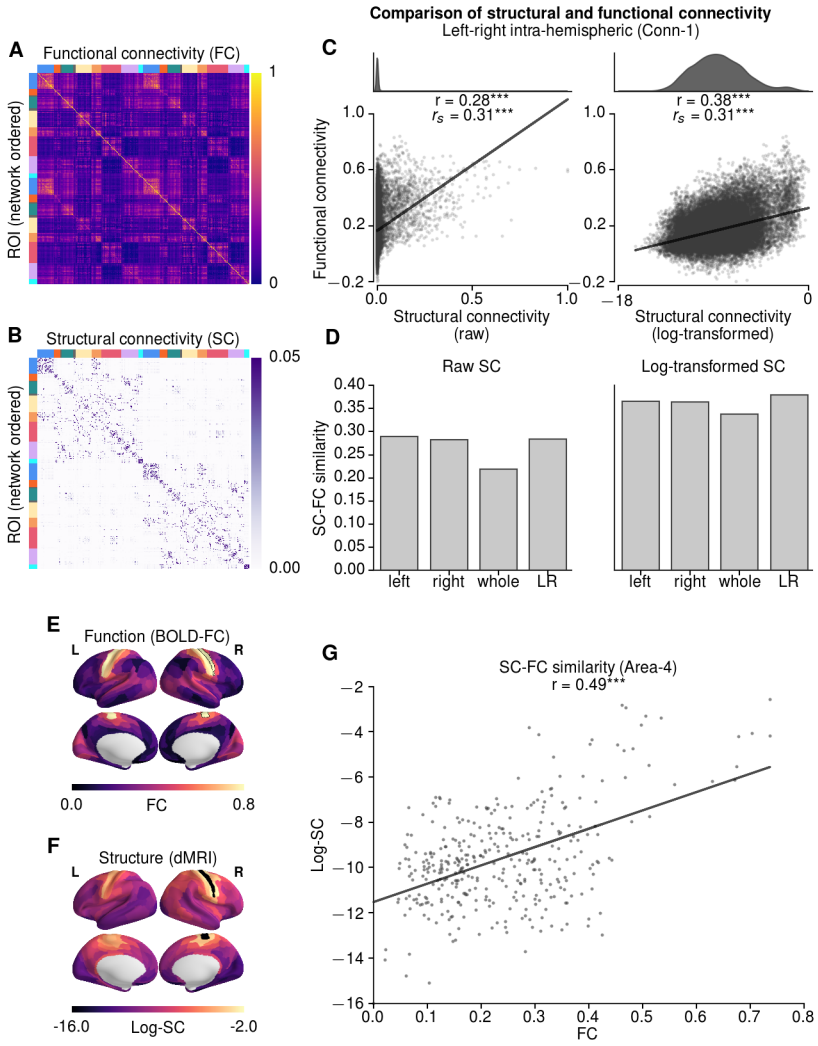

Figure S1: Correspondence Between Empirical Structural and Functional Connectivity. (Related to Figures 1 and 2)

(A,B) Group-averaged bilateral structural connectivity (SC) derived from dMRI and functional connectivity (FC) derived from resting-state fMRI. The matrices are ordered by resting-state networks within each hemisphere (color-coded as in Figure 1).

(C) Correlation between average FC and SC for intra-hemispheric connections, for raw SC weights (left) and log-transformed SC weights (right). Consistent with previous literature, we found a significant relationship between dMRI-derived SC and rs-FC ($r_s = 0.31$, $p < 10^{-5}$).

(D) Correlation between group-averages FC and SC for intra-hemispheric (left, right), whole-brain (i.e., including intra- and inter-hemispheric connections), and left-right concatenated matrices. The SC-FC correlation is lower for the whole-brain matrix which includes inter-hemispheric connections.

(E,F) FC and SC topographies with area 4 in the right hemisphere as an example seed seed region.

(G) Correlation between whole-brain FC and log-transformed SC for the example seed region, right-hemisphere area 4 ($r = 0.49$, $p < 10^{-5}$). *** indicates $p < 0.0001$.

