Effective dynamics analysis of fMRI data points to reconfiguration of cortical information processing during visual and auditory perception

Matthieu Gilson[1], Gustavo Deco[1,2], Karl Friston[3], Patric Hagmann[4,5], Dante Mantini[6,7], Viviana Betti[8], Gian Luca Romani[8], Maurizio Corbetta[8,9]

[1]Center for Brain and Cognition, Computational Neuroscience Group, Department of Information and Communication Technologies, Universitat Pompeu Fabra, Roc Boronat 138, Barcelona, 08018, Spain;

[2]Institució Catalana de la Recerca i Estudis Avançats (ICREA), Universitat Pompeu Fabra, Passeig Lluís Companys 23, Barcelona, 08010, Spain;

[3]Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, 12 Queen Square, London WC1N 3BG, United Kingdom;

[4]Department of Radiology, Lausanne University Hospital and University of Lausanne (CHUV-UNIL), Rue du Bugnon 46, 1011 Lausanne, Switzerland;

[5]Signal Processing Lab 5, École Polytechnique Fédérale de Lausanne (EPFL), Station 11, 1015 Lausanne, Switzerland

[6]Research Center for Motor Control and Neuroplasticity, KU Leuven, 101 Tervuursevest, 3001 Leuven, Belgium;

[7]Department of Health Sciences and Technology, ETH Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland;

[8]Institute of Advanced Biomedical Technologies - G. d'Annunzio University Foundation, Department of Neuroscience and Imaging, G. d'Annunzio University, Via dei Vestini 31, Chieti, 66013, Italy;

[9]Departments of Neurology, Radiology, Anatomy of Neurobiology, School of Medicine, Washington University, St. Louis, St Louis, USA.

· Corresponding author:
Matthieu Gilson, Universitat Pompeu Fabra, C/ Tanger, 122-140, 08018 Barcelona, Spain.
matthieu.gilson@upf.edu

· Conflicts of interest: the authors declare no conflict of interest.

ABSTRACT

Our behavior entails a flexible and context-sensitive interplay between brain areas to integrate information according to goal-directed requirements. Nevertheless, the neural mechanisms governing the integration and entrainment of functionally specialized brain areas remain poorly understood. Here, we introduce a theoretical framework that discloses these mechanisms – at the whole brain level – based on fMRI data. Our approach relies on a model of functional connectivity that allows for the estimation of the strengths of (extrinsic) cortico-cortical interactions, as well as the input or baseline activity (intrinsic) to each cortical region. We examine quantitatively the contributions of effective connectivity and intrinsic input to changes in functional integration during movie viewing. Our results show how the interactions between cortical areas are moderately, but significantly, modified to select pathways that integrate sensory influxes from the visual and auditory systems to high-level areas in the frontal and central parts of the brain. Meanwhile, opposing changes in (intrinsic) input and (extrinsic) connectivity preserve a balance of distributed responses. To characterize this balance, we define an ‘effective drive’ that quantifies the propagation of neuronal activity (measured by the amplitude of fMRI fluctuations) between cortical regions. Finally, homotopic regions in opposite hemispheres increased their effective connectivity and drive during movie viewing, whereas they are effectively disconnected when watching a black screen. These findings speak to a dynamic and balanced functional integration that underlies hierarchical processing in the brain.

INTRODUCTION

The brain comprises a large number of functionally distinct areas in which information and computational processes are highly segregated (Van Essen et al., 1992; Biswal et al., 1995). This structure shows a high degree of complexity and diversity, where a massive number of inter-area connections coordinate the activity of large networks. Accordingly, empirical evidence from functional magnetic resonance imaging (fMRI), electroencephalography (EEG), magnetoencephalography (MEG) in humans, as well as cell recordings in animals, supports the notion that brain functions emerge from distributed processing across multiple brain areas (e.g., Cabeza and Nyberg, 2000). The study of neural dynamics over the whole cortex is the key to understand the distributed processing of information in distinct functional regions. Long-range synchronization of oscillatory activity has been proposed as a dynamical mechanism for mediating the interaction between brain areas in a task-dependent manner (Engel et al., 2001; Fries, 2005; Chawla et al., 1999). Furthermore, it was recently shown that neuronal synchronization mediates neuronal communication within large-scale cortical subnetworks both at rest (Grecius et al., 2003; Fox and Raichle, 2007; Brookes et al., 2011; de Pasquale et al., 2012) and during task performance (Hipp et al., 2011; Betti et al. 2013).

A fundamental question in system neuroscience is how information can be processed in a distributed fashion. Global cortical dynamics – as observed at rest – exhibit a default pattern of communication, which is reconfigured depending on attentional or task set (de Pasquale et al., 2015). Experimental evidence of this sort suggests that long-range functional interactions may be established through the coordination of the dynamical states of cortical nodes, which is referred to as functional connectivity (FC). Presumably, both sensory-driven and cognitive-driven processes are involved in shaping FC. Descriptive characterizations of distributed brain processing are usually based on estimates of FC using brain imaging (Cordes et al., 2000).

In this paper, we examine the specific role played by neuro-anatomical connections among brain areas in shaping intracortical communication, via the study of FC. In particular, we focus on the well-established hypothesis that the coordination of different regions or neuronal populations depends on changes of the strengths of intracortical or intrinsic connectivity. Based on dynamic models for blood oxygen level dependent (BOLD) activity at the level of a cortical region, techniques to estimate those strengths have been developed. These techniques are based on models of effective connectivity (McIntosh and Gonzalez-Lima, 1994; Friston, 2002; Honey et al., 2007; Friston and Dolan, 2010; Deco et al., 2011; Cabral et al., 2011). The distinction between functional and effective connectivity is crucial here: FC is defined as the statistical dependence between remote neuro-physiological dynamics, whereas effective connectivity (EC) is defined as the influence one neural system exerts over another (Friston et al., 2003). In neurophysiology, effective connectivity was originally defined as the simplest circuit diagram that could replicate

observed patterns of functional connectivity (Aertsen et al., 1989). Here, we assume that effective connectivity within and between cortical areas mediates the statistical dependencies among their activities, i.e., FC.

We use a noise-diffusion network to model BOLD signals and a recently established estimation procedure to infer changes in both intrinsic EC and region-specific inputs (Gilson et al., 2016). The key is to use the spatiotemporal information in FC; namely, BOLD covariances with and without time shifts to inform estimates of EC. The ensuing model integrates anatomical data obtained from diffusion spectrum imaging (DSI) to constrain extrinsic (between-region) connectivity. This model offers a new perspective on measured data: as the amplitudes of BOLD fluctuations (i.e., variances) convey information about the task – and appear consistent with neural computations performed by the cortical areas – the key parameters of this model are the amplitude (i.e. variance) of region-specific inputs. The model parameters estimated from the fMRI data thus provide a dynamic connectivity architecture for information propagation and integration in the brain.

RESULTS

We analyze BOLD signals recorded in 19 healthy volunteers who were watching either a black screen or a movie, to determine significant changes in their spatiotemporal covariances. We use a model-based approach to extract information at the subject level, where model fitting provides estimates of (the variance of) region specific inputs Σ and effective connectivity C of a cortical network of 66 ROIs (or cortical regions). These parameters are then analyzed to characterize condition-specific changes in functional integration, which was found to be more pronounced in the early visual and auditory systems – as expected for watching movies.

