**Subject- and behavior-specific signatures extracted from fMRI data using whole-brain effective connectivity**

AUTHORS:

Pallares V\*1, Insabato A\*1, Sanjuán A1, Kühn S2,3, Mantini D4,5,6, Deco G\*\*1,7, Gilson M\*\* 1

\* Equal contribution

\*\* Equal contribution

1. Center for Brain and Cognition, Computational Neuroscience Group, Department of Information and Communication Technologies, Universitat Pompeu Fabra, Carrer de Ramon Trias Fargas, 25-27, Barcelona, 08005, Spain

2. Max Planck Institute for Human Development, Center for Lifespan Psychology, Lentzeallee 94, 14195 Berlin, Germany

3. University Clinic Hamburg-Eppendorf, Clinic and Policlinic for Psychiatry and Psychotherapy, Martinistraße 52, 20246 Hamburg, Germany

4. Research Center for Motor Control and Neuroplasticity, KU Leuven, 101 Tervuursevest, 3001 Leuven, Belgium

5. Department of Health Sciences and Technology, ETH Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

6. Department of Experimental Psychology, Oxford University, 15 Parks Road, Oxford OX1 3PH, United Kingdom

7. Institució Catalana de la Recerca i Estudis Avanats (ICREA), Universitat Pompeu Fabra, Passeig Lluı́s Companys 23, Barcelona, 08010, Spain

**Main text: ~4600 words (limit 4000)**

**References: 53 (limit 50)**

**INTRODUCTION**

Functional magnetic resonance imaging (fMRI) has been widely used to observe human brain activity with high spatial resolution via the blood-oxygen-level dependent (BOLD) signals. The context of the present study is about extracting (e.g., behavioral) information from fMRI data. From experiments with participants performing motor, sensory and cognitive tasks, brain areas have been mapped to functions [Cordes 2000]. However, even at rest, the brain exhibits patterns of correlated activity between neighboring and distant areas [Biswall 1995; Raichle 2001]. The functional connectivity (FC), which measures the statistical dependencies between the BOLD activities of brain regions, has then been studied for subjects performing tasks and compared with the resting state [Sala-Llonch 2012]. Recent interest has also focused on the temporal BOLD structure: using the frequency spectrum of individual regions [He 2011], measuring the cross-covariance lags (at the scale of seconds) between areas [Mitra 2015] and ‘dynamic FC’ to quantify the BOLD correlations at the scale of minutes [Gonzalez-Castillo 2017]. Following fundamental discoveries about brain functions, fMRI has been also increasingly used to complement clinical diagnostic for neuropathologies [Matthews 2016]. Resting-state fMRI has also been found to be informative about neuropsychiatric disorders [Greicius 2008]; for instance, alterations in FC correlate with the severity of Alzheimer's disease [Kurth 2015]; FC can also be used to predict clinical disease scores for individual patients [Rahim 2017].

Recent studies have focused on the reliability of these FC measures recorded from the same subject over successive sessions [Shehzad 2009; Mueller 2015; Chen 2015; Pannunzi 2017]. Consistent differences between subjects (with individual stability) allow subject identification using recorded FC as a “fingerprint” [Finn 2015]. Moreover, this subject specificity may even be enhanced in task-evoked activity [Finn 2017]. Another model-based approach used linear-regression coefficients of BOLD signals instead of FC to identify the subjects [Miranda-Dominguez 2014]. A recent prospective study about the evolution of psychiatric disorders emphasized individual traits (FC stabilization during childhood) irrespective of the condition [Kaufmann 2017], whereas traditional group-averaging aims to remove the individual differences to obtain task-specific [Xie 2017] or pathology-specific [Drysdale 2016] signatures. The mixture of session-to-session, subject-specific and condition-related variability in FC is a crucial issue for real-life applications where only a few sessions can be recorded for a single subject, such as clinical diagnostic. This may dramatically impair the he generalization capability of prediction methods to future (unseen) data. However, a strong limitation for previous studies [Finn 2015; Miranda-Dominguez 2014; Finn 2017] is the use of datasets with at most 3 resting-state sessions per subject. Our study aims to specifically address this point.

Distributed signatures in FC across the whole brain have been observed in memory tasks [Rissman 2012] or when the subject experiences psychological pain [Chang 2015]. Moreover, the etiology of many mental disorders is unknown: they are suspected to arise from network dysfunction, as reported for large-scale FC alterations in patients with schizophrenia [Hoptman 2012]. These examples strongly point in favor of whole-brain approaches to study high-level cognition [Deco 2011] and brain diseases [Deco 2014]; in contrast, focusing on a few cortical areas only to test hypotheses [Goebel 2003; Bastos-Leite 2015] may not capture sufficient information and network effects. However, such whole brain approaches typically involve a large number of parameters to estimate, which may impair the robustness. The aim of the present study is to bring a practical answer to this trade-off dilemma.