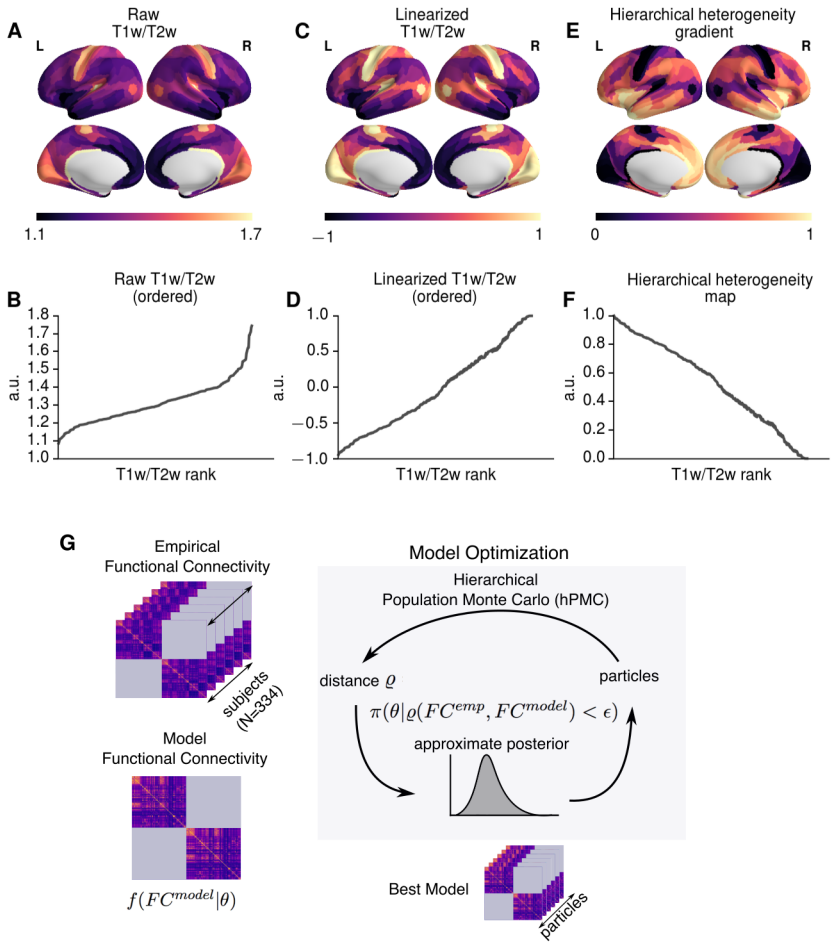

Figure S2: Calculation of the T1w/T2w-Derived Hierarchical Heterogeneity Map. (Related to Figure 2)
(A,B) Rank ordering of T1w/T2w map values.
(C,D) Rank ordering of linearized T1w/T2w map. Linearization is performed by transforming values with the error function.
(E,F) Rank ordering of hierarchical heterogeneity map values used in the model. After linearizing T1w/T2w map values (T_i), were normalized and inverted between 0 and 1 according to: $1 - \frac{max\{T\}-T_i}{max\{T\}-min\{T\}}$.
(G) Schematic for Approximate Bayesian Computation (ABC) via hierarchical Population Monte Carlo (hPMC). After sampling particles (a set of model parameters θ) from the proposal distribution and estimating the model FC, we calculated the average correlation (r) between subject FCs and the model FC. The distance measure δ was defined as $1-(r+c)$, where c is an additional cost which controls for the mismatch between the model and empirical FC. A particle was accepted if the distance was smaller than the threshold (ε), which was decreased at each iteration. The particles were sampled until 1000 particles satisfied the threshold. Accepted particles were used to update the proposal distribution for the next iteration. The convergence criterion for terminating the optimization was an acceptance rate being lower than 0.001.

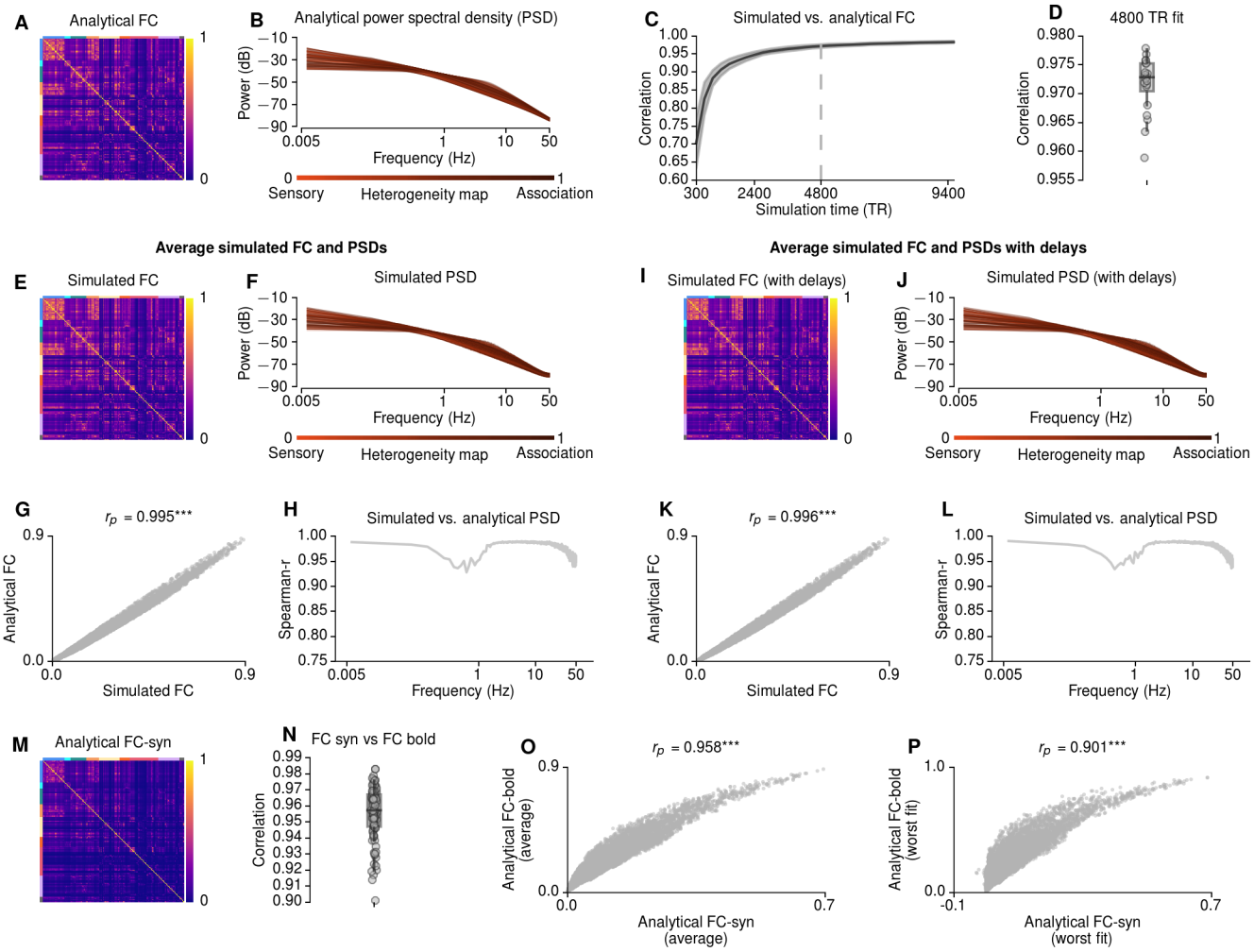

Figure S3: [Caption on next page]

