroscience* **9,** 23–25 (2006).

26. Schroeder, C. E. & Lakatos, P. Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences* **32,** 9–18 (2009).

27. Liégeois, R. *et al.* Resting brain dynamics at different timescales capture distinct aspects of human behavior. *Nature Communications* **10,** 2317 (2019).

28. Zhang, D. & Raichle, M. E. Disease and the brain's dark energy. *Nature Reviews Neurology* **6,** 15–28 (2010).

29. Breakspear, M. Dynamic models of large-scale brain activity. *Nature Neuroscience* **20,** 340–352 (2017).

30. Deco, G., Jirsa, V. K. & McIntosh, A. R. Resting brains never rest: Computational insights into potential cognitive architectures. *Trends in Neurosciences* **36,** 268–274 (2013).

31. Stephan, K. E., Iglesias, S., Heinzle, J. & Diaconescu, A. O. Translational perspectives for computational neuroimaging. *Neuron* **87,** 716–732 (2015).

32. Kelso, J. A. S. Multistability and metastability: Understanding dynamic coordination in the brain. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367,** 906–918 (2012).

33. Kuramoto, Y. *Chemical Oscillations, Waves, and Turbulence.* (Springer, Berlin, Heidelberg, 1984).

34. Laurent, M. & Kellershohn, N. Multistability: a Major means of differentiation and evolution in biological systems. *Trends in Biochemical Sciences* **24,** 418–422 (1999).

35. Zhang, M., Beetle, C., Kelso, J. A. S. & Tognoli, E. Connecting empirical phenomena and theoretical models of biological coordination across scales. *Journal of The Royal Society Interface* **16,** 20190360 (2019).

36. Buzsáki, G., Logothetis, N. & Singer, W. Scaling brain size, keeping timing: Evolutionary preservation of brain rhythms. *Neuron* **80,** 751–764 (2013).

37. Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences* **79,** 2554–2558 (1982).

38. Amit, D. J. & Brunel, N. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral Cortex* **7,** 237–252 (1997).

39. Hansen, E. C., Battaglia, D., Spiegler, A., Deco, G. & Jirsa, V. K. Functional connectivity dynamics: Modeling the switching behavior of the resting state. *NeuroImage* **105,** 525–535 (2015).

40. Buzsáki, G. & Draguhn, A. Neuronal oscillations in cortical networks. *Science* **304,** 1926–1929 (2004).

41. Buzsáki, G. & György, B. *Rhythms of the Brain* (Oxford University Press, New York, 2006).

42. Tognoli, E. & Kelso, J. A. S. The metastable brain. *Neuron* **81,** 35–48 (2014).

43. Uhlhaas, P. J. & Singer, W. Neuronal dynamics and neuropsychiatric disorders: Toward a translational paradigm for dysfunctional large-scale networks. *Neuron* **75,** 963–80 (2012).

44. Deco, G., Jirsa, V., McIntosh, A. R., Sporns, O. & Kötter, R. Key role of coupling, delay, and noise in resting brain fluctuations. *Proceedings of the National Academy of Sciences* **106,** 10302–10307 (2009).

45. Deco, G. & Jirsa, V. K. Ongoing cortical activity at rest: Criticality, multistability, and ghost attractors. *Journal of Neuroscience* **32,** 3366–3375 (2012).

46. Golos, M., Jirsa, V. & Daucé, E. Multistability in large scale models of brain activity. *PLoS computational biology* **11,** e1004644 (2015).

47. Wilson, H. R. & Cowan, J. D. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal* **12,** 1–24 (1972).

48. Deco, G. *et al.* Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *Journal of Neuroscience* **33,** 11239–11252 (2013).

49. Deco, G. *et al.* How local excitation–inhibition ratio impacts the whole brain dynamics. *Journal of Neuroscience* **34,** 7886–7898 (2014).

50. Ermentrout, G. B. & Cowan, J. D. Temporal oscillations in neuronal nets. *Journal of Mathematical Biology* **7,** 265–280 (1979).

51. Campbell, S. & Wang, D. Synchronization and desynchronization in a network of locally coupled Wilson-Cowan oscillators. *IEEE Transactions on Neural Networks* **7,** 541–554 (1996).

52. Borisyuk, R. & Hoppensteadt, F. Oscillatory models of the hippocampus: a Study of spatio-temporal patterns of neural activity. *Biological Cybernetics* **81,** 359–371 (1999).

53. Frank, T., Daffertshofer, A., Peper, C., Beek, P. & Haken, H. Towards a comprehensive theory of brain activity: Coupled oscillator systems under external forces. *Physica D: Nonlinear Phenomena* **144,** 62–86 (2000).

54. Atasoy, S., Donnelly, I. & Pearson, J. Human brain networks function in connectome-specific harmonic waves. *Nature Communications* **7,** 10340 (2016).

55. Muldoon, S. F. *et al.* Stimulation-based control of dynamic brain networks. *PLOS Computational Biology* **12,** e1005076 (2016).

56. Anticevic, A. *et al.* NMDA receptor function in large-scale anticorrelated neural systems with implications for cognition and schizophrenia. *Proceedings of the National Academy of Sciences* **109,** 16720–5 (2012).

57. Murray, J. D. *et al.* Linking microcircuit dysfunction to cognitive impairment: Effects of disinhibition associated with schizophrenia in a cortical working memory model. *Cerebral Cortex* **24,** 859–872 (2014).

58. Brunel, N. & Sergi, S. Firing frequency of leaky intergrate-and-fire neurons with synaptic current dynamics. *Journal of Theoretical Biology* **195,** 87–95 (1998).

59. Wang, X.-J. Synaptic basis of cortical persistent activity: the Importance of NMDA receptors to working memory. *Journal of Neuroscience* **19,** 9587–9603 (1999).

60. Brunel, N. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of Computational Neuroscience* **8,** 183–208 (2000).

61. Brunel, N. & Wang, X.-J. Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *Journal of Computational Neuroscience* **11,** 63–85 (2001).

62. Demirtaş, M. *et al.* Hierarchical heterogeneity across human cortex shapes large-scale neural dynamics. *Neuron* **101,** 1181–1194.e13 (2019).

63. Hagmann, P. *et al.* Mapping the structural core of human cerebral cortex. *PLoS Biology* **6,** e159 (2008).

64. Van Essen, D. C. *et al.* The WU-Minn Human Connectome Project: An overview. *NeuroImage* **80,** 62–79 (2013).

65. Civier, O., Smith, R. E., Yeh, C.-H., Connelly, A. & Calamante, F. Is removal of weak connections necessary for graph-theoretical analysis of dense weighted structural connectomes from diffusion MRI? *NeuroImage* (2019).

66. Grillner, S., Kozlov, A. & Kotaleski, J. H. Integrative neuroscience: Linking levels of analyses. *Current Opinion in Neurobiology* **15,** 614–621 (2005).

67. Robinson, P. A., Rennie, C. J., Rowe, D. L., O'Connor, S. C. & Gordon, E. Multiscale brain modelling. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360,** 1043–1050 (2005).

68. Breakspear, M. & Stam, C. J. Dynamics of a neural system with a multiscale architecture. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360,** 1051–1074 (2005).

