Dynamical consequences of regional heterogeneity in the brain's transcriptional landscape

Gustavo Deco[†1,2,3,4], Morten L. Kringelbach[†6,7], Aurina Arnatkevičiūtė[4], Stuart Oldham[4], Kristina Sabaroedin[4], Nigel C. Rogasch[5], Kevin M. Aquino[#4], and Alex Fornito[#*4]

[†] equal author contribution

[#] equal author contribution

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

2. Institució Catalana de la Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, Barcelona, 08010, Spain

3. Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany

4. Turner Institute for Brain and Mental Health, School of Psychological Sciences, and Monash Biomedical Imaging, Monash University, Melbourne, Clayton VIC, Australia

5. Discipline of Psychiatry, Adelaide Medical School, University of Adelaide, and Hopwood Centre for Neurobiology, Lifelong Health Theme, South Australian Health and Medical Research Institute, Adelaide SA, Australia

6. Department of Psychiatry, University of Oxford, Oxford, UK

7. Center for Music in the Brain, Department of Clinical Medicine, Aarhus University, DK

Corresponding author: alex.fornito@monash.edu

Short title: Dynamical effects of brain heterogeneity

Abstract

Brain regions vary in their molecular and cellular composition, but how this heterogeneity shapes neuronal dynamics is unclear. Here, we investigate the dynamical consequences of regional heterogeneity using a biophysical model of whole-brain functional magnetic resonance imaging (MRI) dynamics in humans. We show that models in which transcriptional variations in excitatory and inhibitory receptor (E:I) gene expression constrain regional heterogeneity more accurately reproduce the spatiotemporal structure of empirical functional connectivity estimates than do models constrained by global gene expression profiles and MRI-derived estimates of myeloarchitecture. We further show that regional transcriptional heterogeneity is essential for yielding both ignition-like dynamics, which are thought to support conscious processing, and a wide variance of regional activity timescales, which supports a broad dynamical range. We thus identify a key role for E:I heterogeneity in generating complex neuronal dynamics and demonstrate the viability of using transcriptomic data to constrain models of large-scale brain function.

Brain dynamics are often described as complex, displaying properties that are interposed between order and disorder (*1*, *2*). These complex dynamics arise from two cardinal factors: (1) the properties of local population activity within each brain region; and (2) the mutual influences that these populations exert on each other, as mediated by their interconnecting anatomical projections.

Biophysically plausible non-linear models of brain dynamics have been used to clearly demonstrate the importance of anatomical connectivity in supporting complex patterns of network activity. For example, when simulated on human inter-regional structural connectivity (SC) matrices (connectomes) generated with diffusion magnetic resonance imaging (MRI), these models can reproduce empirically observed patterns of inter-regional functional connectivity (FC) (reviewed in (*2*)). In macaque, a biophysical model simulated on a connectome derived from viral tact tracing indicated that both feedback projections and a heterogeneous distribution of connection strengths across projections are required for the emergence of ignition-like dynamics (*3*), which are thought to be a necessary condition for conscious processing (*4*).

By comparison, variations of local circuitry within each region have received less attention. Precise coordination of firing across spatially distributed neuronal ensembles depends on a delicate balance between excitatory and inhibitory activity. The first wave of biophysical models of large-scale brain activity treated all local population dynamics as homogenous, driven by sub-populations of inhibitory and excitatory neurons with uniform properties across all brain regions (*5*–*7*). However, heterogeneity in regional cytoarchitecture has been noted for over a century, as famously mapped by Brodmann (*8*), and subsequent work has provided mounting evidence for large-scale cortical gradients in gene expression, cellular composition, connectivity, and function (*9*, *10*). Thus, quantitative variations of local circuit properties across the cortex also play an important role in shaping complex neural dynamics, likely through their influence on the local ratio, or balance, of excitatory and inhibitory cell activity (E:I balance) (*9*).