1. Changes in FC induced by movie viewing

As shown in Fig. 1A, the BOLD signals do not exhibit changes in their mean across the two viewing conditions, but some ROIs show different variances (blue crosses). Evaluating the mean over all ROIs for each subject, we observe a larger average BOLD variance for the movie than for the black screen: the black line indicates equality in the left panel of Fig. 1B. The right panel of Fig. 1B displays time constants τ_x estimated from BOLD autocovariance functions; they indicate the "memory depth" of corresponding time series; namely, how long fluctuations persist over successive TRs (see Eq. 4 in Methods). Here, we assume a single time constant for all ROIs and did not find a consistent change between the two conditions. From the two plots in Fig. 1B, we discarded three individuals (in red) with extreme values: two for the variances and one for the time constant. This left 19 subjects for subsequent analysis, for which we calculate the FC along all ROIs: in our framework, FC0 corresponds to an instantaneous covariance (with no time shifts are zero like), whereas FC1 corresponds to a life of one repetition time (TR) (see Eq. 3 in Methods for details). The significance of changes in each matrix element was evaluated using Welch's t-test and is reported for variances in Fig. 1C (in order of the smallest p-values) and in Fig. 1D for all matrix elements of FC0 and FC1. Fig. 1C shows the most significant changes are seen in several areas in the early visual and auditory systems. Changes with p-values smaller than 0.01 (i.e., score larger than 2 in Fig. 1D) are plotted in Fig. 1E: both FC0 and FC1 exhibit significant increases and decreases during movie viewing. Changes in BOLD variances were seen mainly the early visual and auditory pathways (rLOCC, lLOCC, rCUN, lCUN, lPCAL, rST, lST, lMT and lIT) and multimodal integration regions (rFUS, lFUS, rBSTS, lBSTS, rIP, lIP and lTP), as well as two more central regions rPARH and lENT.

To complete the empirical FC data analysis, we compared the significance of changes in BOLD covariances with that of BOLD correlations in Fig. 1F, which are commonly used in the literature. Fig. 1G shows that variances are quite informative in discriminating between the two conditions, which supports our approach based on covariances (as opposed to correlations, which remove variances). Although these FC results are interesting, our aim is not to characterize FC

changes between conditions, but to use the FC as data to estimate the underlying EC – and region-specific activity that best explains the observed functional connectivity.

2. A (noise-diffusion) model of EC that explains spatiotemporal FC

The results in Fig. 1 show that BOLD (co)variances discriminate between the two conditions. To make sense of the mixture of increases and decreases observed in the empirical FC, we used a recently developed generative model (Gilson et al., 2016) that explains or predicts spatiotemporal FC. This model is schematically represented in Fig. 2, for a few cortical regions. The skeleton of extrinsic connectivity in Fig. 2A is determined by (thresholded) DTI data, which estimates the density of white matter fibers between the 66 ROIs considered here. We only retain the information about the presence of connections (and discard quantitative details of the DTI values). It is known that DTI can miss inter-hemispherical connections, so we added these connections *post hoc* (in red). This increased the density of structural connectivity (SC) from 27% to 28% of all possible connections.

In our noise-diffusion model, each ROI receives a fluctuating input that is propagated by recurrent extrinsic EC throughout the network, generating covariances that predict FC, as illustrated in Fig. 2B. The strengths of connections constitute the matrix C and the input variances comprise the leading diagonal of matrix Σ. These EC parameters are estimated from the data by an iterative optimization procedure that tunes the model so as to reproduce empirical FC matrices, see Fig. 3C. The model is initially fitted by estimating the time constants τ_x in Fig. 1B, for each subject and condition (but applied to all ROIs). As demonstrated in Gilson et al. (2016), both the EC and input parameters can be estimated from FC0 and FC1. The uniqueness of this estimation follows from the bijective mapping from the model parameters C and Σ to the FC pair (FC0,FC1). The precision of the estimated EC is limited by the number of time points in the BOLD signals, as with all models. The advantage of our approach as compared to MVAR estimation is that we can incorporate information from DTI in our model – and that intrinsic (within region) nonlinear dynamics can be incorporated in the model.

The goodness of fit is shown in Fig. 3A for a single subject in both conditions, and summarized in Fig. 3B for all subjects. The Pearson correlation between the model and empirical FC matrix elements is larger than 0.7 for most subjects and conditions, synonymous with relatively high accuracy. More importantly, the model captures the change in FC between the two conditions, as illustrated in the right panel of Fig. 3B with Pearson correlation larger than 0.6 four changes in FC (ΔFC: movie minus blank screen). To verify our results, we repeated the same estimation procedure, but using FC2 with a time shift of 2 TRs instead of FC1. We find nearly identical Σ estimates and very similar C estimates, see Fig. 3C.

3. Movie viewing induces greater changes in (intrinsic) activity than (extrinsic) connectivity

Over the 19 subjects, changes in the structure of both Σ and C were observed, as exemplified for a single subject in Fig. 4A and measured by the Pearson correlation in Fig. 3B. Significant changes with p-value<0.01 (uncorrected) for Σ in Fig. 3C identify 15 nodes (29% of the 66 in total), while C showed 156 significant changes (14% of 1148 connections). In the model, $\Delta\Sigma$ reports the change of empirical BOLD variances, as shown in Fig. 4D: the model M/BS that combines C from the movie condition and Σ from the blank screen reproduces the FC during movie viewing much more poorly the than M/M, which is the model estimated from the movie data. Interestingly, ΔC is constrained by the FC0 covariances and FC1, as indicated by the comparison between the model BS/M and M/M reported in the middle and right panels. This can be understood from the Lyapunov equation (based on our noise-diffusion model), which suggests that changes in FC0 only cannot disentangle the contributions from input and connectivity. Our approach resolves this indeterminacy by considering FC0 and FC1 (or FC2), enabling us to estimate both.

Fig. 5A shows that changes in the distributions of Σ values are greater than C values, for both conditions. To characterize the nature of those changes, we report the most significant changes Σ in Fig. 5B, which list ROIs that showed significant changes in BOLD variances in Fig. 1C. Crucially, the ranking is not the same, which means that the model captures the causes of the changes observed in the BOLD variances, and not the observed changes *per se*. The mean of Σ is plotted in Fig. 5C for each ROI and condition, showing that the most significant changes ROIs are also the strongest in amplitude. The SC for the early visual and auditory systems, featuring most ROIs in Fig. 5B, is displayed in Fig. 5D: this dense connectivity exhibits a hierarchy, represented along the diagonal, with regions FUS and BSTS in the middle integrating the inputs from visual and auditory ROIs at each end. Finally, we calculated normalized scores for C and Σ in Fig. 5E based on the histograms in Fig. 5A, in order to indicate the statistical strength of estimated values over the all subjects: these normalized scores (the average over subjects, relative to the median over all regions or connections, divided by their standard deviation over subjects) are shown as the dashed-dotted line in Fig. 5A. As expected from previous observations in Fig. 4, changes in Σ are larger than those in C in Fig. 5E. Nevertheless, the details of $\Delta\Sigma$ in the bottom panel shows that the strongest increases are for the visual LOCC and auditory ST, as well as the integrative BSTS. Concerning ΔC, the EC within the auditory ST, TT and MT increase, whereas the effective connectivity within the visual system is decreased. Importantly, effective connections from and to the integrative FUS and BSTS are increased.

4. Sensory influx and cortical path selection

To better understand the implications of those distributed changes, we define in Fig. 6A an

effective drive (ED), which measures, for each connection, the input amplitude (i.e., standard deviation of BOLD signal) that is transferred from the source ROI to the target ROI. It is defined over subjects as above as the normalized effective connectivity times the standard deviation of bold activity in each source region. The ED is a canonical measure for noise-diffusion processes that describes the propagation of large-amplitude fluctuations in the network. The effective drive estimated in Fig. 6A clearly shows that the hierarchical processing of visual and auditory information for the movie from the sensory ROIs to the associative or integrative ROIs; namely, FUS and BSTS. When watching a black screen, however, this integration is much lower and only the visual system appears to distribute variance, probably in waiting for something to occur.

