How local excitation-inhibition ratio impacts the whole brain dynamics

Gustavo Deco  
*Universitat Pompeu Fabra*

Adrian Ponce-Alvarez  
*Universitat Pompeu Fabra*

Patric Hagmann  
*University of Lausanne*

Gian Luca Romani  
*G. d'Annunzio University*

Dante Mantini  
*University of Oxford*

See next page for additional authors

Follow this and additional works at: http://digitalcommons.wustl.edu/open_access_pubs

Recommended Citation  
Deco, Gustavo; Ponce-Alvarez, Adrian; Hagmann, Patric; Romani, Gian Luca; Mantini, Dante; and Corbetta, Maurizio, "How local excitation-inhibition ratio impacts the whole brain dynamics." *The Journal of Neuroscience*. 34,23. 7886-7898. (2014).  
http://digitalcommons.wustl.edu/open_access_pubs/2973

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact engeszer@wustl.edu.
Authors
Gustavo Deco, Adrian Ponce-Alvarez, Patric Hagmann, Gian Luca Romani, Dante Mantini, and Maurizio Corbetta

This open access publication is available at Digital Commons@Becker: http://digitalcommons.wustl.edu/open_access_pubs/2973
How Local Excitation–Inhibition Ratio Impacts the Whole Brain Dynamics

Gustavo Deco,1,2* Adrián Ponce-Alvarez,3,4 Patric Hagmann,3,4 Gian Luca Romani,5 Dante Martini,6,7 and Maurizio Corbetta6

1Center for Brain and Cognition, Computational Neuroscience Group, Department of Information and Communication Technologies, and 2Institució Catalana de la Recerca i Estudis Avançats (ICREA), Universitat Pompeu Fabra, Barcelona, 08010, Spain, 3Department of Radiology, Lausanne University Hospital and University of Lausanne (CHUV-UNIL), 1011 Lausanne, Switzerland, 4Signal Processing Laboratory 5, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland, 5Institute of Advanced Biomedical Technologies, G. d’Annunzio University Foundation, Department of Neuroscience and Imaging, G. d’Annunzio University, Chieti, 66013, Italy, 4Department of Health Sciences and Technology, ETH Zurich, 8057, Zurich, Switzerland, 7Department of Experimental Psychology, University of Oxford, OX1 3UD, Oxford, United Kingdom, and 6Department of Neurology, Radiology, Anatomy of Neurobiology, School of Medicine, Washington University in St Louis, St Louis, Missouri 63110

The spontaneous activity of the brain shows different features at different scales. On one hand, neuroimaging studies show that long-range correlations are highly structured in spatiotemporal patterns, known as resting-state networks, on the other hand, neurophysiological reports show that short-range correlations between neighboring neurons are low, despite a large amount of shared presynaptic inputs. Different dynamical mechanisms of local decorrelation have been proposed, among which is feedback inhibition. Here, we investigated the effect of locally regulating the feedback inhibition on the global dynamics of a large-scale brain model, in which the long-range connections are given by diffusion imaging data of human subjects. We used simulations and analytical methods to show that locally constraining the feedback inhibition to compensate for the excess of long-range excitatory connectivity, to preserve the asynchronous state, crucially changes the characteristics of the emergent resting and evoked activity. First, it significantly improves the model’s prediction of the empirical human functional connectivity. Second, relaxing this constraint leads to an unrealistic network evoked activity, with systematic coactivation of cortical areas which are components of the default-mode network, whereas regulation of feedback inhibition prevents this. Finally, information theoretic analysis shows that regulation of the local feedback inhibition increases both the entropy and the Fisher information of the network evoked responses. Hence, it enhances the information capacity and the discrimination accuracy of the global network. In conclusion, the local excitation–inhibition ratio impacts the structure of the spontaneous activity and the information transmission at the large-scale brain level.

Key words: anatomical connectivity; functional connectivity; large-scale brain model; local feedback inhibition; resting-state activity

Introduction

The spontaneous activity of the brain, i.e., not stimulus- or task-driven, shows different features at different spatial scales. Indeed, although long-range activity correlations between brain regions are highly and strongly structured in spatiotemporal patterns, known as resting-state networks, on one hand, short-range correlations within local circuits are low or even negligible, on the other hand.

Numerous neuroimaging experiments demonstrate the existence of spontaneous long-range correlations, i.e., resting functional connectivity (FC), by fMRI (Biswal et al., 1995; Greicius et al., 2003; Fransson, 2005; Fox and Raichle, 2007), MEG (Li et al., 2010; Brookes et al., 2011; Luckhoo et al., 2012; de Pasquale et al., 2012), and EEG-fMRI techniques (Martini et al., 2007). The topology of human FC patterns has been studied in detail across different conditions of rest or cognitive tasks (Achard et al., 2006, Bassett et al., 2006). In humans, a consistent set of functional networks (default, dorsal attention, ventral attention, vision, somatomotor, auditory, frontoparietal) have been identified in the resting-state activity across subjects and across multiple sessions using a variety of methods, which can be used to generate whole-brain functional parcellations of the cerebral cortex (Douet el al., 2011; Power et al., 2011; Yeo et al., 2011; Hacker et al., 2013). Of particular interest is the default network: a set of correlated brain regions that are more activated at rest than during the performance of cognitive goal-directed tasks (Shulman et al., 1997; Raichle et al., 2001; Fox and Raichle, 2007).
In contrast, single-cell recordings in monkeys indicate that short-range spontaneous correlations between neighboring neurons are low (Ecker et al., 2010), despite a large amount of shared presynaptic inputs. Different mechanisms have been proposed to account for this decorrelation including chaotic dynamics (van Vreeswijk and Sompolinsky, 1996), balance of excitation and inhibition (Renart et al., 2010), and feedback inhibition (Tetzlaff et al., 2012). In theoretical models, these mechanisms all produce the observed low (~3 Hz for excitatory neurons) and irregular (Poisson-like) cortical activity observed in vivo (Burns and Webb, 1976; Softky and Koch, 1993; Wilson et al., 1994).

The aim of this study is to investigate the effect of local decorrelation on both the resting and evoked activity of a large-scale brain model. Resting-state brain models assume that long-range functional correlations emerge from the embedding of local dynamics in the underlying anatomical connectivity structure (Honey et al., 2007, 2009; Ghosh et al., 2008; Deco et al., 2009, 2013a,b; Deco and Hughes, 2012). In general, biophysical resting-state models use single-area dynamical models that, when isolated, emulate spontaneous state activity, i.e., uncorrelated low firing rate activity. However, the long-range anatomical couplings break down the asynchronous state, producing intra-area correlations which are not consistent with the empirical findings. Here, we show that the regulation of the local feedback inhibition provides a mechanism for reconciling the two levels of spontaneous correlation at different spatial scales and we evaluate its functional consequences for both the resting and the evoked global brain dynamics.

Materials and Methods

Structural connectivity matrix. Neuroanatomical structure was obtained using diffusion spectrum imaging (DSI) data and tractography from five healthy right-handed male human subjects (Hagmann et al., 2008; Honey et al., 2009). The gray matter was subdivided into 998 regions-of-interest (ROIs) which are grouped into 33 cortical regions per hemisphere (66 areas in total) according to anatomical landmarks (Table 1). White matter tractography was used to estimate the fiber tract density connecting each pair of ROIs, averaged across subjects. Anatomical connectivity among the 66 cortical regions was calculated by summing all incoming fiber strengths to the corresponding ROIs of the target region, and dividing it by its region-dependent number of ROIs, resulting in a nonsymmetric connectivity matrix. This normalization by the number of ROIs, which have approximately the same surface on the cortex, i.e., the same number of neurons, is required because neuronal activity is sensitive to the number of incoming fibers per neuron in the target region. As the dynamical model of one region already takes into account the effect of its internal connectivity (see below), the connection of a region to itself was set to 0 in the connectivity matrix for the simulations.