Figure S3: [Figure on previous page] **Comparison Between Numerically Simulated and Analytical BOLD FC and Power Spectral Density (PSD) in the Heterogeneous Model.** (Related to Figures 2 and 7)
(A-D) Analytically approximated BOLD FC (**A**) and PSD (**B**) for 20 particles drawn from posterior distribution, for the left hemisphere. FC matrices are ordered by resting state networks (marked by colored bands on top and left of matrices). The evolution of the correlation between approximated and simulated BOLD FC as a function of time **(C)**, and for 4800-TR simulations **(D)**. Therefore, the analytical BOLD FC is a very accurate approximation of the simulated BOLD FC, given the long scanning times of the empirical BOLD time series.
(E-F) Numerically simulated BOLD FC (**E**) and PSD (**F**) for the same particles. The correlations between simulated and approximated values are very high for both FC ($r = 0.999$) (**G**) and PSD ($r > 0.95$) (**H**). These results demonstrate the robustness of the analytical linearization approximation of FC. The values are averaged across 10 simulations and across all particles. *** indicates $p < 0.0001$. **(I-J)** Numerically simulated BOLD FC (**I**) and PSD (**J**) for the same particles including synaptic delays in long-range interactions due finite axonal transmission speed. The FC and PSD correlations were unaffected by long-range synaptic delays. At low frequencies related to the BOLD signal, the synaptic delays are much shorter than the characteristic timescale of the signals (see also Deco et al., 2014), and therefore have little impact on FC (**K**). Synaptic delays also have little impact on the PSD patterns, at higher frequencies, in the model (**L**). The values are averaged across 10 simulations and across all particles. *** indicates $p < 0.0001$.
(M–P) The relationship between analytical BOLD and synaptic FC. **(M)** Analytically approximated synaptic FC. **(N)** The correlations between BOLD and synaptic FC. **(O,P)** The correlations between average **(O)** and the worst fit **(P)** between BOLD and synaptic FC. These results indicate that for the heterogeneous model, with its larger dynamical repertoire, the assumption that the synaptic FC is a good approximation of the BOLD FC does not generally hold.

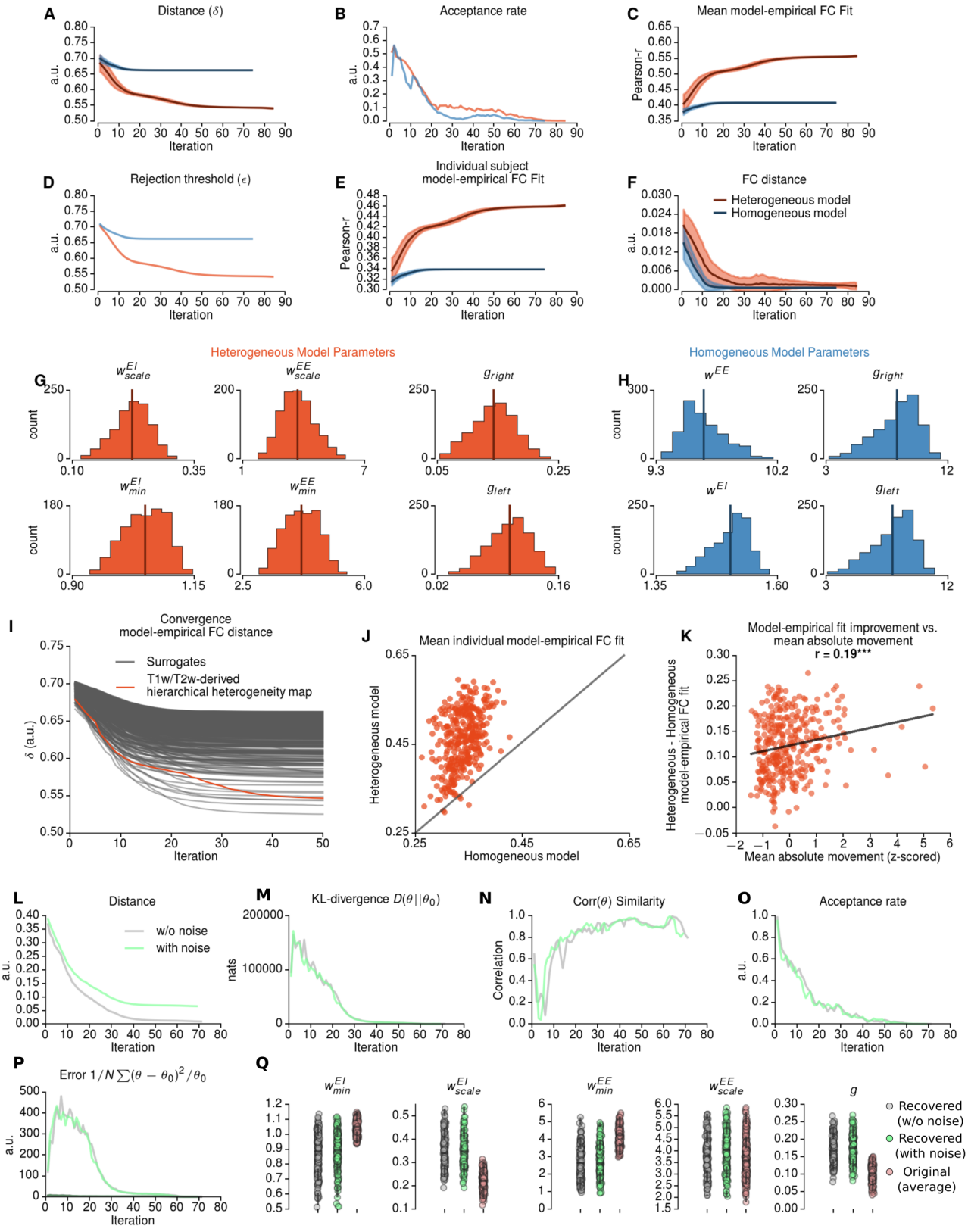

Figure S4: [Caption on next page]