69. Sporns, O. Network analysis, complexity, and brain function. *Complexity* **8,** 56–60 (2002).

70. Somogyi, P., Tamás, G., Lujan, R. & Buhl, E. H. Salient features of synaptic organisation in the cerebral cortex. *Brain Research Reviews* **26,** 113–135 (1998).

71. Attneave, F. Multistability in perception. *Scientific American* **225,** 62–71 (1971).

72. Arthur, W. B. Positive feedbacks in the economy. *Scientific American* **262,** 92–99 (1990).

73. Angeli, D., Ferrell, J. E. & Sontag, E. D. Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proceedings of the National Academy of Sciences* **101,** 1822–1827 (2004).

74. Orser, B. A., Pennefather, P. S. & MacDonald, J. F. Multiple mechanisms of ketamine blockade of N-methyl-D-aspartate receptors. *Anesthesiology* **86,** 903–917 (1997).

75. Wang, M., Wong, A. H. & Liu, F. Interactions between NMDA and dopamine receptors: a Potential therapeutic target. *Brain research* **1476,** 154–63 (2012).

76. Verma, A. & Moghaddam, B. NMDA receptor antagonists impair prefrontal cortex function as assessed via spatial delayed alternation performance in rats: Modulation by dopamine. *Journal of Neuroscience* **16,** 373–379 (1996).

77. Golubitsky, M., Stewart, I., Buono, P.-L. & Collins, J. J. Symmetry in locomotor central pattern generators and animal gaits. *Nature* **401,** 693–695 (1999).

78. Golubitsky, M. & Stewart, I. *The Symmetry Perspective - From Equilibrium to Chaos in Phase Space and Physical Space* (Birkhäuser, Basel, 2002).

79. Kelso, J. A. S., Dumas, G. & Tognoli, E. Outline of a general theory of behavior and brain coordination. *Neural Networks* **37,** 120–131 (2013).

80. Pillai, A. S. & Jirsa, V. K. Symmetry breaking in space-time hierarchies shapes brain dynamics and behavior. *Neuron* **94,** 1010–1026 (2017).

81. Sporns, O. in *Coordination Dynamics: Issues and Trends* (eds Jirsa, V. K. & Kelso, J. A. S.) 197–215 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2004).

82. Sporns, O. The human connectome: a Complex network. *Annals of the New York Academy of Sciences* **1224,** 109–125 (2011).

83. Wang, X.-J. & Buzsáki, G. Gamma oscillation by synaptic inhibition in a hippocampal interneuronal network model. *Journal of Neuroscience* **16,** 6402–6413 (1996).

84. Bartos, M., Vida, I. & Jonas, P. Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. *Nature Reviews Neuroscience* **8,** 45–56 (2007).

85. Middleton, S. *et al.* NMDA receptor-dependent switching between different gamma rhythm-generating microcircuits in entorhinal cortex. *Proceedings of the National Academy of Sciences* **105,** 18572–7 (2008).

86. Atallah, B. V. & Scanziani, M. Instantaneous modulation of gamma oscillation frequency by balancing excitation with inhibition. *Neuron* **62,** 566–577 (2009).

87. Mann, E. O. & Mody, I. Control of hippocampal gamma oscillation frequency by tonic inhibition and excitation of interneurons. *Nature Neuroscience* **13,** 205–212 (2009).

88. Zhang, M., Kalies, W. D., Kelso, J. A. S. & Tognoli, E. Topological portraits of multiscale coordination dynamics. *Journal of Neuroscience Methods* **339,** 108672 (2020).

89. Baria, A. *et al.* Linking human brain local activity fluctuations to structural and functional network architectures. *NeuroImage* **73,** 144–155 (2013).

90. Honey, C. J., Kötter, R., Breakspear, M. & Sporns, O. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences* **104,** 10240–10245 (2007).

91. Uddin, L. Q., Supekar, K. S., Ryali, S. & Menon, V. Dynamic reconfiguration of structural and functional connectivity across core neurocognitive brain networks with development. *Journal of Neuroscience* **31,** 18578–18589 (2011).

92. Yeomans, J. S. The absolute refractory periods of self-stimulation neurons. *Physiology & Behavior* **22,** 911–919 (1979).

93. Andersen, P., Silfvenius, H., Sundberg, S. & O. Sveen, H. W. Functional characteristics of unmyelinated fibres in the hippocampal cortex. *Brain Research* **144,** 11–18 (1978).

94. Abbott, L. F. & Chance, F. S. Drivers and modulators from push-pull and balanced synaptic input. *Progress in Brain Research* **149,** 147–155 (2005).

95. Renart, A., Brunel, N. & Wang, X.-J. in *Computational Neuroscience: A Comprehensive Approach* (ed Feng, J.) 432–490 (CRC Press, 2003).

Supplementary Materials

S1 Computation of bifurcation diagrams

The computation of bifurcation diagrams (Figure 5) was carried out in MATLAB, utilizing the build-in function `fsolve`. Given a proper initial guess, `fsolve` provides the coordinates of a nearby fixed point and the Jacobian matrix at the fixed point. The spectrum $\{\lambda_k\}_{k=1}^{2N}$ of the Jacobian matrix is used to classify the fixed point, where N is the number of brain regions in the model. The fixed point is a stable equilibrium if λ_k is real and negative for all k. The fixed point is associated with damped oscillation if $\mathrm{Re}\,\lambda_k < 0$ for all k and $\mathrm{Im}\,\lambda_k \neq 0$ for some k. The fixed point is associated with a limit cycle if $\mathrm{Re}\,\lambda_k > 0$ and $\mathrm{Im}\,\lambda_k \neq 0$ for some k with the additional criteria that after a small perturbation from the fixed point, the time-average of the solution remains close to the fixed point. All other types of fixed points are classified as unstable. For damped oscillation and limit cycles in the local model, the frequency of the oscillation (Figure 2) is defined as $|\,\mathrm{Im}\,\lambda_k|/(2\pi)$.

For the local model, a 2D dynamical system, the complete characterization of all fixed points is relatively easy by searching exhaustively through a grid of initial guesses (as for Figure 5a-c). This approach becomes unfeasible when it comes to the global model due to the high dimensionality. Thus, for the global model, we implemented a recursive search: for each value of G, (1) find zeros of equation 4-6 (main text) given a set of initial guesses that includes, if any, the zeros for $G - \delta G$ ($\delta G = 0.01$ for the present study); (2) sort the list of zeros obtained from (1) by the average of $S_E^{(i)}$'s; (3) use the middle points between consecutive zeros in the sorted list as initial guesses; (4) continue to use middle points between past initial guesses as new initial guesses recursively until at least one new zero is found or the recursion has reached a certain depth; (5) append the new zero(s) to the list of zeros and repeat (2)-(5) until the number of identified zeros exceeds a certain value. In the present study, we limit the maximal depth in (4) to 8 and the maximal number of zeros in (5) to 200.

S2 Computation of BOLD signal and low-frequency power

In the present study, we are interested in not only the high-frequency activity measurable by, for example, EEG recordings but also low-frequency fluctuations that are often a subject of investigation in fMRI studies. Therefore, we simulated the BOLD activities induced by the underlying neural dynamics and examine their low-frequency properties.