The idea underlying the study of FC – in the broad sense – lies in that it reflects how brain areas dynamically bind to exchange and process information [Fries 2005; Hipp 2011; Betti 2013]. To move beyond a phenomenological description of FC, our method relies on a model inversion to interpret FC [Gilson 2016]: changes in FC between movie viewing and rest are decomposed into changes in network connectivity and local fluctuating activity. Note that we borrow the terminology of effective connectivity (EC) to describe the interactions between brain regions from dynamic causal model (DCM) [Friston 2011], although our model implies simplifications compared to DCM. As with FC, a crucial issue for EC is whether the estimated model parameters are reliable across several sessions for the same subject [Frässle 2015], which determines whether they can predict the subjects' identities in practice [Miranda-Dominguez 2014].

The present study proposes a quantitative framework to extract multivariate signatures from fMRI data, discriminative against subjects or behavioral conditions (as a first step toward clinical conditions). The goal is double. First, we couple whole-brain EC estimation with adequate machine learning tools to identify subjects from resting state fMRI (i.e., classify single sessions to the corresponding participant) capitalizing on previous studies relying on FC [Finn 2015, Miranda-Dominguez 2014]. Here the focus is on the comparison between EC and FC in their generalization capabilities. To do so, we rely on specific datasets with large numbers of sessions and subjects to identify, performing a strict benchmark with respect to session-to-session variability. Second, we predict both the subject's identity and condition (rest versus movie viewing) to verify that EC can disentangle the two types of signatures. Meanwhile, we compare the areas to which belong the EC links supporting the twofold classification. Although we test the method on behavioral conditions, we benchmark it keeping in mind that our future goal is an application to clinical diagnostic.

**RESULTS**

In this study we used fMRI data from three datasets with healthy participants:

- The first dataset, from the Day2day study, was acquired at the Max Planck Institute for Human Development in Berlin for two different resting-state studies [Filevich et al., 2017]: 6 subjects were scanned 40-50 times with 5-minute sessions over six months (Dataset A1); 50 participants underwent one scanning session using the same MRI sequences (Dataset A2). This dataset was used to test the robustness of subject identification to session-to-session variability.

- Dataset B has been made publicly available by the Consortium for Reliability and Reproducibility (CoRR) [Zuo et al., 2014] and contains resting-state fMRI sessions from 30 participants. Each subject underwent 10 times a 10-minute scanning session every three days for one month. This dataset was used to test the generalization capability of the identification procedure for a larger number of subjects than Dataset A.

- Dataset C includes 22 subjects with 5 sessions of 10 minutes each in two different conditions, two sessions at rest and three sessions watching a movie [Gilson biorxiv]. This dataset was used to obtain both individualized and behavioral signatures, irrespective of the between-session variability.

After applying standard preprocessing pipeline to fMRI BOLD signal (see Methods for details), we parcellated the brain into 116 regions of interest (ROIs) by using a standard anatomical space [Tzourio-Mazoyer et al., 2002] for Datasets A1, A2 and B (see Figure 1A). Dataset C was parcellated into 66 ROIs covering the cortex [Hagmann et al., 2008].

**Functional and effective connectivity as measures of the brain network dynamics**

Classical functional connectivity (corrFC) was obtained for both Datasets A and B by calculating the pairwise Pearson correlation coefficient (PCC) between ROI time courses, obtaining a 116x116 symmetric matrix for each recorded session; see Eq. 2 in Methods. We used the whole-brain dynamic model [Gilson et al., 2016] illustrated in Figure 1B: each ROI is a node in a noise-diffusion network whose topology (skeleton) is determined by the structural connectivity (SC) obtained from diffusion tensor imaging (DTI) or similar techniques. In the model, the global pattern of FC arises from the local variability Σi that propagates via the network connections ECij (from j to i). To fit each fMRI session, all relevant ECij and Σi parameters are iteratively tuned such that the model spatio-temporal FC – as measured by FC0 (0-lag covariances) and FC1 (1-lag shifted covariances) – best reproduces the empirical counterpart. A detailed description of the model and the maximum-likelihood estimation procedure is provided in Methods. In essence, the model inversion decomposes the empirical matrices (FC0,FC1) into two estimates EC and Σ, which can be seen as multivariate biomarkers for the brain dynamics in each fMRI session.

**FIGURE 1**

**Similarity of connectivity measures across sessions and subjects**

Using Datasets A and B, we compared the capability of the 4 connectivity measures (corrFC, FC0, FC1 and EC), as well as Σ, in terms of within- and between-subject similarity (WSS and BSS, respectively), as a first step toward subject identification. For each pair of sessions, the similarity SX was calculated using the PCC between two z-scored vectorized connectivity measures X (non-zero elements for EC, low-triangle elements for corrFC; see Figure 1C, Eqs. 12 and 13 in Methods). In the matrix of SEC values for Dataset A1 (Figure 2A), 6 diagonal blocks with larger values corresponding to the WSS can be noticed. The remaining matrix elements correspond to BSS. Figure 2B compares the distributions of SEC and ScorrFC: WSS and BSS distributions are better separated for EC than for corrFC. In other words, sessions from the same subject are more similar to one another, and more different from those of other subjects, viewed from the EC than the corrFC viewpoint. This suggests a better capability of EC to discriminate between subjects. Note that the BSS from Datasets A1 (6 subjects) and A2 (50 subjects) remarkably overlap for both corrFC and EC, showing that BSS of 6 subjects generalizes well to larger number of subjects.