Figure S4: [Figure on previous page] **Optimization of Model Parameters.** (Related to Figure 2)
(A–F) The evolution of optimization parameters across iterations for the homogeneous (blue) and heterogeneous (red) models: distance measure δ **(A)**, acceptance rate **(B)**, average correlation between model FC and group-averaged FC **(C)**, rejection threshold **(D)**, average correlation between model FC and individual subject FCs **(E)**, and distance between the average model FC and average FC **(F)**. The model fits were stabilized after 20 iterations for homogeneous model and 50 iterations for heterogenous model (**C**). The acceptance rate falls below 0.001 after 70 iterations for both models.
(G) Marginal posterior distributions for heterogeneous model parameters. w^{EI} intercept was 1.05 ± 0.04 with a scaling factor of 0.22 ± 0.04, and w^{EE} intercept was 4.19 ± 0.50 with a scaling factor of 3.71 ± 0.81. The global coupling parameters were 0.10 ± 0.02 for left hemisphere and 0.14 ± 0.03 for right hemisphere. Values are mean $\pm$ std. dev.
(H) Marginal posterior distributions for homogeneous model parameters. $w^{EE} = 9.65 \pm 0.15$ and $w^{EI} = 1.50 \pm 0.04$. The global coupling parameters were 7.92 ± 1.52 for left hemisphere and 8.22 ± 1.58 for right hemisphere.
(I) The evolution of distance between empirical and model FCs for the T1w/T2w map-derived heterogeneity map (red) and surrogate heterogeneity maps (gray). For all surrogate maps, the similarity between model and empirical FC stabilizes within 50 iterations. Shaded regions indicate standard deviations across particles.
(J) Improved individual subject model fit in heterogeneous model compared to homogeneous model. For the majority of the subjects (329 of 334) the heterogeneous model performed better than homogeneous model.
(K) The relationship between model improvement and mean absolute movement across subjects exhibited a weak but statistically significant correlation ($r = 0.19$). The heterogeneous model improved fit, relative to the homogeneous model, across the entire range of absolute movement.
(L–Q) Recovery of model parameter values through the optimization procedure. 334 particles were drawn from the approximate posterior distribution of the heterogeneous model (left hemisphere only). Then, we performed the hPMC approach using the model-generated FC of these 334 particles This procedure was repeated with, and without, addition of moderate observation noise to each FC element independently ($\mathcal{N}(0, 0.05)$). (**L**) The distance between objective and particle FCs converged similarly for with- and without-noise conditions, although the distance was larger with noise. (**M**) We assessed the similarity between objective and particle distributions using Kullback-Leibler (KL) divergence. (**P**) We further characterized the similarity between correlation across particle distributions (**N**), and normalized squared error of the parameter values. (**O**) After approximately 70 iterations, the acceptance rate dropped substantially. **(Q)** Distributions of recovered model parameters (gray, without noise; green, with noise; red, original parameters).