BOLD (Blood-oxygen-level-dependent) activities are computed using the Balloon-Windkessel model [1–4]. The hemodynamic response of the ith brain

area takes the form

$$\dot{s}_i = z_i - \kappa_i s_i - \gamma_i(f_i - 1) \tag{S1}$$

$$\dot{f}_i = s_i \tag{S2}$$

$$\tau_i \dot{v}_i = f_i - v_i^{1/\alpha} \tag{S3}$$

$$\tau_i \dot{q}_i = \frac{f_i}{\rho_i}[1 - (1 - \rho_i)^{1/f_i}] - v_i^{1/\alpha - 1} q_i \tag{S4}$$

$$BOLD_i = V_0[k_1(1 - q_i) + k_2(1 - q_i/v_i) + k_3(1 - v_i)] \tag{S5}$$

where the interpretation and value of the parameters are given in Table S1. The initial condition is

$$[s_i(0), f_i(0), v_i(0), q_i(0)] = [0, 1, 1, 1] \tag{S6}$$

which is a hemodynamic equilibrium state without neural activity. $z_i(t)$ is the simulated neural activity, corresponding to the gating variable of the excitatory populations $S_E^{(i)}(t)$.

parameter	**interpretation**	**value**
z_i	neuronal activity	$S_E^{(i)}$
s_i	vasodilatory signal	variable
f_i	blood inflow	variable
v_i	blood volume	variable
q_i	deoxyhemoglobin content	variable
κ_i	rate of signal decay	$0.65(s^{-1})$
γ_i	rate of flow-dependent elimination	$0.41(s^{-1})$
τ_i	hemodynamic transit time	0.98 (s)
α	Grubb's exponent	0.32
ρ	resting oxygen extraction fraction	0.34
V_0	resting blood volume fraction	0.02
k_1	BOLD weight parameter	$7\rho_i$
k_2	BOLD weight parameter	2
k_3	BOLD weight parameter	$2\rho_i - 0.2$

Table S1: Parameters of the Balloon-Windkessel model of BOLD activities, obtained from [4].

The power spectrum for each simulated BOLD time series is computed using Welch's method [5], after being subsampled at 720ms intervals (matching the TR of resting state fMRI used in the Human Connectome Project [6]). The full power spectrum $P(\omega)$ was first normalized such that

$$\int_0^{\omega_N} P(\omega)\,\mathrm{d}\,\omega = 1 \tag{S7}$$

where ω_N is the Nyquist frequency (approximately 0.7 Hz for the chosen sampling interval). The low-frequency power is defined as

$$p_\ell = \int_{0.01}^{0.1} P(\omega)\,\mathrm{d}\,\omega. \tag{S8}$$

S3 Dependency of spectral properties on local and global structural connectivity

In Figure 6 of the main text, we illustrate with two simulated trials how high-frequency and low-frequency dynamics depend on local excitatory-to-excitatory connectivity w_{EE} and the topology of the global network. To show that these effects are not incidental, we simulated 200 trials for each of the conditions: (1) the global network is uniform but local connectivity w_{EE} is diverse (as in Figure 6a,b), and (2) local connectivity w_{EE} is identical but the global network follows the human connectome (as in Figure 6c,d). We characterize the high-frequency content of a spectrum as its peak frequency, i.e. the frequency at which the spectral power is the highest (e.g. peaks in Figure 6a,c); the low-frequency content as the integral of the power between 0.01 and 0.1 Hz (Section S2). The dependency of these features on local (w_{EE}) and global structural properties (node degree) is quantified using Spearman correlation. The distributions of the correlation coefficients (ρ) and corresponding p-values are shown in Figure S1 and Figure S2 for condition 1 and 2 respectively. Figure S1 shows that local structural connectivity w_{EE} strongly affects the peak frequency of the brain region (a,c) but not so much the low-frequency power (b,d). The stronger the local connectivity, the higher the peak frequency. Figure S2 shows that the node degree of the global network has a strong and negative effect on the low-frequency power, and a weak and positive effect on the peak frequency. Figure 6 illustrates such dependencies using typical trials (median correlation coefficients) from the distributions (Figure S1-S2).

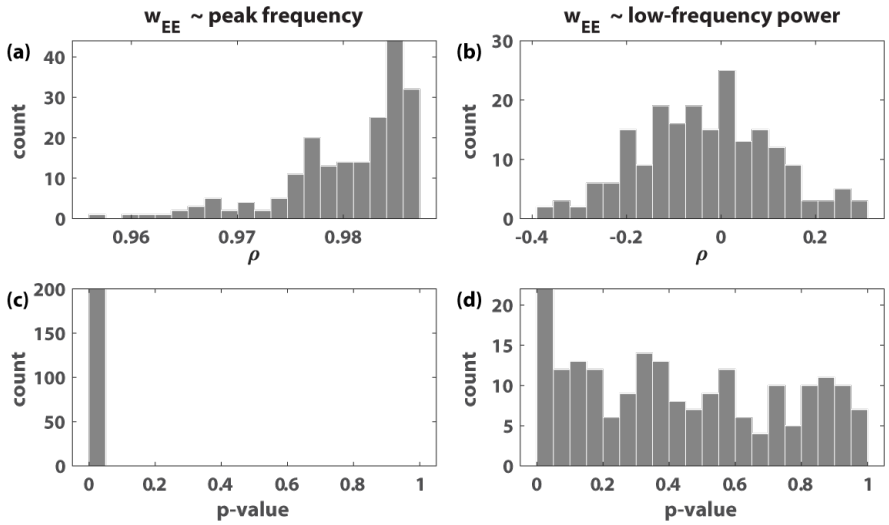

Figure S1: **Dependency of peak frequency and low-frequency power on local excitatory-to-excitatory connectivity**. 200 trials are simulated following the same parameter setting as Figure 6a,b, where the global network is uniform but the local connectivity w_{EE}'s spread between 1 to 2 for different brain regions. The noise terms in equation 4-5 make these trials different realizations of the same noisy process. The peak frequency of the spectra, e.g. from 10 to 30 Hz in Figure 6a, strongly depends on local connectivity w_{EE} (a: ρ's all close to 1; c: p-values all less than 0.05). In contrast, low-frequency power does not significantly depend on w_{EE} (b: ρ's distribute around zero; d: p-values spread between 0 and 1). Figure 6b shows this lack of dependency in an example trial that corresponds to the median of the distribution (b).