These qualitative observations are confirmed by the Kolmogorov-Smirnov (KS) distance between similarity distributions (blue versus red and blue versus green in Figure 2B). As summarized in Table 1, EC gave larger KS distance than corrFC and all other FC-related measures. Note that we also calculated KS distance using only the links in corrFC and FC0 corresponding to the 4056 existing connections in EC (determined by SC), in order to compensate for the (relative) sparsity of EC links as compared to corrFC and FC0; this did not change the results. Last, the diagonal elements of Σ showed the smallest distances. In the rest of the article we focus on connectivity measures (EC and corrFC). Supplementary Figures S1 and S2 show similarity distributions for FC0, FC1, corrFC/SC and Σ using Datasets A1 and B.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Dataset A1 | Dataset A1 & A2 |
| EC | EC weights estimated with the model | 0.6440 | 0.6581 |
| corrFC | FC computed as Pearson correlation | 0.4517 | 0.5477 |
| FC0 | FC computed as 0-lag covariance matrix | 0.3685 | 0.4888 |
| FC1 | FC computed as 1-lag covariance matrix | 0.4139 | 0.5118 |
| corrFC/SC | FC computed as Pearson correlation and masked with SC | 0.4769 | 0.5729 |
| FC0/SC | FC computed as 0-lag covariance matrix and masked with SC | 0.3770 | 0.4644 |
| Σ | Local variability for each ROI estimated with the model | 0.3778 | 0.3287 |

**Table 1.** Kolmogorov-Smirnov (KS) distance between WSS and BSS distributions for connectivity measures (corrFC, FC0, FC1) and model estimates (EC, Σ). The third column corresponds to the distance between the blue and red distributions in Figure 2B, and the fourth column to the blue and green distributions.

**FIGURE 2**

**Structure of individual session-to-session variability for EC and corrFC**

The high dimensionality of our connectivity measures may reduce their predictive power, which is known as Hughes phenomenon [Hughes 1968]. This is especially important in our case where the number p of dimensions for the multivariate measures (for Dataset A1 p=4056 for EC and p=6670 for corrFC) exceeds the number n of session samples (n=6 subjects x 50 sessions=300 samples), thereby constraining the data in a (n-1)-dimensional space. In order to further characterize individual variability over sessions, we performed a reduction of dimensionality of these multivariate measures using principal component analysis (PCA) on the sessions of Dataset A1.By the naked eye, the colored clouds representing all sessions for each subject exhibit smaller overlap for EC than corrFC in Figure 2C, where the space of the first 6 principal components (PCs) is represented in two three-dimensional panels.

We quantified the clustering degree of these clouds using a silhouette coefficient for each session (see Eq. 16 in Methods), ranging from -1 (poor clustering) to 1 (perfect clustering). As shown in Figure 2D, EC produced larger (almost all positive) silhouette values than corrFC, confirming the visual impression of Figure 2C. Silhouette coefficients were calculated on the data projected in the space of first 6 PCs, i.e. the number of PCs that maximized the silhouette coefficient (see Figure S3). As can be seen in Figure 2E, the silhouette coefficients for the data in the original link space (left violin plots) are smaller than those for data in the PC space (right), for both datasets (for dataset B we projected data on the first 27 PCs for which the silhouette is maximized, see Figure S3). PCA preprocessing to reduce the data dimensionality is thus expected to facilitate the identification of subjects. However an important issue is whether the variability captured by first PCs – which account for the largest part of the total variability – carries information about subjects' identity or is more related to session-to-session variability. If the latter is true classification performance might even be worsened by an erroneous choice of PCs.

**Subject identification using EC is more robust than using corrFC**

Next we turn to the main goal of our study: the classification of single sessions – attributing them to subjects – based on EC or corrFC. In classification algorithms the problem of overfitting describes the situation where the algorithm performs very well with the data it is trained with, but fails to generalize to new samples. Due to the high dimensionality of the connectivity measures [Hughes 1968], it is essential to control for overfitting with appropriate training and test data xxx test procedure, as explained below.

Subject identification from fMRI data was pioneered by recent publications [Miranda-Dominguez 2014; Finn 2015], which demonstrated that more than 100 subjects can be identified with high accuracy relying on corrFC. Although not explicitly mentioned in Finn et al. (2015) the identification procedure employed a k-nearest-neighbor (kNN) classifier with k=1 and PCC as metric (see Methods for details). In order to classify one session, called *target* in Finn et al. (2015), the PCC between the target and 1 known reference session for each subject *(*called *database*) is calculated; the predicted identity for the target will be that of the subject corresponding to the closest (most similar) session (see Figure S4). In contrast with these studies using 1NN inn 2015; Kaufmann 2017], our method relies on Multinomial Logistic Regression (MLR) classifier, a classical tool in machine learning. MLR uses a linear model to predict the probability that an input sample belongs to a class (subject here; for a technical comparison of both approaches, see Methods).