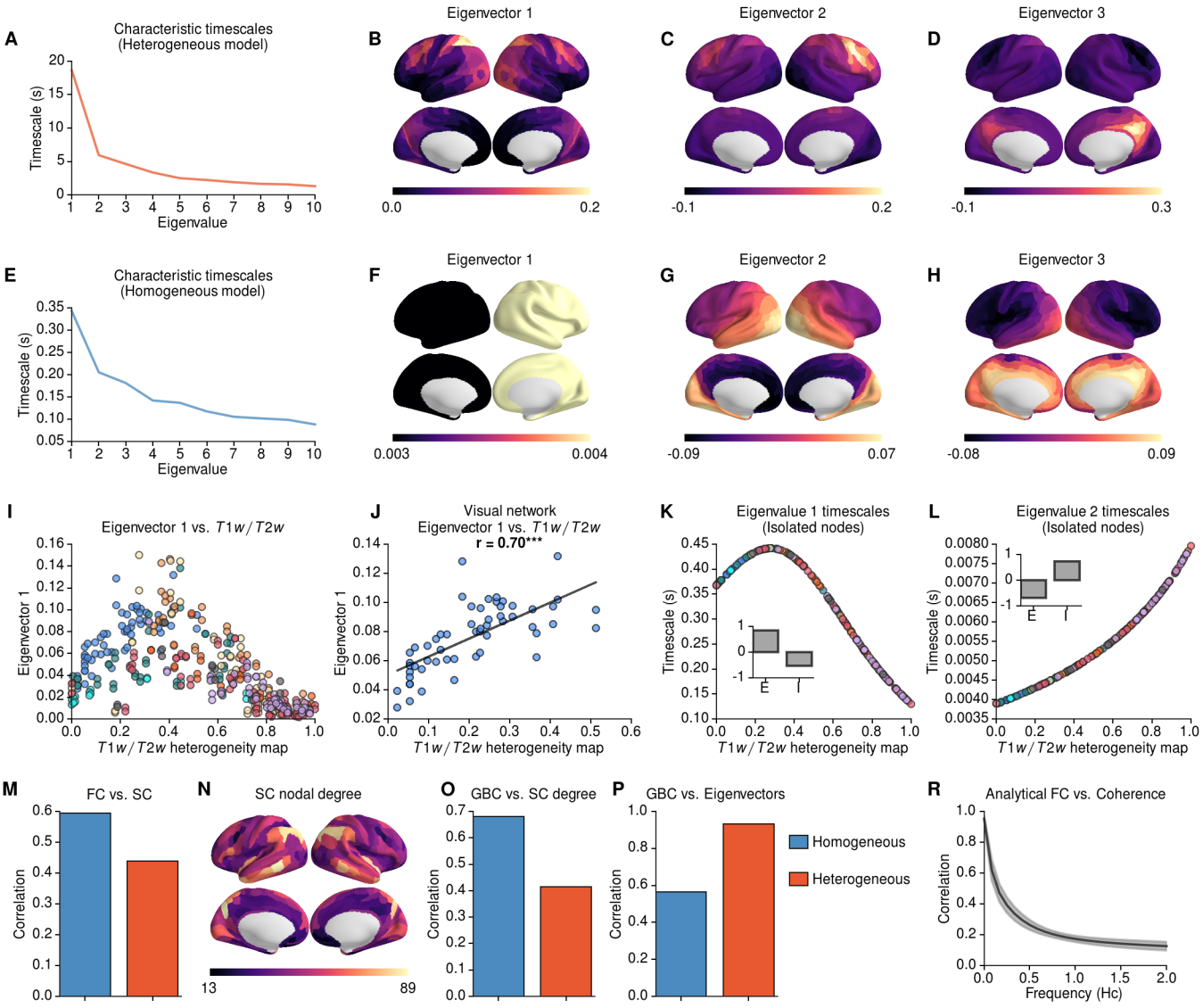

Figure S5: [Caption on next page]

Figure S5: [Figure on previous page] **Eigenvector Analysis.** (Related to Figure 2)
(A) The characteristic timescales ($\tau = \frac{1}{\Re(\lambda)}$) of the heterogeneous model (the first 10 modes ordered by timescale). The first eigenvector has a timescale of $\sim$ 20s, approximately 4 times higher than the subsequent eigenvectors. The values were averaged across 100 samples from the approximated posterior distribution.
(B–D) The corresponding eigenvectors of the three slowest modes of the heterogeneous model. The first eigenvector generally follows the T1w/T2w-derived heterogeneity map across sensory areas, but the largest values are observed in lateral inferior parietal cortex (and to some extent in lateral prefrontal cortex) **(B)** . The subsequent eigenvectors have peaks at lateral prefrontal cortex **(C)**, and medial posterior parts of the cortex (i.e., posterior cingulate and precuneus)**(D)**. Therefore, the heterogeneous model offers spatially structured modes, which might be related to the integration of information across distinct resting-state networks.
(E) The characteristic timescales of the homogeneous model (the first 10 out of 20 modes ordered by timescale. It is important to note that the homogeneous model exhibits complex eigenvalues/eigenvectors, which come in pairs of (eigenvalues of form $\lambda = a \pm ib$). Since the complex conjugate of each mode does not provide additional information about the timescales, conjugate modes were omitted in this analysis. For the homogeneous model, the characteristic timescales decrease smoothly having lower values than the heterogenous model (approximately 0.15–0.3 s).
(F–H) The corresponding eigenvectors of the slowest modes of the homogeneous model. The first eigenvector reflects a global mode (as it has a complex part, this mode generates the global spectral peak at alpha range as seen in Figure 7) **(F)**. The subsequent modes reflect coarse spatial gradients: posterior-to-anterior **(G)** and medial-to-lateral **(H)**. These modes are likely driven by the spectral properties of the SC matrix.
(I) The relationship between the slowest eigenvector **(B)** of the heterogeneous model and T1w/T2w-derived heterogeneity map for all areas. Areas are color-coded by functional network as in **Figure 1D**. There is not a linear relationship between the slowest eigenvector and the T1w/T2w-based hierarchical heterogeneity map. The eigenvector weights follow the hierarchical heterogeneity map in sensory networks, and peak at some of the association areas in fronto-parietal, dorsal- and ventral-attention networks. In contrast, the eigenvector weights of the other higher-order association areas are negatively correlated with the heterogeneity map values.
(J) Within the visual network, there is a strong positive correlation between eigenvector weights and the hierarchical heterogeneity map values ($r = 0.7$).
(K–L) Characteristic timescales for isolated nodes. The heterogeneous model parameters for each area were preserved, but the long-range connections were removed after compensating for the total input driven by global coupling parameter (also see **Figure S7D**). This approach removes the influence of long-range coupling, and allows studying the characteristic timescales of isolated nodes. An isolated local E-I node exhibits two dynamical modes. The first mode timescale exhibits a non-linear relationship to T1w/T2w heterogeneity map values **(K)**. In contrast, the second mode timescale increases monotonously with T1w/T2w heterogeneity map values **(L)**.
(M) Correlations between SC and model FC. Unlike the heterogeneous model, the homogeneous model FC is highly correlated with SC. Therefore, heterogeneity of local properties can cause FC patterns to diverge from SC patterns.
(N) The spatial distribution of the node degrees in SC matrix (threshold $= 10^{-5}$).
(O–P) The relationships between model GBC and nodal degree **(O)** and the slowest eigenvalues **(P)**. The homogeneous model GBC is driven by the contributions of high degree nodes in the SC, whereas the heterogeneous model GBC is driven by the critical slowing of the dynamics through hierarchical heterogeneity map (note that since the first eigenvector of the homogeneous model exhibits a global pattern, the second eigenvector is shown).
(R) The relationship between analytical BOLD FC and coherence across frequencies. The analytical BOLD FC exhibits substantial similarity with low-frequency coherence ($r > 0.97$), which decays quickly as a function of time. This result shows the relationship between low-frequency coherence and BOLD FC. Furthermore, it provides an explanation for why synaptic FC may not correspond to BOLD FC in a larger dynamic range (i.e. 1/f-like power spectrum).