The difference between the effective drive in movie and black-screen conditions is shown in the middle panel of Fig. 6B: it shows an overall increase in the auditory system, but more importantly, an increase to and from the integrative ROIs. In Fig. 5, we show the change in effective drive due to the changes ΔC and $\Delta \Sigma$: if only inputs Σ are modified, the whole area saturates and the integrative nodes are no longer driven by the sensory inputs. The changes in C regulate this sensory drive and boost the connections from the integrative FUS and BSTS to LOCC and ST. In other words, $\Delta \Sigma$ can be related to the stimulus load in our model and its propagation is gated or modulated along specific pathways by ΔC. The opposing effects of $\Delta \Sigma$ and ΔC observed in Fig. 6B are a key finding and suggest a dynamic balance of the cortical information flow across conditions.

Finally, we step back to the whole cortex and analyze the changes in ED for 6 groups of ROIs in Fig. 7A: in addition to the visual, auditory and integration regions in Figs. 5 and 6, we include motor, frontal and central regions. ED at the cortical level exhibits an increase of forward interactions from the visual and auditory groups to the integration, frontal and central areas. A moderate increase of feedback from integration to the visual system can also be seen. To further analyze this network propagation, we use the Louvain method (Blondel et al., 2008) on the ED matrix in both conditions to cluster ROIs: in Fig. 7B, darker pixels indicate a participation index in the same cluster for each pair of ROIs. When watching the black screen, we observe that the visual regions are strongly bound to themselves and a few central ROIs, whereas the auditory group is strongly clustered with the integration, motor and frontal groups. It is worth noticing that the two hemispheres are poorly connected. In contrast, inter-hemispherical clustering is much stronger when watching the movie, the integration ROIs divide to cluster with both visual and auditory groups. Likewise, individual ROIs within the frontal and central groups cluster with the entire visual/auditory/integration groups. Similar results were obtained when applying the Louvain method to the connectivity matrix C, instead of ED. This underlines selective coordination of paths to implement a distributed processing of information during movie watching, whereas the cortical connectivity is stronger within groups of functionally specialized regions in the black-screen condition.

DISCUSSION

Our results shed light on a fundamental issue in neuroscience: how do inputs and connectivity interact to enable large-scale integration of information in the brain? To address this question, we apply a recently developed generative model to task-evoked activity. This model and provides quantitative estimates of (extrinsic) effective connectivity (intrinsic) regional activity, to assess the information flow in the cortical network. In turn, this allows for the comparison between task-evoked BOLD time series in a recurrently connected network of ROIs. Our main finding is that the extrinsic connectivity is adjusted to select specific pathways for the integration of information (Figs. 5, 6 and 7). This includes a down-regulation of forward connections of the auditory and visual systems in a compensatory manner, such that regional inputs do not saturate the network: meanwhile, backward connections are boosted, enabling high-level areas to influence sensory areas, despite their activity increase. Here, the dynamic balance between these effects depends on the task and effective drive – as a canonical measure of fluctuation – that speaks to functional synchronization at the network level. Our results are in line with previous studies that observed such balanced activity from the neuronal level (Destexhe et al., 2003; Dehghani et al., 2016) to the cortical level (Deco and Corbetta, 2011).

Although the connectivity matrices corresponding to the black screen and movie may appear rather similar when measured by Pearson correlation (Fig. 4), they induce very different community pooling (Fig. 7). At the level of the whole cortex, we observe that the two hemispheres are rather isolated from each other, when watching the black screen. In contrast, homotopic areas exchange information via interhemispheric effective connections. Moreover, the cortex becomes specialized when engaging in a task: as revealed by the community analysis based on the effective drive or connectivity: specific high-level ROIs from the frontal and central lobes bind themselves to sensory and integration areas in the movie condition. This illustrates an elaborate scheme of functional segregation and integration. In particular, network effects seem to generate a multiplicity of inter-areal interaction patterns supported by the same structural network, as shown by Battaglia et al. (2012).

Importantly, our framework is based on the propagation of fluctuating activity (second-order statistics): the noise in the model is “functional” and the input variances in Σ represent spontaneous activity, as well as stimulus-related activity. Therefore, we interpret the FC variance as a proxy for the neuronal processing at each ROI, while C governs the corresponding information flow between ROIs that generates the FC covariances. The results in Fig. 4D show that inputs must be estimated as well as effective connectivity, to explain observed changes in FC0. As shown in Fig. 1, lagged covariances FC1 conveys crucial information that discriminates between the states of functional integration.

The effective drive in Fig. 6 is tied to the noise-diffusion model and measures the propagation of fluctuations in the network, like an entrainment degree between nodes. Beyond the

present model, our study suggests that BOLD (co)variances convey important information about the task engaged by – or the state of – subjects. Moreover, the temporal component of FC should be considered, not only the spatial component in FC0. The notion of effective dynamics developed in the present study is related to the approach of dynamic causal modelling, in the sense that statistical dependences of the ROIs' activities depend on both dynamic variables (inputs Σ here) and connection strengths. The change of perspective is that the variances of inputs are stimulus-dependent and contribute to a larger extent to the shaping of FC.

ACKNOWLEDGMENTS

This work was partly supported by the 7th Framework Programme of the European Commission (grant PCIG12-334039 to DM) and the KU Leuven Special Research Fund (grant C16/15/070 to DM).

MATERIAL AND METHODS

Study design for Empirical fMRI data during Resting and Passive Movie Viewing:

We re-analyzed BOLD imaging data reported in our previous papers (Hlinka et al., 2011; Mantini et al., 2012, 2013). Twenty-four right-handed young, healthy volunteers (15 females, 20–31 years old) participated in the study. They were informed about the experimental procedures, which were approved by the Ethics Committee of the Chieti University, and signed a written informed consent. The study included a resting state and a natural viewing condition. In the resting state, participants fixated a red target with a diameter of 0.3 visual degrees on a black screen. In the natural viewing condition, subjects watched (and listened) to 30 minutes of the movie "The Good, the Bad and the Ugly" in a window of 24x10.2 visual degrees. Visual stimuli were projected on a translucent screen using an LCD projector, and viewed by the participants through a mirror tilted by 45 degrees. Auditory stimuli were delivered using MR-compatible headphones.

Data acquisition:

Functional imaging was performed with a 3T MR scanner (Achieva; Philips Medical Systems, Best, The Netherlands) at the Institute for Advanced Biomedical Technologies in Chieti, Italy. The functional images were obtained using T2*-weighted echo-planar images (EPI) with BOLD contrast using SENSE imaging. EPIs comprised of 32 axial slices acquired in ascending order and covering the entire brain (32 slices, 230 x 230 in-plane matrix, TR/TE=2000/35, flip angle = 90°, voxel size=2.875×2.875×3.5 mm3). For each subject, 2 and 3 scanning runs of 10 minutes duration were acquired for resting state and natural viewing, respectively. Each run included 5 dummy volumes – allowing the MRI signal to reach steady state and an additional 300 functional volumes that were used for analysis. Eye position was monitored during scanning using a pupil-corneal reflection system at 120 Hz (Iscan, Burlington, MA, USA). A three-dimensional high-resolution T1-weighted image, for anatomical reference, was acquired using an MP-RAGE sequence (TR/TE=8.1/3.7, voxel size=0.938x0.938x1 mm3) at the end of the scanning session.

Data processing:

Data were preprocessed using SPM8 (Wellcome Department of Cognitive Neurology, London, UK) running under MATLAB (The Mathworks, Natick, MA). The preprocessing steps involved: (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 mm^3. 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 (Sui et al., 2009). This BOLD artifact removal procedure was performed by means of the GIFT toolbox (Medical Image

Analysis Lab, University of New Mexico). No global signal regression or spatial smoothing was applied.

