Synchrony, metastability, dynamic integration, and competition in the spontaneous functional connectivity of the human brain

Vincent Wens[a,b,*], Mathieu Bourguignon[a,c,d], Marc Vander Ghinst[a], Alison Mary[e], Brice Marty[a], Nicolas Coquelet[a], Gilles Naeije[a], Philippe Peigneux[f], Serge Goldman[a,b], Xavier De Tiège[a,b]

[a]*LCFC – Laboratoire de Cartographie fonctionnelle du Cerveau, UNI – ULB Neuroscience Institute, Université libre de Bruxelles (ULB), Brussels, Belgium*
[b]*Magnetoencephalography Unit, Department of Functional Neuroimaging, Service of Nuclear Medicine, CUB – Hôpital Erasme, Brussels, Belgium*
[c]*Laboratoire Cognition Langage et Développement, UNI – ULB Neuroscience Institute, Université libre de Bruxelles (ULB), Brussels, Belgium*
[d]*BCBL – Basque Center on Cognition, Brain and Language, 20009 San Sebastian, Spain*
[e]*UNICAEN, PSL Research University, EPHE, INSERM, U1077, CHU de Caen, Neuropsychologie et Imagerie de la Mémoire Humaine, Caen, France*
[f]*UR2NF – Neuropsychology and Functional Neuroimaging Research Unit at CRCN – Centre de Recherches Cognition et Neurosciences, and UNI – ULB Neuroscience Institute, Université libre de Bruxelles (ULB), Brussels, Belgium*

Abstract

The human brain is functionally organized into large-scale neural networks that are dynamically interconnected. Multiple short-lived states of resting-state functional connectivity (rsFC) identified transiently synchronized networks and cross-network integration. However, little is known about the way brain couplings covary as rsFC states wax and wane. In this magnetoencephalography study, we explore the synchronization structure among the spontaneous interactions of well-known resting-state networks (RSNs). To do so, we extracted modes of dynamic coupling that reflect rsFC synchrony and analyzed their spatio-temporal features. These modes identified transient, sporadic rsFC changes characterized by the widespread integration of RSNs across the brain, most prominently in the β band. This is in line with the metastable rsFC state model of resting-state dynamics, wherein our modes fit as state transition processes. Furthermore, the default-mode network (DMN) stood out as being structured into competitive cross-network couplings with widespread DMN-RSN interactions, especially among the β-band modes. These results substantiate the theory that the DMN is a core network enabling dynamic global brain integration in the β band.

Keywords:
Connectivity state transition; Dynamic functional connectivity; Independent component analysis; Magnetoencephalography; Network mixture model; Resting state

Highlights:
— We study synchronization among dynamic functional connectivity at rest.
— We identify transient modes of dynamic coupling using an ICA of connectivity.
— Some α-band modes are longer lived and β-band modes are more widespread.
— Modes of the DMN are structured into competitive cross-network couplings.
— Results concur with a metastable dynamics and the core network model of the DMN.

1. Introduction

The large-scale organization of the human brain is based on the existence of functional networks, which reflect the preferential integration of distant neural assemblies needed to support various functions, from sensory perception and motor behaviors to complex cognitive processes. At the macroscopic level, spatial patterns of interaction between segregated brain areas can be disclosed during the performance of dedicated tasks, but also in the absence of any explicit task (i.e., the so-called *resting state*). These background couplings are thought to represent a fingerprint of this organization into functional networks and to reflect the intrinsic functional architecture of the human brain (for a review, see, e.g., Deco and Corbetta, 2011).

One method to map functional brain networks from experimental measurements is functional connectivity (FC) analysis, where the coupling among network nodes is estimated via measures of statistical similarity between their activities (Bastos and Schoffelen, 2015; Friston, 2011; O'Neill et al., 2015a). This approach generally assumes that networks are temporally stable. Multiple static interaction patterns—conventionally referred to as *resting-state networks* (RSNs)—have been identified in this way, some overlapping primary systems such as the sensorimotor, the auditory, and the visual networks and others involving higher-level systems such as the attentional, the executive-control, and the default-mode networks. These observations have been consistently reproduced across neuroimaging studies and modalities, from functional magnetic resonance

Corresponding author. Address: Magnetoencephalography Unit, Department of Functional Neuroimaging, Service of Nuclear Medicine, CUB – Hôpital Erasme, 808 route de Lennik, 1070 Brussels, Belgium. *Telephone:* +3225554884. *E-mail address:* vwens@ulb.ac.be.

imaging (fMRI) (Beckmann et al., 2005; Biswal et al., 1995; Damoiseaux et al., 2006; Fox et al., 2005; Smith et al., 2009) to magnetoencephalography (MEG) (Brookes et al., 2011, 2012a,b; Hall et al., 2013; Hipp et al., 2012; Liu et al., 2010; Luckhoo et al., 2012; Wens et al., 2014a,b, 2015) and electroencephalography (EEG) (Knyazev et al., 2016, 2017; Liu et al., 2017; Siems et al., 2016; Sockeel et al., 2016).

However, this static picture of RSNs—i.e., fixed networks without time-dependent coupling modulations within and across them—hardly mirrors the functional landscape of the human brain. Rather, it has been hypothesized that RSNs represent an average of temporally varying interaction patterns, which would transiently fragment into sub-networks to bind different systems together, hence bringing about a richer repertoire of functional integration (Deco et al., 2011). This hypothesis was supported by dynamic FC studies, which disclosed a large body of time-dependent interaction patterns fluctuating over relatively short timescales (from seconds to tens of seconds), including transient cross-network couplings (Allen et al., 2014; Brookes et al., 2014; Chang and Glover, 2010; de Pasquale et al., 2010, 2012; Handwerker et al., 2012; Hutchison et al., 2012; Kiviniemi et al., 2011; O'Neill et al., 2015b, 2017b; Zalesky et al., 2014). Still, many aspects of dynamic FC remain to be investigated (for reviews, see Hutchison et al., 2013; O'Neill et al., 2017a).

In this work, we explore the existence and the properties of spontaneous temporal synchrony among dynamic brain couplings and consider the hypothesis that coupling synchronization is a correlate of the generation of sub-networks and cross-network integration. We started with resting-state FC measured via MEG power envelope correlation from key nodes of well-established RSNs, and submitted the time-dependent FC data to an independent component analysis (ICA) so as to disclose "*modes of dynamic coupling*" that reflect patterns of synchronous fluctuation among brain interactions. This approach has already been used to disclose task-related transiently synchronized networks (O'Neill et al., 2017b). An extra challenge here was to identify physiologically relevant modes of dynamic coupling in the absence of any controlled, goal-directed task. To do this and test our main hypothesis, we sought to relate each mode to a combination of template RSNs (henceforth referred to as "*network mixture models*"). This would establish an association between coupling synchrony and the phenomenon of cross-RSN integration. The framework of network mixture modeling also allows to investigate the key hypothesis that resting-state activity emerges from a spontaneous switching among diverse network configurations, which can be interpreted effectively as a "*dynamical competition*" among RSNs (Deco and Corbetta, 2011). We sought to identify competitive modes of dynamic coupling by comparing their mixture model to a similar model explicitly constrained to display no competition. Establishing their existence would provide empirical support to the theory of Deco and Corbetta (2011).

2. Material and Methods

2.1. Data acquisition

The dataset used in this work consists of MEG resting-state recordings of 100 healthy adult volunteers (48 females and 52 males, mean age: 26.5 years, age range: 18–41 years) gathered from nine experiments containing a rest session (5 minutes, eyes open) intermingled with task-driven sessions. Results related to task-positive phenomena have been published previously (Bourguignon et al., 2011, 2013; Clumeck et al., 2014; Marty et al., 2015; Mary et al., 2015; Vander Ghinst et al., 2016). Participants were all right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971), had no history of neurologic or psychiatric disease, and signed a written informed consent prior to data acquisition. All nine studies were approved by the CUB – Hôpital Erasme Ethics Committee.

Neuromagnetic activity at rest was acquired using a 306-channel whole-scalp-covering MEG system (Vectorview, Elekta Oy, Helsinki, Finland) placed in a lightweight magnetically shielded room (Maxshield™, Elekta Oy, Helsinki, Finland). Signals were band-pass filtered at 0.1–330 Hz and sampled at 1 kHz. Subjects were sitting comfortably in the MEG armchair with the head inside the MEG helmet, and were asked to relax and fixate the gaze at a point on the wall or on a screen. Their head position was tracked with four indicator coils. An electromagnetic digitalization system (Fastrack, Polhemus, Colchester, VT, USA) was used before MEG data acquisition to locate these coils relative to anatomical fiducials, as well as at least 150 head-surface points. A high-resolution 3D T1-weighted cerebral magnetic resonance image (MRI) of each subject was also acquired using a 1.5 T MRI scanner (Intera, Philips, The Netherlands) after the MEG recordings.

2.2. Data preprocessing

Resting-state MEG data were preprocessed offline for noise reduction and extraction of band-limited activity. The temporal extension of signal space separation was first applied using the Maxfilter software (Maxfilter™, Elekta Oy, Helsinki, Finland; version 2.2 with default parameters) in order to suppress external magnetic interferences and correct for head movements (Taulu et al., 2005). Remaining cardiac, ocular, and system artifacts were then removed via an ICA (FastICA algorithm with dimension reduction to 30 and nonlinearity *tanh*, see Hyvärinen and Oja, 2000) applied to band-pass filtered (0.5–45 Hz) MEG signals. Components corresponding to these artifacts were visually selected (number of identified components per subject: 5.0 ± 1.4, mean ± SD) and projected out of the full-rank data (Vigario et al., 2000). Of note, this ICA decomposition did not isolate high-frequency muscle artifacts, but their contribution below 45 Hz was subdominant. The cleaned MEG data were finally filtered in four frequency bands (θ: 4–8 Hz, α: 8–12 Hz, β: 12–21 Hz, β': 21–30 Hz) and Hilbert transformed to obtain their analytic signals.

Individual MRIs were also preprocessed in order to build the MEG forward model needed for source reconstruction. First, the coordinate system associated with MEG was coregistered

manually to that of MRI using the digitized fiducials for initial estimation and the head-surface points for manual refinements. The MRI was then segmented using the FreeSurfer image analysis suite (Fischl, 2012). Sources in three orthogonal directions were also placed at each node of a cortically-constrained grid (inter-sources distance: 5 mm, total number of nodes: 13229). The grid was built on the Montreal Neurological Institute (MNI) template MRI and mapped onto each individual MRI via a non-linear spatial deformation algorithm implemented in the SPM8 toolbox (Friston et al., 2007). Finally, the MEG forward model was computed at these sources using the one-layer boundary element method of the MNE-C software suite (Gramfort et al., 2014).

2.3. Sliding-window connectivity estimation

We evaluated FC between the main nodes of RSNs and the rest of the cortex using sources envelope correlation, which was chosen here for its established ability to uncover all RSNs typically disclosed with fMRI (de Pasquale et al., 2010; Brookes et al., 2011; Hipp et al., 2012; Wens et al., 2014b). The FC analysis used here relied on minimum norm estimation (MNE) for source reconstruction and the geometric correction scheme for the suppression of spatial leakage effects, which yield spurious contributions to both static and dynamic FC (Wens, 2015). The pipeline has been described in Wens et al. (2015), to which we refer for further details. The main difference is that FC was estimated here within short time windows sliding across the recording (i.e., dynamic FC) rather than within a single window covering the entire recording (i.e., static FC). Specifically, the correlation between source envelopes (low-pass filtered at 2 Hz to improve the connectivity-to-noise ratio; see de Pasquale et al., 2010; Hipp et al., 2012) was estimated within 10 s-long windows (de Pasquale et al., 2010) sliding with a step of 5 s, leading to 57 windows per subject. Note that leakage-corrected dynamic FC was computed using one-dimensional projected sources (as in O'Neill et al., 2015b, 2017b) rather than the Euclidean norm of three-dimensional sources (as in Betti et al., 2018).

The seed sources used for the computation of seed-based dynamic FC were chosen as key nodes of six well-known RSNs: the default-mode (DMN), the sensorimotor (SMN), the auditory (AN), the visual (VN), and the left and right fronto-parietal (ʟ/ʀFPN) networks (see Table 1). The resulting RSN-specific sets of seed-based envelope correlation maps fluctuating from window to window represent the time development of couplings between each RSN node and the rest of the cortex.

2.4. Identification of modes of dynamic coupling

To detect modes of dynamic coupling that reflect synchronous fluctuations in the dynamic FC of each RSN, we assumed these modes to be mutually temporally independent (asynchronous) and thus estimated them using group ICA decomposition, as in O'Neill et al. (2017b). The method is illustrated in the top part of Fig. 1. Further background is provided in the Supplementary Methods S1 to S3.

To design a group-level analysis, we first standardized individual FC time series to zero mean (to ensure that the sought

seed RSN	key nodes	MNI coordinates (mm)
DMN	PCC	$-3, -54, 31$
	MPFC	$-2, 51, 2$
	ʟTPJ	$-43, -76, 35$
	ʀTPJ	$51, -64, 32$
SMN	ʟSM	$-42, -26, 54$
	ʀSM	$38, -32, 48$
AN	ʟA	$-54, -22, 10$
	ʀA	$52, -24, 12$
VN	ʟV	$-20, -86, 18$
	ʀV	$16, -80, 26$
LFPN	ʟFEF	$-26, -12, 53$
	ʟIPS	$-25, -67, 48$
RFPN	ʀFEF	$30, -13, 53$
	ʀIPS	$23, -69, 49$

Table 1: *List of seed locations used in dynamic FC mapping.* Coordinates were taken from de Pasquale et al. (2012) and Hipp et al. (2012). PCC: posterior cingulate cortex; MPFC: mesio-prefrontal cortex; TPJ: temporo-parietal junction; SM: sensorimotor; A: auditory; V: visual; FEF: frontal eye field; IPS: intra-parietal sulcus; ʟ: left; ʀ: right.

dynamical effects are not confounded with inter-individual variability in static FC) and unit variance (to ensure that each individual subject is given a similar weight) and concatenated them temporally across subjects. We also concatenated spatially the seed-based maps associated with the nodes of the RSN under consideration (see Table 1) in order to consider all its interactions at once. The resulting FC dataset was then decomposed into temporally independent components (ICs), each IC thus representing one mode of dynamic coupling. Parameters of the ICA were similar to those used in static FC analyses of MEG resting-state data (Brookes et al., 2011; Wens et al., 2014b). The dimensionality was first reduced to 50, which removed approximately 50% of the total variance in all cases. Of note, a similar proportion of variance was discarded as well in the previous MEG envelope ICA that successfully disclosed RSNs (see Supplementary Results S1). The ICA itself was performed using the FastICA algorithm with nonlinearity *tanh* and the number of ICs to compute was set to 20 (Hyvärinen and Oja, 2000).

We next proceeded with a detailed characterization of the modes of dynamic coupling. Each mode was associated with one IC time series and multiple IC spatial maps (one per seed belonging to the RSN under consideration). The maps were obtained by temporal correlation between FC and IC time series (see the raw IC maps in Fig. 1). We then applied various statistical analyses to

(i) identify for each mode the spatial location and the temporal characteristics of the couplings involved,

(ii) decompose its IC maps into combinations of template RSNs using a network mixture model, and

(iii) determine whether the mode is indicative of a "competitive" behavior between RSNs.
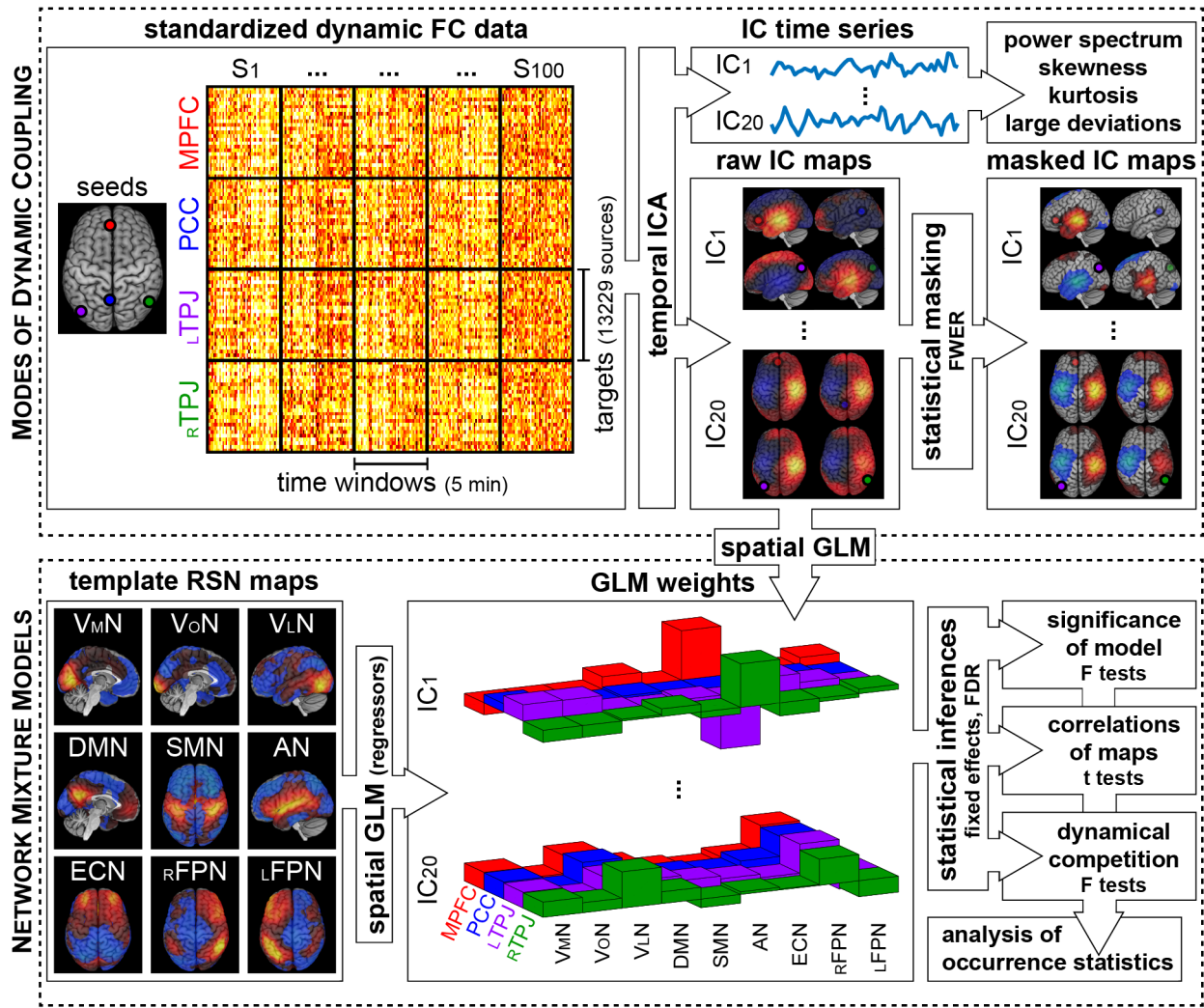
Figure 1: *Schematic illustration of the analysis pipeline.* (Example for the dynamic integration of the DMN) **Top: Identifying modes of dynamic coupling.** The standardized seed-based FC time series are concatenated temporally across subjects ($S_1, S_2, \ldots, S_{100}$) and spatially across seeds (indicated by colored discs in the left insert) and then submitted to a temporal ICA. This outputs several IC time series, whose power spectrum, higher-order temporal statistics, and large-deviation events are then analyzed, as well as associated IC maps (one per seed), which are masked statistically. **Bottom: Network mixture modeling.** Nine template RSN maps are used as regressors for a spatial GLM applied to each raw (unmasked) IC map. This results into a set of GLM weights defining a mixture model for each mode of dynamic coupling. Statistical inferences are then derived to establish model significance, univariate maps correlations, and dynamical competition. The significance of detection rates across modes of dynamic coupling is further quantified using occurrence statistics.

2.5. Spatial mapping and temporal properties

To locate the couplings involved in each mode, we masked all IC maps statistically using a parametric correlation test based on the null hypothesis that Fisher-transformed correlation values follow a normal distribution with mean zero and SD $1/\sqrt{\nu - 3}$. Here, the number of temporal degrees of freedom ν was estimated as the total number of time windows across subjects divided by two to take into account windows overlap. The significance level was set to $p < 0.05$ with the family-wise error rate (FWER) controlled via Bonferroni correction for the S seeds (see Table 1), the twenty ICs, and the effective number (here, $\rho = 46$) of independent cortical sources estimated as the rank of the MEG forward model (see Wens et al., 2015). Both positive and negative correlations were considered (two-tailed test), leading to significance thresholds of ± 0.081 (when $S = 4$) and ± 0.078 (when $S = 2$) used to mask the IC maps (see the masked IC maps in Fig. 1). The volume fraction (across the S maps) of these statistical masks was used to assess the cortical coverage of those couplings identified for each mode. Note that our simple approach might *a priori* overestimate the parameter ν (because of temporal autocorrelations in FC series) and lead to lenient statistical masking. However, this was not the case, as estimates based on Fourier-phase surrogates, which preserve temporal autocorrelation, led to lower statistical thresholds (about ± 0.06).

We also examined some temporal characteristics of the IC time series (see Fig. 1). To detect possible oscillatory dynamics, their power spectral density was estimated by applying the discrete Fourier transform to the single-subject parts of these time series and group averaging of the resulting magnitude-squared Fourier coefficients. We also considered higher-order statistics of IC time series, i.e., their skewness and kurtosis. Significance was assessed using a parametric test based on the null, approximately normal distribution obtained in the case of Gaussian signals. The null mean and SD were respectively zero and $\sqrt{6/\nu}$ for the skewness, and 3 and $\sqrt{24/\nu}$ for the kurtosis. The significance level was set to $p < 0.05$ with Bonferroni correction for the twenty ICs. We tested for both positive and negative skewness (two-tailed test, threshold: ± 0.14) and for the excess kurtosis (i.e., one-tailed test above 3, threshold: 3.26).

It is useful to explain how the skewness and the kurtosis inform the interpretation of the modes of dynamic coupling. The IC skewness elucidates the meaning of maps sign. Indeed, positivity and negativity have no absolute meaning because of the sign ambiguity inherent to ICA (Hyvärinen and Oja, 2000) but they do have a relative meaning among the four maps and the skewness. Within a given mode, two connections (i.e., two seed-target pairs) with the same sign in the corresponding IC maps represent two couplings that fluctuate in synchrony. If they have opposite sign, the two couplings are anticorrelated, which hints at a "dynamical competition" in this mode. A positive (negative) skewness indicates a tendency of the mode to increase connections with positive (negative) map values and concomitantly decrease those with opposite sign, whereas a vanishing skewness indicates the absence of such a preferential direction in FC changes.