Empirical functional connectivity. The empirical resting state MRI FC was measured in 48 scanning sessions from 24 right-handed healthy young volunteers (15 females, age range 20–31 years). Subjects were informed about the experimental procedures and provided written informed consent. The study design was approved by the local Ethics Committee of Chieti University. Subjects lay in a supine position and viewed a black screen with a centered red fixation point of 0.3 visual degrees, through a mirror tilted by 45 degrees. Each volunteer underwent two scanning runs of 10 min in a resting-state condition. Specifically, they were instructed to relax, but to maintain fixation during scanning. The eye position was monitored at 120 Hz during scanning using an ISCAN eye tracker system. Scanning was performed with a 3T MR scanner (Achieva; Philips Medical Systems) at the Institute for Advanced Biomedical Technologies in Chieti, Italy. Functional images were obtained using T2-weighted echo-planar imaging (EPI) with blood oxygenation level-dependent (BOLD) contrast using SENSE imaging. EPIs (TR/TE = 2000/35 ms) comprised 32 axial slices acquired continuously in ascending order covering the entire brain (voxel size = 3 × 3 × 3.5 mm³). For each scanning run, initial five dummy volumes allowing the MRI signal to reach steady-state were discarded. The next 300 functional volumes were used for the analysis. A three-dimensional high-resolution T1-weighted image (TR/TE = 9.64/4.6 ms, voxel size = 0.98 × 0.98 × 1.2 mm³) covering the entire brain was acquired at the end of the scanning session and used for anatomical reference. Initial data preprocessing was performed using the SPM5 software package (Wellcome Department of Cognitive Neurology, London, UK) running under MATLAB (MathWorks). The preprocessing steps involved the following: (1) correction for slice-timing differences, (2) correction of head-motion across functional images, (3) coregistration of the anatomical image and the mean functional image, and (4) spatial normalization of all images to a standard stereotaxic space (Montreal Neurological Institute; MNI) with a voxel size of 3 × 3 × 3 mm³. Furthermore, the BOLD time series in MNI space were subjected to spatial independent component analysis (ICA) for the identification and removal of artifacts related to blood pulsation, head movement and instrumental spikes and those that correlate with the white matter and CSF patterns (Sui et al., 2009; Mantini et al., 2013). The BOLD artifact removal procedure was performed using the GIFT toolbox (http://mialab.mrn.org/software/gift/index.html). No global signal regression was performed. For each recording session (subject and run), the BOLD time series from the 998 ROIs of the brain atlas (Hagmann et al., 2008) were averaged over the corresponding 66 brain regions. Finally, we concatenated in time the remaining sessions for each parcel and calculated the correlation matrix (FC).

Computational model. In this study we modeled a cortical area as a canonical local network composed of interconnected excitatory and inhibitory neurons coupled through NMDA, AMPA, and GABA synapses. In the large-scale brain model, we assumed that the inter-area connections are constrained by the empirical anatomical matrix that expresses

---

**Table 1. Names and abbreviations (alphabetically) of the brain regions considered in the human connectome from Hagmann et al. (2008)**

<table>
<thead>
<tr>
<th>Brain region</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bank of the superior temporal sulcus</td>
<td>BSTS</td>
</tr>
<tr>
<td>Caudal anterior cingulate cortex</td>
<td>CACC</td>
</tr>
<tr>
<td>Caudal middle frontal cortex</td>
<td>CMF</td>
</tr>
<tr>
<td>Cuneus</td>
<td>CUN</td>
</tr>
<tr>
<td>Entorhinal cortex</td>
<td>ENT</td>
</tr>
<tr>
<td>Frontal pole</td>
<td>FP</td>
</tr>
<tr>
<td>Fusiform gyrus</td>
<td>FUS</td>
</tr>
<tr>
<td>Inferior parietal cortex</td>
<td>IPC</td>
</tr>
<tr>
<td>Isthmus of the cingulate cortex</td>
<td>ISTC</td>
</tr>
<tr>
<td>Inferior temporal cortex</td>
<td>IT</td>
</tr>
<tr>
<td>Lingual gyrus</td>
<td>LING</td>
</tr>
<tr>
<td>Lateral occipital cortex</td>
<td>LOC</td>
</tr>
<tr>
<td>Lateral orbitofrontal cortex</td>
<td>LOF</td>
</tr>
<tr>
<td>Medial orbitofrontal cortex</td>
<td>MOF</td>
</tr>
<tr>
<td>Middle temporal cortex</td>
<td>MT</td>
</tr>
<tr>
<td>Paracentral lobule</td>
<td>PARC</td>
</tr>
<tr>
<td>Parahippocampal cortex</td>
<td>PARH</td>
</tr>
<tr>
<td>Posterior cingulate cortex</td>
<td>PC</td>
</tr>
<tr>
<td>Pericalcarine cortex</td>
<td>PICAL</td>
</tr>
<tr>
<td>Precuneus</td>
<td>PCUN</td>
</tr>
<tr>
<td>Pars opercularis</td>
<td>POPE</td>
</tr>
<tr>
<td>Pars orbitalis</td>
<td>PORB</td>
</tr>
<tr>
<td>Precentral gyrus</td>
<td>PREC</td>
</tr>
<tr>
<td>Postcentral gyrus</td>
<td>PSC</td>
</tr>
<tr>
<td>Pars triangularis</td>
<td>PTRY</td>
</tr>
<tr>
<td>Rostral anterior cingulate cortex</td>
<td>RAC</td>
</tr>
<tr>
<td>Rostral middle frontal cortex</td>
<td>RMF</td>
</tr>
<tr>
<td>Superior frontal cortex</td>
<td>SF</td>
</tr>
<tr>
<td>Supramarginal gyrus</td>
<td>SMAR</td>
</tr>
<tr>
<td>Superior parietal cortex</td>
<td>SP</td>
</tr>
<tr>
<td>Superior temporal cortex</td>
<td>ST</td>
</tr>
<tr>
<td>Temporal pole</td>
<td>TP</td>
</tr>
<tr>
<td>Transverse temporal cortex</td>
<td>TT</td>
</tr>
</tbody>
</table>
the density of white matter fiber tracts connecting two given brain areas. However, because the large-scale spiking model is composed of thousands of coupled nonlinear differential equations corresponding to each single neuron and synapse, some approximations need to be done to simplify the model. For this, we used a dynamic mean field (DMF) technique (Wong and Wang, 2006) that approximates the average ensemble behavior, instead of considering the detailed interactions between individual neurons. This allows us varying the parameters and, furthermore, to greatly simplify the system of stochastic differential equations by expressing it in terms of the first and second-order statistical moments, means and covariances, of network activity. In the following we present the details of the model and its simplified versions.

Single area spiking model. To show the effect of feedback inhibition on the correlations between intra-area neurons and on the firing rate activity we used a network of integrate-and-fire (IF) spiking neurons with excitatory (AMPA and NMDA) and inhibitory (GABA-A) synaptic receptor types (Deco and Jirsa, 2012). Each local cortical area is modeled by a fully connected recurrent network of a population of \( N_E \) excitatory pyramidal neurons and a population of \( N_I \) inhibitory neurons. The dynamical evolution of the IF neuron’s membrane potential \( V(t) \) is driven by the incoming local excitatory and inhibitory inputs within the same cortical area, long-range excitatory inputs from all other cortical areas, and external inputs. The time evolution of the membrane potential of a given neuron obeys the following differential equation:

\[
C_m \frac{dV(t)}{dt} = -g_N(V(t) - V_s) - g_{AMPA,ext} (V(t) - V_s) s_{AMPA,ext}(t) - g_{AMPA,rec} (V(t) - V_s) \sum_{j=1}^{N_E} w_{ij} s_{AMPA,rec}(t) - g_{NMDA}(V(t) - V_s) \sum_{j=1}^{N_I} w_{ij} NMDA(t) - g_{GABA}(V(t) - V_s) \sum_{j=1}^{N_I} w_{ij} GABA(t)
\]

(1)

if the membrane potential is below a given threshold \( V_{th} \). The neuron generates a spike, when the membrane potential reaches the threshold \( V_{th} \). The spike is transmitted to other neurons and the membrane potential is instantaneously reset to \( V_{reset} \) and maintained there for a refractory time \( \tau_{ref} \) during which the neuron is unable to produce further spikes. In Equation 1, \( g_m \) is the membrane leak conductance, \( C_m \) is the capacity of the membrane, and \( V_s \) is the resting potential. The membrane time constant is defined by \( \tau_m = C_m/g_m \). The synaptic input current is given by the last four terms on the right hand side of Equation 1. The spikes arriving at the synapse produce a postsynaptic excitatory or inhibitory potential given by a conductance-based model specified by the synaptic receptors, corresponding to: glutamatergic AMPA external excitatory currents, AMPA and NMDA recurrent excitatory currents, and GABAergic recurrent inhibitory currents. The respective synaptic conductances are \( g_{AMPA,ext} \), \( g_{AMPA,rec} \), \( g_{NMDA} \), and \( g_{GABA} \) and \( V_s \) and \( V_f \) are the excitatory and inhibitory reversal potentials, respectively. The dimensionless parameters \( w_{ij} \) of the connections are the synaptic weights defined as following: the recurrent self-excitation within each excitatory population is given by the weight \( w_{EE} = w_{+}, \) for both AMPA and NMDA recurrent synapses; recurrent inhibition and connections from excitatory to inhibitory neurons and vice versa have the weight \( w_{IP} \) and \( w_{IR} \), respectively. The gating variables \( s(t) \) are the fractions of open channels of neurons and are given by the following:

\[
\frac{ds(t)}{dt} = \frac{s(t)}{\tau_s} + \sum \delta(t - t_i), \text{ for } J = AMPA \text{ or GABA} \quad (2)
\]

\[
\frac{ds_NMDA(t)}{dt} = -\frac{s_NMDA(t)}{\tau_{NMDA,decay}} + \alpha s_NMDA(t)(1 - s_NMDA(t)) \quad (3)
\]

\[
\frac{dx_NMDA(t)}{dt} = -\frac{x_NMDA(t)}{\tau_{NMDA,rise}} + \sum \delta(t - t_i) \quad (4)
\]