For each recording session (subject and run), we extracted the mean BOLD time series from the 66 regions of interest (ROIs) of the brain atlas used in Hagmann et al. (2008); see Table 1 for details. For each ROI, we calculated the overall percent signal change value in the natural viewing condition using the resting state condition as baseline. For resting state and natural viewing sessions separately, we concatenated the BOLD time series for each region, and calculated the 66x66 correlation matrix representing the FC between each pair of cortical areas.

Structural Connectivity Matrix:

Anatomical connectivity was estimated from Diffusion Spectrum Imaging (DSI) data collected in five healthy right-handed male participants (Hagmann et al., 2008; Honey et al., 2009). The grey matter was first parcellated into 66 ROIs, using the same low-resolution atlas used for the FC analysis (Hagmann et al., 2008). For each subject, we performed white matter tractography between pairs of cortical areas to estimate a neuro-anatomical connectivity matrix. In our method, the DSI value between two brain areas is only used to indicate the existence of an structural connection, yielding a binary matrix of structural connectivity (SC) – obtained by averaging the matrices over subjects. The strength of individual intracortical connections is optimized relation to the functional data, as explained below.

Many important links are missing in the DSI tractography (Hagmann et al., 2008); for example, the inter-hemispheric connections between homologous cortical areas are severely under-estimated. Therefore, we enhanced the DSI-based matrix by adding connections that were tuned during the optimization.

Empirical covariances:

For each of the two sessions for black-screen watching and for movie matching, the BOLD time series s_i^t for each region $1 \le i \le N$ with discrete time $1 \le t \le 180$ are centered to remove the mean signal. Following Gilson et al. (2016), the empirical covariances are calculated as:

$$\hat{Q}_{ij}^0 = \frac{1}{T-1}\sum_t s_i^t s_j^t \text{ and } \hat{Q}_{ij}^1 = \frac{1}{T-2}\sum_t s_i^t s_j^{t+1} . \quad (3)$$

For each individual and session, we calculate the time constant τ_x associated with the exponential decay of the autocovariance function $\hat{Q}_{ii}^\tau$ averaged over all regions:

$$\tau_x = \frac{1}{\log\left(\hat{Q}_{ii}^0\right) - \log\left(\hat{Q}_{ii}^1\right)}. \quad (4)$$

This is used to calibrate the model, which is then optimized to reproduce the empirical covariances.

Cortical model:

The model comprised $N = 66$ interconnected cortical regions, which each receive a stochastic input that induces fluctuating activity. The latter is shaped by the recurrent affective connectivity C to generate the functional connectivity matrix Q.

In practice, the network model is based on a multivariate Ornstein-Uhlenbeck process with nonlinear interactions. Each population indexed by $1 \le i \le N$ has an activity variable x_i, which decays exponentially with time constant τ_x and receives contributions from other populations in addition to random situations or region-specific input: $\frac{dx_i}{dt} = \frac{-x_i}{\tau_x} + \sum_{j \neq i} C_{ij} x_j + \sigma_i dB_i$. Here, dB_i is unitary white noise scaled by the variance σ_i. The extrinsic (between-region) effective connectivity is encoded by the matrix C, for which DTI determines the sparsity structure (non-zero elements). In practice, its connectivity density is close to 30%. In our model, the spatiotemporal covariances of the variables x_i arise from the noisy inputs dB_i – they are defined as $Q_{ij}^{\tau} = \left\langle \left(x_i^t - \bar{x}_i \right) \left(x_j^{t+\tau} - \bar{x}_j \right) \right\rangle$, where $\bar{x}_i = \left\langle x_i^t \right\rangle$ is the mean of the activity variable and the angular brackets denote averaging over random fluctuations. In practice, we use two time shifts: zero and τ equal to 1 TR, as this is sufficient to characterize the network parameters.

The matrices C and Q^0 are related via a Lyapunov equation: $JQ^0 + Q^0 J^T + \Sigma = 0$, where the Jacobian of the dynamical system $J_{ij} = C_{ij} - \tau_x^{-1}$ depends on the mean activity of the network, and Σ is a diagonal matrix with the variances of the random fluctuations, namely $\Sigma_{ii} = \sigma_i^2$. The time shifted covariance is given by $Q^{\tau} = Q^0 \mathrm{expm}(J^T \tau)$. These two consistency equations allow for the quick estimation of the predicted FC matrices; i.e., without simulating the network with noisy inputs and averaging over a long recording period.

Parameter estimation procedure:

We tune the model such that its covariance matrices Q^0 and Q^1 reproduce the empirical FC, namely empirical $\hat{Q}^0$ and $\hat{Q}^1$. We summaries here the essential steps of the procedure described in Gilson et al. (2016) that iteratively optimizes the network parameters C and Σ. At each step, the noiseless network is simulated to evaluate the mean activity $\bar{x}_i$ for the current values of C and Σ, which allows for the calculation of the Jacobian J. Then, the model FC matrices Q^0 and Q^1 are calculated from the consistency equations, using the Bartels-Stewart algorithm to solve the

Lyapunov equation. The Jacobian update is $\delta J^T = \left(Q^0\right)^{-1}\left[\delta Q^0 + \delta Q^1 \text{expm}\left(-\text{J}^\text{T}\right)\right]$, which involves the FC error between the empirical and model matrices, namely $\delta Q^0 = \hat{Q}^0 - Q^0$ and $\delta Q^1 = \hat{Q}^1 - Q^1$; the superscript T denotes the matrix transpose. Finally, the connectivity update is $\delta C_{ij} \propto \delta J_{ij}$ for existing connections. We impose non-negativity of the EC values during the optimization. The input variances are tuned according to the heuristic optimization $\delta\Sigma_{ii} \propto \delta Q^0_{ii}$, which enforces an equality between the model and empirical FC variances.

REFERENCES

Aertsen AM, Gerstein GL, Habib MK, Palm G (1989) Dynamics of neuronal firing correlation: modulation of “effective connectivity”. J Neurophysiol 61: 900-917

Battaglia D, Witt A, Wolf F, Geisel T (2012) Dynamic effective connectivity of inter-areal brain circuits. PLoS Comp Biol 8: e1002438

Betti V, Della Penna S, de Pasquale F, Mantini D, Marzetti L, Romani GL, Corbetta M (2013) Natural scenes viewing alters the dynamics of functional connectivity in the human brain. Neuron 79: 782-797

Biswal B, Yetkin F, Haughton V, Hyde J (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn Reson Med 34: 537-541

Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. J Stat Mech 10: P10008

Brookes MJ, Woolrich M, Luckhoo H, Price D, Hale JR, Stephenson MC, Barnes GR, Smith SM, Morris PG (2011) Investigating the electrophysiological basis of resting state networks using magnetoencephalography. Proc Natl Acad Sci USA 108: 16783-16788

Cabeza R, Nyberg L (2000) Imaging cognition II: empirical review of 275 PET and fMRI studies. J Cogn Neurosci 12: 1-47

Cabral J, Hugues E, Sporns O, Deco G (2011) Role of local network oscillations in resting-state functional connectivity. Neuroimage 57: 130-139

Chawla D, Lumer E, Friston K (1999) The relationship between synchronization among neuronal populations and their mean activity levels. Neural Comput 11: 1389-1411