The IC excess kurtosis estimates to what extent these FC modulations arise in sharp peaks but does not inform on the temporal unfolding of these peaks *per se*. To quantify their transient character, we used the mean lifetime of large-deviation events in IC time series. The identification of those large deviations driving the IC kurtosis was based on a threshold defined heuristically as follows. We increased candidate thresholds from 1 by step of 0.5 (recall here that IC time series are standardized to zero mean and unit SD), computed for each mode, RSN, and frequency band the probability of supra-threshold events within IC time series, and retained the smallest threshold for which there was a significantly positive ($p < 0.05$) Pearson correlation (across the $20 \times 9 \times 4$ ICs) between this tail probability and the kurtosis. This approach led to a threshold of 2.5. To assess the recurrent character of IC large deviations and their tendency to happen in isolation (notwithstanding the temporal independence constraint of ICA that penalizes simultaneous events, see also Supplementary Methods S1), we also estimated the distribution of (co-)occurrences, i.e., the number of events identified in each time window.

The effects of the frequency band ($\theta, \alpha, \beta, \beta'$) and seed RSN (Table 1) on the distributions across modes of the IC kurtosis, the IC events lifetime, and the volume fraction of IC masks were assessed using a non-parametric two-way ANOVA (Friedman test) and *post-hoc* Wilcoxon rank tests.

2.6. Network mixture modeling

To identify modes of dynamic coupling that involve cross-network integration between RSNs, we sought to establish a relationship between each raw IC map and a mixture of one or more static RSN maps.

Template RSN maps were derived from ten standard RSNs identified via a twenty-component ICA of fMRI resting-state data available at https://www.fmrib.ox.ac.uk/datasets/brainmap+rsns (Smith et al., 2009). These fMRI maps were smoothed with a Gaussian kernel of 8 mm full-width-at-half-maximum using SPM8 (Friston et al., 2007) and resampled to the MEG cortical grid to obtain same size and comparable spatial smoothness across the two modalities. Of notice, the cerebellar RSN of Smith et al. (2009) was discarded since it does not involve cortical areas, resulting in nine template RSN maps: the visuo-medial (VmN), visuo-occipital (VoN), and visuo-lateral (VlN) networks, the DMN, the SMN, the AN, the executive-control network (ECN), and the r/lFPN (see the template RSN maps in Fig. 1).

We then used general linear modeling (GLM) with these template RSN maps as regressors (bottom part of Fig. 1). This analysis differs from the standard use of GLM in neuroimaging (typically established voxel by voxel using sampling over subjects or time, see, e.g., Friston et al., 1994) by the fact that sampling is here over the cortical sources. Estimation of the GLM weights followed a standard procedure without regularization since the resulting design matrix was full rank (see Supplementary Methods S4). Statistical inference was based on our assessment of the effective number of spatial degrees of freedom in MNE source maps, i.e., $\rho = 46$. This setup takes into account the autocorrelation structure due to spatial leakage (Wens

et al., 2015) and thus avoids wrongly inflating significance (see, e.g., Monti, 2011, for a review of the consequences of sample autocorrelation for GLMs in the context of fMRI time series).

The network mixture model of a mode of dynamic coupling was then defined as the set of GLM weights associated with the S seed nodes and the nine template RSNs (see the GLM weights in Fig. 1). For each mode and seed, model significance was established using F tests (Friston et al., 2007) at $p < 0.05$ with false discovery rate (FDR) correction to control for multiple comparisons (i.e., 20 modes $\times$ S seeds). Modes associated with significant mixture models for at least one seed were interpreted as indicative of a cross-network interaction. To obtain a clearer picture of which RSNs are involved in significant mixtures, we also investigated *post hoc* the univariate (partial) correlations between each raw IC map and each template RSN map via two-tailed t tests (Friston et al., 2007) at $p < 0.05$ corrected for the same amount of multiple comparisons as done for the F tests. All statistical inferences were based on a fixed-effect design since the regressors were built upon deterministic template RSN maps.

Full details on this spatial GLM and the associated statistical tests are provided in the Supplementary Methods S4.

2.7. Dynamical competition testing

We also used the framework of network mixture modeling to investigate the idea that different RSNs "compete dynamically" to establish functional connections. Based on our discussion above about the relative sign of IC maps, we reasoned that such competitive behavior must involve substantial temporal anticorrelations among the cross-network couplings identified in a given mode of dynamic coupling. This situation is reflected by the presence of both positive and negative weights in a mixture model. To investigate formally the existence of dynamical competition in this sense, we thus compared statistically each network mixture model to a "non-competitive" model built under the constraint that all GLM weights have the same sign (i.e., weights are all positive or all negative), which idealizes the absence of competition. Importantly, imposing the same sign across the seeds considered (Table 1) was necessary to include the effect of relative sign across the IC maps and detect at once any instance of competition (i.e., among RSNs for fixed seed, among seeds for fixed template RSN, or between different seed-RSN pairs). Further background is provided in the Supplementary Methods S5 and S6.

We derived the exact solution of these sign-constrained GLMs using a semi-analytical approach detailed in the Supplementary Methods S7. Our technique combines (i) the observation that the sign-constrained problem corresponds to an ordinary, analytically solvable GLM under certain (but *a priori* unknown) zero-weight constraints, with (ii) a numerical approximation of the solution via a projected gradient descent algorithm (whose convergence can be analyzed using fixed-point methods; see, e.g., Jung, 2017) to determine the zero-weight constraints.

For each mode and seed, we then assessed statistically whether the original (unconstrained) GLM provided a significantly better model than the sign-constrained GLM. The model

comparison was performed using F tests (at $p < 0.05$ corrected for the same amount of multiple comparisons as done for the model assessment F tests) since the two GLMs are nested (Friston et al., 2007). Modes associated with significantly better unconstrained GLM for at least one seed were deemed to exhibit significant RSN competition. The only exception to this claim was when the unconstrained GLM did satisfy the sign condition since the associated F statistic was then ill-defined. In this case, however, the mode was clearly non-competitive.

The technical details are developed fully in the Supplementary Methods S4, S7, and S8.

2.8. Occurrence of cross-network integration and competition

We quantified the tendency of modes to be structured into cross-network interactions, to target specific RSNs, and to exhibit dynamical competition. The occurrence rate for cross-network coupling was estimated as the proportion of modes with a significant network mixture model, that for RSN detection, as the proportion of significant mixture models with a significant correlation t test with the corresponding template RSN for at least one seed, and that for dynamical competition, as the proportion of significant mixture models with significant model comparison F test for at least one seed. Significance testing for these proportions was based on the binomial distribution associated with the expected false positive rate (i.e., the FDR-corrected critical p value) of these tests. We report one-tailed significance at $p < 0.05$ with Bonferroni correction, but also consider uncorrected significance in cases of small sample sizes where the sensitivity of these statistics is limited (see, e.g., Krzywinski and Altman, 2014).

We further applied Friedman and Wilcoxon tests to identify possible effects of frequency band or seed RSN on the proportion of cross-network coupling occurrence.

3. Results

To examine temporal synchrony among the cortical interactions of each RSN (Table 1), we used an ICA that decomposed the band-limited FC data into twenty modes of dynamic coupling. Each mode was associated with a number of IC maps that locate synchronously varying couplings with the seed nodes of the RSN under consideration (Table 1) and one IC time series encoding their temporal dynamics. Their spatial and temporal signatures were then analyzed statistically as outlined in Fig. 1.

To ease the understanding and interpretation of the quantitative results, we start by describing qualitatively two example modes obtained from the α-band FC data of the DMN. We then consider the distribution of key spatio-temporal characteristics of the modes across all RSNs and frequency bands. An analysis of the modes is developed with more detail in the Supplementary Results S2 for the case of the α- and the β-band DMN.

3.1. Example modes of dynamic coupling

Figure 2 (left) summarizes the characteristics of the mode α_2 (i.e., the 2nd IC obtained in the α band) for the DMN. The positive part of its raw IC maps (Fig. 2, left, top) was strongest

over the occipital, the occipito-temporal, and the posterior parietal cortices bilaterally and also peaked over the pre-frontal area. Their negative part covered the central sulci bilaterally. These FC topographies could also be explained in terms of cross DMN-RSN integration using network mixture modeling. The GLM weights (Fig. 2, left, bottom left) disclosed a positive contribution of visual networks (VNs)—most prominently the VʟN—and of the DMN (cross DMN-VN and intra-DMN couplings) mainly, and a negative contribution of the SMN (cross DMN-SMN coupling). This corresponds to a situation where cross DMN-SMN integration competes with intra-DMN and DMN-VN integration. The sign of the IC maps and their GLM weights could be interpreted by considering the skewness of the IC time series, which was positive (see the arrow on the maps scale in Fig. 2, left). This means that the mode α_2 is associated with increases of cross DMN-VN and intra-DMN couplings and concomitant decreases of cross DMN-SMN couplings, rather than the opposite. Further analysis of the IC time series showed that these FC modulations were non-periodic and occurred as recurrent, sporadic transient events (Fig. 2, left, bottom right).

As a second example, we consider the DMN mode α_{20} (Fig. 2, right). Spatially, the SMN emerged clearly from the four IC maps but with opposite signs in the left and the right hemispheres. Temporally, this mode also exhibited aperiodicity, transience, and recurrence, but now IC skewness was close to zero (indicated by a double-headed arrow on the maps scale in Fig. 2, right). The mode α_{20} thus corresponds to an non-periodic, competitive mode of cross DMN-SMN integration wherein transient events of increased coupling with the right-hemispheric part of the SMN and decreased coupling with the left-hemispheric part alternate with events of decreased coupling with the right-hemispheric SMN and increased coupling with the left-hemispheric part. However, network mixture modeling failed to detect this cross DMN-SMN integration. In fact, this example illustrates the two caveats of this approach. First, the GLM weights corresponding to the SMN were close to zero because the IC maps asymmetry destroyed any correlation with the full SMN map. Generally, the sensitivity to template RSN maps identification is lessened in cases where the target RSN is split into sub-networks. Second, the GLM weights associated with the FPNs emerged with opposite signs because the SMN peaks leaked towards the frontal and the posterior parietal cortices, leading to a spurious interpretation of the mode in terms of alternating cross DMN-ʀ/ʟFPN integration. This is due to the intrinsic blurriness of MEG-based FC mapping that persists after leakage correction (see, e.g., Wens et al., 2015) and limits the specificity of RSN identification.

3.2. Temporal dynamics of the modes

The aperiodicity and the transient character of IC time series observed in the two preceding examples (Fig. 2) generalized to all the modes. The IC skewness was significant for about half of the modes and the IC kurtosis, for all of them. The latter result confirms that each mode exhibited large deviations. Further, comparing the distributions of IC kurtosis across modes obtained with different RSNs and frequency bands (Fig. 3, top

left), we identified an effect of the band ($p = 2.2 \times 10^{-19}$, Friedman test) but not of the RSN ($p = 0.09$). The band effect was due to a larger median across the α-band modes, indicating that some of these modes exhibited higher IC kurtosis than in the other bands, irrespectively of the seed RSN.

The transience of the large-deviation events (implied by high IC kurtosis) was confirmed by considering their lifetime distribution (Fig. 3, bottom left), whose median was only slightly above the windows step size. We observed a band effect ($p = 2.3 \times 10^{-7}$) without RSN effect ($p = 0.49$) reflecting a tendency for longer-lived events in the α band and for shorter-lived events in the β band. The latter was explained by the fact that no β-band mode exhibited lifetimes above 6.4 s, whereas longer lifetimes emerged in the other bands (maximum lifetime across θ-band modes: 7.3 s, α: 7.5 s, β': 6.9 s). The former was explained by a higher number of α-band modes with lifetimes above those of the β band (proportion across θ-band modes: 5%, α: 8%, β': 1%). This is illustrated explicitly in the Supplementary Results S3.

Table 2 shows that large-deviation events occur fairly often, since 35% of the time windows contained at least one event. Simultaneous occurrences are scarce, as 77% of large deviations were isolated and 20% consisted of two simultaneous events. The co-occurrence of more than two events turned out to be rare (about 3% of the time windows).

n	0	1	2	3	4	5	6	≥ 7
P_n (%)	65	27	7	1	0.2	0.01	0.003	0

Table 2: *Probability distribution P_n that n large-deviation events occur simultaneously.* The probability was estimated by pooling (co-)occurrence counts across all frequency bands (θ, α, β, β') and all RSNs (see Table 1). The counting process is illustrated in the Supplementary Results S3.

Finally, we note that the combination of higher kurtosis and longer lifetimes for α-band modes also implicates a tendency for larger deviations (see Supplementary Results S3).

3.3. Spatial organization of the modes

We now consider the spatial structure of the modes of dynamic coupling (notwithstanding the specificity/sensitivity issues identified above). All IC maps involved significant contributions to the FC data. Further, they all exhibited positive and negative significant values, which indicates that coupling synchronization is typically accompanied with coupling anticorrelation. We assessed the spatial extent of the significant FC patterns associated to each mode by measuring the fraction of volume filled by the IC statistical masks (Fig. 3, top right). Comparing their distributions showed that the DMN modes tended to be more focused spatially (RSN effect, $p = 2.1 \times 10^{-8}$) and that the β-band modes tended to be more extended (band effect, $p = 3.2 \times 10^{-10}$).

We also determined whether these FC patterns could be understood in terms of cross-network integration using network mixture modeling. The detection rate of significant mixtures was higher than expected by chance in all cases ($p < 10^{-3}$, binomial tests) except for the VN and the ʀFPN in the θ band
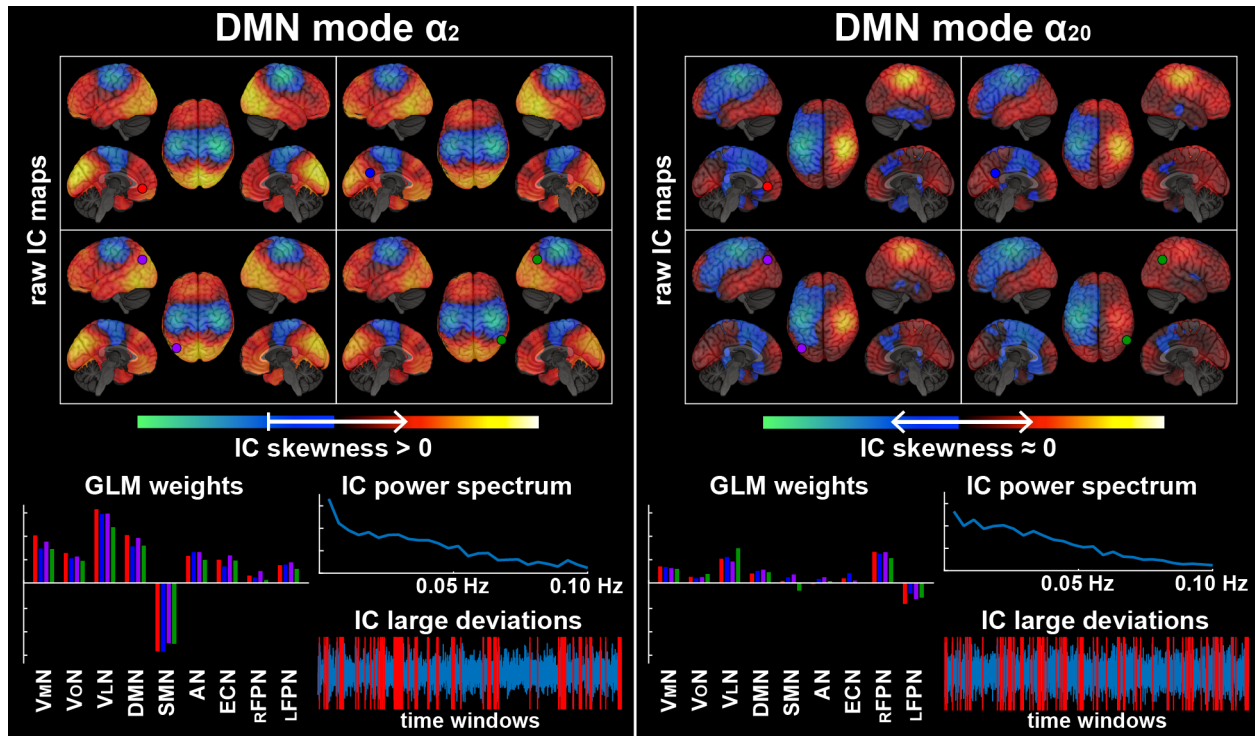
Figure 2: *Example modes of dynamic coupling.* Qualitative characteristics are shown for two ICs of the DMN-based FC data in the α band, here without statistical assessment. Spatial aspects encompass the four raw IC maps (**top**) and the associated network mixture model weights (**bottom left**). Seeds are color coded (red: MPFC, blue: PCC, violet: ʟTPJ, green: ʀTPJ). Temporal properties include the skewness (sign of non-zero skewness indicated by a single arrow on maps scale, approximately zero skewness indicated by a double-headed arrow), the kurtosis and the corresponding large-deviation events (emphasized in red on the IC time course), and the power spectrum (**bottom right**) of the IC time series. Scales and units are set consistently across the two modes but are otherwise arbitrary.

($p > 0.04$). This means that most FC datasets revealed some cross-RSN interactions. This occurrence rate appeared substantial for the DMN, where a large majority of modes could be classified in terms of cross-network coupling patterns, but smaller for the other RSNs (Fig. 3, bottom right). This was confirmed statistically by a RSN effect on these proportions ($p = 2.9 \times 10^{-18}$, Friedman test). We also disclosed a band effect ($p = 5.7 \times 10^{-4}$) due to higher occurrence rates in the α and the β bands compared to the θ and the β' bands. Note in Fig. 3 (bottom right) that the comparison for the β and the β' bands was only barely non-significant ($p = 0.06$, Wilcoxon test).

We further sought to identify which RSNs were predominantly involved among the cross-network couplings obtained in mixture models (Fig. 4, top). For the DMN modes, we detected a wide range of cross DMN-RSN interactions more often than expected by chance, particularly in the β band where RSN occurrence was significant for all the template RSNs. Interestingly, cross-network integration appeared more restricted in the α band where only the DMN-VᴍN, DMN-VʟN, and DMN-SMN couplings were detected with significant rate. Patterns of RSN occurrence were comparably scarcer with the other seed RSNs (in line with Fig. 3, bottom right). It is noteworthy that the VoN occurred significantly often only among the DMN modes in the β band, and not at all in the other cases (bar a non-significant occurrence for the β'-band DMN).

3.4. Competition among cross-network couplings

The detection rates for dynamical competition among the significant mixture models are shown in Fig. 4 (bottom). A significant majority of DMN modes exhibited competitive cross-network couplings, with the highest occurrence rates among the α- and the β-band modes. Competition also emerged with the other seed RSNs, although significance was more variant. This could merely reflect the scarcer occurrence of significant mixture models in these cases (Fig. 3, bottom right), which limits statistical power. No competition was detected for the VN modes in the θ band and for the ʟFPN modes in the θ and the β bands, and the detection rate for the α-band ʟFPN was below the uncorrected significance level.

Interestingly, this widespread detection of dynamical competition contrasts with a similar analysis performed at the level of static FC, where no significant RSN map appeared to display RSN competition (see Supplementary Results S1).