In addition, previous studies [Finn 2015; Finn 2017; Vanderwal 2017] used a small test set and found highly variable classification performance on this test set. This does not allow a reliable assessment of the classification performance, which is crucial for its future application to the clinical domain. Our train-test procedure and the use of large test-retest datasets allow a trustworthy characterization of the quality of the classifiers. Figure 3A shows our train-test procedure for the identification of subjects: 1) fMRI sessions (EC in the figure) are randomly split in training and test datasets; 2) after preprocessing (orange arrows) involving within-session z-score followed – or not – by PCA, the classifier is optimized as illustrated for the MLR with boundaries that best predict the training dataset; 3) test set is used to verify the generalization capability of the classifier (blue arrows), by measuring to which extent the classifier boundaries, estimated with the training set, correctly classify single sessions from the test set.

We first used Dataset A1 with 6 subjects and 40-50 sessions of 5 minutes per subject: we increased the number of training sessions per subject from 1 to 40 in order to evaluate how many sessions are necessary to satisfactorily identify the subjects. As shown in Figure 3B (left panels), EC (red) outperformed corrFC (blue) by more than one standard deviation, for both MLR and 1NN. Moreover, almost perfect classification was reached with MLR for 5 training sessions only, whereas 10-15 were necessary for kNN. The right panels in Figure 3B show the classification accuracy on Dataset B (10 minutes per session), used to test the robustness of the algorithms with respect to the number of subjects to be classified. We trained the algorithms with 1 session per subject and evaluated the classification performance varying the number of subjects from 2 to 30 (test set comprised the remaining 9 sessions per subject). Again, EC is more robust than corrFC: while performance with corrFC rapidly deteriorates as the number of subjects is increased, classification using EC is barely affected by the number of subjects. Also with this dataset MLR confirms its superiority to 1NN. These results show that EC and MLR allow better performance than corrFC and 1NN; importantly, the performance is especially improved between 1 and 10 training sessions per subject, which is the interesting range for (clinical) applications.

As mentioned above PCA might improve the identifiability of subjects by discarding superfluous dimensions (in particular when p>>n), thereby reducing the noise in the data. For Dataset A1, PCA only marginally increases the performance when only 1 session is used with MLR and EC (less than 1% of difference between means, see Figure S9). The performance increase is slightly larger (~3%) with 1NN and EC: 1NN benefited more from the data denoising because it achieved a lower accuracy compared to MLR. Two factors limit the usefulness of PCA: 1) when classification accuracy is already very high (~95% for MLR and EC with 1 session) PCA is unlike to make it better; 2) when the intrinsic noise of the measure is very high (FC has higher variability than EC, see Figure 2B) PCA is not able to reduce much the noise. PCA also allows for the investigation of the distribution of the subject-discriminative information between PCs. As shown in Supplementary Figures S6 and S7, the first PCs are not the most informative, whereas subsequent PCs exhibit a broad distribution of relevant information. This supports the use of proper machine learning tools to make the best use of this distributed information.

An important advantage of the MLR over kNN is its efficiency in characterizing the links that contribute to the classification. We used recursive feature elimination (RFE, described in details in Methods) to rank the links according to their weight in the classification and chose the lowest number of links that achieved the maximum classification performance. In comparison, the same procedure with kNN would require an incommensurate amount of computational power, recalculating the closest neighbor for all combinations of links (here the number of links is p >> 1000; see Methods for further detail). The resulting support network for dataset A1 had 18 links, to be compared with 44 links for dataset B. In both cases, subject identification using only those links achieved perfect accuracy (90% of all available sessions were used for training and 10% for testing, see Figure S8). The adjacency matrix of these two networks is shown in Figure 3C: remarkably, the networks are very sparse and non-uniformly distributed across the whole brain. This is the signature of the most subject-discriminative ROIs: frontal and cingulate cortices, as well as the temporal and occipital regions, seem to play a major role here. It is worth noting that the adjacency matrix is not symmetric, which implies different roles for nodes as receivers (especially frontal ROIs) or senders (cingulate).

The sparsity of the signature in Figure 3C hides the fact that the rankings for dataset A1 and B are to some extent aligned (PCC=0.59, p-value<<10-50), indicating that similar neural networks characterize individuals in two disjoint sets of subjects. To further measure the overlap between these networks, we selected for each dataset the subset of links with the highest ranking and computed the number of common links. Figure 3D shows that the proportion of common links exceeds by far its expectation under the hypothesis of random rankings (shaded gray area). This indicates a good agreement between the two datasets even at the single-link level.