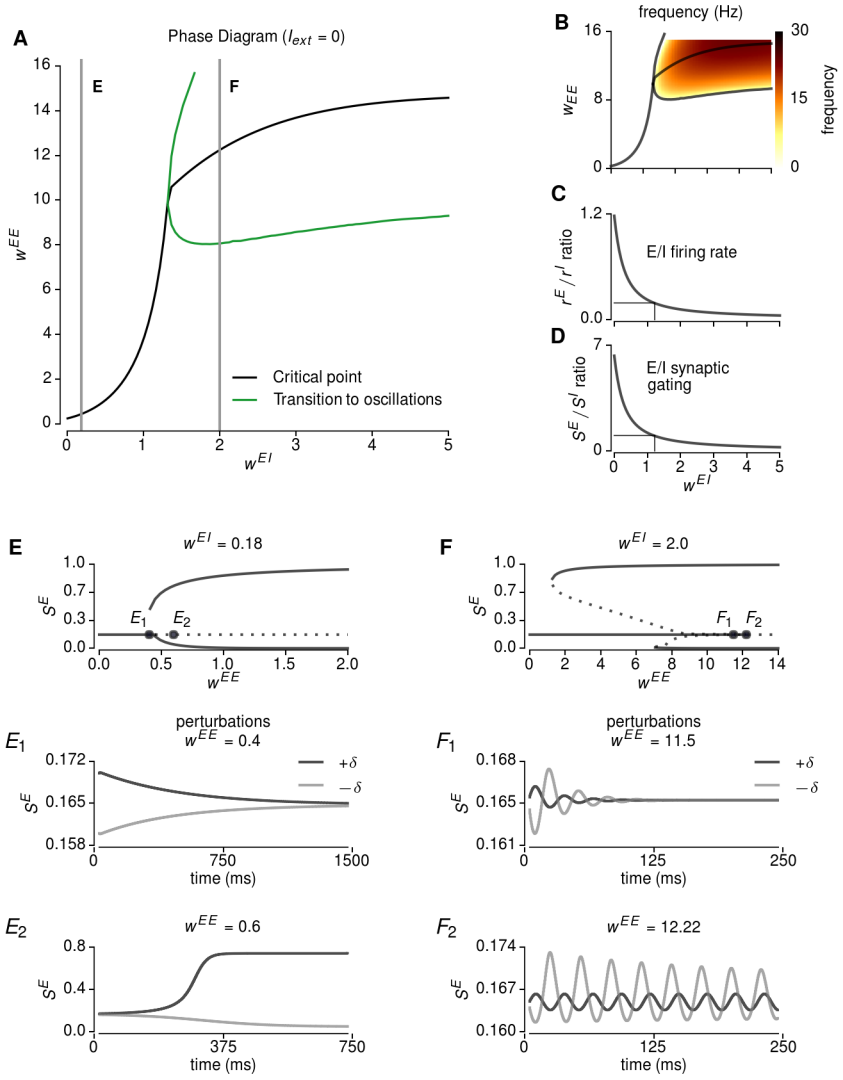

Figure S6: Dynamical Systems Analysis of a Local Excitatory-Inhibitory Node. (Related to Figure 7)
(A) The extended phase diagram for w^{EE} and w^{EI}. For low values of w^{EI} the model exhibits a single pitchfork bifurcation, whereas for high values the system exhibits oscillatory activity.
(B) Intrinsic frequencies of the system calculated from imaginary parts of the eigenvalues of the Jacobian matrix. For high values of w^{EI} and w^{EE}, the system generates oscillations with intrinsic frequencies between 0 and 30 Hz.
(C-D) The ratio of excitatory to inhibitory firing rates (**C**) and synaptic gating variables (**D**) as a function of excitatory-to-inhibitory strength w^{EI}. For both models, the optimal parameter range was near the critical point at which the system exhibits a oscillatory activity. In this regime, the ratio between excitatory and inhibitory synaptic gating variables is approximately equal to 1 (i.e., the excitatory and inhibitory synaptic activities are balanced), while the firing rate of inhibitory neurons is higher than excitatory neurons, consistent with cortical physiological recordings.
(E) The bifurcation diagram for $w^{EI} = 0.18$ (i.e. the value proposed in Deco et., 2014). The perturbations ($\pm\delta$) around fixed point for $w^{EE} = 0.4$ E_1 and $w^{EE} = 0.4$ E_2. Before bifurcation, the synaptic gating variable returns to its steady state value E_1. After bifurcation, the synaptic gating variable moves towards up- and down-attractor states E_1.
(F) The bifurcation diagram for $w^{EI} = 2.0$ (i.e. after emergence of oscillations). The perturbations ($\pm\delta$) around fixed point for $w^{EE} = 11.5$ F_1 and $w^{EE} = 12.22$ F_2. Before bifurcation, the synaptic gating variable exhibits damped oscillations around the steady state value F_1. The oscillations are sustained around the bifurcation point F_2.