4. Discussion

4.1. Summary of findings

This paper investigates spontaneous synchronization patterns of dynamic FC with RSN nodes, or in short, modes of dynamic coupling. These modes corresponded to recurrent and sporadic transient events of FC changes involving varied and widespread coupling patterns. The α band disclosed longer lived FC events
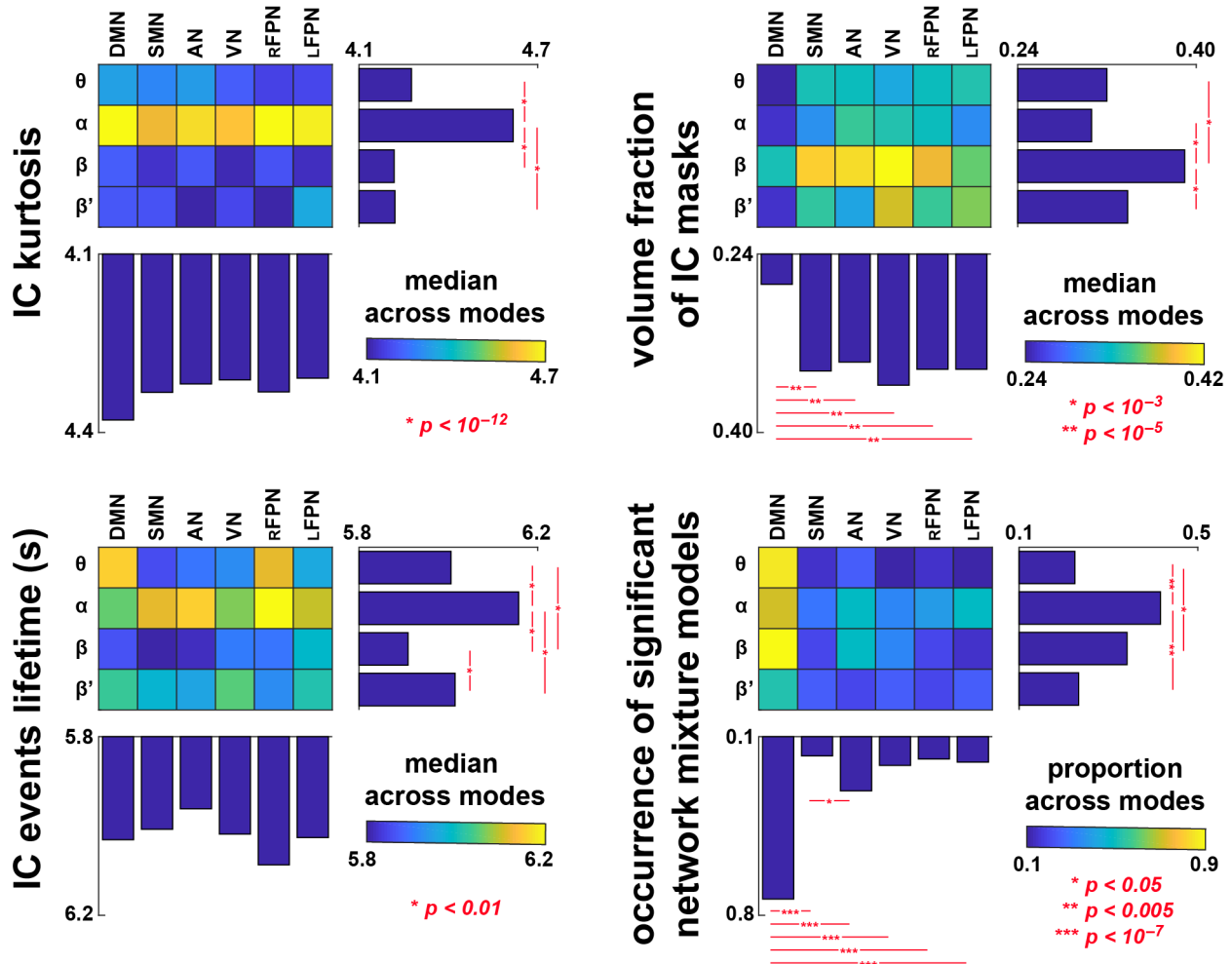
Figure 3: *Effects of seed RSN and frequency band on characteristics of the modes of dynamic coupling.* The distributions across modes of the IC kurtosis (**top left**), the IC large-deviation events lifetime (**bottom left**), the volume fraction of IC statistical masks (**top right**), and the proportion of significant network mixture models (**bottom right**) are compared across frequency bands (θ, α, β, β') and seed RSNs (see Table 1). The bar plots show the median values computed over RSNs (right side) or bands (bottom side). Significant effects identified via post-hoc Wilcoxon rank tests are indicated on these bar plots by red stars.
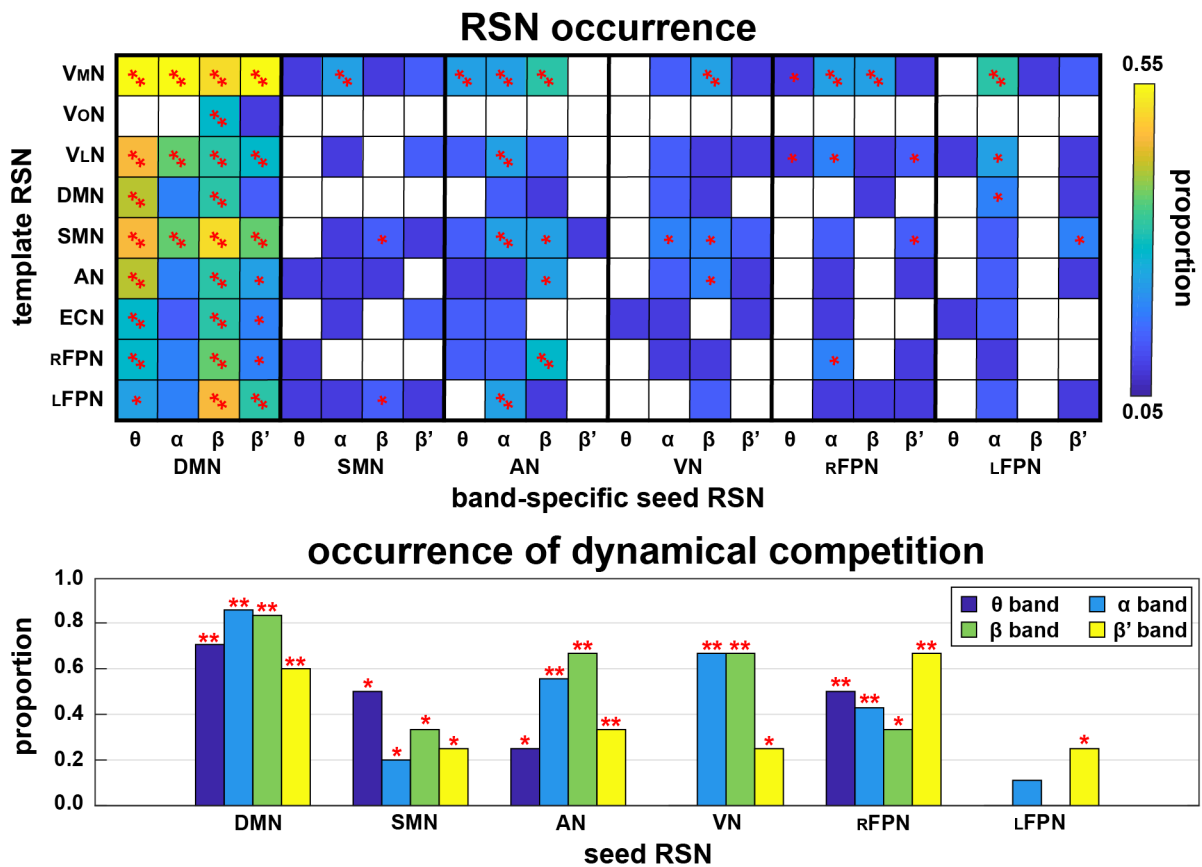
Figure 4: *Occurrence of RSNs and dynamical competition among network mixture models.* **Top:** The proportion of significant models disclosing RSN contribution is shown for each RSN template, frequency band ($\theta, \alpha, \beta, \beta'$), and seed RSN (see Table 1). White entries indicate the absence of occurrence. *: $p < 0.05$ Bonferroni corrected within each column (i.e., for nine comparisons), **: $p < 0.05$ corrected for all factors. **Bottom:** The proportion of significant models disclosing dynamical competition is also shown for each band and seed RSN. *: $p < 0.05$ uncorrected, **: $p < 0.05$ Bonferroni corrected. Significance of occurrence rates is based on binomial tests.

with larger deviations. The β-band modes involved spatially more extended coupling patterns. Cross-network integration was predominant among the DMN modes but scarcer among the others, and it was more common in the α and the β bands. The dynamic integration of the DMN with other RSNs involved mostly the SMN and the VL/MN in the α band, most RSNs in the θ and the β' bands, and all of them in the β band. Importantly, and in contradistinction with static FC, dynamical competition among these cross DMN-RSN couplings was ubiquitous.

4.2. Resting-state dynamics and coupling synchrony

While the nature and function of spontaneous brain interactions remain difficult to investigate experimentally, the picture emerging from empirical FC studies is one of a dynamic integration involving transient RSN fragmentation into subnetworks, cross-network binding, and alternation of core hubs (for reviews, see, e.g., Hutchison et al., 2013; de Pasquale et al., 2018). This entails a resting-state dynamics that has been mainly conceptualized as a jump process between discrete, temporally non-overlapping FC states that recur at random times. Two notions of FC states have been devised: transiently synchronized networks at the supra-second timescale (Allen et al., 2014; O'Neill et al., 2015b) and sub-second states detectable using hidden Markov models (Baker et al., 2014; Vidaurre et al., 2018). The latter have been suggested to underlie the former (Baker et al., 2014), however they are inaccessible to sliding-window FC and will thus not be discussed in detail hereafter. The supra-second jump dynamics is also compatible with the observation (within periods of concomitant strong intra-RSN coupling and high RSN centrality) of a few core networks playing in alternation, and swiftly shifting, the role of global brain integrator (de Pasquale et al., 2012, 2016).

The concept of coupling synchrony differs from these aspects and thus reveals novel properties of dynamic functional integration, but it is also compatible with the existing picture. We make four general inferences about resting-state dynamics:

(i) Spontaneous brain activity is characterized by coordinated coupling fluctuations around its RSN backbone.
(ii) These synchronized couplings reflect transition events between transiently synchronized networks.
(iii) Coupling synchronization is transient and covers large cortical areas, especially in the β band.
(iv) Events of coupling synchronization outside the β band can be longer lived, especially in the α band.

Inference (i) is based on the result that all modes of dynamic coupling disclosed significant IC maps, i.e., they explained a significant fraction of FC variance. This reveals that not only is resting-state activity structured into RSNs, but its functional interactions themselves are spontaneously organized into modes, which may thus be viewed by analogy as "networks of couplings" (O'Neill et al., 2015a). State analysis (Allen et al., 2014; O'Neill et al., 2015b) could not lead directly to this observation since it is geared towards the detection of stable patterns rather than their transition *per se* (because of the strict temporal exclusion constraint).

Inference (ii) asserts that the coupling synchrony we uncovered relates specifically to these FC state transitions. This is inevitable if we take state models literally (since FC changes can only occur at state jumps), but then they represent a simplification of the underlying neural dynamics. Empirically too, the spatio-temporal characterization of our modes depicts them as transition processes between successive transiently synchronized networks. Indeed, all the modes corresponded to transient events of FC changes that were recurrent, aperiodic, and mostly isolated, which fits with random state jumps (Allen et al., 2014). They also involved both increases and decreases as may be expected for transitions reflecting the difference between two successive states. Furthermore, in this context, the modes associated with preserved FC change sign can be interpreted as encoding non-reversible state transitions (i.e., they are more likely to occur in one direction than its opposite) and those with alternating sign, as reversible transitions. Therefore, our results are not only compatible with the jump process theory of resting-state dynamics but actually complements it. Still, this claim should be taken with two provisos. First, our confirmation is not completely unbiased because our ICA seeked transient, temporally sparse dynamics (see Calhoun et al., 2013; Daubechies et al., 2009, for a discussion of this point in the context of fMRI). This bias alone cannot explain the strongly transient character of our modes (see below and Supplementary Methods S2) but, critically, it may render our analysis insensitive to possible modes of dynamic coupling associated with less prominent temporal sparsity. Second, further study is required to prove the link between modes of dynamic coupling and FC state transitions.

The association between modes of dynamic coupling and transient FC events can also be validated with simulations. In the Supplementary Results S4, we used synthetic MEG data generated from a simple two-state system and identified a single mode corresponding to their transition. It also fits well within the framework of Deco et al. (2011) based on large-scale neurocomputational models of the human connectome. Their simulations suggested that resting-state activity is generated by a near-critical, multistable system composed of attractors corresponding to network configurations tightly constrained by anatomy but destabilized by local, stochastic or chaotic fluctuations within neural populations (Deco et al., 2009; Deco and Jirsa, 2012; Hansen et al., 2015). This generates a metastable dynamics characterized by fast transitions between attractors, which we suggest are captured empirically by the modes of dynamic coupling (the stable periods in between transitions being presumably encoded in FC states, see, e.g., Hutchison et al., 2013). Besides, our analysis indicates a mixture of reversible and non-reversible processes. It would be interesting to confront this finding with computational models and see if (non)reversibility is an emergent property of the human connectome.

Inference (iii) refers to the fact that the modes of dynamic coupling involved FC events that are short lived (less than 10 s) and cover a substantial fraction of the cortex, and significantly more so for the β-band modes. This preferential implication of β activity agrees with its suggested role as a functional background facilitating long-range neural synchronization (Bressler

and Richter, 2015; Kopell et al., 2000) and with the "statu-quo signal" theory of Engel and Fries (2010) according to which β activity maintains the current behavioral state, e.g., the exploratory state generated by metastability (Deco et al., 2011). Accordingly, β-band FC appears to be the best (but not the only) electrophysiological correlate of this functional exploration. In fact, Deco and Corbetta (2011) proposed that this exploration ensures responsiveness to future behavioral changes by retaining active representations of possible stimuli or tasks built upon past experiences. This led to two hypotheses that have been corroborated specifically in the β band, i.e., resting-state FC predicts task performance (Mary et al., 2015) and resembles task-positive FC in naturalistic paradigms (Betti et al., 2018). On this ground, we propose to interpret the β-band modes of dynamic coupling as reflecting the exploration at rest of complex, ecological activity patterns. Such a process would presumably require the integration of neural activity within widespread cortical areas. This also fits with the specific role played by the β band in the theory of de Pasquale et al. (2016) whereby global brain integration is sustained dynamically by a few core networks (the DMN, the SMN, and the bilateral FPN).

The tendency of β-band modes to exhibit shorter lifetimes and wider coupling patterns does not preclude contributions from the other bands to the functional exploration generated by metastability. For instance, there is no reason to expect the emergence of β rhythms only in the computational models of Deco et al. (2011), although it would be interesting to see if refined models (Hansen et al., 2015) could account for their special role. Further, large-deviation events were highly transient in all cases, which is indicative of metastability for the other bands as well. Actually, the fact that lifetimes were close to the lowest limit accessible to our sliding-window FC data suggests to extend our findings to finer timescales. In any case, we hypothesize that our functional interpretation above generalizes to the other bands.

Still, our inference (iv) highlights a peculiarity of the other bands and most prominently of the α band, namely, they exhibit modes with slightly longer lived FC events (by a couple of seconds, which is small compared to the timescale of our FC data but was sufficient to generate significant band effects). This suggests smoother state transitions that could reflect an increased stability of some FC states. This hypothesis is in line with, e.g., the proposed function of α rhythms for the top-down modulation of attention, perception, and consciousness level (Jensen et al., 2012; Klimesch, 2012) and that of θ rhythms for memory (Colgin, 2013; Klimesch et al., 2010). Indeed, state stabilisation occurs when the brain departs from its exploratory state and is engaged into cognitive processes (Deco and Corbetta, 2011). As attentional or perceptual drifts happen, we may thus expect certain FC states (i.e., the transiently synchronized networks that support such function) to partially stabilize. Based on this, we surmise that some non-β-band modes (specifically, those exhibiting higher kurtosis or longer lifetimes than the β-band modes) correspond to spontaneous drifts in mentation and the others, to the aforementioned functional exploration of metastable states.

That said, these hypotheses remain to be confirmed in behaviorally controlled experiments. Further, it would be interesting to generalize our analysis using a data-driven frequency band selection, as done in Vidaurre et al. (2018).

4.3. Functional integration of the default-mode network

So far, our interpretations of spontaneous coupling synchrony were not tied up to RSNs. The DMN stood out when we analyzed the relationship between coupling synchrony and cross-network integration. We make two inferences specifically involving the DMN:

(v) The DMN is a core network specialized in the transient integration with other RSNs.
(vi) Several RSNs compete to bind with the DMN.

Inference (v) is based on our observation that the large majority of DMN modes involved cross-network couplings. Some modes derived from the other RSNs shared this feature but many others did not. So the DMN appeared as the sole RSN that is systematically bound to other RSNs, which presumably endows it with a central role in the functional exploration generated by metastability.

This property specific to the DMN has, to our knowledge, never been emphasized. It is closely related to the core network theory of de Pasquale et al. (2016) that highlights the DMN as a global brain integrator, at least half of the time and mostly in the β band. In fact, the alternating dynamics of core networks was partially reflected in our data. Indeed, the variety of cross DMN-RSN couplings appeared widespread among the β-band modes, with significant occurrence of all RSNs (in opposition to the α-band modes). The highest occurrence rates across β-band modes spotted the two other core networks identified by de Pasquale et al. (2016), i.e., the SMN and the R/LFPN (combining both FPNs, the rate for the bilateral FPN even reached 60%), as well as the VMN. The emergence of these two core networks may be explained by the fact that their high-centrality periods must overlap (as each spends 40–50% of their time as hub, de Pasquale et al., 2016), increasing their likelihood to participate in β-band cross-network integration. On the other hand, the substantial involvement of DMN-VMN couplings and the lesser involvement of intra-DMN couplings may seem contradictory, as the VN is not part of the core networks and high DMN centrality periods overlap with strong intra-DMN FC periods (de Pasquale et al., 2016). Sill, given the difficulty with MEG to discriminate the PCC part of the DMN and the VMN, it may be that the DMN-VMN couplings reflect intra-DMN integration instead. One additional possible inconsistency is that neither the SMN nor the FPNs emerged from our data as the DMN did. Further work is needed to clarify this point.

Inference (vi) rests on our finding that a large majority of DMN modes involved a dynamical competition among cross DMN-RSN couplings. In fact, competition was not limited to the DMN and emerged with other RSNs too (although somewhat less consistently). This property fits well with the functional interpretation of metastability as a spontaneous competition among RSNs for the allocation of neural resources and optimum processing of future sensory, motor, or cognitive demands (Deco and Corbetta, 2011). Our data bring two major

inputs to this hypothesis. First, this competition emerges predominantly (but not solely) at the level of DMN-RSN integration. Second, it was observable at the short timescale of dynamic FC but not in static FC (see Supplementary Results S1), which makes sense as cross-network integration is transient by nature. This last observation contrasts with fMRI studies identifying negative static FC between the DMN and the bilateral FPN (Fox et al., 2005), but they are in line with others (see, e.g., Smith et al., 2009) as well as all MEG-based RSN studies (see, e.g., de Pasquale et al., 2010, for a discussion). Actually, detecting this anticorrelation requires a global signal regression (Fox et al., 2009) that was not used here. So we cannot exclude the existence of static RSN competition, but it is presumably subtler.

Functionally speaking, competitive cross DMN-RSN integration may play a major role in the main functions of the DMN, i.e., inner-directed cognition (e.g., spontaneous thoughts, mind wandering, or the mentation of past or hypothetical events, see, e.g., Buckner et al., 2008) and conscious awareness (Baars et al., 2003; Giacino et al., 2014). In this respect, the α-band DMN modes may be particularly informative, given their putative relation with spontaneous drifts in mentation. We speculate that the associated FC changes enable the neural communication between the DMN and other RSNs needed for, e.g., motor (DMN-SMN coupling), auditory (DMN-AN), or visual (DMN-V$_M$/$_L$N) imagery. It is noteworthy that DMN-SMN and DMN-V$_M$/$_L$N interactions stood out as those occurring the most among the α-band modes, usually in competition (see Supplementary Results S2). This hints at a possible predominance at rest of a cognitive alternation between motor and visual mentations. Also of interest is the complete absence of DMN-VoN coupling, which is in line with data suggesting that the primary visual cortex is not necessary for visual imagery (de Gelder et al., 2015). This appears plausible as no direct retinal stimulation is involved, but this remains a matter of debate (see, e.g., Pearson et al., 2015).

4.4. Methodological considerations

Our analysis pipeline (Fig. 1) was split into two separate main steps: (i) an ICA of dynamic FC data to identify spatio-temporal patterns of synchrony among brain couplings, and (ii) network mixture modeling for their classification in terms of cross-network integration and dynamical competition.

A general issue with dynamic FC is that short-time correlations are affected by large random estimation errors, leading to difficulties in distinguishing genuine dynamics from statistical variability (Hindriks et al., 2016). One question about our ICA is thus whether the decomposition into modes was driven by FC noise rather than coupling synchronization. This was not the case because the likelihood that the distribution of excess kurtosis across our modes (which were all significantly positive) emerges from correlation errors only, was extremely small ($p < 10^{-7}$ based on numerical estimates). In fact, ICA entailed a reduction of FC noise because the modes disclosed large patterns of coupling synchrony that cannot be explained by correlation errors (which do not exhibit spatial coordination

beyond the intrinsic blurriness of MEG FC), so that their contribution within each mode averaged out. These two analytical arguments are developed fully in the Supplementary Methods S2 and S3. The ability of our ICA to detect coupling synchrony among noisy dynamic FC data is also illustrated in the simulation reported in the Supplementary Results S4. Finally, note that the seminal study of O'Neill et al. (2017b) provided a proof of concept as they used a similar ICA of sliding-window correlations to identify task-related transiently synchronized networks, which appeared meaningful in view of the brain processes expected in their experimental design.

Another general confound of dynamic FC is that variations in sliding-window correlations could be driven by transient modulations of local activity (i.e., power changes) rather than couplings *per se*. We show in the Supplementary Results S5 that sliding-window power estimates (with sLORETA for depth bias correction, see Pascual-Marqui, 2002) only poorly correlated with the IC time series (absolute Pearson correlations below 0.03 across all modes, RSNs, and frequency bands) and thus could not explain the spatio-temporal features of the modes. A similar analysis focusing on high-frequency power also confirmed that our results were not driven by muscle artifacts.

Yet another difficulty of our ICA of resting-state FC (i.e., without task or stimulation onset) is the subjective selection of relevant ICs, e.g., by visual inspection of their maps. Here, network mixture models allowed us to classify them objectively as linear superpositions of fMRI RSN templates. The validity of using static fMRI as reference for MEG dynamic FC is obviously debatable. The benefit is that fMRI provides clear-cut RSN atlases, which have guided both the design (de Pasquale et al., 2010, 2012, 2016; Sockeel et al., 2016) and the interpretation (Brookes et al., 2011; Liu et al., 2017) of MEG/EEG FC analyses. Resting-state fMRI atlases are also used as functional brain parcellations in MEG studies (see, e.g., Vidaurre et al., 2018). The standard and reproductible character of fMRI RSNs is a crucial aspect because the ensuing interpretations are tied up to the choice of templates. Of note, the RSN atlas of Smith et al. (2009) miss a couple of known systems, e.g., the ventral attentional and the language networks. This might *a priori* incur a lack of classification sensitivity for these two RSNs, but this is presumably mitigated by the fact that they have been scarcely disclosed with MEG resting-state FC. A more fundamental drawback lies in the different nature of the signals, the distinct spatial and temporal resolutions, and in the usage of static RSNs. So our mixture models should not be over-interpreted as genuine models of dynamic integration. They merely provide a crude approximation, but one that still allowed detection of several cross-network coupling patterns. We reported a lessened sensitivity to patterns involving sub-networks (our prototype was the mode α_{20}, Fig. 2 right) but this issue appeared relatively limited and only a few modes were totally misclassified (see Supplementary Results S2). This is because spatial correlation between a full RSN and a sub-network map may be sensitive to their partial overlap. We also noted a restricted specificity related to the intrinsic blurriness of MEG FC rather than to the use of fMRI templates.

Technically, we constructed the mixture models using GLM

weights, which provide a multivariate measure of maps correlations. However, their values must be interpreted cautiously. First, effect sizes are strongly inflated if maps autocorrelation (dominated by MEG spatial leakage) is not controlled. That is why we considered standardized weights wherein the reduction in spatial degrees of freedom was built in (see Supplementary Methods S4). Second, cross-correlations among GLM regressors lead to various suppression effects, which are beneficial to the GLM as a whole but may render individual weight values ambiguous (Watson et al., 2013). Specifically, cross-over suppression entails a magnitude boost of some weights while dampening or even reversing the sign of others (Watson et al., 2013). In particular, the sign-reversal possibility challenges the validity of our novel dynamical competition test. However, there is no such issue in our case because the RSN templates were weakly correlated (absolute value of pairwise Pearson correlations: 0.04 ± 0.02) so our GLM design was well conditioned (see Supplementary Methods S5 and Supplementary Results S6 for full justification). Actually, this provides another argument for using a fMRI-based RSN atlas in our mixture models. Last, we emphasize that GLMs do not provide causal information (Weichwald et al., 2015). So mixture models cannot inform us on whether superpositions of static RSNs underlie modes of dynamic coupling, whether these modes generate RSNs, or whether both have a common factor (e.g., hidden brain states or metastable attractors).

Despite its limitations, some of which may hopefully be overcome in future developments, our current approach allowed to uncover new features of the metastable resting-state dynamics and the specialized function of the DMN for dynamic cross-network integration. It could also come in handy to reveal the impact of behavioral manipulations or brain disorders on the intrinsic functional organization of the human brain.

Acknowledgements

Supplementary Methods

S1. More on the independent component analysis to identify modes of dynamic couplings

We provide here further background on our use of ICA to disclose modes of dynamic couplings from sliding-window FC. From an abstract perspective, the problem of identifying synchronization patterns among dynamic FC data can be viewed as a "metaconnectivity" analysis, i.e., a connectivity analysis of connectivity values. Considering the analogy with usual FC, two natural options emerge: using a seed-based-like approach (wherein temporal correlations of FC time series would be estimated between a chosen "seed" connection and all other connections) or an ICA. Here, we focused on the latter choice for several reasons.