**FIGURE 3**

**Joint classification of subject identity and behavioral condition**

Finally, we used Dataset C to extract from EC, one signature for the subject's identity and another for the behavioral condition. This is schematically depicted in Figure 4B, with two fictive dimensions: the dimension on the abscissa carries information about subject identity, while the dimension on the ordinate carries information about the condition. In this idealized 2D scenario, it is possible to classify a session with respect to both subjects and conditions using different dimensions of the data. In the high dimensional case, we expect that different sets of links will support the double classification. Using MLR and EC, we achieved very high performance (accuracy > 90%) for subject identification and perfect classification for the condition. As done before with Datasets A1 and B, we used RFE to rank the links according to their contribution to the classification. We computed the number of common links for the subject and condition identifications, which fell within the expected values with the null hypothesis (Figure 4C). This indicated that distinct subsets of links are relevant for the subjects' identity and behavioral condition.

We then sought the smallest subsets of links that achieve the maximum performance of each classification, as done before in Figure 4D (see also Methods for details). Both support networks were again very sparse and distributed across the brain, as can be seen in their adjacency matrix in Figure 4A. More links are necessary to identify the subjects (57) than the behavioral conditions (13), indicating a higher complexity for the former. Note that the quality of the prediction is poorer compared to the other datasets due to the smaller number of sessions available, which limits the size of the train set.

**FIGURE 4**

Similar to Datasets A1 and B, subject identification of Dataset C largely concerns the frontal and cingulate systems. Condition identification is also supported by occipital and temporal cortices, which are expected to have the strongest activity modulations during movie viewing. The top panels in Figures 5A and B represent the two support networks such that the directed nature of links can be easily appreciated. Apart from two small components, the subject network is almost fully connected with several central nodes (hubs, indicated by their large size) located in frontal and cingulate regions. On the contrary, the network for condition is segregated with small isolated components. The bottom plots in Figure 5 show the lateralization of the support links, stressing the asymmetries between the two hemispheres: most of the important links are ipsilateral (i.e. within the same hemisphere) and many belong to the left hemisphere for the subject network, whereas they are mainly contralateral for the condition network.

**FIGURE 5**

**DISCUSSION:**

In this study, we have proposed a framework to predict the identity of subjects as well as their condition from fMRI time series, by robustly extracting discriminative signatures about subject/condition differences. We obtain very sparse signatures, supposedly because of the datasets used (30 subjects maximum; 2 conditions). Their size is expected to increase with the complexity of the “environment” to represent (many subjects, many tasks); resources are becoming available to test this quantitatively [Zuo 2014 ; Gordon 2017]. Importantly, we have proven that such EC-based signatures are robust to the session-to-session variability, and can be obtained relying on a limited number of sessions (4-5 recordings of 5 minutes each). Proper machine-learning tools such as MLR are necessary to efficiently extract those signatures. We now discuss specific points.

The fundamental advancement of our study is the development of a reliable and well-benchmarked method, extending the previously published proofs of concept [Miranda- Dominguez 2014; Finn 2015]. EC discriminates individuals better than corrFC regardless of the number of sessions per subject used, for the two classifiers presented here, MLR and kNN (Figures 2 and 3). In particular, the generalization ability for EC is much more robust than FC when the classification becomes harder (few sessions per subject or many subjects to identify, see Figure 3B). This confirms that fMRI covariance structure – captured by the EC – reflects the identity of the subject [Miranda-Dominguez 2014], as previously shown for a task involving (or not) attention [He 2011] or for wake versus sleep [Mitra 2015]. Here the focus was on EC because it performed better than Σ estimates in subject identification, but it has been recently shown that Σ is strongly affected when engaging a task condition [Gilson biorxiv], so it could help to improve the classification for conditions, in particular involving sensory stimuli.

Our model is a continuous-time network with linear feedback and incorporates topological constraints from SC. EC corresponds to a maximum-likelihood estimate and can be very efficiently calculated for the whole brain with ~100 ROIs and each session with ~300 time points per ROI [Gilson 2016; Gilson biorxiv]. Our results show that, although the dynamic model and estimation procedure are a simplification compared to the sophisticated dynamic causal model (DCM) with hemodynamics and Bayesian machinery [Stephan et al. 2014; Friston 2014], it nonetheless provides powerful signatures that can be used for discrimination between subjects and conditions. Our study has focused on two coarse parcellations for the whole brain [Tzourio-Mazoyer 2002] and the cortex [Hagmann et al., 2008]. Although the two parcellations where applied to different datasets we didn't observe significant difference in the performance of the classifiers. Much work has been done recently to correct the bias due to the use of specific parcellations [Da Mota 2014]; for our purpose, more refined parcellations may entail better discriminability in higher-dimensional spaces, but raise issues for the EC estimation robustness. PCA, as a preprocessing step, was not found to significantly enhance the performance here. Nonetheless, PCA may be useful for datasets with larger number of subjects and conditions. In the end, the generalization capability is *the* criterion for the classification performance and further work is needed to define a suitable level of detail for applications with many subjects and conditions.