Ideally, one could model regional heterogeneity using empirically observed estimates of areal variations in excitatory and inhibitory cell counts, or a related measure of cell-type specific activity. No such data are available on a large scale for primate brains, but some early modelling work has used alternative methods to constrain biophysical models. Deco et al. (*11*) showed that heterogeneity in local feedback inhibition, in which E:I balances were algorithmically adjusted to achieve a uniform firing rate of 3 Hz across all regions (in line with experimental data; (*12*–*14*),

yields networks with more stable dynamics and improves model fits to static measures of FC when compared with a strictly homogeneous model. Using data in macaque, Chaudhuri et al. (*15*) scaled the excitatory input strength of each of 29 cortical areas according to its assumed hierarchical position, as determined by laminar patterning of inter-regional projections (*16*). They found that this heterogeneity was essential for generating dynamics characterized by a hierarchy of regional activity timescales, in which areas with a higher hierarchical position show a prolonged response to a simulated sensory input. This result mirrors previous single unit findings in macaque cortex (*17*) and aligns with other evidence for regional differences in activity timescales (see also (*18–20*)). Two subsequent studies modelling human functional MRI found that, when compared to a regionally homogeneous model, scaling a region's excitatory activity according to its assumed hierarchical position improves fits to measures of edge-level and node-level functional connectivity, more faithfully reproduces empirically observed regional variations in activity timescales, and better explains FC disturbances in people with schizophrenia (*21*, *22*). A complementary study used a stochastic optimization procedure to invert local dynamical parameters such as recurrent synaptic connectivity strengths to optimally fit functional MRI data (*23*). Regional variations in the estimated local dynamical parameters showed hierarchical ordering and correlated with variations in myelin and layer-specific neuronal density, presumed cognitive functions (derived from meta-analysis of task-based fMRI studies), and a principal gradient of resting-state FC (*24*).

These initial models of regional heterogeneity demonstrate the critical role of local variations in E:I balance for generating realistic patterns of brain-wide dynamics. However, in the absence of a gold-standard biological constraint, there has been considerable variability in how heterogeneity has been instantiated within a given model. There is also substantial variability in the dynamical properties that have been modelled. For example, modelling work in macaques shows that it is, in principle, possible to find a parameter regime that results in ignition-driven and hierarchical dynamics (*3*, *15*), but the parameters were not fitted to empirical data making it difficult to evaluate the model's capacity to reproduce other key features, such as realistic patterns of static or dynamic FC. Similarly, studies in humans have modelled specific outcome measures, such as static or dynamic FC at the level of individual connections (*6*, *7*, *22*, *25*), the mean static FC of each region (*22*), or timescale hierarchies (*3*, *22*), meaning that it is unclear whether the working point for a model fitted to one property can also capture other properties. Developing a coherent approach that can account for each of these diverse features is an essential step towards

developing a unified model that can parsimoniously explain diverse empirical properties of complex whole-brain dynamics.

Here, we seek to address this aim by introducing a model that leverages brain-wide transcriptional data (*26*) to constrain regional heterogeneity and tune each region's dynamics according to regionally-specific measures of inhibitory and excitatory receptor gene expression. We evaluate the performance of this model with respect to some of the major outcome measures evaluated thus far in the literature; namely, edge-level static FC, node-level static FC, FC dynamics (FCD), ignition capacity, and timescale hierarchies. We show that our model more faithfully reproduces empirical properties of static and dynamic FC than other models in which heterogeneity is imposed using plausible biological constraints, such as regional variations in the ratio of T1-to-T2-weighted MRI signal (T1w:T2w), which is thought to index intracortical myelin content (*27*) and has been proposed as a marker of a cortical region's hierarchical position (*28*), and regional loadings on the dominant mode of gene expression (*22*, *28*). We then show that transcriptomically-constrained models show the greatest capacity for ignition-like dynamics and display a broad range of regional activity timescales. Our results indicate that regional variations in the transcriptional activity of inhibitory and excitatory receptor genes provide a viable means for constraining biophysical models of large-scale neural dynamics and, when coupled with an empirically-derived connectivity matrix, can offer a parsimonious account of diverse properties of large-scale human brain activity.

Results

Overview

Our overall analysis strategy is outlined schematically in Figure 1 and the details of our approach are provided in the Methods. Briefly, we use diffusion MRI to generate a group-level representative structural connectome for a sample of 293 healthy individuals, representing SC between each pair of 68 regions defined according to an extensively used neuroanatomical atlas (Figure 1A). We also estimate, for a partially overlapping sample of 389 individuals, various empirical FC properties using task-free resting-state functional MRI to probe spontaneous blood-oxygenation-level-dependent (BOLD) dynamics (Figure 1B).