First, ICA is more sensitive to the detection of coupling synchrony, which may be hard to identify via direct temporal correlations because of its complex, transient nature (this observation motivated, e.g., the maximum correlation window approach in de Pasquale et al., 2010). On the other hand, ICA relies on a linear decomposition into modes of dynamic FC with maximally independent (hence asynchronous) temporal variations, so that one mode of dynamic coupling gets by design concentrated into a single IC. In fact, the typical ICA algorithms being based on the maximization of non-gaussianity indices closely related to excess kurtosis (Hyvärinen and Oja, 2000), they promote temporally sparse ICs rather than independence *per se* (Calhoun et al., 2013; Daubechies et al., 2009). The ICA is thus biased towards the detection of modes of dynamic couplings driven by transient FC synchronization events. One noteworthy consequence of this bias is that modes associated with Gaussian or thin-tailed temporal processes are missed altogether by ICA.

The spatial leakage effect affecting FC (Brookes et al., 2012b; Wens, 2015; Wens et al., 2015) also represents a difficulty for the seed-based-like option, because it is bound to induce spurious temporal correlations among dynamic FC estimates with substantial spatial extent. In the absence of principled correction methods for such "connectivity-level spatial leakage", ICA represents a better choice since it is less prone to this issue, as noted in Brookes et al. (2011).

Finally, the seed-based-like approach would require the prior selection of one or a few connections of interest. Given that this is the first study investigating patterns of coupling synchrony at rest and in the absence of specific expectations on which functional connections would be involved, we favored the data-driven framework provided by ICA.

S2. Distribution of modes kurtosis without coupling synchrony

A serious challenge for dynamic FC analysis is the large statistical variability in short-time correlation estimates, which can be misinterpreted as dynamics (Hindriks et al., 2016). For example, our FC data contained 20 time samples per window so the SD due to FC estimation errors was of the order of $1/\sqrt{20-3} \approx 0.24$, which is considerable. In this section and the next, we argue that our ICA was not dominated by FC noise and justify the validity of our results. A simulation-based proof of concept is also developed in the Supplementary Results S4.
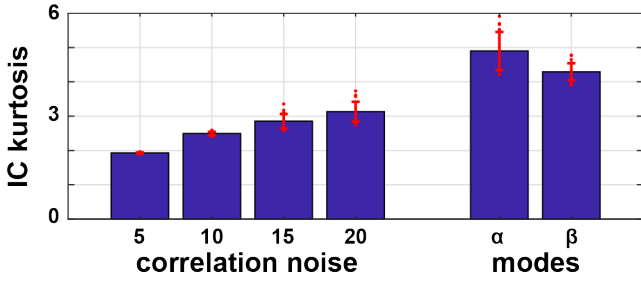
Figure S1: *Kurtosis for correlation noise ICs.* The plot depicts the distribution of kurtosis across 20 ICs (blue bins: mean, red bars: SD, red dots: samples) obtained from correlation noise data generated with ω = 5, 10, 15, and 20 temporal degrees of freedom within windows (**left**). For comparison, the distributions obtained from the modes of dynamic coupling for the DMN in the α and the β bands (see also Fig. S9, middle) are shown (**right**).

Our first evidence is based on the strongly positive excess kurtosis of the modes of dynamic coupling. We estimate here numerically the distribution of IC kurtosis obtained from an ICA of correlation errors and show its incompatibility with our data (Fig. S1).

The distributions in Fig. S1 (left) were obtained by modeling correlation noise as a random variable whose Fisher transform follows a Gaussian with zero mean and SD $1/\sqrt{\omega - 3}$. Here, ω denotes the number of temporal within-window degrees of freedom used in correlation estimates. In our case, $\omega \leq 20$ (2 Hz-envelope signals within 10 s-long windows) so we probed ω = 5, 10, 15, and 20. We randomly generated FC noise data with similar numbers of spatial and temporal degrees of freedom than in our analysis (i.e., 50 independent FC time series corresponding to our pre-ICA dimension reduction, each with ν independent samples corresponding to the effective number of non-overlapping windows) and applied the same ICA (nonlinearity *tanh* and 20 ICs to compute). This procedure was repeated 100 times.

Figure S1 shows that, for all considered ω, the IC kurtosis was smaller in FC noise than in the FC data (Fig. S1, right; $p < 10^{-7}$ when comparing to the α- and the β-band DMN modes, Wilcoxon rank tests). We also observe in Fig. S1 (left) that IC kurtosis was below 3 for $\omega \leq 15$. This reflects that fact that correlation noise is characterized by negative excess kurtosis (its kurtosis being less than 2.65 for $\omega \leq 20$), although ICA entails an overestimation due to its bias towards positive excess kurtosis. What Fig. S1 illustrates is that this bias is not sufficient to explain the high IC kurtosis observed in the modes of dynamic coupling, justifying that they are not driven by FC noise and that their transient character is genuine.

S3. *Error analysis for the modes of dynamic coupling*

The preceding argument justifies that our ICA of dynamic FC was not driven by the substantial FC noise but does not explain how its contribution was reduced within ICs. The basic reason is that ICA looks for coordinated signals among the FC data, whereas correlation errors do not covary beyond what is entailed by spatial leakage. We develop here the argument analytically and estimate the effect of FC noise on IC time series

and IC maps. We show that ICA approximates well the temporal and spatial features of the true modes. To our knowledge, the following development is not available in the literature.

Preliminaries on the ICA model. We start from the assumption that the standardized FC time series $\boldsymbol{Z}$ are genuine linear combinations of M temporally independent modes, i.e., $\boldsymbol{Z}(t) = \sum_{m=1}^{M} \boldsymbol{A}_m S_m(t) + \boldsymbol{\varepsilon}(t)$. Here, t indexes the sliding windows, $\boldsymbol{A}_m$ and $S_m(t)$ denote the spatial pattern and time series of the mth mode, and $\boldsymbol{\varepsilon}(t)$ represents correlation noise. Without loss of generality, we may choose the $\boldsymbol{A}_m$ to be linearly independent and the $S_m(t)$ to have unit variance. To understand how noise affects their estimability, it is useful to express correlation errors as

$$\boldsymbol{\varepsilon} = \sum_{m=1}^{M} \boldsymbol{A}_m \varepsilon_m + \boldsymbol{\varepsilon}_\perp \qquad \text{with} \qquad \boldsymbol{A}_m^{\mathrm{T}} \boldsymbol{\varepsilon}_\perp = 0, \qquad (1)$$

so that $\boldsymbol{Z}(t) = \sum_{m=1}^{M} \boldsymbol{A}_m \hat{S}_m(t) + \boldsymbol{\varepsilon}_\perp(t)$ with

$$\hat{S}_m(t) = S_m(t) + \varepsilon_m(t). \qquad (2)$$

We thus find that ICA algorithms such as FastICA cannot determine the time series $S_m(t)$ but only the noisy version (2).

The associated spatial pattern $\boldsymbol{A}_m$ is then estimated from the data $\boldsymbol{Z}$ by linear regression of $\hat{S}_m(t)$ (Hyvärinen and Oja, 2000), i.e., $\hat{\boldsymbol{A}}_m = \mathrm{Cov}(\boldsymbol{Z}, \hat{S}_m)/\mathrm{Var}(\hat{S}_m)$. Using the independence of the modes and correlation noise, we find

$$\hat{\boldsymbol{A}}_m = \frac{\boldsymbol{A}_m + \mathrm{Cov}(\boldsymbol{\varepsilon}, \varepsilon_m)}{\mathrm{Var}(\hat{S}_m)}. \qquad (3)$$

The IC time series and IC maps considered in the main text correspond to the estimates (2) and (3), up to an irrelevant normalization factor.

Covariance structure of correlation noise. For the next point, we need to evaluate the size of the noise contributions in equations (2) and (3), i.e., the variance $\mathrm{Var}(\varepsilon_m)$ of the projected noise ε_m and its covariance $\mathrm{Cov}(\boldsymbol{\varepsilon}, \varepsilon_m)$ with $\boldsymbol{\varepsilon}$. For notational convenience, let us represent the $\boldsymbol{A}_m$ and $\boldsymbol{\varepsilon}(t)$ as $D \times 1$ vectors ($D = SN$, i.e., S seeds $\times N$ = 13229 sources) and gather the former into a $D \times M$ matrix $\boldsymbol{A}$. It follows from the decomposition (1) that $\varepsilon_m = [(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}]_m \boldsymbol{\varepsilon}$, so we have

$$\mathrm{Var}(\varepsilon_m) = [(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}} \, \mathrm{Cov}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^{\mathrm{T}}) \, \boldsymbol{A}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}]_{m,m} \qquad (4)$$

and

$$\mathrm{Cov}(\boldsymbol{\varepsilon}, \varepsilon_m) = \mathrm{Cov}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^{\mathrm{T}}) \, [\boldsymbol{A}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}]_m. \qquad (5)$$

We discuss first the $D \times D$ covariance $\mathrm{Cov}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^{\mathrm{T}})$ of correlation errors. Because of the geometric structure of spatial leakage in MNE (Wens et al., 2015), correlation noise is restricted to the d-dimensional subspace where the $d = S\rho$ spatial degrees of freedom live and contributes homogeneously to each of them. This is expressed as

$$\mathrm{Cov}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^{\mathrm{T}}) = \sigma^2 \, \boldsymbol{P}, \qquad (6)$$

where $\boldsymbol{P}$ denotes the projection matrix onto this subspace and σ^2, the variance of (standardized) correlation noise. Furthermore, in our context where only leaky reconstructions are accessible, we may assume that the $\boldsymbol{A}_m$ also belong to this subspace, i.e., $\boldsymbol{PA}_m = \boldsymbol{A}_m$. This means that the projector $\boldsymbol{P}$

can effectively be replaced by the identity when plugging (6) into (4) and (5). Consequently, the variance (4) simplifies to $\sigma^2[(\boldsymbol{A}^\mathrm{T}\boldsymbol{A})^{-1}]_{m,m}$ and the covariance (5), to $\sigma^2[\boldsymbol{A}(\boldsymbol{A}^\mathrm{T}\boldsymbol{A})^{-1}]_m$.

Error estimates. Using this to evaluate the estimation errors in the IC time series (2) and the IC spatial maps (3), we find

$$\mathrm{Var}(\hat{S}_m) = \mathrm{Var}(S_m) + \mathrm{Var}(\varepsilon_m) = 1 + \sigma^2\left[(\boldsymbol{A}^\mathrm{T}\boldsymbol{A})^{-1}\right]_{m,m} \quad (7)$$

and

$$\hat{\boldsymbol{A}}_m = \boldsymbol{A}_m + \frac{\sigma^2 \sum_{n\neq m}[(\boldsymbol{A}^\mathrm{T}\boldsymbol{A})^{-1}]_{n,m}\boldsymbol{A}_n}{1 + \sigma^2[(\boldsymbol{A}^\mathrm{T}\boldsymbol{A})^{-1}]_{m,m}}\,. \quad (8)$$

The contribution of correlation noise is modulated by the entries of the matrix $(\boldsymbol{A}^\mathrm{T}\boldsymbol{A})^{-1}$. We can estimate them crudely by noting that, generically, the diagonal entries of $\boldsymbol{A}^\mathrm{T}\boldsymbol{A}$ scale as d (because the $\boldsymbol{A}_m$ live in a subspace of dimension d) and its off-diagonal entries, as $\sqrt{d}$ due to sign cancellations expected from dynamical competition. Assuming $d \gg 1$, it follows that the diagonal and off-diagonal entries of $(\boldsymbol{A}^\mathrm{T}\boldsymbol{A})^{-1}$ scale respectively as $1/d$ and $(1/d)^{3/2}$. On this basis, we find at first approximation

$$\mathrm{Var}(\hat{S}_m) = 1 + O(\sigma^2/d) \quad (9)$$

and

$$\hat{\boldsymbol{A}}_m = \boldsymbol{A}_m\left[1 + O\left(\sigma^2\sqrt{M-1}/d^{3/2}\right)\right]. \quad (10)$$

In the last equation, the extra factor $\sqrt{M-1}$ comes from the linear sum over $M-1$ terms in equation (8).

Numerically, in our case we had $\rho = 46$ so $d = 184$ when $S = 4$ and $d = 92$ when $S = 2$, which is large enough for the above estimates to hold. Further, since we applied ICA to standardized FC data, σ^2 actually represents the noise variance after standardization so it is bounded by one. Also, the linear independence of the $\boldsymbol{A}_m$ imposes $M \leq d$. We conclude that, in the worst case, the contribution of FC noise to the IC temporal variance and IC maps was of the order of $1/d = 0.5\%$ when $S = 4$ and 1% when $S = 2$. This establishes our claim that IC time series and IC maps are good approximations.

Consistency with data. We further check the consistency of these estimates with the observed IC time series and IC maps. For illustrative purposes, we focus on the modes of the DMN in the α and the β bands. Figure S2 plots the relative error matrix

$$\left[(\hat{\boldsymbol{A}}^\mathrm{T}\hat{\boldsymbol{A}})^{-1}\right]_{n,m} \sqrt{\frac{\hat{\boldsymbol{A}}_n^\mathrm{T}\hat{\boldsymbol{A}}_n}{\hat{\boldsymbol{A}}_m^\mathrm{T}\hat{\boldsymbol{A}}_m}}\,. \quad (11)$$

The diagonal entries $[(\hat{\boldsymbol{A}}^\mathrm{T}\hat{\boldsymbol{A}})^{-1}]_{m,m}$ approximate the relative error term $[(\boldsymbol{A}^\mathrm{T}\boldsymbol{A})^{-1}]_{m,m}$ in the IC time series (see equation 7). They were all below 0.6%, which is comparable to our analytical estimate. Likewise, the off-diagonal entries approximately measure (the norm of) each of the $M-1$ contributions to the relative error in the IC spatial maps (see equation 8). They were all below 0.2%.

Interestingly, Fig. S2 indicates that errors in the β band tended to be smaller than in the α band. This relates to the fact that β-band IC maps tended to be more extended spatially (see Fig. 3, top right).
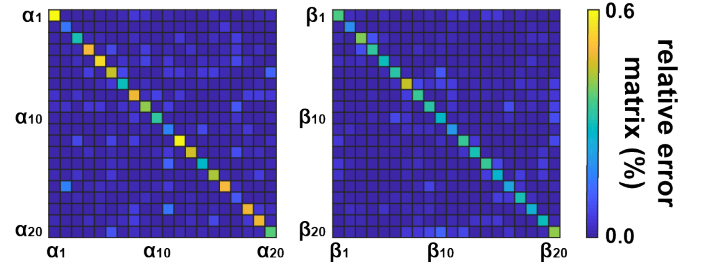

Figure S2: *Relative error estimates for the modes of dynamic coupling (DMN, α and β bands).* The relative error matrix (11) estimated from the IC maps shown in Figs. S10 and S11 is depicted for both the α (**left**) and the β bands (**right**). The diagonal elements approximate the relative contribution of FC noise to the variance of the IC time series and the off-diagonal elements, the relative error on the IC maps.
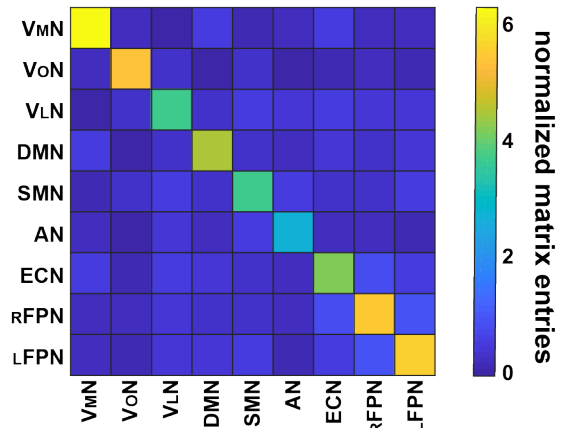

Figure S3: *Good conditioning of the GLM design.* The plot of the matrix $\boldsymbol{X}\boldsymbol{X}^\mathrm{T}$ (here normalized by the number N of sources) shows that it was dominated by its diagonal entries, so its inversion was stable. Correspondingly, the condition number κ of $\boldsymbol{X}\boldsymbol{X}^\mathrm{T}$ (i.e., the ratio of its largest and smallest eigenvalues) was of order one (precisely, $\kappa = 3.15$).

S4. Spatial general linear modeling

We describe here in more detail the spatial GLM applied to the raw IC maps associated to the $M = 20$ modes of dynamic coupling and the S seeds (see Table 1 in the main text). The regressors correspond to the $R = 9$ template RSN maps. We represent all maps as $1 \times N$ row vectors, where $N = 13229$ corresponds to the number of cortical sources.

If $\boldsymbol{Y}_{m,s}$ denotes the Fisher-transformed raw IC map for the mode m ($1 \leq m \leq M$) and seed s ($1 \leq s \leq S$) and if $\boldsymbol{X}$ denotes the $R \times N$ design matrix whose R rows represent the template RSN maps, then we modeled $\boldsymbol{Y}_{m,s}$ as the linear combination

$$\boldsymbol{Y}_{m,s}^{\mathrm{model}} = \boldsymbol{\beta}_{m,s}\boldsymbol{X} \quad \text{with} \quad \boldsymbol{\beta}_{m,s} = \boldsymbol{Y}_{m,s}\boldsymbol{X}^\mathrm{T}(\boldsymbol{X}\boldsymbol{X}^\mathrm{T})^{-1}, \quad (12)$$

i.e., with the $1 \times R$ weights $\boldsymbol{\beta}_{m,s}$ estimated via least-squares minimization (Friston et al., 2007). In our specific case, the matrix $\boldsymbol{X}\boldsymbol{X}^\mathrm{T}$ was full rank so its inversion was performed without regularization (Fig. S3).

Fundamental to statistical inference on these GLMs is the $R \times R$ sample covariance matrix of the weights $\boldsymbol{\beta}_{m,s}$. Taking into account the fact that the number of spatial degrees of freedom is $\rho \ll N$ (because of the smoothness of these maps, see Wens

et al., 2015), the sample covariance was taken as

$$\Sigma_{m,s} = S_{m,s}^2 (XX^{\mathrm{T}})^{-1},\tag{13}$$

where

$$S_{m,s}^2 = \frac{Y_{m,s}[\mathbf{1}_{N\times N} - X^{\mathrm{T}}(XX^{\mathrm{T}})^{-1}X]Y_{m,s}^{\mathrm{T}}}{\rho - R}\tag{14}$$

is an estimator of the true regression error variance $\sigma_{m,s}^2$. The central formulas (13) and (14) can be derived by revisiting the classical theory of GLMs (see, e.g., Friston et al., 2007) using the noise covariance structure (6) with $d = \rho$. The variables $Y_{m,s}$ and X are approximately Gaussian, as $Y_{m,s}$ contains Fisher-transformed correlation values and X contains z scores (Smith et al., 2009). Therefore, the weights $\beta_{m,s}$ were considered Gaussian with true population mean $\mu_{m,s}$ and, in a fixed-effect design, with true covariance $\sigma_{m,s}^2(XX^{\mathrm{T}})^{-1}$.

In this framework, the model significance F tests for each mode m and seed s were established using the goodness-of-fit statistic

$$\mathcal{F}_{m,s} = \tfrac{1}{R}\beta_{m,s}\Sigma_{m,s}^{-1}\beta_{m,s}^{\mathrm{T}},\tag{15}$$

which follows the $F_{R,\rho-R}$ distribution under the multivariate null hypothesis $\mu_{m,s} = 0$ that all R weights are zero, i.e., no linear mixture of the template RSNs explains the associated raw IC map (Friston et al., 2007). The univariate weight significance t tests for each mode m, seed s, and template r ($1 \le r \le R$) were based on the standarized version

$$\mathcal{T}_{m,s,r} = \frac{(\beta_{m,s})_r}{\sqrt{(\Sigma_{m,s})_{r,r}}}\tag{16}$$

of the weights (the subscript indicating that the rth component is taken). This statistic follows Student's $t_{\rho-R}$ distribution under the univariate null hypothesis $(\mu_{m,s})_r = 0$ that this weight vanishes, i.e., zero partial correlation between the associated raw IC map and the rth template RSN map (Friston et al., 2007).

Finally, the nested models comparison F tests used to disclose RSN competition was based on the statistic

$$\mathcal{G}_{m,s} = \tfrac{1}{C_{m,s}}\left(\beta_{m,s}\Sigma_{m,s}^{-1}\beta_{m,s}^{\mathrm{T}} - \tilde{\beta}_{m,s}\Sigma_{m,s}^{-1}\tilde{\beta}_{m,s}^{\mathrm{T}}\right),\tag{17}$$

which measures the gain in goodness-of-fit when using the original rather than the sign-constrained model. Here, $\tilde{\beta}_{m,s}$ denotes the weights of the constrained GLM and $C_{m,s}$, the number of those weights that vanish ($0 \le C_{m,s} \le R$). These quantities are detailed in the dedicated section S7. The statistic (17) follows the $F_{C_{m,s},\rho-R}$ distribution under the null hypothesis that the two models have similar explanatory power (Friston et al., 2007). It is ill-defined whenever the unconstrained GLM does satisfy the sign condition, since then $\beta_{m,s} = \tilde{\beta}_{m,s}$ and $C_{m,s} = 0$. Further justification for this dynamical competition test is provided in section S6.

S5. Interpretability of weights in well-conditioned general linear models

Cross-correlations among regressors lead to suppression effects that may obfuscate the direct meaning of GLM weights

magnitude and sign (Watson et al., 2013), leading to possible issues in interpreting results of network mixture models and their dynamical competition test. We review here why there are no such ambiguities in our well-conditioned design (see Fig. S3).

For illustrative purposes, let us consider a toy example where $Y_{m,s}$ is genuinely a linear superposition of R maps X with true weights $\mu_{m,s}$ and an additive noise $\varepsilon_{m,s}$, i.e.,

$$Y_{m,s} = \mu_{m,s}X + \varepsilon_{m,s}.\tag{18}$$

For conciseness, we shall assume that the regressors used in the GLM estimation (12) are correctly identified, so

$$\beta_{m,s} = \mu_{m,s} + \varepsilon_{m,s}X^{\mathrm{T}}(XX^{\mathrm{T}})^{-1}.\tag{19}$$

Therefore, in the case where XX^{T} is invertible (and only in this case), GLM estimation is unbiased and each weight $(\beta_{m,s})_r$ coincides with the true value $(\mu_{m,s})_r$ up to a random error of order $\sqrt{(\Sigma_{m,s})_{r,r}}$ (i.e., the SD of the projected noise contribution, see equations 13 and 14). This leads to two interpretation rules for GLM weights:

(i) Their magnitude must always be considered with regard to the estimated error, or equivalently to their effect size as measured by the normalized weights (16).