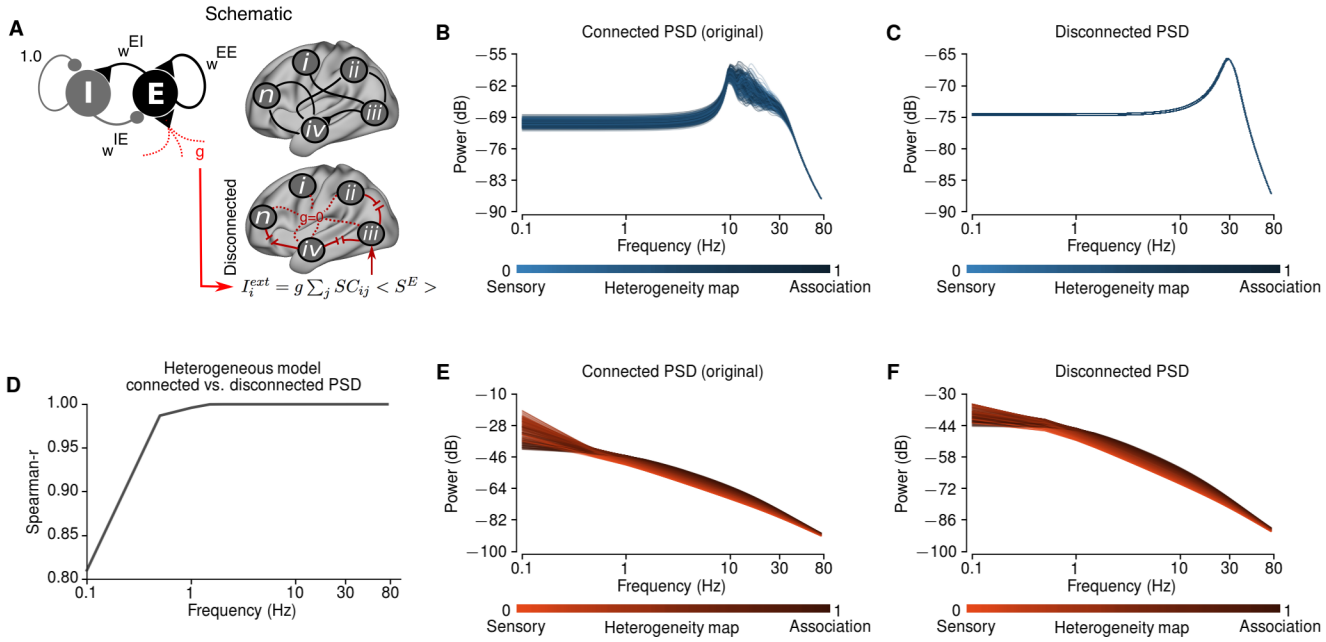

Figure S7: Long-Range Disconnection Analyses. (Related to Figure 7)

(A) Schematic illustrating the disconnection analysis. To study the role of long-range connectivity on power spectral densities (PSDs), we calculated the power spectral density (PSD) after setting the global coupling parameter to 0 (i.e., after removing long-range connections). Since the strength of feedback inhibition (w^{IE}) depends on the total synaptic input to each node, we added compensatory external input currents to each node such that the local microcircuit parameters were preserved. **(B–C)** PSD of the homogeneous model for the full connected model **B** and the disconnected model **C**. In the homogeneous model, the spatial patterns in the PSD were completely destroyed and collapsed into a single pattern after disconnecting the long-range connections. **(D–F)** PSD of the heterogeneous model for full connected model **E** and disconnected model **F**. Unlike the homogeneous model, the spatial patterns in the PSD were preserved in high-frequency bands (i.e. the correlation between two maps were close to 1 (**D**). The correlation between connected and disconnected PSDs were lower in very low-frequencies (i.e. $<$1 Hz). This shows that the regional patterns of high-frequency power emerged as a local property in the model.