We have found that very few links (<4%) were sufficient to classify perfectly 30 subjects from Dataset B (Figure 3C) and both subjects and conditions in Dataset C (Figure 4D). For a larger cohort [Finn 2015] and more tasks [Zuo 2014], we expect this number to grow and the infra/supra-linear dependency with the subject number should be addressed carefully. Those support networks for the twofold classification (subject and condition) show several noticeable differences (Figure 5). The subject network is large, almost fully connected, distributed over the two hemispheres (with more links within the left one) and concentrated in the cingulate and frontal areas. This suggests subject-specific dynamics within areas involved in high-level functions and overlapping with the default mode network [Raichle 2001]. This interpretation of EC in terms of brain communication comes from the directed nature of EC, which considers the propagation of BOLD activity. It follows that the discriminative EC patterns may reflect heterogeneities in the interactions between the different neural subsystems (e.g., frontal to cingulate in Figure 5) and the propagation of information between them [Heeger 2002; Ekstrom 2010; Engel 2013]. We also found a much higher percentage of contralateral links for condition than subject. This is in line with strong inter-hemispheric interactions observed for the same dataset with community analysis [Gilson subm biorxiv]. As expected with the movie viewing condition studied here, links in the visual and temporal areas are discriminative.

To provide a direction toward clinical applications of neuroimaging data [Matthews and Hampshire 2016], an fMRI-based classification method requires disentangling signatures related to the subject and condition, while properly conditioning out the day-to-day variability (as uninformative intrinsic noise). The idea of personalized medicine to characterize brain disorders at the patient level is emerging [Yahata 2017], using neuroimaging techniques similarly to molecular and genetic approaches. The development of tailored therapeutic protocols [Shen 2014], aiming to optimize recovery and minimizing adverse secondary effects, requires quantitative tools that allow for a precise diagnostic of the patient's evolution. The generalization capability of prediction methods to future (unseen) data [Hughes 1968] is crucial in such clinical conditions. In particular, the interpretation of the condition signatures in terms of interactions between brain regions aims to bring a practical solution to the recent criticism that “a major reason for disappointing progress of psychiatric diagnostics and nosology is the lack of tests which enable mechanistic inference on disease processes within individual patients” [Stephan 2014].. The general scheme we have in mind for diagnostic is thus to follow a patient's trace over time in the (high-dimensional) EC space: as schematized in Figure 4B, the classification for condition would correspond to healthy opposed to various pathologies, from which one (or several) signature(s) have been extracted from resting-state [Greicius 2008] or task-evoked fMRI. To do so, we expect the proposed EC-based method to generalize and ignore the day-to-day variability to focus on important relevant signatures (better than FC-based methods). Importantly, the goal is not so much to discriminate between subjects as to prevent individual signatures from mingling with those for pathologies. We expect the latter to be much more complex [Chang et al., 2015] than the very sparse signature for movie viewing in Figure 4. The framework could be extended to the three-fold space (subject, pathology, task) as specific tasks may reveal powerful signatures for certain pathologies, such as memory exercises for Alzheimer [Kurth 2015]. The model can also be useful when pathologies are associated by alterations in SC, like stroke [**REF**].