(ii) Their sign correctly identifies the true weight sign at large effect sizes but are more prone to be randomly flipped by projected noise at small effect sizes.

In ill-conditioned cases, the matrix inversion in equation (19) is unstable in some directions and effectively blows up the projected noise for the weights comprised in these directions. So the GLM is contaminated by false negatives (small effect size and random sign), although this problem can be alleviated by proper regularization. In any case, it is noteworthy that the rule (ii) still works to identify the interpretable signs.

Because it is key in our analysis of dynamical competition (see next section), we further justify the sign rule (ii) with a zero-order correlation analysis (see Supplementary Results S6).

S6. Rationale for the dynamical competition test

The sign rule (ii) of section S5 is a key element of the rationale behind the dynamical competition test. Goodness-of-fit measures such as (15) are dominated by the weights with large effect size, and are thus mostly sensitive to the interpretable weight signs and relatively impervious to the random weight signs. Therefore, *the sign constraint causes a substantial goodness-of-fit change* (17) *when it affects some of the interpretable signs, but a small change when only random signs are involved.*

More specifically, consider first a non-competitive situation where all nonzero true weights have identical sign. All large effect-size weights share this sign by (ii) and the constraint will mainly affect the small effect-size weights (to keep the total sum of squared errors close to the unconstrained minimum). The constrained and unconstrained GLMs will thus be comparable, likely resulting in a nonsignificant test. Alternatively, in a competitive situation, the sign of some large effect-size weights will be forced to change, leading to substantially different GLMs and a likely significant test.

S7. Solution of sign-constrained general linear models

The sign-constrained mixture model for the mth mode of dynamic coupling was obtained by finding the set of GLM coefficients $\tilde{\beta}_{m,s}$ ($1 \le s \le S$) that minimizes the total sum of squared errors (as in the unconstrained case) given that the weights sign is fixed to ϵ_m across all $1 \le r \le R$ and $1 \le s \le S$, i.e., $\epsilon_m (\tilde{\beta}_{m,s})_r \ge 0$. This problem was solved for $\epsilon_m = +$ and $\epsilon_m = -$ separately, and the sign yielding the least total sum of squared errors was then selected.

In each case, we approached the solution using a projected gradient descent algorithm. Explicitly, successive approximations $\tilde{\beta}_{m,s}^{(N)}$ were built iteratively for $N = 1, 2, \ldots$ using

$$\left(\tilde{\beta}_{m,s}^{(N)} \right)_r = \epsilon_m \left| \left(\tilde{\beta}_{m,s}^{(N-1)} \right)_r - \frac{\tau}{\Lambda} \left[\left(\tilde{\beta}_{m,s}^{(N-1)} X - Y_{m,s} \right) X^{\text{T}} \right]_r \right|, \quad (20)$$

where Λ denotes the maximum eigenvalue of XX^{T}. This corresponds to an unconstrained gradient descent with step size τ together with a projection enforcing the sign constraint on each iterate. The convergence analysis of this sequence is developed in section S8. Global convergence (i.e., for arbitrary initial conditions $\tilde{\beta}_{m,s}^{(0)}$) is ensured for any $0 < \tau < 2$. The limit as $N \to \infty$ does not coincide with the exact solution but their difference scales with the step size τ, so this projected gradient descent required $\tau \ll 1$. The algorithm was stopped once the Euclidean norm $\left\| \tilde{\beta}_{m,s}^{(N)} - \tilde{\beta}_{m,s}^{(N-1)} \right\|$ of the Nth step reached below a pre-defined constant δ controlling the precision of the convergence. In the context of statistical inference, a natural scale for δ is the minimum SD of the GLM weights distribution.

In practice, we defined $\tilde{\beta}_{m,s}^{(0)}$ using $\left(\tilde{\beta}_{m,s}^{(0)} \right)_r = \epsilon_m |(\beta_{m,s})_r|$ and we set

$$\tau = 10^{-2}, \qquad \delta = 10^{-3} \min_{m,s,r} \sqrt{(\Sigma_{m,s})_{r,r}}. \quad (21)$$

Given such small step size and precision parameters, the approximation $\tilde{\beta}_{m,s}^{(N)}$ was fairly close to the exact solution $\tilde{\beta}_{m,s}$. However, the latter could also be derived analytically by observing that it must contain a number $C_{m,s} > 0$ of zero weights (unless the unconstrained solution does satisfy the sign condition). A consequence of the convergence analysis in section S8 is that the vanishing weights can be identified unambiguously using the numerical criterion that

$$\left(\tilde{\beta}_{m,s} \right)_r = 0 \quad \text{if and only if} \quad \left| \frac{2 \Lambda \left(\tilde{\beta}_{m,s}^{(N)} \right)_r}{\left[\left(\tilde{\beta}_{m,s}^{(N)} X - Y_{m,s} \right) X^{\text{T}} \right]_r} \right| < 1. \quad (22)$$

On this basis, the exact solution could then be written as

$$\tilde{\beta}_{m,s} = Y_{m,s} \tilde{X}^{\text{T}} \left(\tilde{X} \tilde{X}^{\text{T}} \right)^+, \quad (23)$$

where $\tilde{X}$ denotes a modified design matrix whose $C_{m,s}$ rows corresponding to the zero weights $(\tilde{\beta}_{m,s})_r = 0$ are set to zero and the superscript $^+$ indicates pseudoinversion.

S8. Convergence analysis of the projected gradient descent

For completeness, we justify mathematically the algorithm (20) and the zero-weight criterion (22) used to solve the sign-constrained GLM. To simplify notations, we drop the subscripts m, s and the tilde symbol so we use the $1 \times R$ weight variable β with components β_r. Each step of the iterative scheme (20) can thus be written as $\beta \to \beta'$ with

$$\beta'_r = \epsilon \left| (f(\beta))_r \right|, \quad (24a)$$

$$f(\beta) = \beta - \frac{\tau}{\Lambda} (\beta X - Y) X^{\text{T}}. \quad (24b)$$

Here, ϵ is a fixed sign, the step size τ is a scalar, X is a $R \times N$ matrix, $\Lambda > 0$ is the maximum eigenvalue of XX^{T}, and Y is a $1 \times N$ row vector. We demonstrate that

(i) any sequence $\beta \to \beta' \to \beta'' \to \ldots$ (i.e., for any initial condition) converges towards a unique weight configuration β^* when $0 < \tau < 2$, and

(ii) the limit β^* obtained at fixed $\tau > 0$ tends to the solution (23) of the sign-constrained GLM when $\tau \to 0$.

For the part (i), we show that the function $\beta \to \beta'$ is contracting, i.e., for any pair of weights β_1, β_2,

$$\| \beta'_1 - \beta'_2 \| \le \ell \, \| \beta_1 - \beta_2 \| \qquad \text{with} \qquad \ell < 1. \quad (25)$$

This claim is an adaptation of Lemma 3 in Jung (2017) to our sign-projected case. We start from the elementary observation that $(\epsilon |a| - \epsilon |b|)^2 = (|a| - |b|)^2 \le (a - b)^2$, so using equation (24a)

$$(\beta'_{1\,r} - \beta'_{2\,r})^2 \le \left[(f(\beta_1))_r - (f(\beta_2))_r \right]^2. \quad (26)$$

Summing over r, taking the square root, and using the definition (24b) then leads to

$$\| \beta'_1 - \beta'_2 \| \le \| f(\beta_1) - f(\beta_2) \| = \left\| (\beta_1 - \beta_2)(1 - \frac{\tau}{\Lambda} XX^{\text{T}}) \right\|. \quad (27)$$

The last term is bounded by $\ell \, \| \beta_1 - \beta_2 \|$, where ℓ is the absolute value of the largest-magnitude eigenvalue of the $R \times R$ matrix $1 - \frac{\tau}{\Lambda} XX^{\text{T}}$. Its extreme eigenvalues being $1 - \tau$ and $1 - \tau/\kappa$ (with $\kappa \ge 1$ denoting the condition number of XX^{T}, see also Fig. S3), we find

$$\ell = \max \{ |1 - \tau|, \, |1 - \tau/\kappa| \}. \quad (28)$$

The contraction property (25) is thus verified if $\ell < 1$, which requires $0 < \tau < 2$. It then follows from the contraction mapping theorem (see, e.g., Lemma 2 in Jung, 2017) that any sequence $\beta \to \beta' \to \beta'' \to \ldots$ converges to the unique fixed point β^* of the iterate $\beta \to \beta'$.

To prove the part (ii), we solve explicitly the fixed-point equations $\beta^*_r = \epsilon \left| (f(\beta^*))_r \right|$. Using the definition (24b), we have

$$\beta^*_r = \pm \left(\beta^*_r - \frac{\tau}{\Lambda} [(\beta^* X - Y) X^{\text{T}}]_r \right) \quad (29)$$

with the sign chosen to ensure the constraint $\epsilon \beta^*_r \ge 0$. Thus

$$[(\beta^* X - Y) X^{\text{T}}]_r = 0 \qquad \text{for the + sign}, \quad (30a)$$

$$2 \beta^*_r = \frac{\tau}{\Lambda} [(\beta^* X - Y) X^{\text{T}}]_r \qquad \text{for the − sign}. \quad (30b)$$

When $\tau \to 0$, we find that either $[(\beta^* X - Y) X^{\text{T}}]_r = 0$ or $\beta^*_r = 0$, which corresponds to the solution (23).

For small but nonzero step size $0 < \tau \ll 1$, the fixed point differs slightly from the exact solution because the would-be
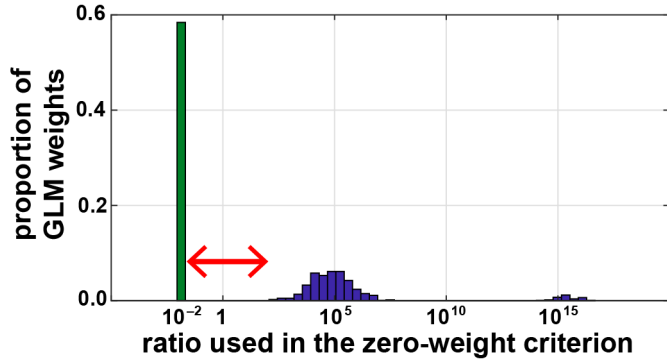
Figure S4: *Illustration of the separation of scales principle.* The histogram depicts the distribution of ratios (31) for the weights obtained from the projected gradient descent algorithm (20, 21) across all the DMN modes m, seeds s, and components r in the α and the β bands. A base-10 logarithmic scale is used on the ratio axis. The values below one are concentrated around $\tau = 10^{-2}$ (green bin) and those above one spread over a large range (blue bins). The gap between them is emphasized by a red double-headed arrow.

zero weights are of order τ (equation 30b). These can be identified unambiguously by considering the expression

$$\left| \frac{2\Lambda\beta_r}{[(\boldsymbol{\beta X} - \boldsymbol{Y})\boldsymbol{X}^{\mathrm{T}}]_r} \right| \tag{31}$$

used in the criterion (22). At the fixed point $\boldsymbol{\beta} = \boldsymbol{\beta}^*$, this ratio is infinite for the components corresponding to the nonzero weights (equation 30a) and equals to τ for the approximately zero weights (equation 30b). For a sufficiently accurate approximation such as $\boldsymbol{\beta} = \tilde{\boldsymbol{\beta}}_{m,s}^{(N)}$, this ratio is proportional to $1/\delta \gg 1$ (where δ measures the convergence precision, see equation 21) in the former case and to the step size $\tau \ll 1$ in the latter case. In other words, the distribution of values for the expression (31) must exhibit a substantial gap around one. In our data, this gap spanned about four orders of magnitude (see Fig. S4). This separation of scales justifies the zero-weight criterion (22).

Supplementary Results

S1. Application to static α- and β-band connectivity

To test our network mixture model paradigm and illustrate the dynamical nature of the RSN competition discussed in the main text, we examined the static FC maps using an ICA similar to that used to identify modes of dynamic coupling, and then applied network mixture modeling to disclose ICs corresponding statistically to RSN templates.

To derive static FC maps in the α and the β bands, we applied a group ICA to temporally concatenated individual source Hilbert envelope time series (low-pass filtered to 2 Hz and standardized to zero mean and unit variance for each subject separately) using FastICA with parameters published previously for MEG-based ICA investigations of RSNs (Brookes et al., 2011; Wens et al., 2014b), i.e., dimension reduction to the 25 first principal components, nonlinearity *tanh*, and number of components to compute set to 20. Of note, the dimension reduction discarded about 50% of envelope data variance, a fact that was

used in the main text to set the dimension reduction parameter for dynamic FC data. The IC maps were then derived by means of temporal correlation between the source envelope and each IC time series and further statistically thresholded as done for the modes of dynamic coupling.

To classify IC maps automatically, we applied network mixture modeling to this case. Of note, there is only one map per IC so each mixture model is actually equivalent to a single spatial GLM. Although the multivariate setup was designed specifically to investigate linear superpositions of several RSN templates in dynamic FC, we naturally expected the resulting classification to associate IC maps to no more than one RSN.

Figures S5 (α band) and S6 (β band) depict all the resulting IC maps after statistical masking. It is interesting to note that the masks covered a large majority of the cortex for most ICs, indicating that most connections appear significant with static FC. This is presumably not a mere effect of spatial leakage (see, e.g., the simulations in Wens et al., 2015) but may rather reflect the subdominant contribution of the transient cross-network couplings studied in this paper.

Figures S7 and S8 report on their network mixture model analysis. Their top row establishes GLM significance, i.e., whether the IC map resembles one or a combination of RSNs, and their second and third rows identify the RSN patterns. By and large, this analysis confirmed previous findings (Brookes et al., 2011; Wens et al., 2014b). It is noteworthy that several successfully classified ICs involved more than one RSN template, but this fact merely illustrates the specificity limitation associated with the spatial blurriness of IC maps. The bottom row of Figs. S7 and S8 shows that dynamical competition testing failed to detect any significant competitive behavior.

S2. Analysis of DMN modes in the α and the β bands

In the main text, our analysis of key spatio-temporal features across the modes of dynamic coupling highlighted the DMN in the α and the β bands. Here, we focus on these special cases and detail the modes of dynamic coupling.

Temporal characteristics. Some temporal properties of the DMN modes are illustrated in Fig. S9 (top: α band, bottom: β band). The plots confirm the aperiodicity of IC times series (Fig. S9, left), the significance of IC excess kurtosis for all the modes (Fig. S9, middle), and the transience of the IC large deviations (Fig. S9, right). The distributions of kurtosis and lifetime also confirm the observation that the higher median values in the α band compared to the β band (see Fig. 3, left) was due to the presence of a few α-band modes with higher values than the β-band modes.

Spatial localization and cross-network integration. The masked IC maps locating the couplings significantly involved in the DMN modes are shown in Figs. S10 (α band) and S11 (β band).

Their description in terms of cross DMN-RSN coupling based on network mixture models is depicted in Figs. S12 (α band) and S13 (β band). Their top row establishes model significance of the spatial GLM for each mode and each seed. There were 14/20 significant mixture models in the α band and 18/20 in the β band (see also Fig. 3, bottom right). These 32 modes
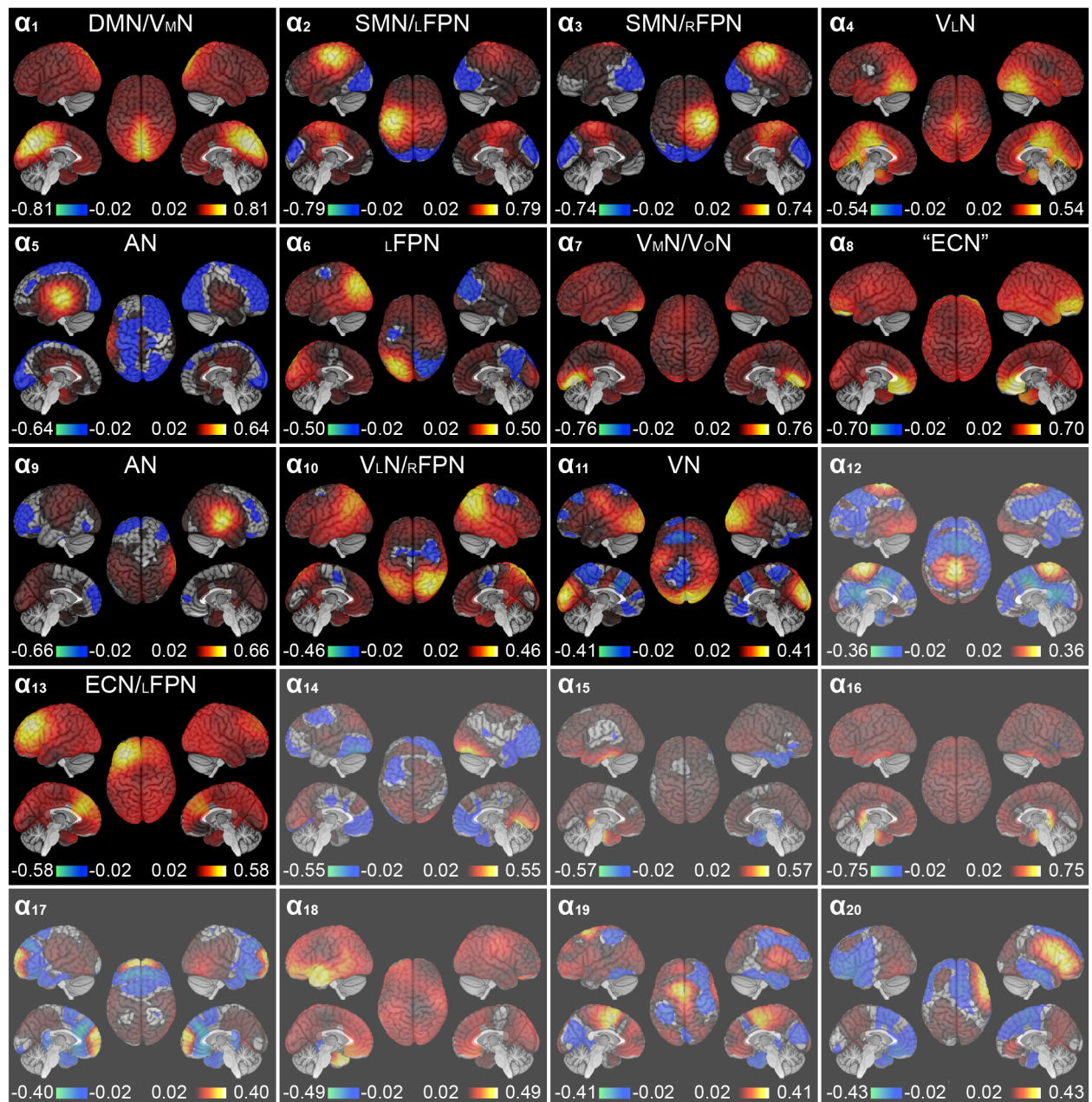
Figure S5: *Static α-band connectivity maps.* The spatial maps of all ICs are shown for the *α*-band static ICA. All maps were statistically masked at $p < 0.05$ with FWER correction. For each map, only the relative sign matters. Those ICs associated with a significant GLM are named according to the univariate maps correlation *t* tests (Fig. S7, third row). The others are shaded.
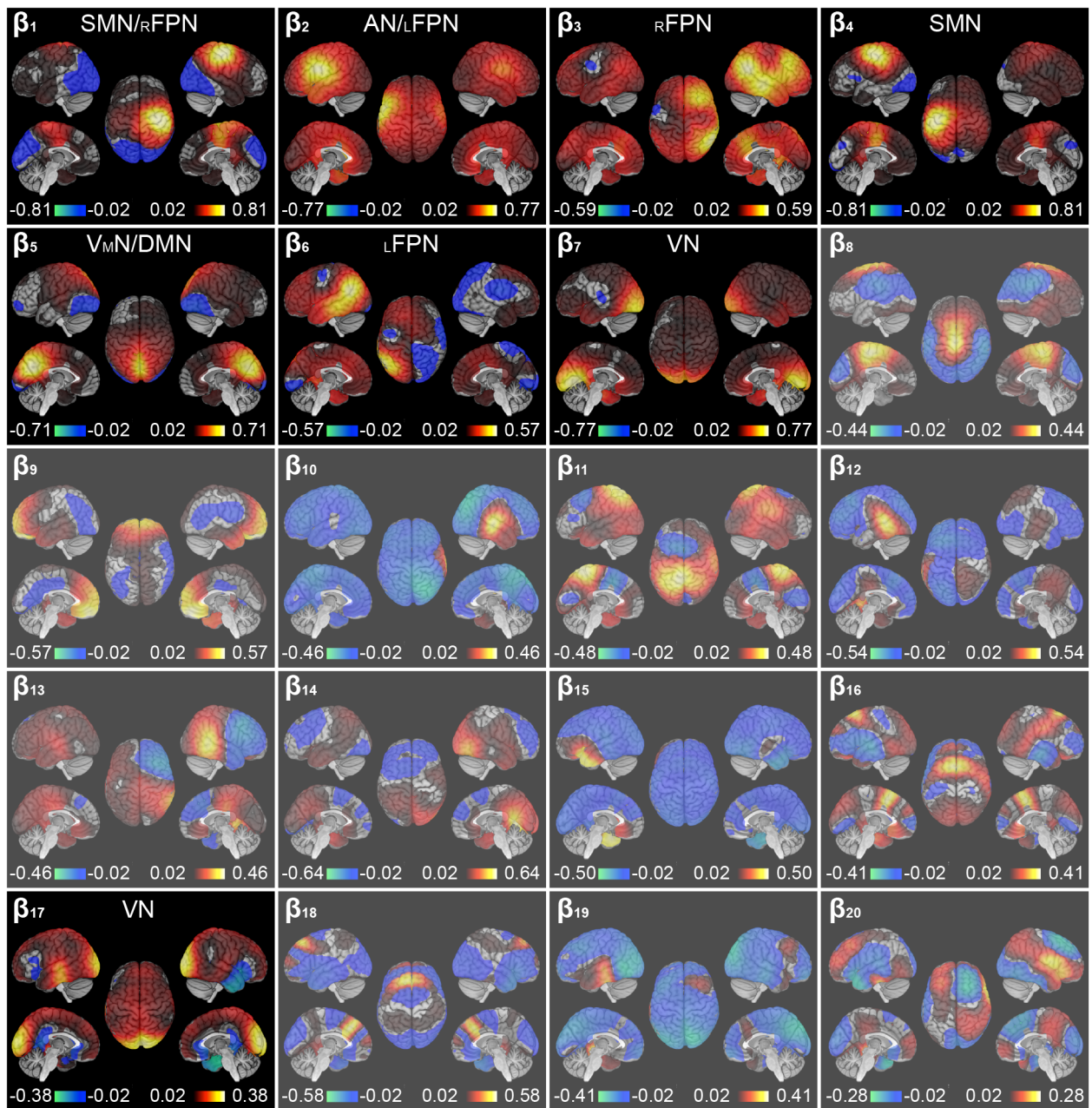
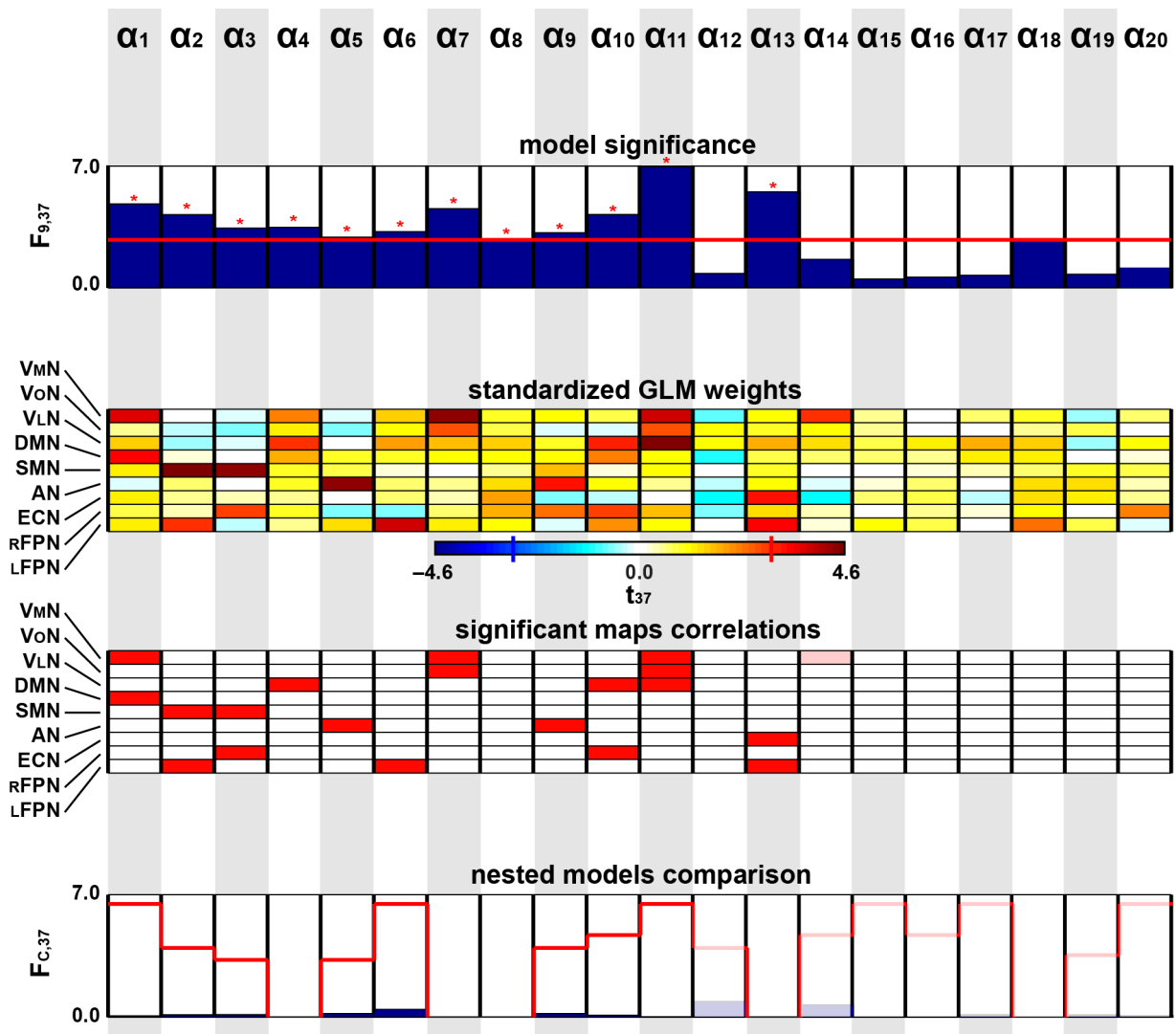Figure S6: *Static β-band connectivity maps.* All is as in Fig. S5.