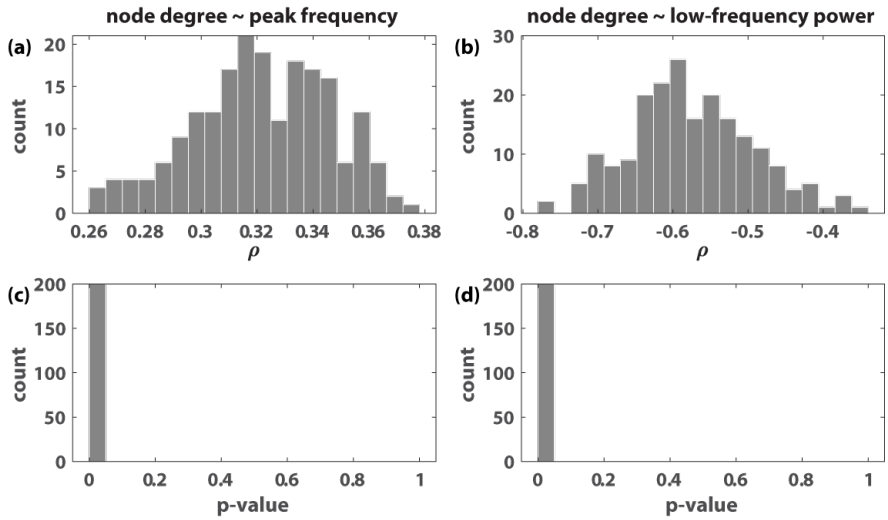

Figure S2: **Dependency of peak frequency and low-frequency power on node degree in the global network**. 200 trials are simulated following the same parameter setting as Figure 6c,d, where the local connectivity w_{EE}'s are the same across brain regions but the global-network reflects the human connectome (see main text). The peak frequency of the spectra, e.g. between 0 and 30 Hz in Figure 6c, moderately increases with the node degree of each region (a: positive ρ's around 0.32; c: p-values all less than 0.05). Low-frequency power decreases more significantly with node degree (b: ρ's distribute around -0.6; d: p-values all less than 0.05). Figure 6d illustrates this dependency with an example trial that corresponds to the median of the distribution (b).

S4 Analysis of the local model

We can see from the numerical analysis that the nullclines (dashed lines in Figure 3) crucially constraint the dynamics of the local model (equation 1-3). Here we take a closer look at their shapes. Red nullcline indicates where there is only vertical flow,

$$\begin{aligned}
&\frac{\mathrm{d}\,S_E}{\mathrm{d}\,t} = 0 \\
\Rightarrow &-\frac{S_E}{\tau_E} + (1-S_E)\gamma_E\, H_E(w_{EE}S_E - w_{IE}S_I + I_E) = 0 \\
\Rightarrow &S_I = f(S_E) = \frac{w_{EE}}{w_{IE}}S_E - \frac{1}{w_{IE}}H_E^{-1}\left(\frac{S_E}{\tau_E\gamma_E(1-S_E)}\right) + \frac{I_E}{w_{IE}},
\end{aligned} \tag{S9}$$

and blue nullcline indicates where there is only horizontal flow,

$$\begin{aligned}
&\frac{\mathrm{d}\,S_I}{\mathrm{d}\,t} = 0 \\
\Rightarrow &-\frac{S_I}{\tau_I} + (1-S_I)\gamma_I\, H_I(w_{EI}S_E - w_{II}S_I + I_I) = 0 \\
\Rightarrow &S_E = g(S_I) = \frac{w_{II}}{w_{EI}}S_I + \frac{1}{w_{EI}}H_I^{-1}\left(\frac{S_I}{\tau_I\gamma_I(1-S_I)}\right) - \frac{I_I}{w_{EI}}.
\end{aligned} \tag{S10}$$

What is common between the two nullcines, $S_I = f(S_E)$ and $S_E = g(S_I)$, is that their shape crucially depends on a linear term S_p and the inverse of the transfer function H_p^{-1} for $p \in \{E, I\}$. Both terms are monotonically increasing with S_p ($H_p^{-1}(\bullet)$ and $S_p/(1-S_p)$ are both monotonically increasing function; so is their composition). H_p^{-1} is only defined on a domain between 0 and r_{max}, for which the nullclines are confined within the interval

$$\left[0, 1 - \frac{1}{r_{max}\tau_p\gamma_p}\right]. \tag{S11}$$

Within this interval $S_I = f(S_E)$ (equation S9; red nullcline), overall, goes down from $+\infty$ to $-\infty$, while $S_E = g(S_I)$ (equation S10; blue nullcline) goes up from $-\infty$ to $+\infty$. This results from the dominant effect of H_p^{-1} for a very large or very small input.

In between these extremes, the effect of the linear term is more pronounced. This is especially the case for $S_I = f(S_E)$ (red nullcline): the linear term monotonically increases with S_E, counteracting the descending trend of $-H_E^{-1}$. Given a sufficiently strong excitatory-to-excitatory connection w_{EE} (self-excitation), the linear term "twists" the nullcline counterclockwise, creating an ascending branch in the middle. If we balance the level of self-excitation with inhibitory-feedback—let $w_{EE} = w_{IE}$—equation S9 becomes

$$S_I = f(S_E) = S_E - \frac{1}{w_{EE}}H_E^{-1}\left(\frac{S_E}{\tau_E\gamma_E(1-S_E)}\right) + \frac{I_E}{w_{EE}}. \tag{S12}$$

In this simplified case, increasing self-excitation w_{EE} reduces the influence of H_E^{-1} such that the slope of middle branch approaches 1.

For $S_E = g(S_I)$ (equation S10), the linear term and the H_I^{-1} term increase together, so that $S_E = g(S_I)$ (blue nullcline) is always monotonically increasing. Given a fixed w_{II}, $S_E = g(S_I)$ increases with S_I at an overall slower rate for larger w_{EI}, or more conveniently seen as $S_I = g^{-1}(S_E)$ increasing faster with S_E for larger w_{EI}. Intuitively, increasing w_{EI} twists $S_E = g(S_I)$ counterclockwise, seen as the middle segment of the blue nullcline becoming more vertical.

We have discussed above how local connectivity w_{EE} and w_{EI} influence the gross geometry of the nullclines—twisting the middle segment of the curve counterclockwise. But how are these geometric changes going to affect the dynamics? We show below that they critically control the multistability and oscillation in the local model.

Multistability. For the local model to be multistable, $S_I = f(S_E)$ (red nullcline) must have an ascending branch, i.e. $f(S_E)$ cannot be monotonically decreasing.

Proof. Suppose that $f(x)$ and $g^{-1}(x)$ are monotonic functions for $x \in [0, 1]$. Specially, $g^{-1}(x)$ is monotonically increasing; $f(x)$ is monotonically decreasing. Assume that $f(x)$ and $g^{-1}(x)$ intersect at two points $x_1 \leqslant x_2$, i.e. $f(x_1) = g^{-1}(x_1)$ and $f(x_2) = g^{-1}(x_2)$. Since $g^{-1}(x)$ is monotonically increasing, we have $g^{-1}(x_1) \leqslant g^{-1}(x_2)$, which implies $f(x_1) \leqslant f(x_2)$. Meanwhile, since $f(x)$ is monotonically decreasing, $f(x_1) \geqslant f(x_2)$. Thus, we have $f(x_1) = f(x_2)$, and by monotonicity, $x_1 = x_2$. In other words, if the two functions intersect, there must be a unique intersection.