We incorporate different forms of regional heterogeneity (Figure 1C and Figure S1) into a whole-brain model in which the local neuronal dynamics of each brain region evolve according to a

dynamic mean field reduction analytically derived from the collective behaviour of empirically validated integrate-and-fire (80% excitatory and 20% inhibitory) neurons (*13*, *14*). Regional dynamics in the model are driven by a pool of excitatory neurons, a pool of inhibitory neurons, the interactions between them, and with the net output of other areas, as mediated by the SC matrix (Figure 1D). In this model, local feedback inhibition is adjusted separately for each region to ensure a consistent firing rate of 3 Hz across regions (*11*). We thus refer to this model as the balanced excitation-inhibition (BEI) model. Although the local adjustments in this model introduce some degree of regional heterogeneity, the firing rates are constrained to be uniform across regions so we consider this BEI model as a homogeneous benchmark against which we evaluate more sophisticated heterogeneous models that allow intrinsic dynamical properties to vary across regions (see (*22*) for a similar approach). More specifically, the activity levels of excitatory and inhibitory populations in each region, *i*, of the BEI model are given by independent sigmoid functions, regulated by a single gain parameter M_i. In our benchmark BEI model, this parameter is set to a fixed value for all regions, imposing uniform excitability across brain regions (Figure 1C).

We compare the performance of this homogeneous BEI model to three different heterogeneous models. In the first, we use an approach similar to (*22*) and adjust M_i according to variations in the regional mean T1w:T2w, which has been proposed as a proxy for cortical hierarchy (*28*). In the second heterogeneous model, we adjust M_i according to the first principal component (PC1) of transcriptional activity for 1,926 brain-specific genes (*28*) that survived our quality control criteria, as quantified in the Allen Human Brain Atlas (AHBA) (*26*, *29*, *30*). The first PC of these genes correlates with the T1w:T2w ratio and other measures of cortical hierarchy (*28*). In the final model, we use a more hypothesis-driven approach and adjust M_i according to variations in the ratio of regional AHBA expression values for genes specifically coding for AMPA, NMDA, and GABA receptors (see Methods). We refer to this model as the E:I model. Cortical surface renderings displaying regional variations in T1w:T2, PC1, and E:I values are shown in Figure 1C.

In all three heterogeneous models, we assume a linear scaling between the regional biological measures of heterogeneity, R_i, and the effective gain within a region (*22*), given by $M_i = 1 + B + ZR_i$, where the two unknown free parameters correspond to a bias, B, and a scaling factor, Z (see *Methods*). Studying how these two free parameters affect the global dynamics of the model allows us to investigate the role of regional heterogeneity.

For the homogeneous BEI and three (T1w:T2w, PC1, E:I) heterogeneous models, we assume that all diffusion MRI-reconstructed streamline fibres have the same conductivity and thus the coupling between different brain areas is scaled by a single global parameter, G. We first tune the G parameter of the BEI model to adjust the strength of effective coupling in the model and identify the brain's dynamic working-point by fitting the model to three empirical FC properties (Figure 1E): (1) the Pearson correlation between model and empirical estimates of static (i.e., time-averaged) FC estimated across all pairs of brain regions (more specifically, the correlation between the values in the upper triangles of the model and empirical FC matrices); (2) the Pearson correlation between model and empirical node-level estimates of average FC (*22*); and (3) similarity in functional connectivity dynamics (FCD), estimated as the distance between model and empirical distributions of frame-by-frame FC properties (as quantified using the Kolmogorov-Smirnov distance, D_{KS})(*25*).

After finding the working point of each model, we evaluate its dynamical properties in two ways. First, we consider the ignition capacity of the model, quantified by examining how activity propagates through the network following a simulated focal perturbation (Figure F-G). Second, we examine regional heterogeneity in intrinsic timescales in two ways, following (*15*), by quantifying the signal decay rate of regional activity following a simulated focal perturbation (Figure F,H) and by studying regional signal autocorrelations during continuous, focal white noise input. Our approach thus allows us to determine the performance of each model in capturing diverse dynamical phenomena at a single working point.