Figure S7: *Statistical inferences on network mixture models for static α-band connectivity.* Statistical results of the spatial GLM analysis are depicted for the twenty α-band static IC maps. The **first row** shows the F statistics values to establish significance of each spatial GLM (i.e., for each static IC $\alpha_1, \ldots, \alpha_{20}$). The significance threshold computed at $p < 0.05$ with FDR correction is emphasized by a red line and significant models, by red stars. The **second row** details all GLM weights after noise normalization (i.e., their t values). The statistical threshold (shown on scale) was computed at the same significance level than the F tests on the top. The **third row** identifies the significant t tests (red: positive maps partial correlation, blue: negative maps partial correlation). Of notice, results within a non-significant GLM (first row) were not considered and thus appear shaded. The **fourth row** shows the F statistics for the nested GLM comparison tests used to disclose RSN competition, together with the significance thresholds derived at the same level than the F tests on the top (red curve). As above, red stars indicate significance and the cases associated with non-significant models are shaded.
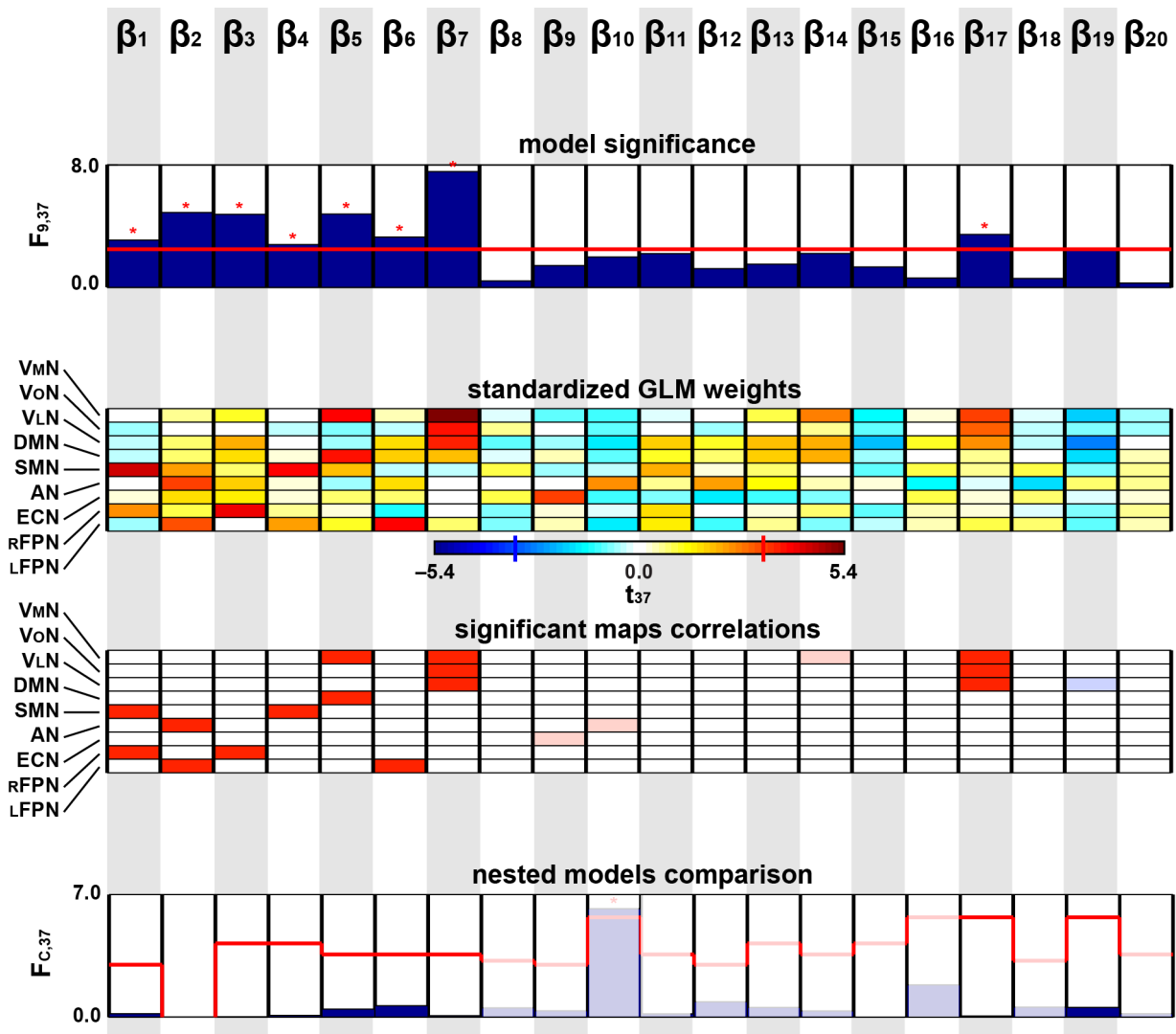
Figure S8: *Statistical inferences on network mixture models for static β-band connectivity.* All is as in Fig. S7.
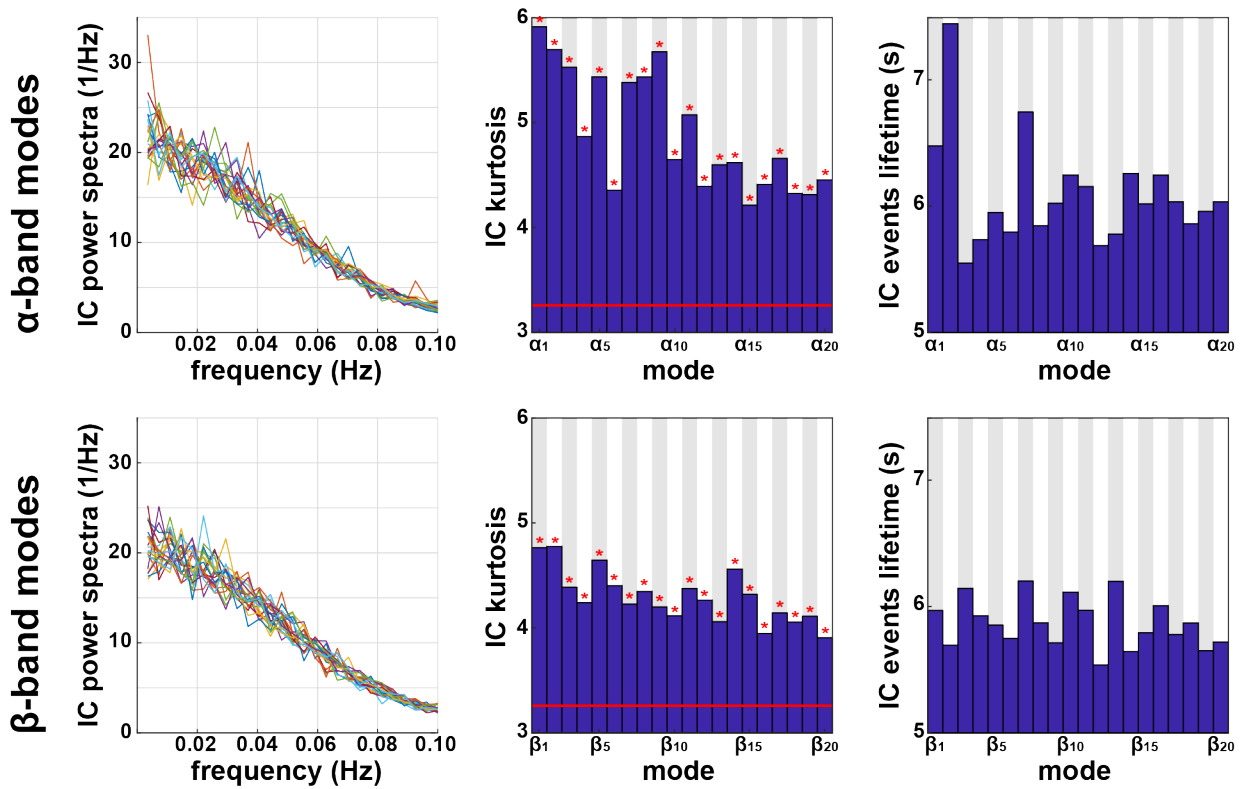
Figure S9: *Temporal properties of the modes of dynamic couplings for the DMN in the α and the β bands.* The power spectra (**left**), kurtosis (**middle**) and lifetime of large-deviation events (**right**) are shown for both the *α*-band (**top**) and the *β*-band (**bottom**) IC time series. The twenty power spectra were normalized to a unit area under the curve (corresponding to the unit variance of IC time series) and superimposed on each other. For the kurtosis, the significance thresholds at $p < 0.05$ (Bonferroni corrected for twenty comparisons) are emphasized by red lines and significant values, by red stars.
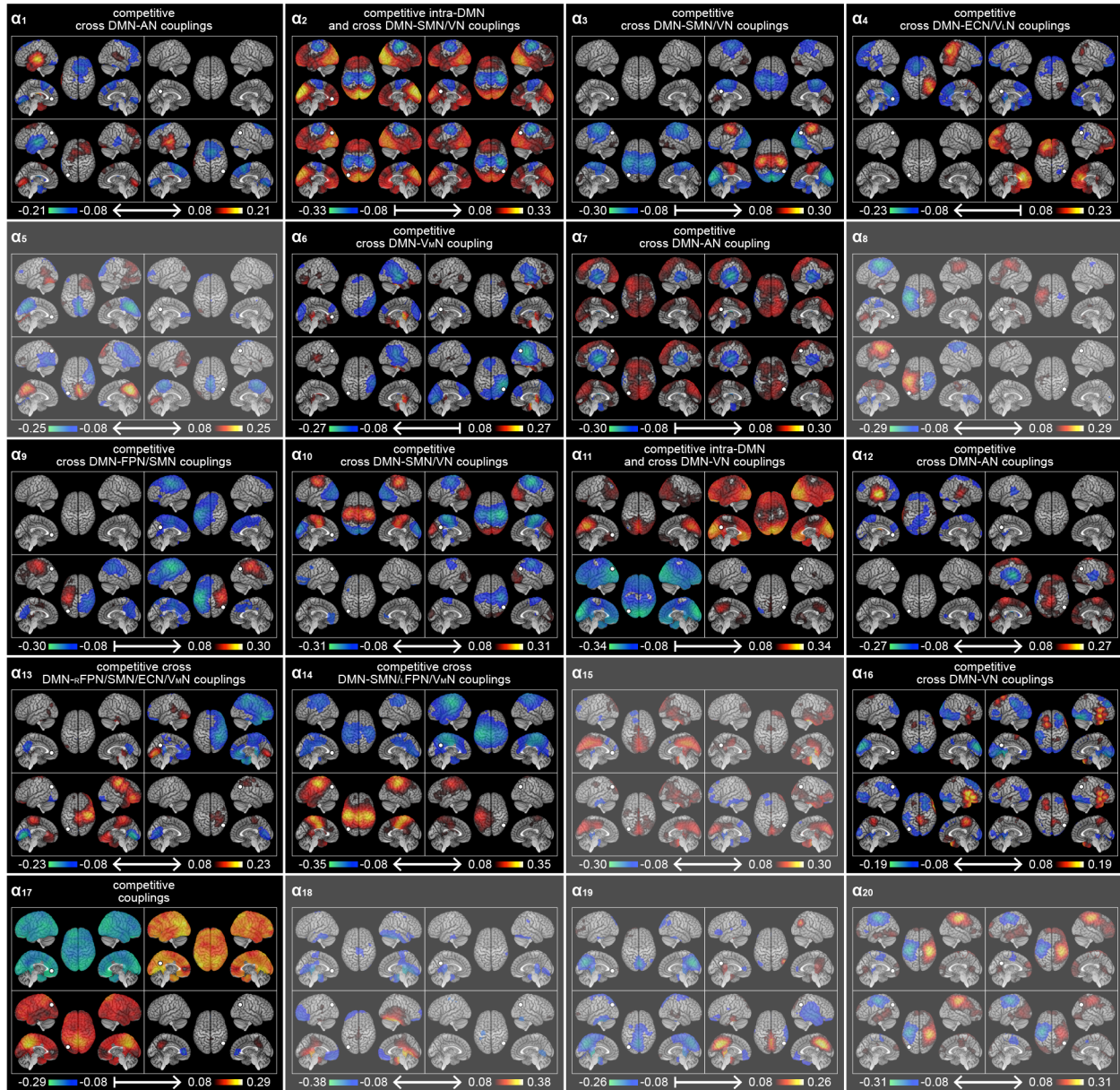
Figure S10: *Spatial signature of the modes of dynamic coupling for the α-band DMN.* All IC maps are shown with statistical masking at $p < 0.05$ with FWER correction. The right (left) arrows between correlation scales indicate significant skewness towards positive (negative) values and double-headed arrows, non-significant skewness. Those ICs associated with a significant network mixture model are named according to the univariate maps correlation t tests (Fig. S12, third row) and the dynamical competition F tests (Fig. S12, fourth row). The others are shaded.
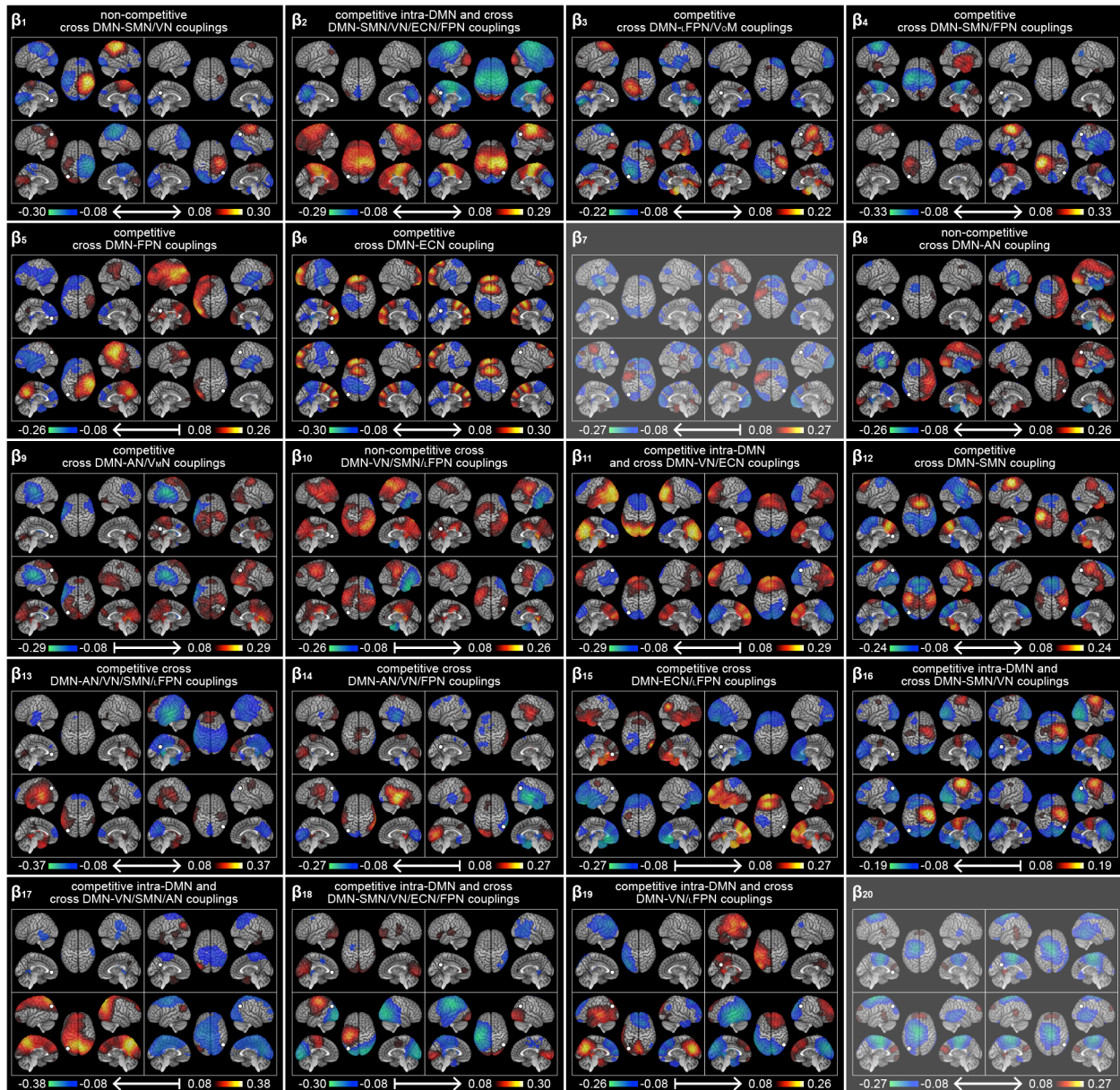
Figure S11: *Spatial signature of the modes of dynamic coupling for the β-band DMN.* All is as in Fig. S10.

could thus be classified in terms of cross DMN-RSN interactions, which were then quantified using the standardized GLM weights (Figs. S12 and S13, second row). To help identifying the RSNs predominantly involved, we also show a statistically thresholded version (Figs. S12 and S13, third row; red: significantly positive partial correlation, blue: significantly negative partial correlation). All but one significant mixture model identified dynamic integration between the DMN and at least one RSN.

Detailed description of cross-network patterns. We identified cross DMN-SMN integration (i.e., with the SMN only) in the mode β_{12} as well as α_9 and β_4. The correlation tests for the latter two also reported contributions from DMN-FPN couplings, but they appeared to be false positives due to IC maps blurriness (see Figs. S10, S11 and the specificity issue mentioned in the main text). Likewise, the mixture models disclosed cross DMN-AN (modes α_1, α_7, α_{12}, and β_8), cross DMN-VN (α_6 and α_{16}), cross DMN-FPN (β_5), and cross DMN-ECN (β_6) interactions. We observed DMN-Vм/ʟN couplings accompanied with intra-DMN couplings as well (α_{11}), but here too IC maps blurriness prevented clear discrimination of the posterior part of the DMN and the Vм/ʟNs.

We also identified cross-network integration involving clearly distinguishable RSNs. Cross DMN-SMN and DMN-VN integration emerged for the modes α_3, α_{10}, and β_1 as well as α_{14} and β_{18} (again with spurious FPN contributions, see Figs. S10, S11). Such DMN-SMN/VN couplings also appeared in association with intra-DMN couplings for α_2, β_{16}, and β_2, the latter with DMN-ECN integration too. As above, these extra contributions may be related to a lack of specificity in clearly disambiguating the anterior/posterior DMN from the ECN/VN in our IC maps. We also observed cross DMN-AN/VN (mode β_9 as well as β_{13} and β_{14} with spurious SMN/FPN contributions, see Fig. S11), DMN-VN/ECN (α_4 and β_{11} with possibly spurious intra-DMN contribution, see Fig. S11), DMN-FPN/VN (β_3, α_{13} and β_{10} with spurious SMN/ECN contributions, see Figs. S10, 4, and β_{19} with possible intra-DMN contribution, see Fig. S11), and DMN-FPN/ECN (β_{15}). The mode β_{17} mixed cross DMN-AN, DMN-SMN, DMN-VN, and intra-DMN couplings, the latter two contributions being hard to discriminate (Fig. S11).