Cordes D, Haughton VM, Arfanakis K, Wendt GJ, Turski PA, Moritz CH, Quigley MA, Meyerand ME (2000) Mapping functionally related regions of brain with functional connectivity MR imaging. AJNR Am J Neuroradiol 21: 1636-1644

de Pasquale F, Della Penna S, Snyder AZ, Marzetti L, Pizzella V, Romani GL, Corbetta M (2012) A cortical core for dynamic integration of functional networks in the resting human brain. Neuron 74: 753-764

de Pasquale F, Della Penna S, Sporns O, Romani GL, Corbetta M (2015) A Dynamic Core Network and Global Efficiency in the Resting Human Brain. Cereb Cortex (online first)

Deco G, Corbetta M (2011) The dynamical balance of the brain at rest. Neuroscientist 17: 107–123

Deco G, Jirsa V, McIntosh A (2011) Emerging concepts for the dynamical organization of resting-state activity in the brain. Nat Rev Neurosci 12: 43-56

Dehghani N, Peyrache A, Telenczuk B, Quyen MLV, Halgren E, Cash SS, Hatsopoulos NG, Destexhe A (2016) Sci Rep 6: 23176

Destexhe A, Rudolph M, Pare D (2003) The high-conductance state of neocortical neurons in vivo. Nat Rev Neurosci. 4: 739–751

Engel A, Fries P, Singer W (2001) Dynamic predictions: oscillations and synchrony in top-down processing. Nat Rev Neurosci 2: 704-716

Fox M, Raichle M (2007) Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nat Rev Neurosci 8: 700-711

Fries P (2005) A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. Trends Cogn Sci 9: 474-480

Friston K (2002) Beyond phrenology: what can neuroimaging tell us about distributed circuitry? Annu Rev Neurosci 25: 221-250

Friston K, Harrison L, Penny W (2003) Dynamic causal modelling. Neuroimage 19: 1273-1302

Friston K, Dolan R (2010) Computational and dynamic models in neuroimaging. Neuroimage 52:752-765

X Gilson M, Moreno-Bote R, Ponce-Alvarez A, Ritter P, Deco G (2016) Estimation of Directed Effective Connectivity from fMRI Functional Connectivity Hints at Asymmetries of Cortical Connectome. PLoS Comput Biol 12: e1004762

Greicius MD, Krasnow B, Reiss A, Menon V (2003) Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc Natl Acad Sci USA 100: 253-258

Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, Sporns O (2008) Mapping the structural core of human cerebral cortex. PLoS Biol 6: e159

Hipp JF, Engel AK, Siegel M (2011) Oscillatory synchronization in large-scale cortical networks predicts perception. Neuron 69: 387-396

Hlinka J, Palus M, Vejmelka M, Mantini D, Corbetta M (2011) Functional connectivity in resting-state fMRI: is linear correlation sufficient? Neuroimage 54: 2218-2225

Honey CJ, Kötter R, Breakspear M, Sporns O (2007) Network structure of cerebral cortex shapes functional connectivity on multiple time scales. Proc Natl Acad Sci USA 104: 10240-10245

Mantini D, Hasson U, Betti V, Perrucci MG, Romani GL, Corbetta M, Orban GA, Vanduffel W (2012) Interspecies activity correlations reveal functional correspondence between monkey and human brain areas. Nat Methods 9: 277-282

Mantini D, Corbetta M, Romani GL, Orban GA, Vanduffel W (2013) Evolutionarily novel functional networks in the human brain? J Neurosci 33: 3259-3275

McIntosh R, Gonzalez-Lima F (1994) Structural equation modeling and its application to network analysis in functional brain imaging. Hum Brain Mapp 2: 2-22

Sui J, Adali T, Pearlson GD, Calhoun VD (2009) An ICA-based method for the identification of optimal FMRI features and components using combined group-discriminative techniques. Neuroimage 46: 73-86

Van Essen D, Anderson C, Felleman, D (1992) Information processing in the primate visual system: An integrated systems perspective. Science 255: 419-423

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

Abbreviation	Brain region
BSTS	**Bank of the superior temporal sulcus**
CAC	**Caudal anterior cingulate cortex**
CMF	**Caudal middle frontal cortex**
CUN	**Cuneus**
ENT	**Entorhinal cortex**
FP	**Frontal pole**
FUS	**Fusiform gyrus**
IP	**Inferior parietal cortex**
ISTC	**Isthmus of the cingulate cortex**
IT	**Inferior temporal cortex**
LING	**Lingual gyrus**
LOCC	**Lateral occipital cortex**
LOF	**Lateral orbitofrontal cortex**
MOF	**Medial orbitofrontal cortex**
MT	**Middle temporal cortex**
PARC	**Paracentral lobule**
PARH	**Parahippocampal cortex**
PC	**Posterior cingulate cortex**
PCAL	**Pericalcarine cortex**
PCUN	**Precuneus**
POPE	**Pars opercularis**
PORB	**Pars orbitalis**
PREC	**Precentralgyrus**
PSTC	**Postcentralgyrus**
PTRI	**Pars triangularis**
RAC	**Rostral anterior cingulate cortex**
RMF	**Rostral middle frontal cortex**
SF	**Superior frontal cortex**
SMAR	**Supramarginalgyrus**
SP	**Superior parietal cortex**
ST	**Superior temporal cortex**
TP	**Temporal pole**
TT	**Transverse temporal cortex.**