**REFERENCES**

1. Bastos-Leite AJ, Ridgway GR, Silveira C, Norton A, Reis S, Friston KJ (2015) Dysconnectivity within the default mode in first-episode schizophrenia: a stochastic dynamic causal modeling study with functional magnetic resonance imaging. Schizophr Bull, 41: 144-153; doi: 10.1093/schbul/sbu080
2. 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
3. Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn Reson Med, 34:537-541; doi: 10.1002/mrm.1910340409
4. Chang LJ, Gianaros PJ, Manuck SB, Krishnan A, Wager TD (2015) A Sensitive and Specific Neural Signature for Picture-Induced Negative Affect. PLoS Biol, 13: e1002180; doi:10.1371/journal.pbio.1002180
5. Chen B, Xu T, Zhou C, Wang L, Yang N, Wang Z, Dong HM, Yang Z, Zang YF, Zuo XN, Weng XC (2015) Individual Variability and Test-Retest Reliability Revealed by Ten Repeated Resting-State Brain Scans over One Month. PLoS One, 10: e0144963. doi: 10.1371/journal.pone.0144963
6. 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
7. Da Mota B, Fritsch V, Varoquaux G, Banaschewski T, Barker GJ, Bokde ALW, Bromberg U, Conrod P, Gallinat J, Garavan H, Martinot J-L, Nees F, Paus T, Pausova Z, Rietschel M, Smolka MN, Ströhle A, Frouin V, Poline J-B, Thirion B (2014) Randomized parcellation based inference. NeuroImage 89: 203-215; http://dx.doi.org/10.1016/j.neuroimage.2013.11.012
8. 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; doi:10.1038/nrn2961
9. Deco G, Kringelbach ML (2014) Great expectations: using whole-brain computational connectomics for understanding neuropsychiatric disorders. Neuron, 84: 892-905; doi: 10.1016/j.neuron.2014.08.034
10. Ekstrom A (2010) How and when the fMRI BOLD signal relates to underlying neural activity: the danger in dissociation. Brain Res Rev, 62: 233-244. doi: 10.1016/j.brainresrev.2009.12.004
11. Engel AK, Gerloff C, Hilgetag CC, Nolte G (2013) Intrinsic coupling modes: multiscale interactions in ongoing brain activity. Neuron, 80: 867-886. doi: 10.1016/j.neuron.2013.09.038.
12. Filevich E, Lisofsky N, Becker M, Butler O, Lochstet M, Martensson J, Wenger E, Lindenberger U, Kühn S (2017) Day2day: investigating daily variability of magnetic resonance imaging measures over half a year. BMC Neurosci, 18: 65. doi: 10.1186/s12868-017-0383-y
13. Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, Papademetris X, Constable RT (2015) Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat Neurosc, 18: 1664-1671; doi: 10.1038/nn.4135
14. Finn ES, Scheinost D, Finn DM, Shen X, Papademetris X, Constable RT (2017) Can brain state be manipulated to emphasize individual differences in functional connectivity? Neuroimage, S1053-8119. doi: 10.1016/j.neuroimage.2017.03.064
15. Fries P (2005) A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. Trends Cogn Sci, 9: 474-480
16. Friston KJ (2011) Functional and effective connectivity: A review. Brain Connect, 1: 8; doi: 10.1089/brain.2011.0008
17. Friston KJ, Kahan J, Biswal B, Razi A (2014) A DCM for resting state fMRI, Neuroimage 94: 396-407
18. 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; doi: 10.1371/journal.pcbi.1004762
19. Gilson M, Deco G, Friston K, Hagmann P, Mantini D, Betti V, et al. (2017): Effective connectivity inferred from fMRI transition dynamics during movie viewing points to a balanced reconfiguration of cortical interactions. BioRxiv.110015
20. Goebel R, Roebroeck A, Kim D, Formisano E (2003) Investigating directed cortical interactions in time- resolved fMRI data using vector autoregressive modeling and Granger causality mapping. Magn Reson Imaging, 21: 1251-1261; doi: 10.1016/j.mri.2003.08.026
21. Gonzalez-Castillo J, Bandettini PA (2017) Task-based dynamic functional connectivity: Recent findings and open questions. Neuroimage. 2017 Aug 3. pii: S1053-8119(17)30653-5. doi: 10.1016/j.neuroimage.2017.08.006
22. Gordon EM, Laumann TO, Gilmore AW, Newbold DJ, Greene DJ, Berg JJ, Ortega M, Hoyt-Drazen C, Gratton C, Sun H, Hampton JM, Coalson RS, Nguyen AL, McDermott KB, Shimony JS, Snyder AZ, Schlaggar BL, Petersen SE, Nelson SM, Dosenbach NUF (2017) Precision Functional Mapping of Individual Human Brains. Neuron, xxx; doi: 10.1016/j.neuron.2017.07.011
23. Greicius MD (2008) Resting-state functional connectivity in neuropsychiatric disorders. Curr Opin Neurology, 21: 424
24. Hansen ECA, Battaglia D, Spiegler A, Deco G, Jirsa VK (2015) Functional connectivity dynamics: modeling the switching behavior of the resting state. Neuroimage, 105: 525-535; doi: 10.1016/j.neuroimage.2014.11.001
25. He BJ (2011) Scale-free properties of the functional magnetic resonance imaging signal during rest and task. J Neurosci 31 (2011) 13786–13795Honey CJ, Thivierge JP, Sporns O (2010) Can structure predict function in the human brain? Neuroimage 52: 766-76. doi: 10.1016/j.neuroimage.2010.01.071
26. Heeger D, Ress D (2002) What does fMRI tell us about neuronal activity? Nat Rev Neurosci. 3: 142-151; doi: 10.1038/nrn730
27. Hipp JF, Engel AK, Siegel M (2011) Oscillatory synchronization in large-scale cortical networks predicts perception. Neuron, 69: 387-396