The only mode with significant mixture model without a significant contribution from a single RSN was α_{17}. This is because its IC maps spread across the whole brain (Fig. S10) so the GLM coefficients were roughly similar and a mixing of RSNs was genuinely required.

Eight mixture models did not reach significance, but this does not preclude physiological meaningfulness in view of the sensitivity issue associated with RSNs splitting into sub-networks. In fact, the modes α_8 and α_{20} could be understood in terms of dynamic integration between the DMN and the intra-hemispheric parts of the SMN, and the modes α_5, α_{15}, and α_{19}, in terms of intra-DMN interactions specifically targeting the PCC (Fig. S10). Only the modes α_{18}, β_7, and β_{20} remained difficult to interpret on the basis of classical RSNs (Figs. S10, S11).

Dynamical competition. Results of the dynamical competition tests in the α and the β bands are depicted in the fourth row of Figs. S12 and S13, respectively. We identified 12/14 significant α-band mixture models and 15/18 significant β-band mixture models that exhibited significant dynamical competition (see also Fig. 4, bottom).

S3. More on the large-deviation events

The recurrence of IC large deviations and their tendency to occur in isolation rather than simultaneously is shown in Fig. S14 (top left) for the β-band DMN modes. The occurrence count used to obtain Table 2 is illustrated in Fig. S14 (bottom left).

The inference made in the main text that higher IC kurtosis and longer lifetime of IC large deviations for the α-band modes (Fig. 3, left) implies that the corresponding deviations must be larger, was confirmed directly by considering the mean absolute value of IC times series during their large-deviation events (Fig. S14, right). We observed a band effect (Friedman test, $p = 9.1 \times 10^{-14}$) explained by larger deviations among the α-band modes, and no RSN effect ($p = 0.28$), hence establishing our claim.

S4. Simulated bistable system of two connectivity states

In this section, we present a simulation of dynamic FC data and examine whether our analysis leads to meaningful modes of dynamic coupling. The aim is twofold:

(i) support our interpretation of the modes in terms of FC state transitions (see inference ii in the main text), and

(ii) show the ability of our ICA to extract content from noisy dynamic FC (see also Supplementary Methods S2, S3).

We generated 100 synthetic MEG recordings (one per subject of our experimental dataset) based on sources exhibiting time-varying cross-network couplings. Specifically, we considered a two-state configuration where one state of cross DMN-SMN coupling alternates with another state of cross DMN-AN coupling. Each state was first built as a configuration of static FC among six point sources chosen at the key nodes of the DMN and of the SMN or the AN (see Table 1 in the main text). Their orientation was fixed along the direction of maximum MEG response (as assessed by the individual forward model), and their time course was obtained from independent realizations of a band-filtered Gaussian white noise (α band, sampling rate: 200 Hz, duration: 5 min). Envelope correlation among the nodes was imposed by pointwise multiplication of the source signals with a common 1-Hz sinusoid. Time-dependent FC was then obtained by splitting the 5 minutes into windows, to which state-specific source time courses were assigned in alternation. Transition times between successive windows were random, but their total number was fixed so as to reproduce the occurrence rate of Table 2. The synthetic MEG data were finally obtained by individual forward projection of the resulting source configuration and addition of realistic sensor noise taken from empty-room MEG data (rescaled to reproduce the signal-to-noise ratio of our experimental dataset).
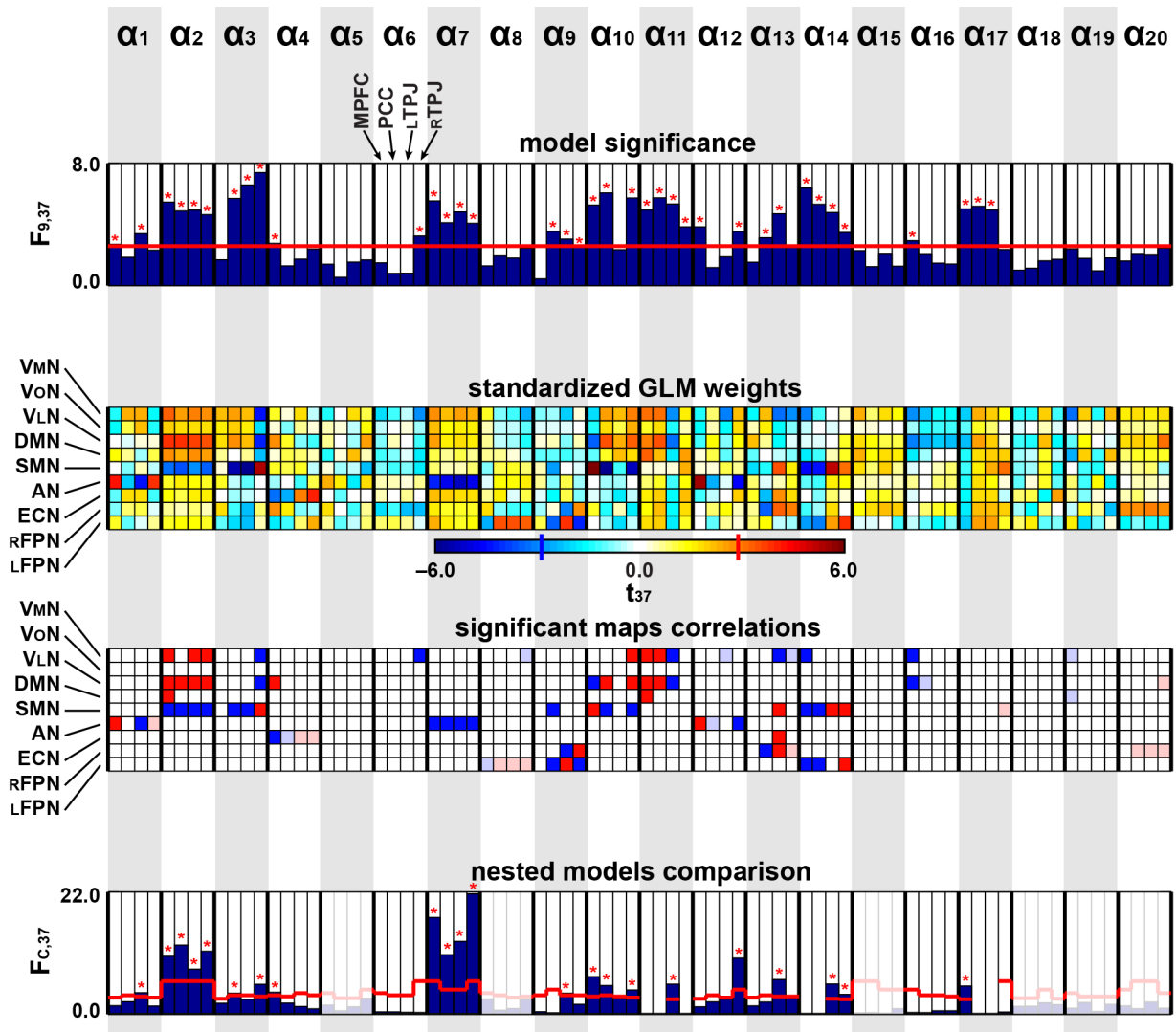
Figure S12: *Statistical inferences on network mixture models for the α-band DMN.* Statistical results of network mixture modeling on the twenty modes of dynamic coupling are depicted. The **first row** shows the *F* statistics values to establish significance of each spatial GLM (i.e., for each mode $\alpha_1, \ldots, \alpha_{20}$, and each seed MPFC, PCC, ʟTPJ, ʀTPJ). The significance threshold computed at $p < 0.05$ with FDR correction is emphasized by a red line and significant models, by red stars. The **second row** details all GLM weights after noise normalization (i.e., their *t* values). The statistical threshold (emphasized on scale) was computed at the same significance level than the *F* tests on the top. The **third row** emphasizes the significant *t* tests (red: positive spatial correlation, blue: negative spatial correlation). Of notice, results within a non-significant GLM (first row) were not considered and thus appear shaded. The **fourth row** shows the *F* statistics for the nested GLM comparison tests used to disclose dynamical competition, together with the significance thresholds derived at the same level than the *F* tests on the top (red curve). As above, red stars indicate significance and the cases associated with non-significant models are shaded.
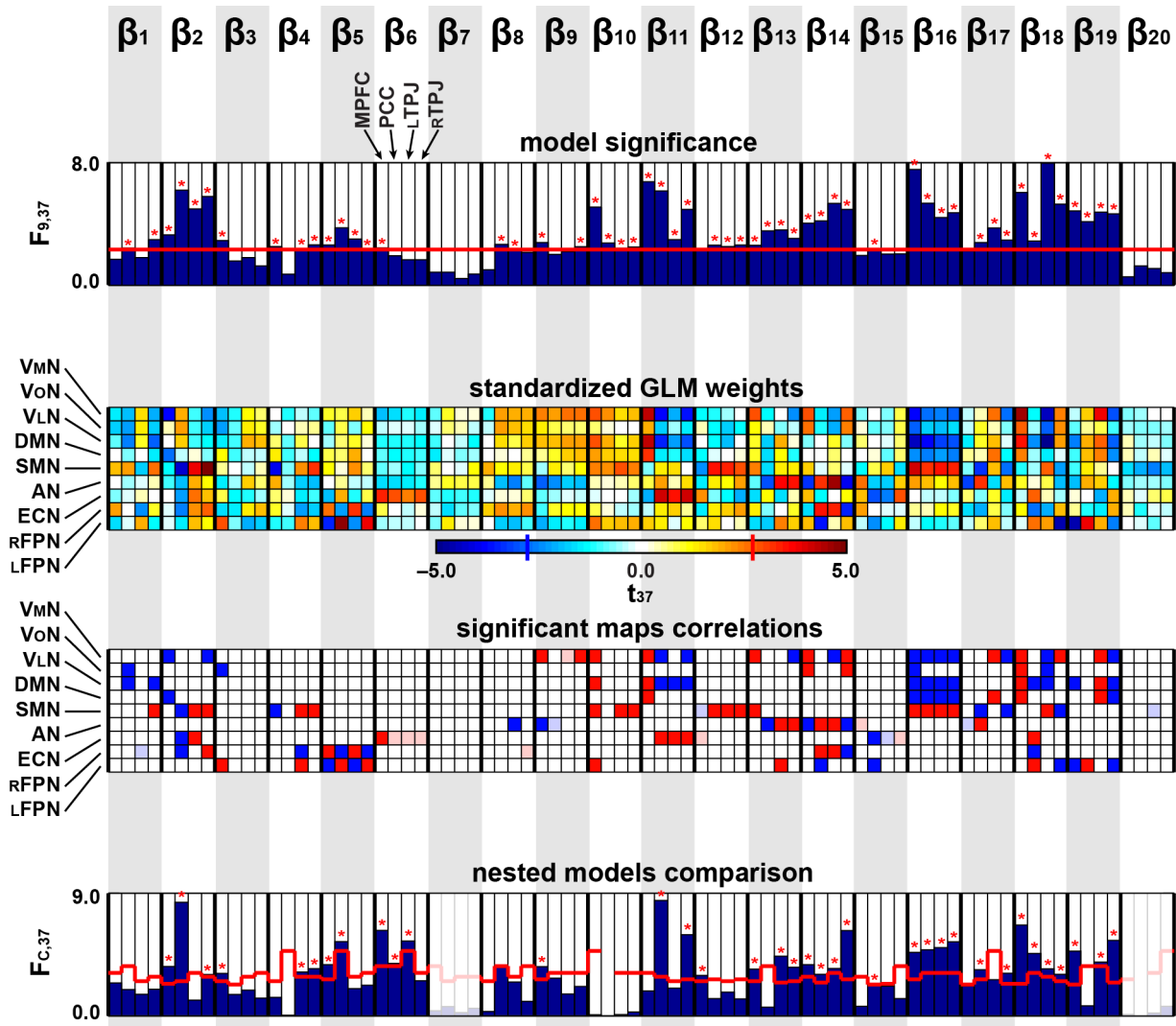
Figure S13: *Statistical inferences on network mixture models for the β-band DMN.* All is as in Fig. S12.
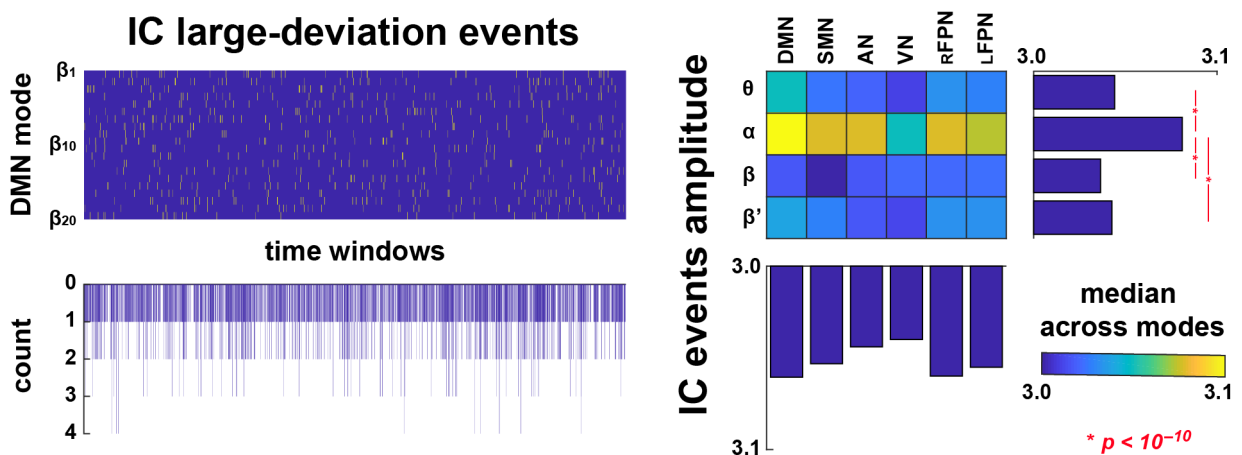

Figure S14: *Extra characteristics of IC large deviations.* The large-deviation events are illustrated in the case of the DMN modes in the β band (**top left**), as well as the associated co-occurrence count of those events (**bottom left**). The distributions across modes of the mean amplitude of IC large deviations (**right**) are compared across frequency bands (θ, α, β, β') and seed RSNs (see Table 1). The bar plots show the median values computed over RSNs (right side) or bands (bottom side). Significant effects identified via post-hoc Wilcoxon rank tests are indicated on these bar plots by red stars.

This dataset was then analyzed exactly as described in the main text. Given our interpretation in terms of state transitions, we expected that one IC would identify competitive DMN-SMN and DMN-AN couplings. This prediction was confirmed by the results summarized in Fig. S15. To detect relevant modes, we first plotted the mean IC maps value across the state-specific connections that were simulated (Fig. S15, left). Only IC_5 (i.e., the fifth IC obtained) exhibited substantial correlations, indicating that the expected state transition process was mostly captured by a single IC. The corresponding raw IC maps reproduced the pattern of FC change between DMN-SMN and DMN-AN states (Fig. S15, right). This provides proof of concept for the two claims (i) and (ii).

S5. Absence of dynamic power biases and muscle artifacts

To exclude possible contributions of dynamic power variations to our modes, we considered the temporal correlation between the IC time series and the sources power time courses. Dynamic power was estimated within the same sliding windows than dynamic FC. We used sLORETA noise normalization (Pascual-Marqui, 2002) to correct for the depth bias of MNE. For conciseness, we focus here on power fluctuations at key source locations: (i) the seeds used to compute seed-based FC (see Table 1 in the main text) and (ii) the source exhibiting the highest value in each IC map (i.e., maximum absolute IC-FC correlation).

Figure S16 shows the maximum absolute correlation value across the modes and seeds, for each RSN and frequency band. Only small values were disclosed, showing that IC time series are not driven by power fluctuations. In fact, the correlations appearing in this analysis were well below the significance threshold derived for the IC maps (about 0.08, see main text).

Similarly, since our data preprocessing did not explicitly control for muscle artifacts, we estimated the correlation between the IC time series and the high-frequency power time course of the MEG recordings, which is a proxy for muscle activity. The latter was estimated within each sliding window as the mean (across sensors) variance of the MEG signals filtered between 110 and 140 Hz. The resulting correlation values did not exceed 0.05, so we conclude that IC time series were not driven by muscle activity.

S6. Comparison to zero-order correlation analyses

To confirm the absence of interpretation ambiguities and suppression effects in our GLM (see Supplementary Methods S5 for a theoretical argument), we repeated our spatial correlation analyses using zero-order (i.e., univariate) correlation t tests. Formally, we applied the spatial GLM with one regressor at a time, and set significance thresholds at $p < 0.05$ with FDR correction. For illustrative purposes, we focus again on the α- and the β-band modes of the DMN.

Figure S17 (left) shows that the normalized weight values were globally similar in the univariate and the multivariate modalities. The statistical results of the zero-order analysis are detailed in Fig. S18, and comparison with the second and third rows of Figs. S12 and S13 further confirms the global similarity of significance patterns. The only notable exception concerns the mode α_{17}, whose GLM was significant but without significant GLM weights (see Supplementary Results S2 and Fig. S12) while several zero-order weights appeared significant. This illustrates the difference between plain and partial correlations.

Given the importance of weights sign for the dynamical competition test, we checked explicitly the consistency of weight signs across the two modalities. The histograms in Fig. S17 (right) show that only weights with small effect size ($|t| < 1.25$) disclosed a sign reversal. For comparison, note that significance thresholds were above 2. About one third of the weights falling in the central bin ($-0.25 \leq t \leq 0.25$) disclosed sign reversal due to their very small effect sizes. This proportion decreased in the next bins and vanished for $|t| > 1.25$, indicating that this randomness is gradually replaced by a stable sign estimate at larger effect sizes. These results corroborate the interpretation rule (ii) derived theoretically in the Supplementary Methods S5.

References

Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., Calhoun, V. D., 2014. Tracking whole-brain connectivity dynamics in the resting state. Cereb Cortex 24, 663–676.

Baars, B. J., Ramsoy, T. Z., Laureys, S., 2003. Brain, conscious experience and the observing self. Trends Neurosci 26, 671–675.

Baker, A. P., Brookes, M. J., Rezek, I. A., Smith, S. M., Behrens, T., Probert Smith, P. J., Woolrich, M. W., 2014. Fast transient networks in spontaneous human brain activity. Elife 3, e01867.

Bastos, A. M., Schoffelen, J. M., 2015. A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. Front Syst Neurosci 9, 175.

Beckmann, C. F., DeLuca, M., Devlin, J. T., Smith, S., 2005. Investigations into resting-state connectivity using independent component analysis. Philos Trans R Soc Lond B Biol Sci 360, 1001–1013.

Betti, V., Corbetta, M., de Pasquale, F., Wens, V., Della Penna, S., 2018. Topology of functional connectivity and hub dynamics in the beta band as temporal prior for natural vision in the human brain. J Neurosci 38 (15), 3858–3871.

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

Bourguignon, M., De Tiège, X., Op de beeck, M., Ligot, N., Paquier, P., Van Bogaert, P., Goldman, S., Hari, R., Jousmäki, V., 2013. The pace of prosodic phrasing couples the listener's cortex to the reader's voice. Hum Brain Mapp 34, 314–326.

Bourguignon, M., De Tiège, X., Op de beeck, M., Pirotte, B., Van Bogaert, P., Goldman, S., Hari, R., Jousmäki, V., 2011. Functional motor-cortex mapping using corticokinematic coherence. Neuroimage 55, 1475–1479.

Bressler, S. L., Richter, C. G., 2015. Interareal oscillatory synchronization in top-down neocortical processing. Curr Opin Neurobiol 31, 62–66.

Brookes, M. J., Liddle, E. B., Hale, J. R., Woolrich, M. W., Luckhoo, H., Liddle, P. F., Morris, P. G., 2012a. Task induced modulation of neural oscillations in electrophysiological brain networks. Neuroimage 63, 1918–1930.

Brookes, M. J., O'Neill, G. C., Hall, E. L., Woolrich, M. W., Baker, A., Corner, S. P., SaE, R., Morris, P. G., Barnes, G. R., 2014. Measuring temporal, spectral and spatial changes in electrophysiological brain network connectivity. Neuroimage 91, 282–299.

Brookes, M. J., Woolrich, M., Luckhoo, H., Price, D., Hale, J. R., Stephenson, M. C., Barnes, G. R., Smith, S. M., Morris, P. G., 2011. Investigating the electrophysiological basis of resting state networks using magnetoencephalography. Proc Natl Acad Sci U S A 108, 16783–16788.