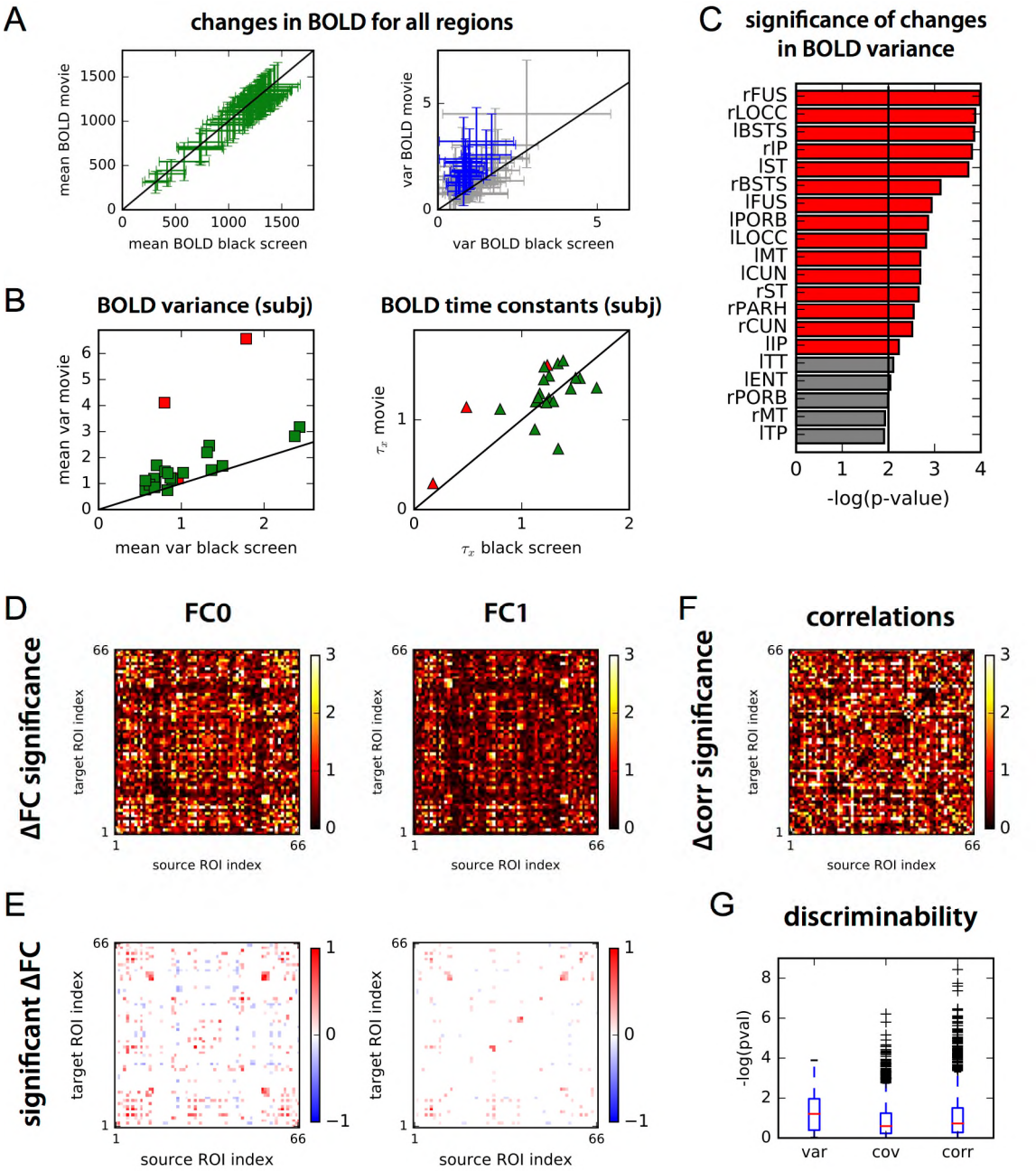

Fig. 1: fMRI data recorded on subjects watching a black screen or a movie. A Changes in BOLD mean and variance between the black-screen and movie conditions. Each cross represents one of the 66 cortical ROIs and the variability corresponds to the distribution over all 22 subjects. For the variances, blue crosses indicate significant changes between black screen and rest (with p-value < 0.01). **B** Comparison of mean variances and time constants τ_x calculated for the two conditions. Each symbol represents a subject, where red indicates three discarded subjects, leaving 19 valid subjects for the following analysis. **C** Areas with most significant changes of BOLD variances between the two conditions. Significance is evaluated using Welch's t-test with unequal variances over valid subjects for each matrix element; p-values smaller than 0.01 correspond log values larger than 2 in red. **D** Significant changes in covariances matrices, FC0 with no time shift and FC1 with a time shift equal to 1 TR. The plotted score are -log(p-value) as in C. **E** Amplitude of significant changes in D: for each pair of regions, red indicates an increase and blue a decrease. **F** Same as D for BOLD correlations instead of covariances. **G** Comparison of p-values for variances (diagonal of FC0), covariances (off-diagonal elements of FC0) and correlations.

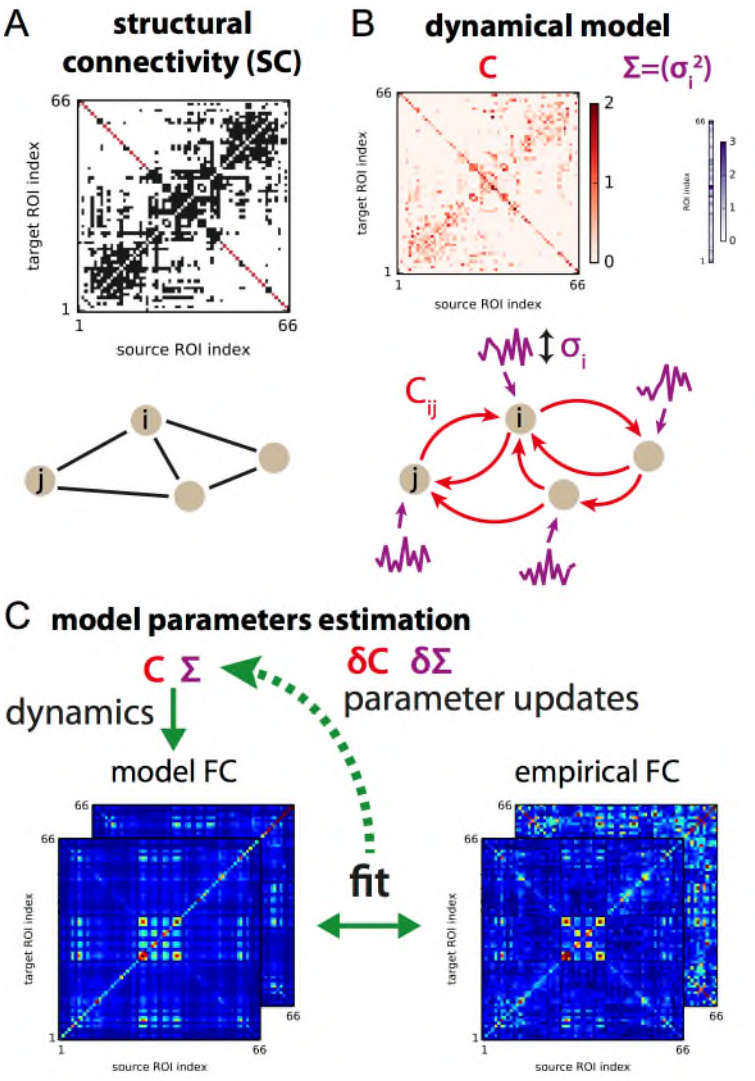

Fig. 2: Dynamical cortical model. A DTI data provide the skeleton of the intracortical connectivity matrix. We had inter-hemispherical connections (red pixels) as they are known to miss from DTI data. **B** The parameters of the model are the recurrent connectivity C and the input variances Σ. **C** From the parameters, the model FC0 and FCτ matrices are calculated and compared to their empirical counterparts, which in turn gives the updates δC and $\delta\Sigma$ for the model. The optimization is performed until the minimal matrix distance is reached between the model and empirical FC matrices (average of both sets).