Since $g^{-1}(x)$ is always monotonically increasing and the existence of multistability requires the existence of multiple intersections between $g^{-1}(x)$ and $f(x)$, a monotonically decreasing $f(x)$ implies that the system cannot be multistable. In other words, if the system is multistable, then $f(x)$ cannot be monotonically decreasing. □

This result highlights the importance of self-excitation w_{EE} in equation S9 and equation S12—multistability can only occur when w_{EE} is sufficiently large. Correspondingly in the numerical result (Figure 2), the region of multistability appears only for larger w_{EE}'s.

Note that the above argument is not restricted to the present model, but applicable to models that share the geometry form of the Wilson-Cowan model in general. Nevertheless, one would hope to know how large a w_{EE} is large enough for multistability to be possible, and this depends on the specific formulation of the transfer function (equation 3) and the underlying assumptions about neuronal level properties (equation 15). Ideally, to know the minimal w_{EE}, one need to find the minimal slope of $H_E^{-1}(u(S_E))$ with respective to S_E, where $u(S_E) := S_E / (\tau_E \gamma_E (1 - S_E))$. The exact solution is, however, rather perplexing to calculate. Here we provide a rough, but simple, estimation instead. The slope of interest is

$$\frac{\mathrm{d}}{\mathrm{d}\, S_E} H_E^{-1}(u(S_E)) = \frac{\mathrm{d}\, H_E^{-1}(u)}{\mathrm{d}\, u} \frac{\mathrm{d}\, u}{\mathrm{d}\, S_E} \tag{S13}$$

$$= \frac{\left(H_E^{-1}\right)'(u)}{\tau_E \gamma_E (1 - S_E)^2}. \tag{S14}$$

Instead of finding the minimum of equation S14, we aim to find a representative point S_E^* such that equation S14 is relatively small.

One option is to use the minimum of the numerator. The minimum of the numerator $\left(H_E^{-1}\right)'(u)$ is simply the reciprocal of the maximum of $H_E'(v)$, where $v = H_E^{-1}(u)$. By design, H_E reaches its maximal slope a_E at the inflection point

$\hat{x}$, where $H_E(v) = r_{max}/2$. That is, we need

$$\frac{S_E^*}{\tau_E \gamma_E (1 - S_E^*)} = \frac{r_{max}}{2}$$
$$S_E^* = \frac{1}{2\, r_{max}^{-1} \tau_E^{-1} \gamma_E^{-1} + 1}.$$

But note here that, in the case where r_{max} is a large number, the representative point S_E^* is very close to one, which further results in a small denominator in equation S14 and a large slope for H_E^{-1}. Thus, the inflection point of $H_E(v)$ is not a very good choice.

To avoid the small denominator problem for equation S14, we need to choose a S_E^* as small as possible while $H_E(v)$ remains close to the line $a_E\, v - b_E$. For this purpose, we take v^* to be the intersection between the line $a_E\, v - b_E$ and the horizontal axis,

$$\begin{aligned} & a_E\, v^* - b_E = 0 \\ \Rightarrow\, & v^* = \frac{b_E}{a_E} \\ \Rightarrow\, & H_E(v^*) \approx \frac{1}{d_E}, \quad H_E'(v^*) \approx \frac{a_E}{2} \end{aligned} \tag{S15}$$

(approximate values can be obtained from the Taylor expansion of $\hat{H}_E$ near v^*). Given equation S15, we need

$$\frac{S_E^*}{\tau_E \gamma_E (1 - S_E^*)} = \frac{1}{d_E} \tag{S16}$$

$$\Rightarrow S_E^* = \frac{1}{d_E \tau_E^{-1} \gamma_E^{-1} + 1} \tag{S17}$$

and

$$\begin{aligned} \left.\frac{\mathrm{d}\, H_E^{-1}(u(S_E))}{\mathrm{d}\, S_E}\right|_{S_E = S_E^*} &= \frac{2}{a_E \tau_E \gamma_E (1 - S_E^*)^2} \\ &= \frac{2(1 + d_E^{-1} \tau_E \gamma_E)^2}{a_E \tau_E \gamma_E} =: h_E. \end{aligned} \tag{S18}$$

Now for the nullcline $S_I = f(S_E)$ to have a positive slope at S_E^*, one simply needs

$$w_{EE} > h_E. \tag{S19}$$

Here h_E is approximately 0.2 based on the present parameter choices, inherited from Wong and Wang's initial derivation [7]. This result is confirmed numerically by the bifurcation diagrams (Figure 5a-c vs. Figure S3a) of the local model—multistability exists for some level of input I_E when $w_{EE} > 0.2$.

Oscillation. Now we look for the conditions for oscillation to emerge. Here we are mainly concerned with the oscillation occurring on the ascending segment of $S_I = f(S_E)$ (red nullcline). Following a similar argument as Wilson and Cowan [8], one notice that for the flow around a fixed point—an intersection between the nullclines—to have consistent rotation, the nullcline $g^{-1}(S_E)$ (blue) must have a greater slope than $f(S_E)$ (red nullcline). Qualitatively, one would expect oscillation to be induced by increasing w_{EI}, which twists $g(S_I)$ (blue nullcline) counterclockwise. This expectation is confirmed by the numerical results in Figure 2a-d: oscillation emerges for sufficiently large w_{EI} for fixed points on the ascending branch of $S_I = f(S_E)$ (Figure 3a-d).

Quantitatively, we consider the derivative of the two nullclines at a respective representative point. First, we extend the results in equation S17-S18 to the second nullcline $S_E = g(S_I)$ (blue):

$$S_I^* = \frac{1}{d_I \tau_I^{-1} \gamma_I^{-1} + 1} \tag{S20}$$

$$h_I := \left. \frac{\mathrm{d}\, H_I^{-1}(u(S_I))}{\mathrm{d}\, S_I} \right|_{S_I = S_I^*} = \frac{2(1 + d_I^{-1} \tau_I \gamma_I)^2}{a_I \tau_I \gamma_I}. \tag{S21}$$

For parameters used in the present study, $h_I \approx 0.4$. We have the slope of the two nullclines at their respective representative points,

$$f'(S_E^*) = \frac{w_{EE} - h_E}{w_{IE}} \tag{S22}$$

$$g'(S_I^*) = \frac{w_{II} + h_I}{w_{EI}}, \tag{S23}$$

and we need

$$\frac{1}{g'(S_I^*)} > f'(S_E^*) \tag{S24}$$

$$\Rightarrow \frac{w_{EI}}{w_{II} + h_I} > \frac{w_{EE} - h_E}{w_{IE}}$$

$$\Rightarrow w_{EI} > \frac{(w_{EE} - h_E)(w_{II} + h_I)}{w_{IE}}. \tag{S25}$$

With balanced inhibitory feedback $w_{IE} = w_{EE}$, we have

$$w_{EI} > (1 - h_E / w_{EE})(w_{II} + h_I). \tag{S26}$$

For very large w_{EE}, one simply need

$$w_{EI} > w_{II} + h_I. \tag{S27}$$

Given the present parameter choices, we need $w_{EI} > 0.45$ to induce oscillation for some level of input I_E and I_I. This is in line with the numerical results in Figure 2. For $h_E > 0$, as assumed here, lowering w_{EE} also lowers the threshold for oscillation.