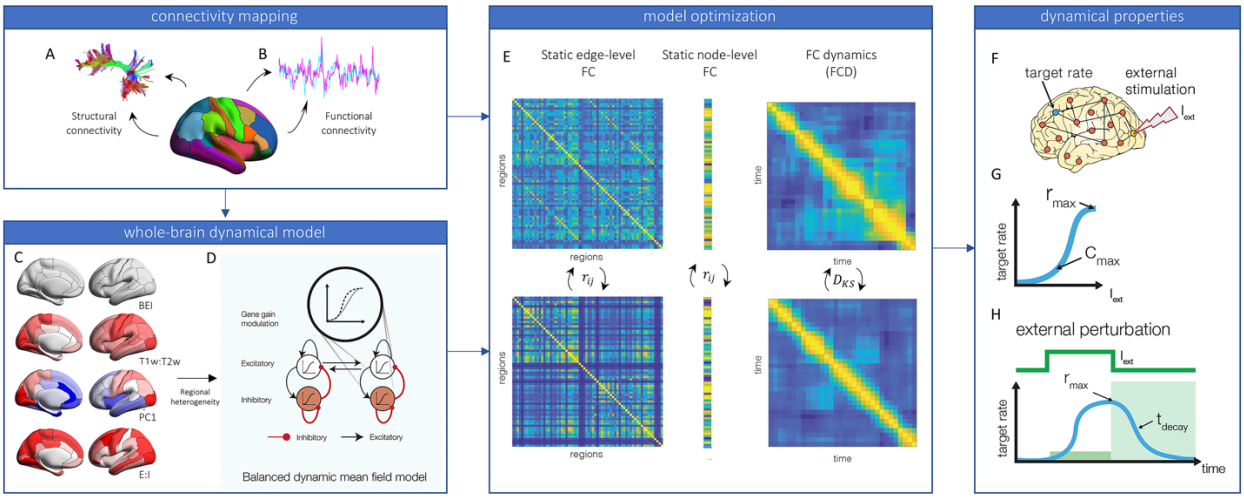

Figure 1. Workflow for model fitting and evaluation. We use empirical diffusion and functional MRI data to generate sample-averaged structural connectivity (**A**) and functional

connectivity (**B**) matrices, respectively. The structural connectivity matrix, and different forms of regional heterogeneity (**C;** for full maps, see Figure S1), are used to respectively define the coupling between, and modulate the gain of, neuronal populations within a large-scale balanced excitation-inhibition (BEI) dynamical model (**D**). For each model, the global coupling parameter, G, is tuned to optimize the fit to three empirical FC properties (**E**): static edge-level FC (left), static, node-averaged FC (middle), and FC dynamics (FC; right). Gain parameters B and Z are then tuned to further improve fits to these properties. **F:** After finding the optimal working point for each model, we simulate a focal perturbation of activity in occipital cortex and measure two dynamical properties of the network: ignition and activity timescales. **G**: To measure ignition, we characterize the response of each area to focal stimulation as the product of two quantities: the maximum firing rate achieved at the highest stimulation intensity, r_{max}, and the concavity of the regional response function (i.e., the maximal second derivative), c_{max}. The ignition capacity of the model is taken as the average of these regional values. **H:** To measure the decay timescale of each region, we fit an exponential curve to each region's firing rate as it returns to baseline after termination of the maximal stimulation intensity (**H**). We also measure intrinsic timescales during a continuous white noise stimulation paradigm (see Methods).

Fitting the Homogeneous Model

We first evaluate the performance of the homogeneous BEI model in reproducing empirical properties of resting-state FC data. We determine and fix the global coupling parameter, G, with all regional gain parameters, $M_i = 1$ (i.e. heterogeneity is not considered). As done previously (*11*)(*31*), we examine how well the model fits, as a function of G, three different properties of empirical resting-state fMRI recordings: edge-level static FC, FCD, and node-level FC (see Figure 1B and Methods for further details). The results of this analysis are shown in Figure 2A-C. In all cases, we consider group-averaged values for both empirical data (across subjects) and the model (across an equivalent number of simulated trials with the same duration as the human experiments) (see *Methods*). Our results align with previous investigations to clearly show [40, 44] that fitting FCD, which captures the rich spatiotemporal structure of the fMRI data, is a stronger constraint on the model. Indeed, where static edge and node FC fits are consistently high across a broad range of G, FCD yields a clear global optimum at $G = 2.1$ (Figure 2A).