Where the sums over the index \( k \) represent all the spikes emitted by the presynaptic neuron \( j \) (at times \( t_i \)); \( \tau_{NMDA,rise} \) and \( \tau_{NMDA,decay} \) represent the decay times for AMPA and GABA synapses, and \( \tau_{NMDA,rise} \) and \( \tau_{NMDA,decay} \) are the rise and decay times for the NMDA synapses. All neurons in the network receive an AMPA-mediated external background input of uncorrelated Poisson spike trains with a mean rate of \( \nu_0 = 2.4 \) kHz. Parameter values of the neurons and the synapses are summarized in Table 2.

Table 2. Parameters for spiking and the DMF model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N_E )</td>
<td>160</td>
<td>( N_I )</td>
<td>40</td>
</tr>
<tr>
<td>( C_m )</td>
<td>0.5 nS</td>
<td>( C_m )</td>
<td>0.2 nS</td>
</tr>
<tr>
<td>( g_N )</td>
<td>25 nS</td>
<td>( g_N )</td>
<td>20 nS</td>
</tr>
<tr>
<td>( V_s )</td>
<td>-70 mV</td>
<td>( V_s )</td>
<td>-70 mV</td>
</tr>
<tr>
<td>( \tau_{ref} )</td>
<td>2 ms</td>
<td>( \tau_{ref} )</td>
<td>1 ms</td>
</tr>
<tr>
<td>( g_{AMPA,ext} )</td>
<td>3.37 nS</td>
<td>( g_{AMPA,ext} )</td>
<td>2.59 nS</td>
</tr>
<tr>
<td>( g_{AMPA,rec} )</td>
<td>0.065 nS</td>
<td>( g_{AMPA,rec} )</td>
<td>0.051 nS</td>
</tr>
<tr>
<td>( g_{NMDA} )</td>
<td>0.2 nS</td>
<td>( g_{NMDA} )</td>
<td>0.16 nS</td>
</tr>
<tr>
<td>( g_{GABA} )</td>
<td>10.94 nS</td>
<td>( g_{GABA} )</td>
<td>8.51 nS</td>
</tr>
</tbody>
</table>

Large-scale cortical dynamic mean field model. As we will see below, the adjustment of the local feedback inhibition of each cortical area involves the recursive adaptation of the inhibitory–excitatory weights. This algorithm requires the computation of the global dynamics during a long period of time, until the FIC constraint is satisfied in all areas. Therefore, we used here a reduced DMF (Wong and Wang, 2006) for computing the global dynamics of the whole cortex. The DMF expresses consistently the time evolution of the ensemble activity of the different excitatory and inhibitory neural populations building up the spiking network. In the DMF approach, each population firing rate depends on the input currents into that population. On the other hand, the input currents depend on the firing rates. Hence, the population firing rate can be determined self-consistently by a reduced system of coupled nonlinear differential equations expressing the population firing rates and the respective input currents. The large-scale model interconnects these local subnetworks according the structural connectivity matrix (SC) defined by the neuroanatomical connections between those brain areas in the human, as obtained by DSI and described above. The inter-area connections are established as long range excitatory synaptic connections either between excitatory pools of different areas only or both between excitatory pools and from excitatory pools to inhibitory the inhibitory pools of different areas (feedforward inhibition). Inter-area connections are weighted by the strength specified in the neuroanatomical matrix \( SC \), denoting the density of fibers between those regions, and by a global scaling factor \( G \). The global scaling factor is a free control parameter that we vary systematically to study the dynamics of the global cortical system.

In brief, the mean field approach considers the diffusion approximation according to which sums of synaptic gating variables (Eq. 1) are replaced by the averaged component and a Gaussian fluctuation term.

Deco, Ponce-Alvarez et al. • Balanced Resting-State
Moreover, the first passage time equation, that gives the mean firing rate of a neuron receiving a noisy input, is approximated by a simple sigmoidal input–output function giving the firing rate as a function of the inputs (Wong and Wang, 2006). Because the synaptic gating variable of NMDA receptors has a much longer decay time constant (100 ms) than the AMPA receptors, the dynamics of the NMDA gating variable dominates the time evolution of the system, while the AMPA synaptic variable instantaneously reaches its steady-state. We thus neglected contributions by the AMPA receptors to the local recurrent excitation. All these approximations greatly simplify the model without compromising its performance. Indeed, it has been previously shown that these approximations conserve the first-order (fixed points of the mean activity) and the second-order (correlation structure) statistics of the original large-scale spiking model (Deco et al., 2013a) which includes AMPA, NMDA, and GABA synapses. Finally, in the present study long-range–mediated connections are assumed to be instantaneous, thus conduction delays between distant cortical areas are neglected.

The global brain dynamics are described by the following set of coupled nonlinear stochastic differential equations:

$$I_i^{(E)} = W_{Ei}b_i + w_iJ_{NMDA}S_i^{(E)} + GL_{NMDA} \sum_j C_{ij} \cdot S_j^{(E)} - I_i^{(E)} - I_{external}$$

(5)

$$I_i^{(I)} = W_{Ii}b_i + J_{NMDA}S_i^{(I)} - S_i^{(I)} - \lambda GL_{NMDA} \sum_j C_{ij} S_j^{(I)},$$

(6)

$$r_i^{(E)} = H_i^{(E)}(I_i^{(E)}) = \frac{1}{1 - \exp(-d_i(a_i I_i^{(E)} - b_i))},$$

(7)

$$r_i^{(I)} = H_i^{(I)}(I_i^{(I)}) = \frac{1}{1 - \exp(-d_i(a_i I_i^{(I)} - b_i))},$$

(8)

$$\frac{dS_i^{(E)}}{dt} = -\frac{S_i^{(E)}}{\tau_E} + (1 - S_i^{(E)})y_{r_i^{(E)}} + \sigma v_i(t),$$

(9)

$$\frac{dS_i^{(I)}}{dt} = -\frac{S_i^{(I)}}{\tau_I} + r_i^{(I)} + \sigma v_i(t),$$

(10)

where $r_i^{(E)}$ denotes the population firing rate of the excitatory (E) or inhibitory (I) population in the brain area i. $S_i^{(E)}$ denotes the average excitatory or inhibitory synaptic gating variable at the local area i. The input currents to the excitatory or inhibitory population i is given by $I_i^{(E)}$. $I_{external}$ encodes external stimulation for simulating task evoked activity (it is zero for all pools under resting state condition, and 0.02 for those pools excited in a task condition). Furthermore, $w_i = 1.4$ is the local excitatory recurrency, and $C_{ij}$ is the structural connectivity matrix expressing the neuroanatomical links between the areas i and j. The parameter $\lambda$ can be equal to 1 or 0 and indicates whether long-range feedforward inhibition (FFI) is considered ($\lambda = 1$) or not ($\lambda = 0$). Note that in the case of FFI the proportion of excitatory–inhibitory long-range connections and excitatory–excitatory long-range connections is the same. $H_i^{(E)}$ and $H_i^{(I)}$ denotes the neuronal input–output functions of excitatory pools and inhibitory pools, respectively. The input–output function converts incoming inputs into firing rates. The kinetic parameters are $\gamma = 0.641/1000$ (the factor 1000 is for expressing everything in ms), and $\tau_E = \tau_{NMDA}$ and $\tau_I = \tau_{GABA}$. The excitatory synaptic coupling $J_{NMDA} = 0.15$ (nA) and the local feedback inhibitory synaptic coupling $J_i$ is 1 for each brain area i in the no-FFI case, and for the FCI case it is adjusted independently, by the algorithm described below. The overall effective external input is $I_{external} = 0.382$ (nA) scaled by $W_{Ei}$ and $W_{Ii}$ for the excitatory pools and the inhibitory pools, respectively. In Equations 9 and 10 $\nu_i$ is uncorrelated standard Gaussian noise and the noise amplitude at each node is $\sigma = 0.01$ (nA). Equation 9 is derived from Equations 3 and 4, by replacing the sum over presynaptic delta-like spikes by the mean firing rate $r_i^{(E)}$ and noting that the effective time constant of NMDA is $\tau_{NMDA} = \sigma_{NMDA}\cdot\tau_{NMDA,decay} = 100$ ms (Wong and Wang, 2006). Similarly, Equation 10 is derived from Equation 2. See Table 3 for parameter values. The values of $W_{Ei}$, $I_{external}$ and $J_{NMDA}$ were chosen to obtain a low level of spontaneous activity for the isolated local area model. With these parameter values the mean spiking activity of the excitatory pool, for an isolate local network ($G = 0$), is equal to 3.0631 Hz. Note that this tuning is necessary because the system of equations (5–10) results from a succession of approximations, among them the neglect of AMPA receptors, which inclusion would otherwise highly complicate the reduced model (Wong and Wang, 2006).