Linear stability analysis. In addition to the presence of oscillation, one would also want to know if such oscillation is sustainable or damped. Here we extend the above analysis by linearizing the system near a specific fixed point. A fixed point is where the two nullclines (equations S9-S10) intersect. Conveniently, we let them intersect at their respective representative points $(S_E^*, f(S_E^*))$ and $(g(S_I^*), S_I^*)$ (see equation S17 and equation S20),

$$S_E^* = g(S_I^*) \tag{S28}$$

$$f(S_E^*) = S_I^*. \tag{S29}$$

The two equations can be satisfied by the appropriate choice of I_E and I_I. The fixed point of our choice (S_E^*, S_I^*) inherits a couple of properties from the above analysis, which we shall soon see. First, we define

$$\frac{\mathrm{d}\,S_E}{\mathrm{d}\,t} = F(S_E, S_I) := -\frac{S_E}{\tau_E} + (1 - S_E)\gamma_E\, H_E(w_{EE}S_E - w_{IE}S_I + I_E) \tag{S30}$$

$$\frac{\mathrm{d}\,S_I}{\mathrm{d}\,t} = G(S_E, S_I) := -\frac{S_I}{\tau_I} + (1 - S_I)\gamma_I\, H_I(w_{EI}S_E - w_{II}S_I + I_I). \tag{S31}$$

At the fixed points, we have from equation S30

$$H_E(w_{EE}S_E^* - w_{IE}S_I^* + I_E) = \frac{S_E^*}{\tau_E\gamma_E(1 - S_E^*)} \tag{S32}$$

$$= \frac{1}{d_E} \quad \text{(by definition, c.f. equation S16)} \tag{S33}$$

which implies that

$$I_E = \frac{b_E}{a_E} - w_{EE}S_E^* + w_{IE}S_I^* \tag{S34}$$

and

$$H_E'(w_{EE}S_E^* - w_{IE}S_I^* + I_E) = \frac{a_E}{2} \quad \text{(per equation S15).} \tag{S35}$$

Similarly from equation S31, we have

$$H_I(w_{EI}S_E^* - w_{II}S_I^* + I_I) = \frac{S_I^*}{\tau_I\gamma_I(1 - S_I^*)} = \frac{1}{d_I} \tag{S36}$$

$$I_I = \frac{b_I}{a_I} - w_{EI}S_E^* + w_{II}S_I^* \tag{S37}$$

$$H_I'(w_{EI}S_E^* - w_{II}S_I^* + I_I) = \frac{a_I}{2}. \tag{S38}$$

Now we are take the partial derivatives of F and G at (S_E^*, S_I^*),

$$\begin{aligned}
\left.\frac{\partial F}{\partial S_E}\right|_{(S_E^*,S_I^*)} &= -\frac{1}{\tau_E} - \gamma_E H_E(w_{EE}S_E^* - w_{IE}S_I^* + I_E) \\
&\quad + (1-S_E^*)\gamma_E w_{EE} H_E'(w_{EE}S_E^* - w_{IE}S_I^* + I_E) \\
&= -\frac{1}{\tau_E} - \frac{\gamma_E}{d_E} + w_{EE}\gamma_E a_E(1-S_E^*)/2 \qquad \text{(S39)} \\
\left.\frac{\partial F}{\partial S_I}\right|_{(S_E^*,S_I^*)} &= -(1-S_I^*)w_{IE}\gamma_E H_E'(w_{EE}S_E^* - w_{IE}S_I^* + I_E) \\
&= -w_{IE}\gamma_E a_E(1-S_I^*)/2 \qquad \text{(S40)} \\
\left.\frac{\partial G}{\partial S_E}\right|_{(S_E^*,S_I^*)} &= (1-S_I^*)w_{EI}\gamma_I H_I'(w_{EI}S_E^* - w_{II}S_I^* + I_I) \\
&= w_{EI}\gamma_I a_I(1-S_I^*)/2 \qquad \text{(S41)} \\
\left.\frac{\partial G}{\partial S_I}\right|_{(S_E^*,S_I^*)} &= -\frac{1}{\tau_I} - \gamma_I H_I(w_{EI}S_E^* - w_{II}S_I^* + I_I) \\
&\quad - (1-S_I)w_{II}\gamma_I H_I'(w_{EI}S_E^* - w_{II}S_I^* + I_I) \\
&= -\frac{1}{\tau_I} - \frac{\gamma_I}{d_I} - w_{II}\gamma_I a_I(1-S_I^*)/2. \qquad \text{(S42)}
\end{aligned}$$

For simplicity, let parameters

$$\alpha_p := \frac{1}{\tau_p} + \frac{\gamma_p}{d_p} \qquad \text{(S43)}$$

$$\beta_p := \gamma_p a_p(1-S_p^*)/2 \qquad \text{(S44)}$$

$$= \frac{\gamma_p a_p d_p}{2(d_p + \gamma_p \tau_p)} \qquad \text{(S45)}$$

for $p \in \{E, I\}$. Note that by definition, both α_p and β_p are positive. Given parameters used in the present study, we have $\alpha_E \approx 14$, $\alpha_I \approx 111$, $\beta_E \approx 71$, and $\beta_I \approx 276$.

We write the Jacobian matrix as

$$\mathbf{J} = \begin{pmatrix} -\alpha_E + \beta_E w_{EE} & -\beta_E w_{IE} \\ \beta_I w_{EI} & -\alpha_I - \beta_I w_{II} \end{pmatrix}. \qquad \text{(S46)}$$

The eigenvalues of the Jacobian are

$$\lambda_{1,2} = \frac{\operatorname{tr}\mathbf{J} \pm \sqrt{\operatorname{tr}^2\mathbf{J} - 4\det\mathbf{J}}}{2} \qquad \text{(S47)}$$

where

$$\operatorname{tr}\mathbf{J} = -\alpha_E - \alpha_I + \beta_E w_{EE} - \beta_I w_{II} \qquad \text{(S48)}$$

$$\det\mathbf{J} = \alpha_E\alpha_I + \alpha_E\beta_I w_{II} - \alpha_I\beta_E w_{EE} + \beta_I\beta_E(w_{IE}w_{EI} - w_{EE}w_{II}). \qquad \text{(S49)}$$

Assuming that the system is already oscillatory near the fixed point, i.e. $\mathrm{tr}^2\,\mathbf{J} < 4\det\mathbf{J}$, to have sustained oscillation (limit cycle), we need

$$\begin{aligned} &\mathrm{tr}\,\mathbf{J} > 0 \\ \Rightarrow\, &w_{EE} > (\alpha_E + \alpha_I + \beta_I w_{II})/\beta_E. \end{aligned} \tag{S50}$$

Given the parameters used in the present study, the emergence of limit cycles requires $w_{EE} > 2$. Correspondingly in the numerical results (Figure 2), equation S50 provides an estimate of the lower bound of the Hopf bifurcation (gray dashed line). Note that stronger inhibitory-to-inhibitory connection w_{II} increases the minimal w_{EE} required to induce sustained oscillation. Overall, these analyses show that sustained oscillation requires both strong self-excitation and a sufficiently active inhibitory population.