28. Hughes GF (1968) On the mean accuracy of statistical pattern recognizers. IEEE Trans Inf Theory, 14: 55–63. doi: 10.1109/TIT.1968.1054102
29. Kurth S, Moyse E, Bahri MA, Salmon E, Bastin C (2015) Recognition of personally familiar faces and functional connectivity in Alzheimer's disease. Cortex, 67: 59-73. doi: 10.1016/j.cortex.2015.03.013
30. Hoptman MJ, Zuo X-N, D'Angelo D, Mauro CJ, Butler PD, Milham MP, Javitt DC (2012) Decreased interhemispheric coordination in schizophrenia: A resting state fMRI study. Schizophr Res, 141: 1-7. doi: 10.1016/j.schres.2012.07.027
31. Kaufmann T, Alnæs D, Doan NT, Brandt CL, Andreassen OA, Westlye LT (2017) Delayed stabilization and individualization in connectome development are related to psychiatric disorders. Nat Neurosci, 20: 513-515. doi: 10.1038/nn.4511
32. Matthews PM and Hampshire A (2016) Clinical Concepts Emerging from fMRI Functional Connectomics. Neuron, 91, 511-528; doi: 10.1016/j.neuron.2016.07.031
33. Messé A, Rudrauf D, Benali H, Marrelec G (2014) Relating Structure and Function in the Human Brain: Relative Contributions of Anatomy, Stationary Dynamics, and Non-stationarities. PLoS Comput Biol 10: e1003530
34. Miranda-Dominguez O, Mills BD, Carpenter SD, Grant KA, Kroenke CD, Nigg JT, Fair DA (2014) Connectotyping: model based fingerprinting of the functional connectome. PLoS One, 9: e111048. doi: 10.1371/journal.pone.0111048
35. Mitra A, Snyder AZ, Tagliazucchi E, Laufs H, Raichle ME (2015) Propagated infra-slow intrinsic brain activity reorganizes across wake and slow wave sleep. Elife, 4: e10781; doi: 10.7554/eLife.10781
36. Mueller S, Wang D, Fox MD, Pan R, Lu J, Li K, Sun W, Buckner RL, Liu H (2015) Reliability correction for functional connectivity: Theory and implementation. Hum Brain Mapp, 36: 4664-4680. doi: 10.1002/hbm.22947
37. Pannunzi M, Hindriks R, Bettinardi RG, Wenger E, Lisofsky N, Martensson J, Butler O, Filevich E, Becker M, Lochstet M, Kühn S, Deco G (2017) Resting-state fMRI correlations: From link-wise unreliability to whole brain stability. Neuroimage. 157: 250-262; doi: 10.1016/j.neuroimage.2017.06.006
38. Park HJ and Friston KJ (2013) Structural and Functional Brain Networks: From Connections to Cognition. Science, 342: 6158; doi: 10.1126/science.1238411
39. Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001) A default mode of brain function. Proc Nat Acad Sci, 98: 676-682; doi:10.1073/pnas.98.2.676.
40. Rahim M, Thirion B, Bzdok D, Buvat I, Varoquaux G (2017) Joint prediction of multiple scores captures better individual traits from brain images. NeuroImage, 158: 145-154; doi: 10.1016/j.neuroimage.2017.06.072.
41. Rissman J, Wagner AD (2012) Distributed representations in memory: insights from functional brain imaging. Annu Rev Psychol, 63: 101-128. doi: 10.1146/annurev-psych-120710-100344
42. Sala-Llonch R, Peña-Gómez C, Arenaza-Urquijo EM, Vidal-Piñeiro D, Bargalló N, Junqué C, Bartrés-Faz D (2012) Brain connectivity during resting state and subsequent working memory task predicts behavioural performance. Cortex, 48: 1187-1196; doi: 10.1016/j.cortex.2011.07.006
43. Sanz Leon P, Knock SA, Woodman MM, Domide L, Mersmann J, McIntosh AR, Jirsa V (2013) The Virtual Brain: a simulator of primate brain network dynamics. Front Neuroinform, 7:10; doi: 10.3389/fninf.2013.0001
44. Sato JR, Rondinoni C, Sturzbecher M, de Araujo DB, Amaro E Jr (2010) From EEG to BOLD: brain mapping and estimating transfer functions in simultaneous EEG-fMRI acquisitions. Neuroimage. 50: 1416-1426. doi: 10.1016/j.neuroimage.2010.01.075
45. Shen H (2014) Neuroscience: Tuning the brain. Nature, 507: 290-292
46. Shehzad Z, Kelly AM, Reiss PT, Gee DG, Gotimer K, Uddin LQ, Lee SH, Margulies DS, Roy AK, Biswal BB, Petkova E, Castellanos FX, Milham MP (2009) Cereb Cortex, 19: 2209-2229. doi: 10.1093/cercor/bhn256
47. Stephan K, Harrison L, Penny W, Friston K (2004) Biophysical models of fMRI responses, Curr Opin Neurol 14: 629-635
48. Stephan KE, Mathys C (2014) Computational approaches to psychiatry. Curr Opin Neurobiol, 25: 85-92; doi: 10.1016/j.conb.2013.12.007
49. Valdes-Sosa P, Roebroeck A, Daunizeau J, Friston K (2011) Effective connectivity: Influence, causality and biophysical modeling. Neuroimage. 58: 339-361. doi: 10.1016/j.neuroimage.2011.03.058
50. Vanderwal T, Eilbott J, Finn ES, Craddock RC, Turnbull A, Castellanos FX (2017) Individual differences in functional connectivity during naturalistic viewing conditions. Neuroimage (in press)
51. Xie H, Calhoun VD, Gonzalez-Castillo J, Damaraju E, Miller R, Bandettini PA, Mitra S (2017) Whole-brain connectivity dynamics reflect both task- specific and individual-specific modulation: A multitask study, Neuroimage; doi:10.1016/j.neuroimage.2017.05.050
52. Yahata N, Kasai K, Kawato M (2017) Computational neuroscience approach to biomarkers and treatments for mental disorders. Psychiatry Clin Neurosci, 71: 215-237. doi: 10.1111/pcn.12502
53. Zuo X-N, Anderson JS, Bellec P, Birn RM, Biswal BB, Blautzik J, …, Milham MP (2014) An open science resource for establishing reliability and reproducibility in functional connectomics. *Scientific Data*, *1*, 140049. http://doi.org/10.1038/sdata.2014.49