The fixed points of the above dynamical system were calculated by numerically solving the steady-state system of nonlinear equations (Eqs. 5–10, in the absence of noise) using the MATLAB’s solve function.

For comparison to the empirical fMRI BOLD functional connectivity matrix, we transformed the simulated excitatory synaptic activity, $S_i^{(E)}$, to BOLD signals using the Balloon–Windkessel hemodynamic model (Friston et al., 2003; Deco et al., 2013a).

$FIC$. For an isolated node, with the intrinsic parameters described above, an input to the excitatory pool equal to $I_i^{(E)} = b_i/a_i = -0.026$; i.e., slightly inhibitory dominated, leads to a firing rate equal to 3.0631 Hz. Hence, in the large-scale model of interconnected brain areas, we aim to constrain in each brain area (i) the local feedback inhibitory weight $I_i$ such that $I_i^{(E)} = b_i/a_i = -0.026$ is fulfilled (with a tolerance of $\pm 0.005$, implying an excitatory firing rate between 2.63–3.55 Hz). To achieve this, we apply following procedure: we simulate during a period of 10 s the system of stochastic differential Equations 5–10 and compute the averaged level of the input to the local excitatory pool of each brain area, i.e., $I_i^{(E)}$. If $I_i^{(E)} < b_i/a_i = -0.026$ then we upregulate the corresponding local feedback inhibition $I_i = I_i + \Delta$; otherwise, we downregulate $I_i = I_i - \Delta$. We recursively repeat this procedure until the constraint on the input to the local excitatory pool is fulfilled in all 66 brain areas.

$Analytical approximation of the activity covariance$. To estimate the network’s statistics, we approximate deterministic dynamical equations for statistical moments of network’s gating variables. This "moments’ method" avoids extensive simulations of the entire stochastic system, that otherwise would be required for estimation of the network moments. For this, we express the system of stochastic differential equations (5–10) in terms of the first- and second-order moments of the distribution of synaptic gating variables $\mu_{(m)}$, the expected mean gating variable of a given local neural population of type m (where $m = E$ or I) of the cortical area i, and $\sigma_{(m)}^{2}$, the covariance between gating variables of neural populations of types m and n of local cortical areas i and j, respectively. The moments are defined as follows:

$$\mu_{(m)}^{(t)}(t) = \langle S_i^{(m)}(t) \rangle$$

(11)

$$\sigma_{(m)}^{2}(t) = \langle (S_i^{(m)}(t) - \mu_{(m)}^{(t)}(t))(S_j^{(m)}(t) - \mu_{(m)}^{(t)}(t)) \rangle.$$  

(12)

Where the angular brackets denote the average over realizations. In vector form, the system of equations writes as follows:

$$\frac{d}{dt}(S(t)) = \langle H^{(E)}(\bar{S}(t)) \rangle. $$

(13)

Where $\bar{S} = \{\bar{S}_0, S_0\}, \cdots, \{\bar{S}_N, S_N\}$, $\bar{S}_k = S_k - \mu_{(m)}^{(t)} - \delta S_{(m)}^{(t)}$, up to the first order, we get the following:

Table 3. DMF model variables

<table>
<thead>
<tr>
<th>Excitatory model variables</th>
<th>Inhibitory model variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_i = 310$ (nC)</td>
<td>$a_i = 615$ (nC)</td>
</tr>
<tr>
<td>$b_i = 125$ (Hz)</td>
<td>$b_i = 177$ (Hz)</td>
</tr>
<tr>
<td>$d_i = 0.16$ (s)</td>
<td>$d_i = 0.087$ (s)</td>
</tr>
<tr>
<td>$\tau_e = \tau_{NMDA} = 100$ (ms)</td>
<td>$\tau_e = \tau_{GABA} = 10$ (ms)</td>
</tr>
<tr>
<td>$W_{Ei} = 1$</td>
<td>$W_{Ei} = 0.7$</td>
</tr>
<tr>
<td>$W_{Ii} = 1$</td>
<td>$W_{Ii} = 0.01$</td>
</tr>
</tbody>
</table>
\[ f_i^{(m)}(\tilde{S}) = f_k^{(m)}(\tilde{\mu}) + \sum_j \frac{\partial f_k^{(m)}}{\partial S_j^{(E)}}(\tilde{\mu}) \cdot \delta S_j^{(E)} + \sum_j \frac{\partial f_k^{(m)}}{\partial S_j^{(I)}}(\tilde{\mu}) \cdot \delta S_j^{(I)}. \]  

(14)

Using this approximation, taking the average over realizations, and noting that:

\[ \langle \delta S_i^{(m)}(t) \rangle \rangle = \langle u(t) \rangle = 0, \]

(15)

we obtain the motion equations for the means of the gating variables and the covariance of the fluctuations around the mean. For the mean values:

\[ \frac{d \mu_i^{(E)}}{dt} = \frac{d}{dt} \langle S_i^{(E)} \rangle = -\frac{\mu_i^{(E)}}{\tau_E} + (1 - \mu_i^{(E)}) g_i^{(E)}(u_i^{(E)}), \]

(16)

\[ \frac{d \mu_i^{(I)}}{dt} = \frac{d}{dt} \langle S_i^{(I)} \rangle = -\frac{\mu_i^{(I)}}{\tau_I} + H_i^{(I)}(u_i^{(I)}). \]

(17)

Where \( u_i^{(m)} \) is the mean input current to the neural population \( m = E, I \) of cortical area \( i \), defined as follows:

\[ u_i^{(E)} = W_i\lambda + w_i J_{NMDA} \mu_i^{(E)} + G_i J_{NMDA} \sum_j C_{ij} \mu_j^{(E)} - I_{\text{external}}. \]

(18)

\[ u_i^{(I)} = W_i\lambda + J_{NMDA} \mu_i^{(I)} + \lambda G_i J_{NMDA} \sum_j C_{ij} \mu_j^{(I)}. \]

(19)

For the fluctuations:

\[ \frac{d}{dt} \langle \delta S_i^{(m)} \delta S_j^{(m)} \rangle = \sum_{p \in \{E,I\}} \frac{\partial f_i^{(m)}}{\partial S_p^{(m)}} \langle \delta S_p^{(E)} \rangle \delta S_j^{(m)} + \sum_{p \in \{E,I\}} \frac{\partial f_j^{(m)}}{\partial S_p^{(m)}} \langle \delta S_p^{(m)} \delta S_j^{(m)} \rangle + \sigma^2 \langle v_i \rangle. \]

(20)

Let \( P \) being the covariance matrix between gating variables. \( P \) is a block matrix defined as follows:

\[ P = \begin{bmatrix} P^{EE} & P^{EI} \\ P^{EI} & P^{II} \end{bmatrix}, \]

where the superscript \( T \) is the transpose.

Using Equation 20, we obtain the motion equation of the covariance matrix:

\[ \frac{dP}{dt} = AP + PA^T + Q_\alpha. \]

(21)

Where \( Q_\alpha \) is the covariance matrix of the noise (which is diagonal for uncorrelated white noise) and \( A \) is the Jacobian matrix of the system. \( A \) is a block matrix defined as follows:

\[ A = \begin{bmatrix} A^{EE} & A^{EI} \\ A^{EI} & A^{II} \end{bmatrix}, \]

where \( A_i^{(m)} = \left( \frac{\partial f_i^{(m)}}{\partial S_j^{(m)}} \right) \).