In summary, we have shown analytically how structural connectivity w_{EE} and w_{EI} critically shape the dynamics—in this very low-dimensional parameter space, the system can easily switch between qualitatively different behavior. In particular, excitatory-to-inhibitory connectivity w_{EI} controls the emergence of oscillation; excitatory-to-excitatory connectivity w_{EE} controls both the emergence of multistability and sustained oscillation. The qualitative description of the system only depends on the gross geometric form of the Wilson-Cowan model, but the exact boundaries between regimes depend on the specific transfer function and the associated biophysical constraints.

S5 Analysis of the global model

Now we take a look at the deterministic version of the global model,

$$\frac{\mathrm{d}\,S_E^{(i)}}{\mathrm{d}\,t} = -\frac{S_E^{(i)}}{\tau_E} + (1 - S_E^{(i)})\gamma_E H_E \left(w_{EE}^{(i)} S_E^{(i)} - w_{IE}^{(i)} S_I^{(i)} + I_G^{(i)}(\vec{S}_E) \right) \tag{S51}$$

$$\frac{\mathrm{d}\,S_I^{(i)}}{\mathrm{d}\,t} = -\frac{S_I^{(i)}}{\tau_I} + (1 - S_I^{(i)})\gamma_I H_I \left(w_{EI}^{(i)} S_E^{(i)} - w_{II}^{(i)} S_I^{(i)} + I_I \right) \tag{S52}$$

where

$$I_G^{(i)}(\vec{S}_E) = G \sum_{\substack{j=1 \\ j\neq i}}^{N} C_{ij} S_E^{(j)}. \tag{S53}$$

In this case, the nullclines are hyper-surfaces,

$$w_{IE}^{(i)} S_I^{(i)} = f^{(i)}\left(\vec{S}_E\right) := w_{EE}^{(i)} S_E^{(i)} - H_E^{-1}\left(\frac{S_E^{(i)}}{\tau_E \gamma_E (1 - S_E^{(i)})} \right) + G \sum_{j\neq i} C_{ij} S_E^{(j)} \tag{S54}$$

$$w_{EI}^{(i)} S_E^{(i)} = g^{(i)}\left(S_I^{(i)}\right) := w_{II}^{(i)} S_I^{(i)} + H_I^{-1}\left(\frac{S_I^{(i)}}{\tau_I \gamma_I (1 - S_I^{(i)})} \right) - I_I. \tag{S55}$$

From equation S54 one can see that the local effect of global coupling is simply tilting the nullcline $S_I^{(i)} = f^{(i)}\left(\vec{S}_E\right)/w_{IE}^{(i)}$ upwards with respect to $S_E^{(j)}$.

The tilting of the nullcline $S_I^{(i)} = f^{(i)}\left(\vec{S}_E\right)/w_{IE}^{(i)}$ impact its number of intersections with $S_E^{(i)} = g^{(i)}\left(S_I^{(i)}\right)/w_{EI}^{(i)}$ in each level set of $G\sum_{j\neq i} C_{ij}S_E^{(j)}$. The number of intersections consequently constrains the number of stable states. A precise characterization of intersections is beyond the scope of the present work. Nevertheless, we hope to provide a few insights about the global geometry below.

Multistability. Following a similar argument as for the local model, we first show that, without global interaction (i.e. $G = 0$), the system cannot be multistable, if w_{EE} is sufficiently small such that $f^{(i)}(\vec{S}_E)$ monotonically decreases with $S_E^{(i)}$ for all i. As shown above, the monotonicity condition implies that each local node by itself is not multistable.

Proof. Assume there are at least two distinct fixed points of the system: $\vec{S}^*$ and $\vec{S}^{**}$, where $\vec{S} = (\vec{S}_E, \vec{S}_I)$ and $\vec{S}_p = (S_p^{(1)}, \cdots S_p^{(i)}, \cdots S_p^{(N)})$ for $p \in \{E, I\}$. Since they are distinct points, there exists an $0 < i \leqslant N$ such that $S_E^{*(i)} \neq S_E^{**(i)}$ ($S_I^{*(i)} \neq S_I^{**(i)}$ implies $S_E^{*(i)} \neq S_E^{**(i)}$ due to the monotonicity of g). Without loss of generality, we let $S_E^{*(i)} < S_E^{**(i)}$.

Since we know that $g^{-1\,(i)}(w_{EI}^{(i)}S_E^{(i)})$ is always a monotonically increasing function, we have

$$g^{-1\,(i)}(w_{EI}^{(i)}S_E^{*(i)}) < g^{-1\,(i)}(w_{EI}^{(i)}S_E^{**(i)}) \tag{S56}$$

$$\Rightarrow S_I^{*(i)} < S_I^{**(i)}, \tag{S57}$$

which also implies that

$$f^{(i)}\left(\vec{S}_E^*\right) < f^{(i)}\left(\vec{S}_E^{**}\right) \tag{S58}$$

by definition of the nullcline $S_I^{(i)} = f^{(i)}\left(\vec{S}_E\right)/w_{IE}^{(i)}$, for any choice of G and C_{ij}.

Now if $f^{(i)}(\vec{S}_E)$ is monotonically decreasing with respect to $S_E^{(i)}$, we know that at least for $G = 0$,

$$f^{(i)}\left(\vec{S}_E^*\right) > f^{(i)}\left(\vec{S}_E^{**}\right), \tag{S59}$$

which leads to a contradiction. Thus, if the system has multiple fixed points, $f^{(i)}(\vec{S}_E)$ cannot be monotonically decreasing with respect to $S_E^{(i)}$ for all i when $G = 0$. □

However, given a sufficiently large global coupling, especially for $G > 1$, multistability becomes possible.

Proof. Following the above proof, the assumption $S_E^{*(i)} < S_E^{**(i)}$ leads us to equation S58, or $\Delta_G < 0$, where

$$\Delta_G := f_G^{(i)}\left(\vec{S}_E^*\right) - f_G^{(i)}\left(\vec{S}_E^{**}\right) \tag{S60}$$

for any global coupling $G \geqslant 0$.

On the other hand, for the special case of $G = 0$, we can plug equation S54 into the definition S60 and have

$$\Delta_0 = f_{G=0}^{(i)}\left(\vec{S}_E^*\right) - f_{G=0}^{(i)}\left(\vec{S}_E^{**}\right) = w_{IE}^{(i)}\left(S_I^{*(i)} - S_I^{**(i)}\right) \tag{S61}$$

$$= f_{G=0}^{(i)}\left(S_E^{*(i)}\right) - f_{G=0}^{(i)}\left(S_E^{**(i)}\right) > 0, \tag{S62}$$

by our assumption that $f_G^{(i)}(\vec{S}_E)$ is a monotonically decreasing function with respect to $S_E^{(i)}$. Since by definition, the coordinates of each fixed point is bounded between 0 and 1, we have

$$0 < \Delta_0 \leqslant w_{IE}^{(i)}. \tag{S63}$$

In the case of $G = 0$, this leads to a contradiction $\Delta_G > 0$, as we have already shown above.