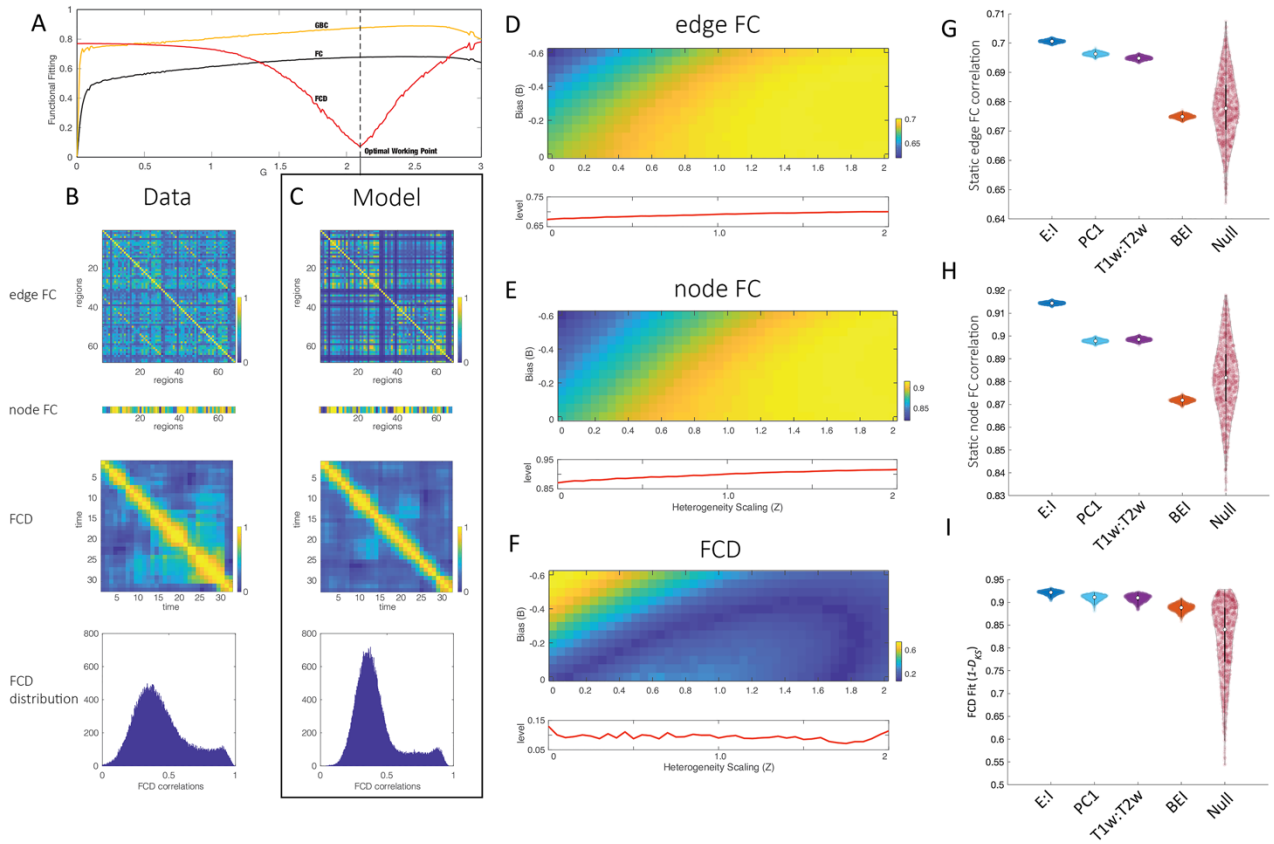

Figure 2. Model optimization and evaluation. The homogeneous BEI model is first fitted to the empirical data. **A:** The correlation between model and empirical static edge-level and node-level FC, as well as the D_{KS} between empirical and model FCD distributions, as a function of the global coupling parameter, G. Only the fit to FCD shows a clear optimum at $G = 2.1$ (note that lower FCD fit values indicate better fit). Panels **B** and **C** show the FC matrix, node-level average FC values, FCD matrices, and FCD distributions for the empirical data (**B**) and model (**C**). Using $G = 2.1$, we introduce regional heterogeneity (Figure 1C) and further optimize the two gain parameters B and Z. Panels **D**, **E**, and **F** show the parameter landscapes for edge-level FC, node-level FC, and FCD, respectively, for the E:I model. Below each 2D space we show the 1D graph of the level heterogeneity scaling (Z) for the optimal bias parameter found as the one corresponding to the optimum FCD fitting (see Methods). This procedure is repeated for each heterogeneous model to find its optimal working point. Lower values indicate better fit in panel **F**. Panels **G**, **H**, and **I** compare the performance of the homogeneous BEI model and heterogeneous T1w:T2w, PC1, E:I, and null expression models according to the edge-level FC, node-level FC, and FCD benchmarks, respectively. Distributions are for values obtained over multiple runs of the optimum models (see Methods). We harmonized the fit statistics so that higher numbers indicate better fit (i.e., for the case of FCD, where a low D_{KS} indicates better fit, we show $1 - D_{KS}$). The E:I model shows the best performance in all cases.

Introducing regional heterogeneity

We now study how regional heterogeneity affects the fitting of static FC, GBC, and FCD. Spatial maps for each form of biological heterogeneity used in our modelling are shown in Figure 1C and Figure S1. The T1w:T2w and PC1 maps are strongly correlated ($\rho = 0.82, p_{spatial} < .001$), as shown previously (*28*), whereas the PC1 and T1w:T2w maps show weaker correlations with the E:I map ($\rho = 0.40, p_{spatial} = .005$ and $\rho = 0.39, p_{spatial} < .011$, respectively). This result indicates that each spatial map introduces a different form of biological heterogeneity to the benchmark BEI model. We introduce this heterogeneity