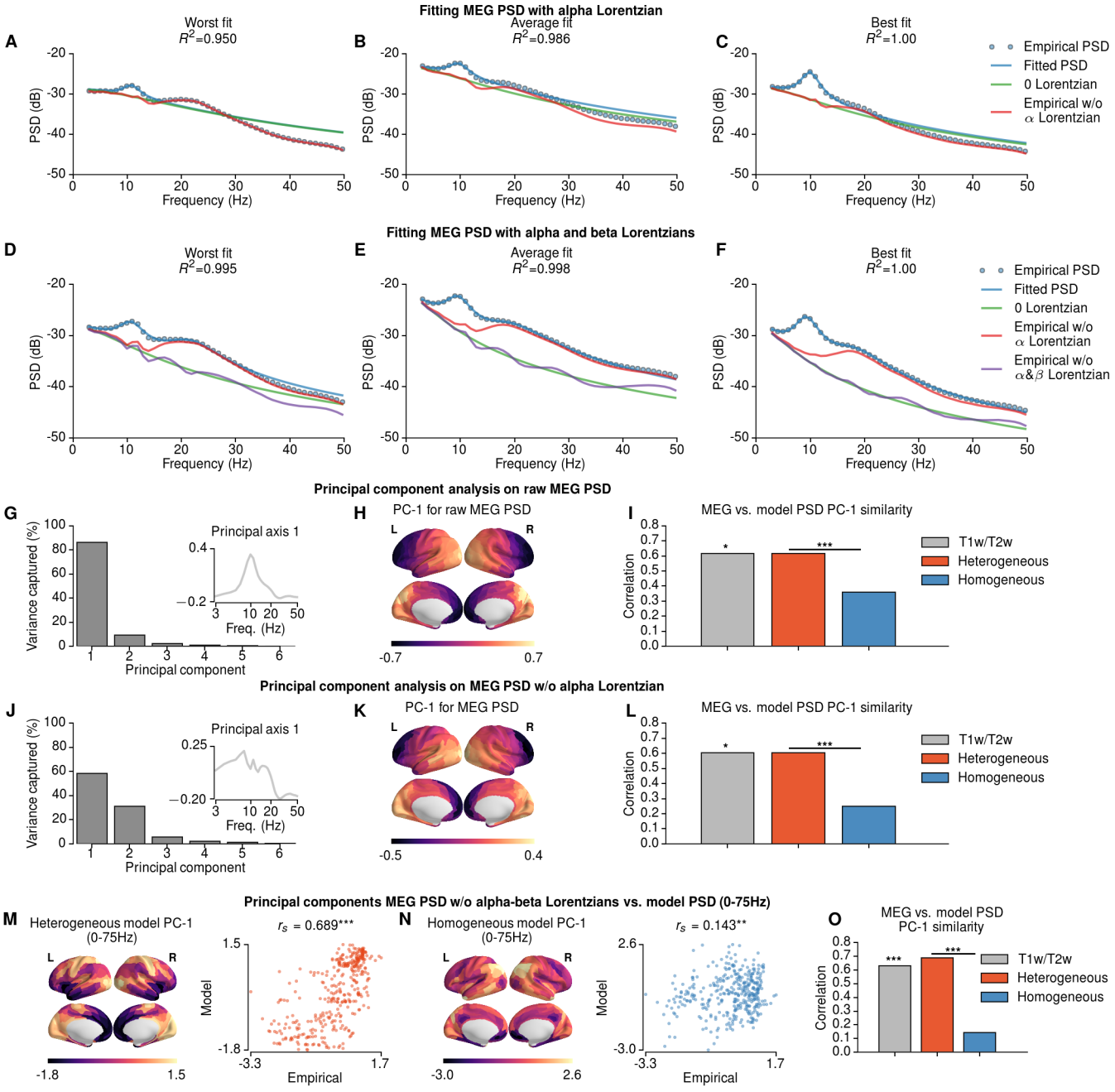

Figure S8: [Caption on next page]

Figure S8: [Figure on previous page] **Comparison of the approaches to fit MEG PSD with Lorentzians.** (Related to Figure 8)

(A–C) Fitting the MEG PSD with a sum of two Lorentzians ($p_{\alpha}(f)$ in Equation 1 in **Experimental Procedures**): a 0 Lorentzian (i.e., Lorentzian with center frequency at 0 Hz), and a Lorentzian with center frequency in the alpha band (α). Plotted are the worst-fit parcel **(A)**, average across parcels, **(B)** and best-fit parcel **(C)**.

(D–F) Fitting the MEG PSD with a sum of three Lorentzians ($p_{\alpha\beta}(f)$ in Equation 2 in **Experimental Procedures**): a 0 Lorentzian, an alpha-band Lorentzian, and a beta-band (β) Lorentzian. Plotted are the worst **(D)**, average **(E)** and best fit **(F)** to the data across areas. Removing alpha and beta Lorentzians performed better than other approaches in removing the spectral peaks (see **G–L**). Note that when only the alpha Lorentzian was fitted and removed (i.e., using $p_{\alpha}(f)$ in Equation 3, without fitting the beta peak), there are prominent residual alpha and beta peaks.

(G–I) Principal component analysis (PCA) applied to the raw MEG PSD relative power maps (i.e., without alpha or beta peaks removed). The first PC (PC-1) is the spatial map that captures the most variance in relative power maps across areas. PC-1 explained 86.39% of the variance across areas **(G)**. The inset shows the PC-1 coefficients across frequencies. Here, the PC1-coefficients show a strong alpha peak, indicating that PC-1 is driven in large part by the alpha-band topography. **(H)** The weights of the PC-1 map exhibit a posterior-to-anterior gradient with a relatively broad spatial autocorrelation (characteristic length scale $d_0 = 16.38$ mm, $\rho = 1.12$, from a spatial lag model). **(I)** The raw MEG PC-1 showed substantial similarity with the T1w/T2w map ($r_s = 0.615, p = 0.048$), PC-1 from the heterogeneous model ($r_s = 0.615$), and PC-1 from the homogeneous model PC-1 ($r_s = 0.359$). The correlation with PC-1 from the heterogeneous model was significantly greater than that from the homogeneous model ($p < 10^{-4}$, dependent correlation test).

(J–L) PCA applied after removing only the alpha Lorentzian (not the beta). **(J)** PC-1 explains 58.40% of the variation. **(K)** PC-1 weights exhibit a posterior-to-anterior spatial topography similar to the raw case. The similarities for the alpha-removed PC-1 remained high for the T1w/T2w map ($r_s = 0.604,\ p = 0.028$) and heterogeneous model PC-1 ($r_s = 0.604$), but was dropped for the homogeneous model PC-1 ($r_s = 0.248$) **(L)**. The correlation with PC-1 from the heterogeneous model was significantly greater than that from the homogeneous model ($p < 10^{-4}$, dependent correlation test).

(M–O) The MEG PSD PC-1 after removing alpha and beta Lorentzians compared to model PSD PC-1 for a broader frequency range (0–75 Hz for the model, compared to 3–50 Hz for the empirical MEG). For the broader frequency range, the similarity between empirical and model PSD was substantially increased for the heterogeneous model ($r = 0.689$ vs. $r = 0.631$ at 3–50 Hz range, which was identical to T1w/T2w similarity) **(M)**, but not for the homogeneous model ($r = 0.143$ vs. $r = 0.138$ at 3–50 Hz range) **(N)**. These results suggest that long-range interactions in very-low frequencies ($<$3 Hz) may enhance predictive power of the model for empirical PSD. However, it is important to note that the model was only informed by the BOLD FC (i.e., not fitted to MEG PSD). Therefore, the model was not predicted to capture all the features of the MEG PSD, apart from the topography of the principal mode of variation. P-values in **(I, L–O)** were calculated using random surrogate maps.