Brookes, M. J., Woolrich, M. W., Barnes, G. R., 2012b. Measuring functional connectivity in MEG: a multivariate approach insensitive to linear source leakage. Neuroimage 63, 910–920.
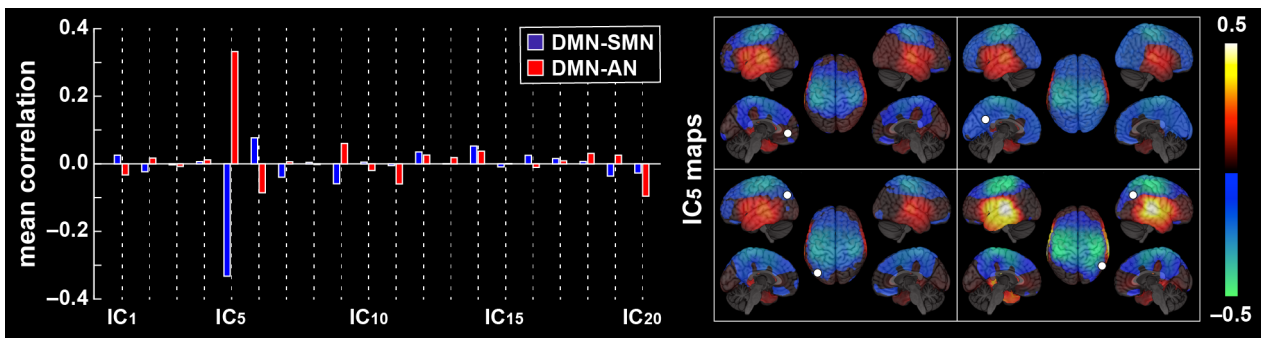
Figure S15: *Mode of dynamic coupling for simulated bistable dynamic FC data.* The IC maps values (i.e., IC-FC temporal correlations) were averaged over the connections corresponding to the two simulated FC states (red: DMN-SMN state, blue: DMN-AN state) for the twenty ICs (**left**). The raw IC maps obtained for IC$_5$ are also shown (**right**).
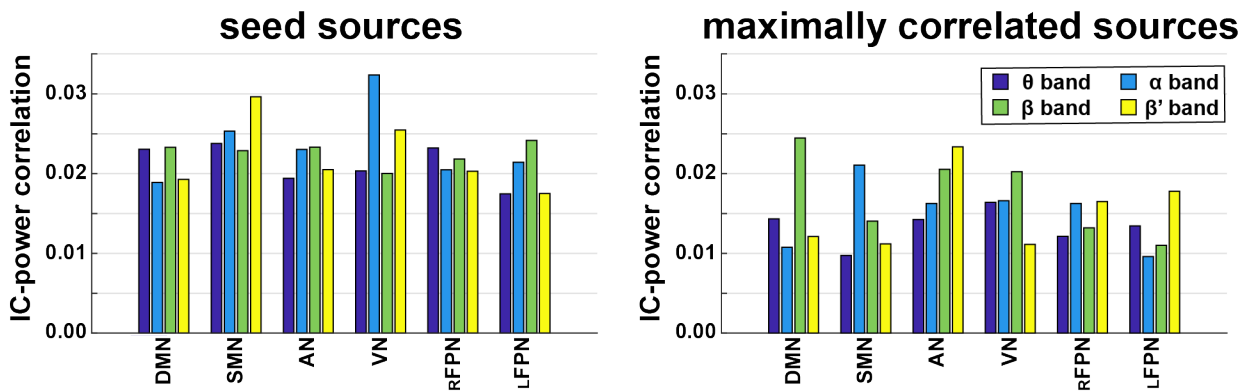

Figure S16: *Correlation analysis with dynamic power.* The maximum (across modes and seeds) absolute correlation between IC times series and power time course at the seed locations (**left**) and at the sources presenting maximum absolute IC-FC correlation (**right**) are shown for all seed RSNs (Table 1) and frequency bands.
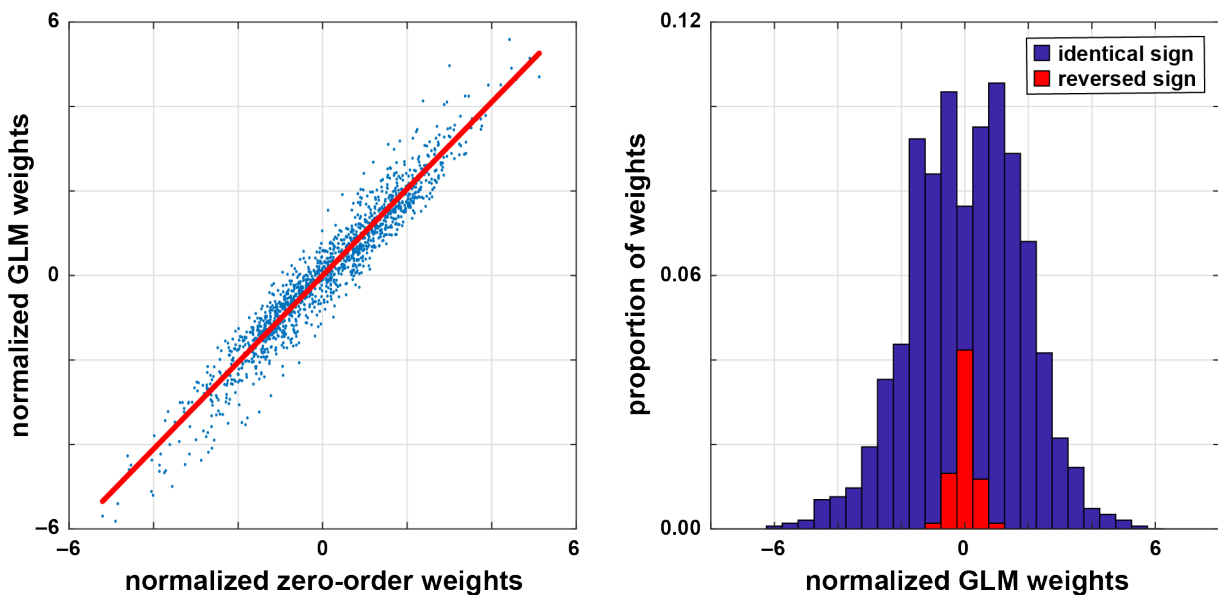

Figure S17: *Comparison between GLMs and zero-order correlations.* The association between the normalized weights (i.e., their *t* values) obtained with the multivariate GLMs and univariate zero-order correlations is illustrated with a scatter plot and a linear fit of estimated intercept 0.00 and slope 1.02 (**left**). The proportion of weights with the same sign in both approaches (blue bins) and of those with reversed sign (red bins) are quantified by two superimposed histograms (**right**). These plots were generated by pooling the weights across all DMN modes, seeds, and RSN templates in the α and the β bands.
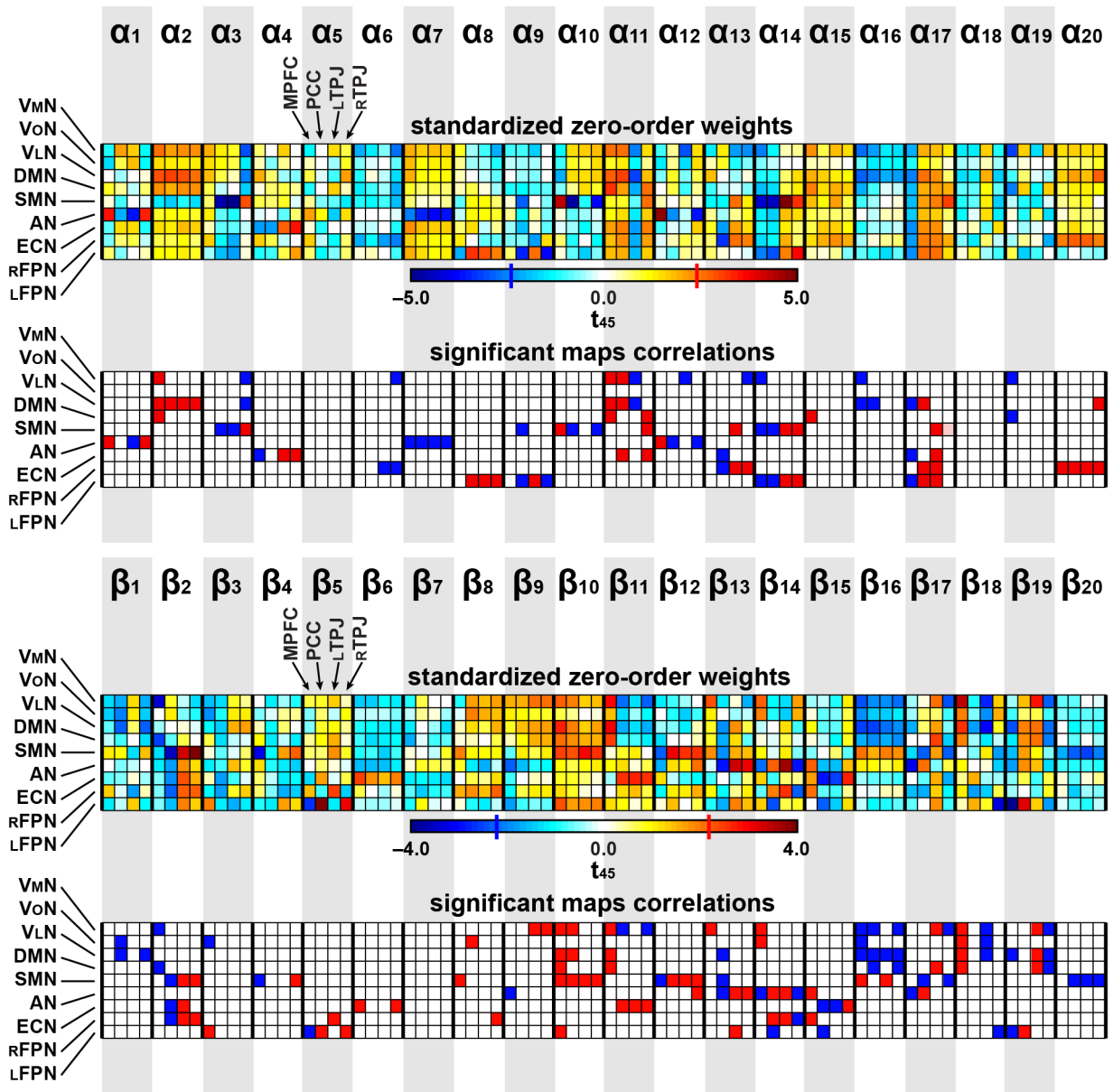
Figure S18: *Statistical inferences on zero-order weights for the DMN modes in the α and the β bands.* Statistical results for univariate, zero-order spatial correlations between each template RSN map and each IC map of the modes of dynamic coupling in the *α* (**top**) and the *β* (**bottom**) bands. The structure is the same as the second and third rows of Figs. S12, S13. Each significance threshold emphasized on scale was computed at *p* < 0.05 with FDR correction.

Buckner, R. L., Andrews-Hanna, J. R., Schacter, D. l., 2008. The brain's default network: anatomy, function, and relevance to disease. Ann N Y Acad Sci 1124, 1–38.

Calhoun, V. D., Potluru, V. K., Phlypo, R., Silva, R. F., Pearlmutter, B. A., Caprihan, A., Plis, S. M., Adali, T., 2013. Independent component analysis for brain FMRI does indeed select for maximal independence. PLoS One 8, e73309.

Chang, C., Glover, G. H., 2010. Time-frequency dynamics of resting-state brain connectivity measured with fmri. Neuroimage 50, 81–98.

Clumeck, C., Suarez Garcia, S., Bourguignon, M., Wens, V., Op de Beeck, M., Marty, B., Deconinck, N., Soncarrieu, M. V., Goldman, S., Jousmäki, V., Van Bogaert, P., De Tiège, X., 2014. Preserved coupling between the reader's voice and the listener's cortical activity in autism spectrum disorders. PLoS One 9, e92329.

Colgin, L. L., 2013. Mechanisms and functions of theta rhythms. Annu Rev Neurosci 36 (1), 295–312.

Damoiseaux, J. S., Rombouts, S. A., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., Beckmann, C. F., 2006. Consistent resting-state networks across healthy subjects. Proc Natl Acad Sci U S A 103, 13848–13853.

Daubechies, I., Roussos, E., Takerkart, S., Benharrosh, M., Golden, C., D'Ardenne, K., Richter, W., Cohen, J. D., Haxby, J., 2009. Independent component analysis for brain fMRI does not select for independence. Proc Natl Acad Sci U S A 106, 10415–10422.

de Gelder, B., Tamietto, M., Pegna, A. J., Van den Stock, J., 2015. Visual imagery influences brain responses to visual stimulation in bilateral cortical blindness. Cortex 72, 15–26.

de Pasquale, F., Corbetta, M., Betti, V., Della Penna, S., 2018. Cortical cores in network dynamics. Neuroimage 180, 370–382.

de Pasquale, F., Della Penna, S., Snyder, A. Z., Lewis, C., Mantini, D., Marzetti, L., Belardinelli, P., Ciancetta, L., Pizzella, V., Romani, G. L., Corbetta, M., 2010. Temporal dynamics of spontaneous MEG activity in brain networks. Proc Natl Acad Sci U S A 107, 6040–6045.

de Pasquale, F., Della Penna, S., Snyder, A. Z., Marzetti, L., Pizzella, V., Romani, G. L., 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, G. L., Corbetta, M., 2016. A dynamic core network and global efficiency in the resting human brain. Cereb Cortex 26 (10), 4015–4033.

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

Deco, G., Jirsa, V. K., 2012. Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. J Neurosci 32 (10), 3366–3375.

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

Deco, G., Jirsa, V. K., McIntosh, A. R., Sporns, O., Kötter, R., 2009. Key role of coupling, delay, and noise in resting brain fluctuations. Proc Natl Acad Sci U S A 106 (25), 10302–10307.

Engel, A. K., Fries, P., 2010. Beta-band oscillations – signalling the status quo? Curr Opin Neurobiol 20 (2), 156–165.

Fischl, B., 2012. FreeSurfer. Neuroimage 62, 774–781.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., Raichle, M. E., 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci U S A 102, 9673–9678.

Fox, M. D., Zhang, D., Snyder, A. Z., Raichle, M. E., 2009. The global signal and observed anticorrelated resting state brain networks. J Neurophysiol 101 (6), 3270–3283.

Friston, K. J., 2011. Functional and effective connectivity: a review. Brain Connect 1, 13–36.

Friston, K. J., Ashburner, J. T., Kiebel, S. J., Nichols, T. E., Penny, W. D., 2007. Statistical Parametric Mapping: The Analysis of Functional Brain Images. Academic Press.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., Frackowiak, R. S. J., 1994. Statistical parametric maps in functional imaging: a general linear approach. Hum Brain Mapp 2, 189–210.

Giacino, J. T., Fins, J. J., Laureys, S., Schiff, N. D., 2014. Disorders of consciousness after acquired brain injury: the state of the science. Nat Rev Neurol 10, 99–114.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Parkkonen, L., Hämäläinen, M. S., 2014. MNE software for processing MEG and EEG data. Neuroimage 86, 446–460.

Hall, E. L., Woolrich, M. W., Thomaz, C. E., Morris, P. G., Brookes, M. J., 2013. Using variance information in magnetoencephalography measures of functional connectivity. Neuroimage 67, 203–212.

Handwerker, D. A., Roopchansingh, V., Gonzalez-Castillo, J., Bandettini, P. A., 2012. Periodic changes in fMRI connectivity. Neuroimage 63, 1712–1719.

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

Hindriks, R., Adhikari, M. H., Murayama, Y., Ganzetti, M., Mantini, D., Logothetis, N. K., Deco, G., 2016. Can sliding-window correlations reveal dynamic functional connectivity in resting-state fMRI? Neuroimage 127, 242–256.

Hipp, J. F., Hawellek, D. J., Corbetta, M., Siegel, M., Engel, A. K., 2012. Large-scale cortical correlation structure of spontaneous oscillatory activity. Nat Neurosci 15, 884–890.

Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., Della Penna, S., Duyn, J. H., Glover, G. H., Gonzalez-Castillo, J., Handwerker, D. A., Keilholz, S., Kiviniemi, V., Leopold, D. A., de Pasquale, F., Sporns, O., Walter, M., Chang, C., 2013. Dynamic functional connectivity: promise, issues, and interpretations. Neuroimage 80, 360–378.

Hutchison, R. M., Womelsdorf, T., Gati, J. S., Everling, S., Menon, R. S., 2012. Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques. Hum Brain Mapp 34, 2154–2177.

Hyvärinen, A., Oja, E., 2000. Independent component analysis: algorithms and applications. Neural Netw 13, 411–430.

Jensen, O., Bonnefond, M., VanRullen, R., 2012. An oscillatory mechanism for prioritizing salient unattended stimuli. Trends Cogn Sci 16 (4), 200–206.

Jung, A., 2017. A fixed-point of view on gradient methods for big data. Front Appl Math Stat 3, 18.

Kiviniemi, V., Vire, T., Remes, J., Elseoud, A. A., Starck, T., Tervonen, O., Nikkinen, J., 2011. A sliding time-window ICA reveals spatial variability of the default mode network in time. Brain Connect 1, 339–347.

Klimesch, W., 2012. Alpha-band oscillations, attention, and controlled access to stored information. Trends Cogn Sci 16 (12), 606–617.

Klimesch, W., Freunberger, R., Sauseng, P., 2010. Oscillatory mechanisms of process binding in memory. Neurosci Biobehav Rev 34 (7), 1002–1014.

Knyazev, G. G., Savostyanov, A. N., Bocharov, A. V., Slobodskaya, H. R., Bairova, N. B., Tamozhnikov, S. S., Stepanova, V. V., 2017. Effortful control and resting state networks: a longitudinal EEG study. Neuroscience 346, 365–381.

Knyazev, G. G., Savostyanov, A. N., Bocharov, A. V., Tamozhnikov, S. S., Saprigyn, A. E., 2016. Task-positive and task-negative networks and their relation to depression: EEG beamformer analysis. Behav Brain Res 306, 160–169.

Kopell, N., Ermentrout, G. B., Whittington, M. A., Traub, R. D., 2000. Gamma rhythms and beta rhythms have different synchronization properties. Proc Natl Acad Sci U S A 97 (4), 1867–1872.

Krzywinski, M., Altman, N., 2014. Points of significance: Nonparametric tests. Nat Methods 11 (5), 467–468.

Liu, Q., Farahibozorg, S., Porcaro, C., Wenderoth, N., Mantini, D., 2017. Detecting large-scale networks in the human brain using high-density electroencephalography. Hum Brain Mapp 38, 4631–4643.

Liu, Z., Fukunaga, M., de Zwart, J. A., Duyn, J. H., 2010. Large-scale spontaneous fluctuations and correlations in brain electrical activity observed with magnetoencephalography. Neuroimage 51, 102–111.

Luckhoo, H., Hale, J. R., Stokes, M. G., Nobre, A. C., Morris, P. G., Brookes, M. J., Woolrich, M. W., 2012. Inferring task-related networks using independent component analysis in magnetoencephalography. Neuroimage 62, 530–541.

Marty, B., Bourguignon, M., Jousmäki, V., Wens, V., Op de Beeck, M., Van Bogaert, P., Goldman, S., Hari, R., De Tiège, X., 2015. Cortical kinematic processing of executed and observed goal-directed hand actions. Neuroimage 119, 221–228.

Mary, A., Bourguignon, M., Wens, V., Op de Beeck, M., Leproult, R., De Tiège, X., Peigneux, P., 2015. Aging reduces experience-induced sensorimotor plasticity. A magnetoencephalographic study. Neuroimage 104, 59–68.

Monti, M. M., 2011. Statistical analysis of fMRI time-series: a critical review of the GLM approach. Front Hum Neurosci 5, 28.

Oldfield, R. C., 1971. The assessment and analysis of handedness: the Edin-

burgh inventory. Neuropsychologia 9, 97–113.

O'Neill, G. C., Barratt, E. L., Hunt, B. A., Tewarie, P. K., Brookes, M. J., 2015a. Measuring electrophysiological connectivity by power envelope correlation: a technical review on MEG methods. Phys Med Biol 60, R271–295.

O'Neill, G. C., Bauer, M., Woolrich, M. W., Morris, P. G., Barnes, G. R., Brookes, M. J., 2015b. Dynamic recruitment of resting state sub-networks. Neuroimage 115, 85–95.

O'Neill, G. C., Tewarie, P., Vidaurre, D., Liuzzi, L., Woolrich, M. W., Brookes, M. J., 2017a. Dynamics of large-scale electrophysiological networks: a technical review. Neuroimage.

O'Neill, G. C., Tewarie, P. K., Colclough, G. L., Gascoyne, L. E., Bae, H., Morris, P. G., Woolrich, M. W., Brookes, M. J., 2017b. Measurement of dynamic task related functional networks using MEG. Neuroimage 146, 667–678.

Pascual-Marqui, R. D., 2002. Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. Methods Find. Exp. Clin. Pharmacol. 24 Suppl D, 5–12.

Pearson, J., Naselaris, T., A., H. E., Kosslyn, S. M., 2015. Mental imagery: Functional mechanisms and clinical applications. Trends Cogn Sci 19 (10), 590–602.

Siems, M., Pape, A. A., Hipp, J. F., Siegel, M., 2016. Measuring the cortical correlation structure of spontaneous oscillatory activity with EEG and MEG. Neuroimage 129, 345–355.

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., Filippini, N., Watkins, K. E., Toro, R., Laird, A. R., Beckmann, C. F., 2009. Correspondence of the brain's functional architecture during activation and rest. Proc Natl Acad Sci U S A 106, 13040–13045.

Sockeel, S., Schwartz, D., Pelegrini-Issac, M., Benali, H., 2016. Large-scale functional networks identified from resting-state EEG using spatial ICA. PLoS One 11, e0146845.

Taulu, S., Simola, J., Kajola, M., 2005. Applications of the signal space separation method. IEEE Trans Sign Proc 53, 3359–3372.

Vander Ghinst, M., Bourguignon, M., Op de Beeck, M., Wens, V., Marty, B., Hassid, S., Choufani, G., Jousmäki, V., Hari, R., Van Bogaert, P., Goldman, S., De Tiège, X., 2016. Left superior temporal gyrus is coupled to attended speech in a cocktail-party auditory scene. J Neurosci 36, 1596–1606.

Vidaurre, D., Hunt, L. T., Quinn, A. J., Hunt, B. A. E., Brookes, M. J., Nobre, A. C., Woolrich, M. W., 2018. Spontaneous cortical activity transiently organises into frequency specific phase-coupling networks. Nat Comm 9, 2987.

Vigario, R., Sarela, J., Jousmäki, V., Hämäläinen, M., Oja, E., 2000. Independent component approach to the analysis of EEG and MEG recordings. IEEE Trans Biomed Eng 47, 589–593.

Watson, D., Clark, L. A., Chmielewski, M., Kotov, R., 2013. The value of suppressor effects in explicating the construct validity of symptom measures. Psychol Assess 25 (3), 929–941.

Weichwald, S., Meyer, T., Özdenizci, O., Schölkopf, B., Ball, T., Grosse-Wentrup, M., 2015. Causal interpretation rules for encoding and decoding models in neuroimaging. NeuroImage 110, 48–59.

Wens, V., 2015. Investigating complex networks with inverse models: analytical aspects of spatial leakage and connectivity estimation. Phys Rev E Stat Nonlin Soft Matter Phys 91, 012823.

Wens, V., Bourguignon, M., Goldman, S., Marty, B., Op de beeck, M., Clumeck, C., Mary, A., Peigneux, P., Van Bogaert, P., Brookes, M. J., De Tiège, X., 2014a. Inter- and intra-subject variability of neuromagnetic resting state networks. Brain Topogr 27, 620–634.

Wens, V., Marty, B., Mary, A., Bourguignon, M., Op de Beeck, M., Goldman, S., Van Bogaert, P., Peigneux, P., De Tiège, X., 2015. A geometric correction scheme for spatial leakage effects in MEG/EEG seed-based functional connectivity mapping. Hum Brain Mapp 36, 4604–4621.

Wens, V., Mary, A., Bourguignon, M., Goldman, S., Marty, B., Op de beeck, M., Van Bogaert, P., Peigneux, P., De Tiège, X., 2014b. About the electrophysiological basis of resting state networks. Clin Neurophysiol 125, 1711–1713.

Zalesky, A., Fornito, A., Cocchi, L., Gollo, L. L., Breakspear, M., 2014. Time-resolved resting-state brain networks. Proc Natl Acad Sci U S A 111, 10341–10346.