Now we consider what happens when $G > 0$. Again, by plugging equation S54 into the definition S60, we have

$$\begin{aligned}\Delta_G &= \Delta_0 + G\sum_{j\neq i} C_{ij} S_E^{*(j)} - G\sum_{j\neq i} C_{ij} S_E^{**(j)} \\ &= \Delta_0 + G\sum_{j\neq i} C_{ij}(S_E^{*(j)} - S_E^{**(j)}).\end{aligned} \tag{S64}$$

We need a bound on the second term in equation S64. Since $G > 0$ and $C_{ij} \geqslant 0$,

$$\begin{aligned}\left|G\sum_{j\neq i} C_{ij}(S_E^{*(j)} - S_E^{**(j)})\right| &\leqslant G\sum_{j\neq i} C_{ij}\left|(S_E^{*(j)} - S_E^{**(j)})\right| \\ &\leqslant G\sum_{j\neq i} C_{ij} && \text{(since } 0 \leqslant S_E^{(j)} \leqslant 1) \\ &\leqslant G && \text{(by equation 7).}\end{aligned}$$

This gives us

$$\Delta_0 - G \leqslant \Delta_G \leqslant \Delta_0 + G. \tag{S65}$$

Thus, contradiction with equation S58 is inevitable if $G < \Delta_0$. On the other hand, by equation S63, we know that for $G > \Delta_0$, there exists some $\vec{S}_E^*$ and $\vec{S}_E^{**}$ for some global network C_{ij} such that $\Delta_G < 0$ consistent with equation S58. Thus it is possible for the global model to be multistable if $G > \Delta_0$, especially if $G > w_{IE}^{(i)}$, or $G > w_{EE}^{(i)}$ given matched inhibitory feedback $w_{EE} = w_{IE}$. This does not mean, however, that the system has to be multistable, due to the dependency on C_{ij}. □

To summarize, the above analyses suggest that a collection of brain regions that have no independent memory capacity (i.e. multistability) can *acquire* memory capacity when connected to each other in a global network, given sufficient global coupling. We further support this claim with numerical analysis (Figure S3). We refer to this kind of memory as *synergistic* memory—it is an emergent property that the parts themselves do not possess.

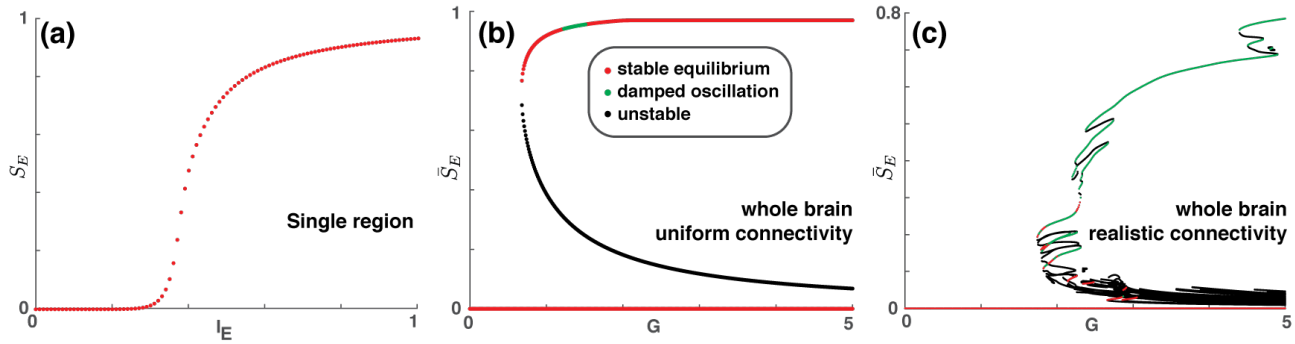

Figure S3: **Synergistic memory between monostable nodes**. Three bifurcation diagrams are shown for local parameters $w_{EE} = 0.1$ and $w_{EI} = 0.35$. They correspond to Figure 5a, d, g but with a lower w_{EE} such that each local node by itself is monostable for any level of input (a). While each local node is completely monostable (no memory capacity), once there is sufficient global coupling G between them, the whole brain acquires memory capacity (b, c) that cannot be attributed to the parts alone—synergistic memory. Nevertheless, the size of the global memory capacity is still fundamentally constrained by the complexity of the local node (42 attractor branches in (c), very small compare to Figure 5g, h, i). See text for further discussion.

What we have not addressed in the above analyses is to what extent the global system is multistable—what is the number of stable states, or the size of the memory capacity—and what are the contributions from local self-excitation and global network connectivity. An analytical approach to this problem is difficult; thus, it is mainly addressed numerically (c.f. Figure S3 and Figure 5). Nevertheless, we provide an intuitive argument below as to how local and global connectivity affects the relevant geometrical properties of the dynamical system.

Local origin of geometrical complexity. At an intuitive level, the number of intersections between these hypersurfaces (equation S54-S55) is likely to increase with the number of folds of each surface. In the present case, the folding of hypersurfaces entails the temporary reversal of the sign of its partial derivative along a certain direction. Observe equation S54 and see that global coupling cannot create any folding of the surfaces. Thus, the geometrical complexity of the nullclines purely depends on the local properties of each node, in particular, the folding effect of self-excitation $w_{EE}^{(i)}$.

The effect of global coupling. Without global coupling ($G = 0$), the number of fixed points of the global model is simply

$$n = \prod_{i=1}^{N} n_0^{(i)} \leqslant \left(\max_i n_0^{(i)}\right)^N \tag{S66}$$

where $n_0^{(i)}$ is the number of fixed points for each corresponding local model when $I_E = 0$. Introducing global coupling ($G \neq 0$) tilts each surface (equation S54) in a way dependent on the structure connectivity C_{ij}. This may remove or introduce new intersections between the surfaces without changing the geometrical complexity of these surfaces. Thus, global coupling allows system-level multistability to be created synergistically, given appropriate structural connectivity C_{ij}.

In summary, local and global coupling produce different geometrical effects on the system and jointly affect the number of possible stable states.

References

1. Buxton, R. B., Wong, E. C. & Frank, L. R. Dynamics of blood flow and oxygenation changes during brain activation: The balloon model. *Magnetic Resonance in Medicine* **39,** 855–864 (1998).
2. Mandeville, J. B. *et al.* Evidence of a Cerebrovascular Postarteriole Windkessel with Delayed Compliance. *Journal of Cerebral Blood Flow & Metabolism* **19,** 679–689 (1998).
3. Friston, K., Mechelli, A., Turner, R. & Price, C. Nonlinear responses in fMRI: the Balloon model, Volterra kernels, and other hemodynamics. *NeuroImage* **12** (2000).
4. Friston, K., Harrison, L. & Penny, W. Dynamic causal modelling. *NeuroImage* **19,** 1273–1302 (2003).
5. Welch, P. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics* **15,** 70–73 (1967).
6. Van Essen, D. C. *et al.* The WU-Minn Human Connectome Project: An overview. *NeuroImage* **80,** 62–79 (2013).
7. Wong, K.-F. & Wang, X.-J. A recurrent network mechanism of time integration in perceptual decisions. *Journal of Neuroscience* **26,** 1314–1328 (2006).
8. Wilson, H. R. & Cowan, J. D. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal* **12,** 1–24 (1972).