Finally, note that the above derivatives can be written as follows:

\[ \frac{\partial f_i^{(m)}}{\partial S_j^{(m)}}(\tilde{\mu}) = -\frac{1}{\tau_m} \gamma g_i^{(E)}(u_i^{(E)}) \delta_{E,i} \delta_{E,j} + \frac{1 + (1 - \mu_i^{(E)}) \gamma g_i^{(I)}(u_i^{(I)}) \delta_{E,i} }{\mu_i^{(E)}} \cdot K_i^{(m,m)}. \]

(22)

Where \( K = \begin{bmatrix} K^{EE} & K^{EI} \\ K^{EI} & K^{II} \end{bmatrix} \), with \( K^{EE} = -J_{\text{spin}} K^{II} = J_{\text{spin}} I + \lambda G J_{NMDA} C, K^{II} = -I, \) and \( K^{EE} = w_i J_{NMDA} + G_i J_{NMDA} C, \) where \( I \) and \( C \) are the \( N \)-by-\( N \) identity matrix and anatomical connectivity matrix, respectively.

Equations 20–22 indicate that the covariance of gating variables is determined by both the underlying connectivity and the dynamics. In the stationary regime, the covariance matrix of fluctuations around the spontaneous state is given by the algebraic equation:

\[ \text{AP} + \text{PA}^T + \text{Q}_\alpha = 0, \]

(23)

which can be solved using the Eigen-decomposition of the Jacobian matrix evaluated at the fixed point: \( A = LDL^{-1} \), where \( D \) is a diagonal matrix containing the eigenvalues of \( A \), denoted \( \lambda_i \), and the columns of matrix \( L \) are the eigenvectors of \( A \). Multiplying Equation 23 by \( L \) from the left and by \( L^{-1} \) from the right (the superscript dagger being the conjugate transpose) we get the following:

\[ P = LML^{-1}, \]

(24)

Where \( M \) is given by the following:

\[ M_{ij} = -\tilde{Q}_i/(\lambda_i + \lambda_j), \]

and \( \tilde{Q} = L^{-1} Q L^{-1} \).

Thus, the covariance matrix of spontaneous fluctuations is determined by the eigenvalues of the Jacobian matrix, which, in turn, is related to the connectivity matrix and the dynamics. Hence, Equation 24 provides a direct link between the correlation structure, the underlying connectivity, and dynamics. Indeed, the interpretation of Equation 24 is that the input covariance \( (M) \) is propagated through the dynamical system and is mapped to its approximated output \( (P) \).

Finally, to compare the results from the above moments’ method to the empirical functional connectivity, we estimated the correlation matrix, noted \( Q \), between the gating variables of the excitatory populations, defined as follows:

\[ Q_{ji} = \frac{E \langle \delta S_i^{(E)} \delta S_j^{(E)} \rangle}{\sigma \langle \delta S_i^{(E)} \rangle \sigma \langle \delta S_j^{(E)} \rangle}. \]

Using Equation 25 we get the following:

\[ -i \omega \delta S(i) = A \delta S(i) + \sigma \tilde{v}(i). \]

(26)

Hence, \( \delta S(\omega) \delta S(\omega) = -i (1 + i \omega)^{-1} \sigma \tilde{v}(\omega) \). The power spectrum being the expectation of the Fourier transformed autocorrelation function, we get the following:

\[ \Phi(\omega) = \langle \delta S(i) \delta S(i) \rangle = \sigma^2 (A + i \omega)^{-1} (A^T - i \omega)^{-1}. \]

(27)

Thus, the power spectrum of \( \delta S \) is as follows:

\[ \Phi_\epsilon(\omega) = \sigma^2 \sum (|J_{ki} + i \omega \delta_{kj}|)^{-1}. \]

(28)

Model prediction of the empirical connectivity. As a measure of similarity between two FC matrices we used the uncentered Pearson correlation coefficient between corresponding elements of the upper triangular part of the matrices, defined as follows:

\[ r = \frac{\langle XY \rangle}{\text{SD}(X) \text{SD}(Y)}. \]

(29)

Where \( X \) and \( Y \) are the vectors containing the upper triangular values of each matrix, respectively, and SD is the standard deviation. For FC matrices, because the sampling distribution of correlation coefficients is not normally distributed, we used the Fisher’s z-transformation to convert the matrix elements before applying any subsequent test. The uncentered correlation takes into account the difference in the mean values of the FC matrices. Confidence intervals were obtained using bootstrap resampling (1000 resamples).

For comparing a structural connectivity matrix and a FC matrix we used the centered Pearson correlation defined as follows:

\[ r_c = \frac{\langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle}{\text{SD}(X) \text{SD}(Y)}. \]

(30)

Entropy of binary evoked patterns. The response of the large-scale network to different stimuli was analyzed to investigate its behavior under...
hypothesised task conditions and its information capacity. Stimuli were constructed by imposing an external input \( I_{\text{external}} = 0.02 \) (1) to the excitatory population of 10% of the brain areas, randomly selected, or (2) to both the excitatory and inhibitory populations of 10% of the brain areas, randomly selected. The procedure was repeated 1000 times and the resulting steady-state evoked response of the network were stored. Evoked patterns were binarized by imposing a threshold above which the excitatory firing rate of a given area was set to 1 and, otherwise, it was set to 0 (only the firing rate of the excitatory pools were used, thus the binary evoked pattern is 66-dimensional). The threshold was defined as \( \theta_{SD} \), where SD is the standard deviation across all evoked patterns and the multiplier \( \theta \) is a scalar parameter. In this way we obtained binary patterns containing 66 entries. The entropy of the set of evoked binary patterns \( R \) is defined as follows:

\[
H(R) = - \sum_{i} p_i \log_2 p_i.
\]  

Where \( n \) is the number of unique patterns and \( p_i \) is the probability that pattern \( i \) is observed. Only patterns with at least one non-null entry were analyzed.

To remove the sampling bias, we corrected the entropy values by using a quadratic extrapolation procedure (Treves and Panzeri, 1995; Shew et al., 2011). This procedure evaluates the entropy for random samples of fractions \( f \) from the full set of \( K \) patterns to estimate the sample entropy \( H(f) \). It has been shown that the bias of the entropy can be accurately approximated as second order expansions in \( 1/f \); i.e.,

\[
H(f) = H_0 + \frac{k_1}{f} + \frac{k_2}{f^2} \quad \text{where} \quad H_0 = \text{the true value of the entropy and} \quad \frac{k_1}{f} + \frac{k_2}{f^2} = \text{the bias (Treves and Panzeri, 1995).}
\]

Fitting (least-square error procedure) the sample entropy \( H(f) \) with a quadratic function of \( 1/(fK) \), provides the estimate of the unbiased entropy \( H_0 \). Additionally, the fitting provides confidence intervals.

**Fisher information.** We quantified the encoding accuracy of the large-scale network response by calculating the Fisher information (FI). Assuming that the network response is well described by a multivariate Gaussian distribution, i.e., for weak noise, FI can be written as the sum of two terms: 

\[
FI = FI_{\text{mean}} + FI_{\text{cov}} (\text{Abbott and Dayan, 1999}),
\]

where

\[
FI_{\text{mean}}(\Delta \lambda) = \frac{1}{2} \text{Trace} (P^{(\Delta \lambda)} (P^{(\Delta \lambda)})^{-1}).
\]

\[
FI_{\text{cov}}(\Delta \lambda) = - \sum_{i,j} (p^{(\Delta \lambda)}_{ij} - p^{(\Delta \lambda)}_{ii} p^{(\Delta \lambda)}_{jj}) p^{(\Delta \lambda)}_{ii} p^{(\Delta \lambda)}_{jj}.
\]

Where \( \bar{\mu}(\Delta \lambda) \) and \( P(\Delta \lambda) \) are the stationary mean synaptic activity and the covariance matrix evoked by an excitatory stimulus of intensity \( \Delta \lambda \); respectively. \( \bar{\mu}(\Delta \lambda) \) and \( P(\Delta \lambda) \) are the first derivatives with respect to the stimulus, evaluated at \( \Delta \lambda \). The values of \( \bar{\mu}(\Delta \lambda) \) and \( P(\Delta \lambda) \) were estimated using the moment’s equations (Eqs. 16–19 for the mean, and Eq. 24 for the covariance), in the presence of an applied stimulus, i.e., \( I_{\text{external}} = \Delta \lambda \). Only the excitatory parts of \( \bar{\mu} \) and \( P \) were considered for calculating the FI. In this study, the stimulus was applied to the excitatory population of both right and left lateral occipital cortex (LOCC), i.e., \( I_{\text{external}} = \Delta \lambda \) for rLOCC and ILOCC, and \( I_{\text{external}} = 0 \) for all other cortical areas. The intensity of the stimulus varied from \( \Delta \lambda = 0.02 \) to \( \Delta \lambda = 0.1 \) in steps of 0.001.