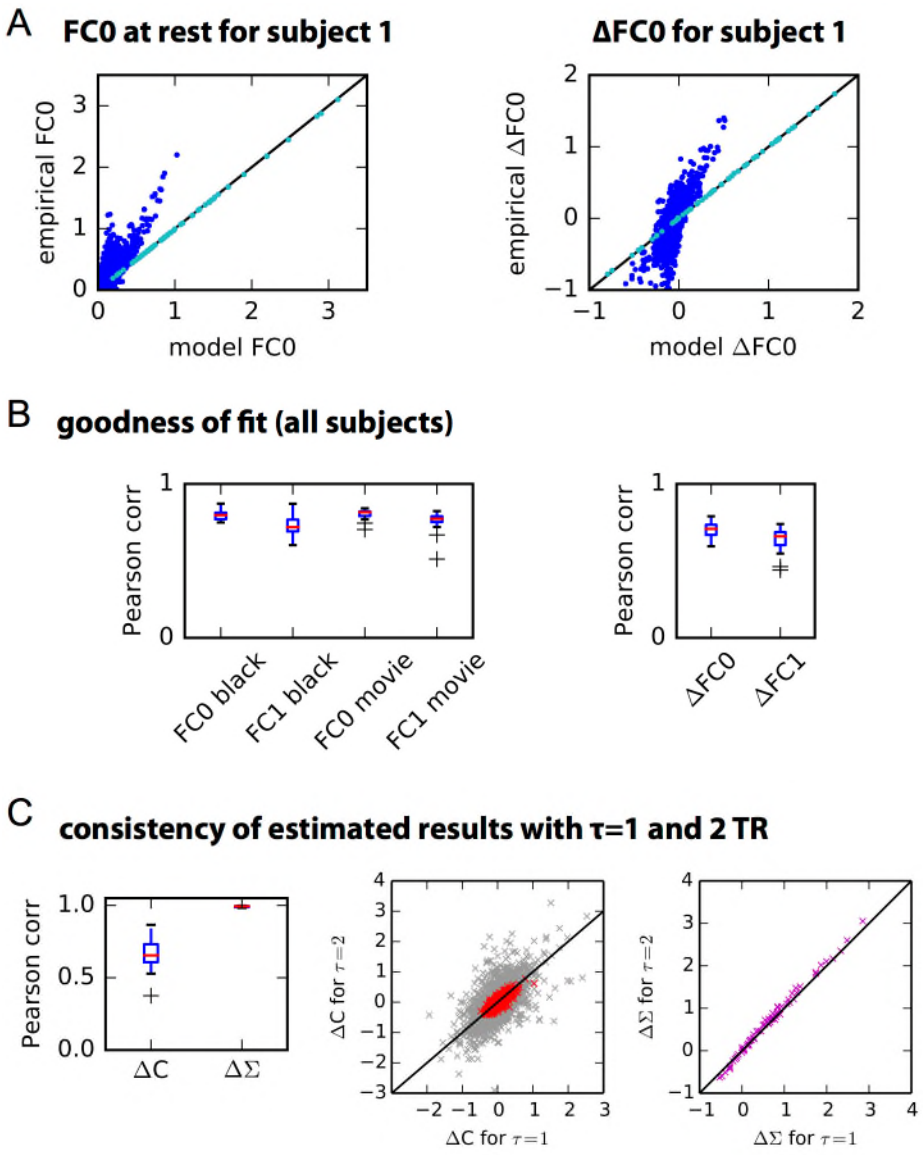

Fig. 3: Goodness of model fit. **A** Changes in FC0 between black screen and movie for a single subject; each dot represent a matrix element. Comparison between the data (left in dark blue) and the model (right in cyan). **B** Goodness of fit for the individual FC0 and FC1 in both conditions over all subjects, as well as $\Delta FC0 = FC0_{movie} - FC0_{rest}$ and $\Delta FC1 = FC1_{movie} - FC1_{rest}$. For each subject, we consider the pair of matrices model/data and the performance is measured by the Pearson correlation coefficient between the matrix elements (cf. plotted dots in A). **C** Consistency between the ΔC and $\Delta \Sigma$ matrices obtained for each subject using two distinct optimizations, namely FC0/FC1 and FC0/FC2. The left panel displays the Pearson correlation between corresponding matrices for τ=1 and 2 TR. The middle and right panels show the correspondence of matrix elements, with the black diagonal indicating a perfect match. Mean values over all subjects are plotted in colors.

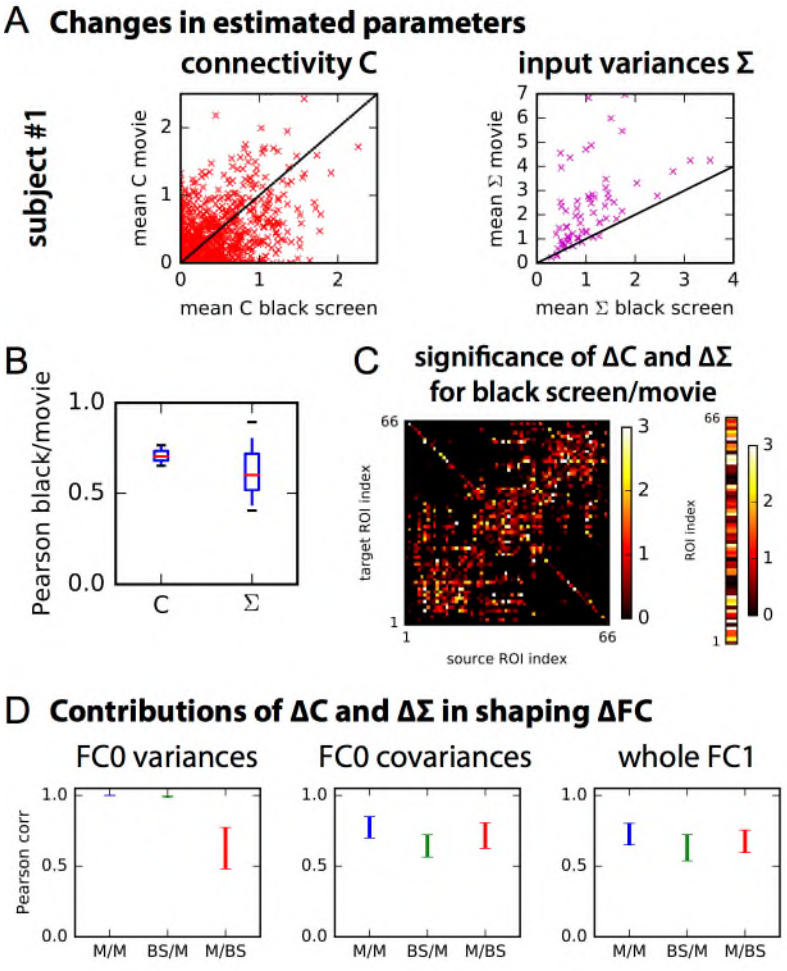

Fig. 4: Global changes in the estimated cortical connectivity *C* and input variances *Σ*. A Changes in *C* and *Σ* for a single subject; each red cross represents a cortico-cortical connection and each purple cross a ROI. **B** Changes in the structure of *C* and *Σ* for all subjects. It is measured by the Pearson correlation between the plotted dots in A. **C** Significance of the changes in C and *Σ* obtained from a t-test with unequal variances. The plotted color corresponds to -log(p-value), estimated with a t-test between the two conditions as in Fig. 1. **D** Comparison of three models that combine the estimated *C* and *Σ* in the following manner: X/Y correspond to C_X and Σ_Y with X and Y indicating either the black screen (BS) or movie (M). The error bars correspond to the variability over the subjects of the Pearson correlation coefficients between the elements of FC matrices generated using C_X and Σ_Y and the empirical counterparts.