**Betweenness centrality:** The centrality of a node within the anatomical network was quantified using the betweenness centrality measure. This network measure expresses the number of shortest paths that pass through a given node. For calculating the betweenness centrality we binarized the anatomical matrix \( C \) by imposing a threshold equal to 0.05 above which the anatomical coupling was set to 1 and, otherwise, it was set to 0.

**Results**

**Local feedback inhibition control**

In the present study we investigated the effect of regulating the local feedback inhibition on the long-range resting correlations between cortical areas. For this, we used a large-scale model, composed of local networks, or “nodes,” representing cortical areas, interconnected by the anatomical connectivity matrix, or “structural connectivity,” between those cortical areas (Deco and Jirsa, 2012; Deco et al., 2013a). Large-scale anatomical connectivity data were obtained using DSI and tractography techniques (Hagmann et al., 2008; see Materials and Methods). Isolated brain area models consist of 80% excitatory neurons and 20% inhibitory IF neurons recurrently connected such that when isolated the network emulates the neuro-physiological characteristics of the empirical observed spontaneous state, i.e., low correlations between the spiking activities of the neurons and low firing rate at 3 Hz. In the FIC scenario, the feedback inhibition is adjusted to compensate the extra excitation the each excitatory pool receive when the brain areas model are connected. b, c, Spiking activity of excitatory units from one brain area without FIC (b, top) and with FIC (c, top). Due to symmetry, the spiking activity in the other brain area has the same statistics. The averaged activities of both excitatory pools are plotted in blue and green, for the model without FIC (b, bottom) and the model with FIC (c, bottom). d, Filled bars, Mean correlation coefficient (±SEM) across all pairwise correlations within one excitatory pool, without and with FIC (red and blue, respectively). Open bars, Correlation between the mean firing rates of both excitatory pools, without and with FIC (red and blue, respectively).
cient is 0.11 ± 0.06 and 0.23 ± 0.05 with and without FIC, respectively, \( p < 10^{-10}, t \) test), but, interestingly, allows for a moderate correlation between the mean firing rates of both excitatory pools (equal to 0.35 and 0.63 with and without FIC, respectively; Fig. 1c,d). In other words, under FIC, “microscopic” activity is partially decorrelated while “mesoscopic” (population level) quantities covary.

This is the key manipulation that we study in the large-scale model. Before proceeding any further, it is important to note that, in the previous example, although the firing rate is maintained constant, manipulation of the inhibitory feedback changes the eigenvalues of the local network, and, thus, modifies the dynamics.

**Large-scale models with and without FIC: stationary states**

In the large-scale cortical network, consisting on coupled cortical areas, we compensated the local feedback inhibition weights \( J_{f,i} \), as previously, to clamp the firing rate \( \approx 3 \) Hz for each local excitatory neural population \( i \) (see Materials and Methods). Because this local compensation mechanism, achieved through recursive adjustments of \( J_{f,i} \), is computationally expensive, we used a DMF to reduce the large-scale spiking network (Wong and Wang, 2006; see Materials and Methods). The DMF ignores the interaction between single neurons within a cortical area and instead considers the ensemble mesoscopic dynamics. By reducing the number of variables, the DMF model allows for the optimization of feedback inhibition weights (Fig. 2a,b).

In the following we compared the fixed points of the model with long-range excitatory-excitatory connections (E–E model) and the model with long-range excitatory-excitatory connections and local feedback inhibition regulation (FIC model). We also consider a third model in which the firing rate among each cortical region is maintained low through long-range connections from the excitatory populations to the inhibitory populations (feedforward inhibition), in the same proportion as long-range excitatory-excitatory connections and according to the anatomical connectivity (see Materials and Methods).

In all models the inter-area coupling is scaled by a single global parameter, \( G \), that changes the network from weakly to strongly connected and determines the dynamical state of the system. We first calculated the bifurcation diagrams characterizing the stationary states of the brain system for the three models (Fig. 2c,d) as a function of \( G \). In all cases we plot, for the different possible states, the maximal firing rate activity across all excitatory populations. The E–E model presents a bifurcation at \( G = 1.47 \) separating two qualitatively different topologies of the attractor landscape. For small values of the global coupling \( G \), only one stable state exists, characterized by a firing activity that increases as \( G \) increases (stable branch). For \( G \) larger than the critical value 1.47 a bifurcation occurs and a second fixed point appears which is not stable (unstable branch). The FIC model has only one stable state of low firing activity in all cortical areas, for \( G < 4.45 \). For larger values of \( G \), long-range interactions are too strong to be compensated by FIC and the activity diverges. In the FFI model the maximum firing rate monotonically increases as a function of \( G \).

**Feedback inhibition control improves the model’s FC prediction**

We next investigate the structure of the emergent correlations from the three global brain models. The predictions of the models were evaluated using the empirical human resting-state FC, obtained by averaging the results across 48 IMRI scanning sessions from 24 healthy human subjects and projected to the same parcellation adopted for the anatomical structural matrix (SC). For comparison with the empirical data, we considered the FC of simulated BOLD signals which are obtained by transforming the model synaptic activity through a hemodynamic model (see Materials and Methods). We calculated the similarity of the empirical FC to the simulated FC as a function of \( G \).

**Figure 2.** The effect of FIC on the spontaneous mean-field activity of the large-scale brain model. a, A dynamic mean field (DMF) reduction of the model was used to study the large-scale model of the brain, composed of \( N \) nodes, each containing one excitatory and one inhibitory neural population. The inter-area connections are established as long range synaptic AMPA-mediated instantaneous connections between the excitatory pools in those areas. In the case of the model with feedforward inhibition (FFI) long-range synaptic AMPA-mediated instantaneous connections from the excitatory pool of a given area to the inhibitory pool of a different area were also considered (dashed arrows). Inter-areal connections are weighted by the strength specified in the neuroanatomical matrix \( SC \), denoting the density of fibers between those regions, and scaled by a global factor \( G \). The local recurrent excitation is \( NMDA \)-mediated. In the FIC condition, the local feedback inhibition weight \( J_{f,i} \) was adjusted such that the excitatory pool of each local brain area has a low firing rate \( 3 \) Hz. b, Adjacency matrix of the neuroanatomical connectivity. c, Attractor landscape as a function of the globally coupling strength \( G \), for the three large-scale models, (E–E: long-range excitatory-excitatory connections; FIC: long-range excitatory-excitatory connections and local feedback inhibition regulation; FFI long-range excitatory-excitatory connections and long-range feedforward inhibition). Each point represents the maximum firing rate activity among all excitatory pools. In the E–E model, for low values of \( G \), the network converges to a single stable state; for \( G > 1.47 \) (vertical line), a bifurcation appears whereby a new unstable state coexists together with the spontaneous state. In the FIC model, for \( G < 4.45 \), the optimization of feedback inhibition weights makes the network to converge to a single stable state of low firing activity; for \( G > 4.45 \), the low firing activity solution becomes unstable. In the FFI model the maximum firing rate monotonically increases as a function of \( G \). d, The stationary-state of excitatory pools (firing rate vector of dimension \( n \approx 66 \)) is shown for \( G = 2.0 \) for the three large-scale DMF models. Blue, FIC model (stable state); the inset shows the stationary firing rates as a function of the local feedback inhibition strength; red, E–E model (stable state); red, open bars, E–E model (unstable state); magenta, FFI model (stable state).
model FC and SC, and the similarity between the model FC and the empirical FC, for each model (Fig. 3a,b; see Materials and Methods). Importantly, note that, for a given $G$, the large-scale underlying connectivity (i.e., $G \times SC$) is the same for the three models, while the local connectivity is different for the FIC model. We found that, in the E–E, the underlying structural connectivity is maximally expressed in the simulated FC at the edge of the critical value of $G$ (Fig. 3a, red trace). Similarly, the empirical FC is optimally fitted by the model near criticality (Fig. 3b, red trace). For the FFI model and for the network with FIC, the region where the SC is maximally expressed by the simulated FC and where the empirical data are optimally fitted is much broader (Figs. 3a,b, blue and magenta traces). Furthermore, the level of agreement between the model and the data reached at the maximum is improved in the FIC case.