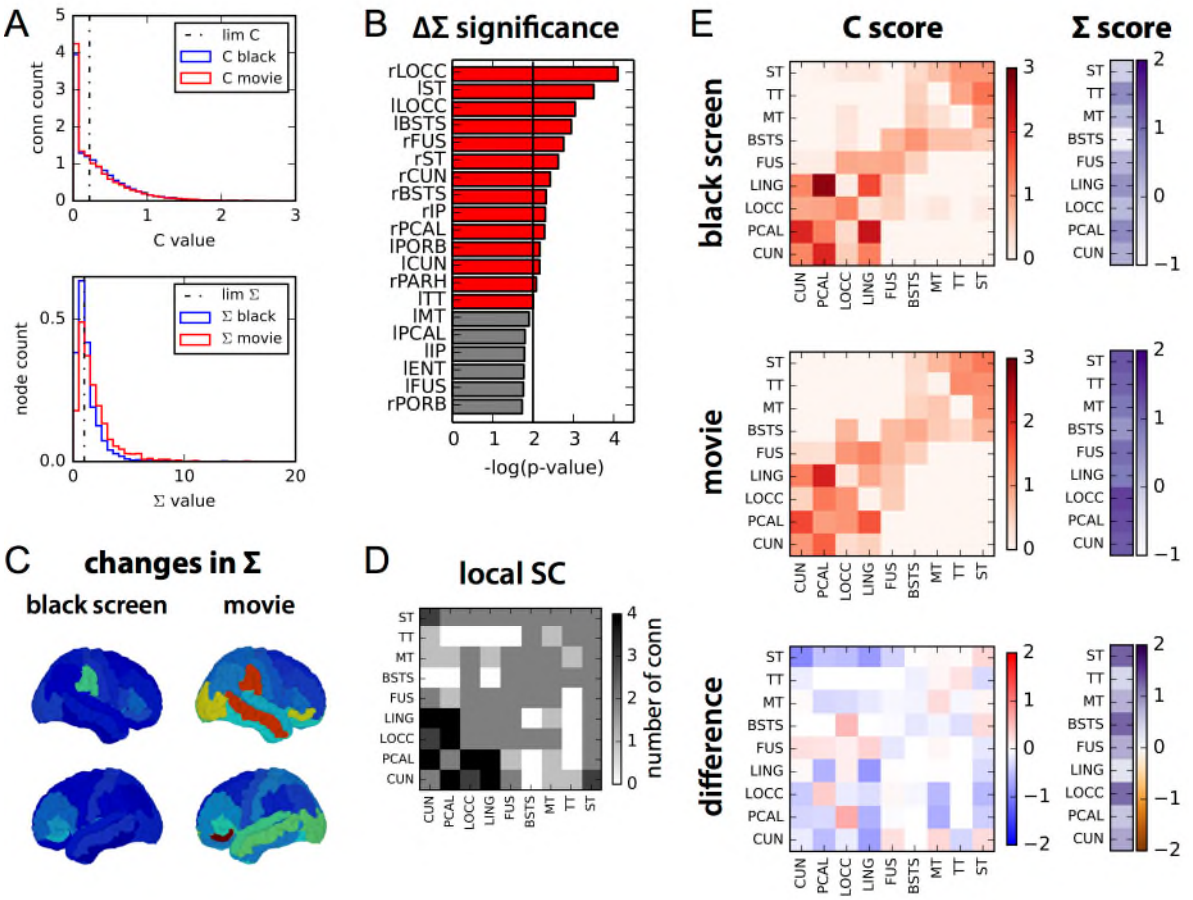

Fig. 5: Changes in the early visual and auditory pathways. A Histograms of C and Σ values in the two conditions. The limit (dashed-dotted line) is chosen to be the median of the distribution for the black screen. **B** Areas with most significant changes in Σ. **C** and score between black-screen and movie conditions in the visual and auditory system. **B** Same as A for significant changes in Σ. **C** Mean input variances Σ mapped on the cortical surface (left and right side views) for the two conditions. **D** Structural connectivity between 18 ROIs in the early visual and auditory pathways. Connections from the left and right hemispheres are grouped together. **E** Scores for the C and Σ values for the same ROIs as in D for each condition, as well as the score difference. The score evaluates the number of large weights across the subjects with respect to the whole distribution (ROIs x subjects): it is calculated for each C/Σ using Eq. 1 with the l_X corresponding to the dashed-dotted line in A (median of the distribution) and the mean/std in Eq. 1 are calculated for each matrix/vector element over the subjects.

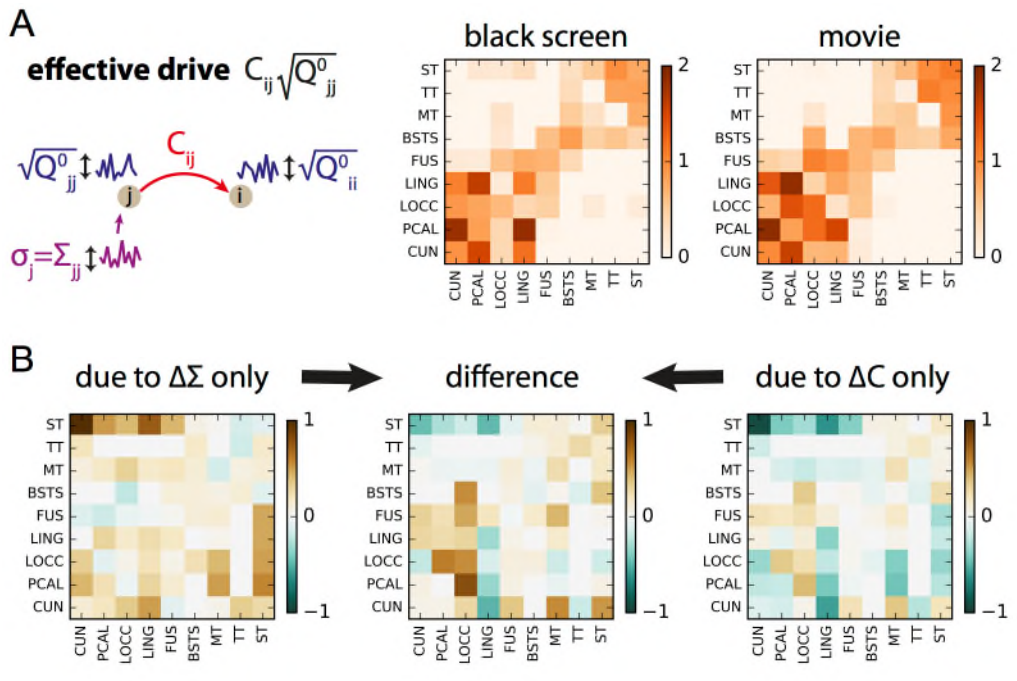

Fig. 6: Effective drive (ED) to quantify the information flow. A Schematic representation of the propagation of fluctuations from node *j* to node *i*, with input variance Σ_j . The effective drive is calculated like the *C* and Σ score, taking into account the variability over subjects (see Eq. 2). **B** Changes in effective drive for ROIs of the early visual and auditory system between black-screen and movie conditions (center), as well as contributions from ΔC (right) and $\Delta\Sigma$ (left).

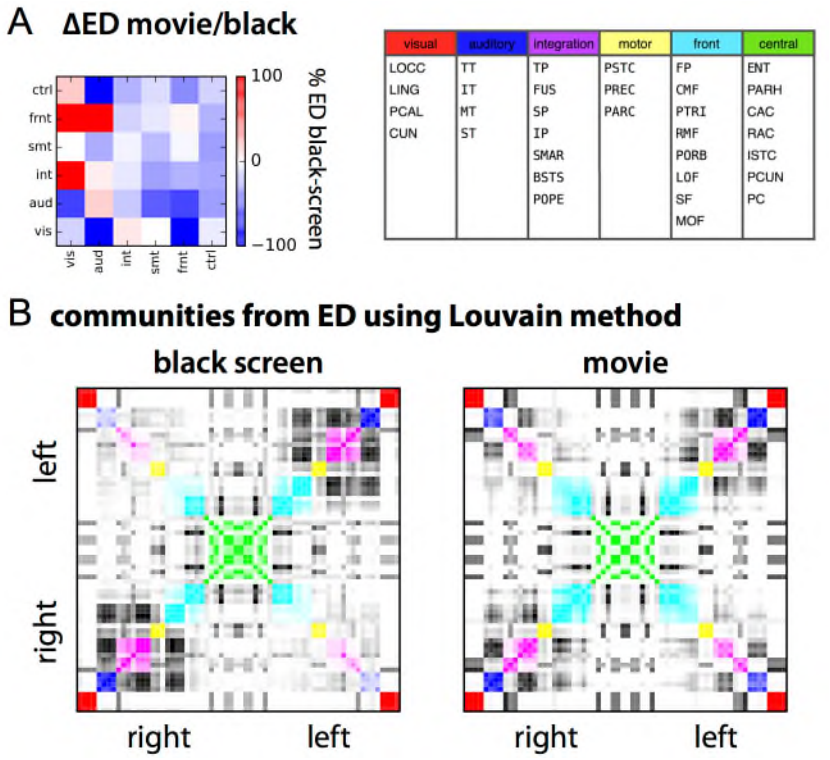

Fig. 7: A Changes in effective drive between ROIs pooled in 6 groups: visual in red, auditory in dark blue, integration in purple, motor in yellow, frontal in cyan and central in green. The change is calculated in % of the value for black screen. The lists of the group concern ROIs from both hemispheres. **B** Communities estimated using the Louvain method from the effective drive in Fig. 6 for each of the two conditions. The connections within the groups are displayed in color. ROIs are ordered to show the left and right hemispheres separately.