We further compared the FC of the two models for the corresponding optimal value of $G$, equal to 1.3, 4.7, and 3.3 for the E–E model, the FFI model, and the FIC model, respectively (Fig. 3c–e). The scatter plot between the model and the empirical correlation coefficients is closer to the identity line for the FIC model than for the two other models, as can be appreciated by linear regression of the data. The agreement between the model FC and the empirical FC is significantly higher for the FIC model than both for the E–E model ($p < 10^{-7}$, Meng’s z test for dependent correlations (Mz-test) and for the FFI model ($p < 10^{-7}$, Mz-test). We also tested the ability of each model to predict the FC of individual fMRI scanning sessions. For this we used the following jackknife procedure: for each of the $n = 48$ scanning sessions we calculated the FC and we computed the agreement between the single-session FC and the FC of each model for which $G$ was optimized using the averaged FC over the remaining $n - 1$ sessions. The procedure was repeated such that each scanning session was held out once. The resulting distributions of similarity values for each model are shown in Figure 3f. The FIC model gives significantly better predictions of the single-session FC matrices than both the E–E model and the FFI model (ANOVA, $p < 10^{-7}$).

Moreover, we further quantified the agreement between the simulated FC matrices and the empirical FC by comparing their first principal component (PC), or “dominant spatial mode”. For a given covariance matrix, the dominant spatial mode is given by the first eigenvector of the matrix. We calculated the vector projection between the first PC of the empirical data and the first PC of the three models, calculated for the corresponding optimal values of $G$, as previously (Fig. 4). We found that, under FIC, the network approximates better the first PC of the empirical data than do the E–E model and the FFI model.

Altogether, the above results show that the quality of the optimal fitting is better under the FIC condition compared with the unconstrained case. One could argue that the increase in the FC prediction for the FIC model is merely due to the optimization of the connection weights, which makes the model more complex than the other two models. However, note that the parameters of the FIC model where optimized to bound the firing rate of each network node only and not in the optimization of the prediction of the empirical FC which uses only one free parameter ($G$) as for the other two models.

**Correlation of linear noise fluctuations**, an analytical approximation

The reduction of the spiking network through dynamic mean field approximation allowed investigating the large-scale model under the FIC condition. However, the nonlinear and stochastic nature of the DMF equations hinders analytical insights and examination of the dynamics needs to be treated numerically. In particular, estimation of second-order statistic of large networks requires long and multiple time-consuming numerical simulations. In the following, we further reduce the dynamical system.
tual fluctuations around the means. This analytical simplification is es-
dynamical equations around the mean values of gating variables (see Materials and Methods). Briefly, the moments' equations for
by deriving deterministic differential equations for these statistics, means and covariances, of the distribution of gating variables and
expressing motion equations for the covariances of the fluctuations around the means. This analytical simplification is es-
tential, because it provides an explicit link between structure, dynamics, and FC. Indeed, within this linear approximation, the covariance matrix of spontaneous fluctuations is determined by
the eigenvalues of the Jacobian matrix (A) of the large-scale net-
work, which in turn is related to the local and large-scale connectivity matrix and the dynamics (see Materials and Methods). Hence, because manipulating the feedback inhibition strength within local networks changes the eigenvalues of A, the correlation structure is modified, even though the firing rates of all model cortical areas remain constant. Using this linear approximation, we calculated the similarity between the correlation matrix of excitatory gating variables, noted Q, and the empirical FC matrix. Note that, in this case, the model correlations matrix is not based on BOLD signals, but on the excitatory gating variable of all cortical areas. We found that the agreement between Q and the empirical FC is higher for the FIC model than for the other two models (Fig. 5a,b). Moreover, the empirical distribution of correlation coefficients is better approximated by the FIC model than for the other two models (Fig. 5c,d).

Note that the decrease of the fitting of the empirical FC for small values of G observed in simulations of BOLD signals (Fig. 3b) is more pronounced that the one obtained using the linear prediction of Q (Fig. 5a). This discrepancy arises from the distribution of functional correlation coefficients $r_{ij}$. Indeed, if the values are narrowly distributed, as for G approaching zero for all models (Fig. 5c), the precise estimation of the correlation structure through stochastic simulations would require a large number of simulation steps, so that the estimation error of $r_{ij}$ becomes very small compared with the variance of the distribution of $r_{ij}$. Opposite to this, the analytic solution does not suffer from the sampling error and, thus, it correctly estimates the fine structure of Q with infinite resolution.

Furthermore, the linear approximation allows for an analyti-
cal estimation of the time-scale of the gating variables’ fluctuations. Since the hemodynamic model used to simulate the BOLD fMRI signal can be seen as a low-pass filter that passes frequencies $<1$ Hz (Robinson et al., 2006), only gating variables correlations at slow time scales are transmitted through the hemodynamic model. We thus calculated the power spectrum of fluctuations around the spontaneous state, by taking the Fourier transform of the equation governing the time evolution of fluctuations (see Materials and Methods; Fig. 5e). We found that in the FIC model the power of low-frequency ($<1$ Hz) fluctuations is maintained high for all values of G, whereas for the other two models slow fluctuations are expected for small values of G only.

Altogether, the linear approximation shows that the FIC model better approximates the functional correlations and maintains the dynamics in a slow time-scale, so that the correlation structure of gating variables is seen through the hemodynam-
cally filtered response.

Feedback inhibition control increases the dynamical repertoire of the network evoked activity
To examine the functional implications of FIC, we analyzed the response of the model network to exogenous stimuli, which are assumed to represent ideal task-related signals. We simulated 1000 different hypothetical task conditions by stimulating exog-
enously 10% of the brain areas, randomly selected. Stimulation was modeled by imposing an external input to the excitatory pools, i.e., for the set of stimulated excitatory pools we imposed $I_{external} = 0.02$. The increased excitatory activity induces an evoked pattern of activity at the stationary state. We first calcu-
lated, for each model, the activity of cortical excitatory pool av-

---

**Figure 4. Prediction of the principal mode.** The first PC of the empirical covariance matrix of BOLD signals (turquoise) and the first PC of the model covariance matrix generated by each of the three models, for the corresponding optimal values of G. The projection (scalar product) between the first PC of the empirical data and the first PC of model data are equal to 0.57 ± 0.05, 0.87 ± 0.02, and 0.71 ± 0.03 for the E–E, FIC, and FFI models, respectively.
erged over all stimuli, for the corresponding optimal values of $G$ (Fig. 6a). Interesting, in both the E–E model and the FFI model, the activity is systematically much higher for some nodes, whether they are directly stimulated or not. These nodes are central nodes of the network, as revealed by their corresponding betweenness centrality (Fig. 6b), a network measure that quantifies the number of shortest paths that pass through a given node. Note that, as previously shown (Hagmann et al., 2008), central cortical areas are members of the default mode network (Fig. 6c).

We next characterized the information capacity of the large-scale model by quantifying the size of the repertoire of different evoked patterns of both models. Indeed, a network that has few degrees of freedom, i.e., few different evoked patterns, has a low capability to transmit information. We thus hypothesized that the networks without FIC, due to their tendency to elevate the firing rate of central nodes in response to distinct stimuli, have less information capacity. We quantify this by calculating the entropy $H_{\text{evoked}}$ of binary evoked patterns (see Materials and Methods) to measure the extent of the repertoire of evoked activity. Binary patterns were constructed by imposing a threshold and setting the steady-state evoked activity of a given area to 0 or 1 whether it is below or above the threshold, respectively. The threshold was defined as $\theta \pm \text{SD}$, where SD is the standard deviation across all possible evoked responses of each model. Several values of the multiplier $\theta$ were tested. For each of the three large-scale models, we compared the entropy of evoked patterns to the maximum entropy corresponding optimal value for each model. We found that the regulation of feedback inhibition increases the FI, thus increasing the accuracy of the stimulus encoding. Relaxing the FIC condition leads to a decrease of the FI because the central nodes are systematically activated in response to the stimulus, thus saturating the network response, although the FIC provides a graded response of the stimulated nodes with practically no activation of the other nodes (Fig. 6g).

**Discussion**

In this study, we investigated the impact of the local FIC on a large-scale model of the brain. We found several important and functionally relevant effects of the local FIC on both the spontaneous and the evoked patterns of activity of the large-scale model. First, we showed that the FIC significantly enhances the model’s prediction of the fMRI human resting functional connectivity, even for single fMRI scanning sessions. Furthermore, we found that the optimal parameter space where the model best approximates the empirical FC is enlarged in the FIC condition. This is important since a common result of all previous resting-state models is that the optimal parameter space is confined around the edge of criticality (Honey et al., 2007, 2009; Ghosh et al., 2008; Deco et al., 2009, 2013a,b; Deco and Hughes, 2012) posing the problem of fine tuning for which self-organization rules remain unknown. Opposite to this, it has been shown that inhibitory synapses can rapidly adapt to changes in the local network activity to scale the activity in cortical circuits (Hartmann et al., 2008), thus providing a natural implementation of the FIC by biological synapses. In
this view the local FIC may provide a homeostatic mechanism that generates the observed resting-state FC. Nevertheless, in the present study we have regulated the inhibitory-to-excitatory coupling, but manipulating other local weights to adjust the ratio between excitation and inhibition in the large-scale model may lead to similar results and would provide a local decorrelation mechanism, providing that excitation and inhibition balance each other (Renart et al., 2010).

Nevertheless, note that the agreement between the model’s FC and the empirical FC is not perfect. This imperfect fitting is expected principally due to the missed of connections of the diffusion tensor/spectrum imaging (DTI/DSI) tractography, used here to estimate the anatomical connectivity that constrains the large-scale connectivity of the model. Indeed, precise prediction of the FC strongly depends on the quality of the estimated structural connectivity. However, interhemispherical connections are not well captured by the DTI/DSI tractography (Hagmann et al., 2008) and the anatomical matrix used here did not include subcortical routes that are known to play an important role in shaping the spontaneous activity of the brain (Robinson et al., 2001; Freyer et al., 2011). Estimation of subcortical connections is a challenging issue because DTI systematically leads to more con...

Figure 6. The effect of FIC on the evoked activity of the large-scale brain model. a, Mean activity in each local brain area (with respect to the activity averaged over all brain areas) for each model obtained in response to 1000 different hypothetical task conditions. Tasks were simulated by imposing an external input $I_{\text{external}} = 0.02$ to the excitatory population of 10% of the brain areas, randomly selected. Dark colors indicate the areas composing the default mode network. b, Betweenness centrality (BC) of each brain area. BC was calculated from the anatomical connectivity matrix. It represents the number of shortest paths passing through a given node. c, The default mode network is composed of midline frontal and parietal areas, posterior inferior parietal lobe, and medial and lateral temporal lobe regions. d, The entropy $H_{\text{max}}$ of evoked binary patterns was calculated for different activity thresholds, defined as $\theta SD$, where SD is the standard deviation across all possible evoked responses. Bars indicate the reduction of entropy $\Delta H$, for the three different models, for different values of $\theta$ ($\theta = 2.5, 3, \text{or} 4$). $\Delta H$ was defined as $\Delta H = (H_{\text{max}} - H_{\text{evoked}})/H_{\text{max}}$ where $H_{\text{max}}$ is the maximum entropy for the set of $n$ patterns, i.e., $H_{\text{max}} = \log(n)$. Error bars indicate estimation errors (50%) given by the quadratic extrapolation procedure, used for the entropy sampling bias correction. e, Entropy reduction when external stimuli are imposed to both the excitatory and the inhibitory populations. The values of $\theta$ are lower than in (d) to avoid patterns with null entries. f, The encoding accuracy of the models was calculated using the FI. The FI was calculated for various stimulus intensities ($\Delta t$). The stimulus was applied to the excitatory population of both right and left LOCC, i.e., $I_{\text{external}} = \Delta t$ for rLOCC and lLOCC, and $I_{\text{external}} = 0$ for all other cortical areas. g, Evoked response of the network for the E–E model (top), the FFI model (middle), and for the FIC model (bottom). The colors indicate the intensity ($\Delta t$) of the applied stimulus. The arrows indicate the stimulated nodes.

and thus, how both the cortex and the subcortical nuclei can be described in a parsimonious model is an open question.

In addition, we have made several simplifying assumptions that may impose an upper bound to the agreement between the model predictions and the empirical observations. Specifically, in this study we consider that all connections between brain areas are instantaneous, thus neglecting the effects of conduction delays which are likely to shape the brain dynamics, giving rise to complex spatiotemporal patterns, oscillations, multistability, and chaos (Roxin et al., 2005; Ghosh et al., 2008; Deco et al., 2009). Indeed, a recent MEG study found robust frequency-specific lagged coherence between interhemispheric functional connections in the dorsal attention network, or between the dorsal attention and other networks (e.g., visual and somatomotor; Marzetti et al., 2013). Considering the delays would make more complex the model and add a new parameter to explore: the propagation velocity.

In this work, we assumed that functional connectivity is a stationary process that arises from the interplay between the underlying anatomical connectivity structure and the neural dynamics on the stationary regime. This assumption is commonly made by the current resting-state models that seek to link the structural and the functional connectivity (Honey et al., 2007, 2009; Ghosh et al., 2008; Deco et al., 2009, 2013a,b; Deco and Hughes, 2012). However, recent studies have demonstrated that the correlations among brain regions evolve over time (Chang and Glover, 2010; Kiviniemi et al., 2011; de Pasquale et al., 2012; Hutchison et al., 2013; Allen et al., 2014). Because nonstationarities are likely to substantially contribute to the resting-state activity (Messé et al., 2014), it is thus crucial to analyze and model the
time-varying behavior of the functional connectivity. How non-stationary functional connectivity emerges in the brain network remains an open question that needs further investigation.

In the present study, we assumed that the spontaneous activity throughout the cortex is low. This is supported by early recordings of cortical activity in vivo (Burns and Webb, 1976; Sofiky and Koch, 1993; Wilson et al., 1994), as well as recent experiments in mammalian neocortex using diverse recording techniques, such as whole-cell patch-clamp, sharp-electrode recordings and Ca$^{2+}$-imaging, in different brain states (awake or under anesthesia), indicating that the spontaneous activity of pyramidal neurons in sensory cortices ranges between 1–5 Hz, with further laminar differences (Sakata and Harris, 2012; for review, see Barth and Poulet, 2012). Nevertheless, more investigation comparing the spontaneous activity of different sensory and association cortices is needed.

By studying the large-scale patterns evoked by external inputs, we showed that, if the FIC constraint is relaxed, external stimulation of the brain model leads to a systematic large activation of cortical areas that are components of the default mode network. This is unlikely to occur in the real brain, because only few experimental conditions leading to activation of the default mode network have been reported. Indeed, a myriad of reports shows that the default network is deactivated during attentionally demanding and goal-directed tasks (Shulman et al., 1997; Raichle et al., 2001; Fox and Raichle, 2007; Thomason et al., 2008; Christoff et al., 2009). Nevertheless, the default mode network is activated by cognitive processes that are internally driven, such as, memory retrieval (Sestieri et al., 2011), internal mentation (Mason et al., 2007; Bar, 2009), and self-reference (Gusnard and Raichle, 2001; D’Argembeau et al., 2005). However, these internal processes are more likely to be initiated spontaneously, possibly due to spontaneous fluctuations in the cortex, as recently proposed in the context of self-initiated movements (Schurger et al., 2012), and thus, they may not be well modeled by imposing external inputs to some cortical areas. Altogether, we believe that the systematic activation of cortical areas belonging to the default mode network, observed in the evoked responses of the brain models without FIC, is not consistent either with current experimental data or with current assumptions about the generation of internal mental states. Here we showed that regulating the local level of feedback inhibition in the brain has an important role at the global level, because it attenuates the response of cortical areas in the default mode network.

A further effect of local FIC on large-scale brain patterns is to increase the entropy of the evoked activity patterns, thus increasing the dynamic repertoire of the network. Hence, the local FIC enhances the information capacity of the global network, i.e., it increases the ability of the network to map different inputs into distinguishable network outputs. This result complements previously reported evidence showing that neural systems tend to maximize the entropy at the level of single neurons (Tsubo et al., 2012) and neural populations (Shew et al., 2011). It has been shown that when excitation is sufficiently restrained by inhibition in local cortical networks, the information capacity of cortical circuits is increased (Shew et al., 2011; Deco and Hughes, 2012). Furthermore, we showed that regulation of the feedback inhibition enhances the stimulus discriminability, i.e., it increases the sensitivity of the network to a small variation of the stimulus. Hence, our results indicate that the control of the excitation–inhibition ratio at the local level has important implications for information transmission at the large-scale brain level. Moreover, because the feedback inhibition acts as an active mechanism for reducing pairwise correlations in local networks (Tetzlaff et al., 2012; Fig. 1), it improves the encoding and readout within local circuits.

In summary, we had identified several effects of local regulation of feedback inhibition on brain dynamics at the large scale: it changes the characteristics of the emergent resting and evoked activity in a crucial way. Regulating the local excitation–inhibition ratio provides a better and more robust prediction of human empirical resting state connectivity, together with more realistic responses to external inputs and higher information capacity and encoding accuracy.

References
Fox MD, Raichle ME. (2007) Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nat Rev Neurosci 8:700–711. CrossRef Medline