by modulating the regional gain functions M_i, at the optimal working point of the homogeneous BEI model ($G = 2.1$), through the bias and scaling parameters introduced above, denoted, B and Z, respectively. Critically, we exhaustively search a broad range of values for B and Z to comprehensively map the parameter landscape and select the corresponding optimal working point for the E:I model ($B = -0.3;\ Z = 1.8$), PC1 model ($B = -0.75;\ Z = 1.6$), and T1w:T2 model ($B = -0.7;\ Z = 1.4$), as shown in Figure 2D-F for the E:I model. As can be seen in this figure, there is some degeneracy in the parameter landscapes but there is also a clear regime in which all three empirical properties are fitted well by the model, particularly for $Z \approx 1.7$.

For these optimal values, we simulate each dynamical model 1000 times to account for the inherent stochasticity of the models and compute the respective measures of model fit. Figures 2G-I show the distributions of fit statistics across runs for the homogeneous and three heterogenous models. In addition, we show results for a null ensemble of models in which the regional E:I receptor gene expression values were spatially rotated (*32*, *33*) to generate surrogates with the same spatial autocorrelation as the empirical data. Across all three benchmark properties to which the data were fitted—static FC, GBC, and FCD—the heterogeneous models perform better than the homogeneous model ($8.64 < z < 38.70$, all $p_{bonf} < .05$). We also find a consistent gradient of performance across all benchmarks, with the E:I model performing best, followed by the PC1 and T1w:T2w models, and the homogeneous model showing the poorest performance. For each benchmark metric, the performance of the E:I model was significantly better than all other models ($33.12 < z < 38.70$, all $p_{bonf} < .05$). Out-of-sample model fit statistics yielded similar findings (Figure S2), supporting the generalizability of our results. Nonetheless, differences in model fit statistics were small. For instance, in Figure 2G, the median

static edge-level FC correlation between model and data was $r = 0.71$ for the E:I model, $r = 0.69$ and $r = 0.70$ for the T1w:T2w and PC1 models, respectively, and $r = 0.67$ for the BEI model. For static node-level FC, the correlations ranged between $r = 0.91$ (E:I model) and $r = 0.87$ (BEI model) (Figure 2H). Similarly, the fits to FCD ranged between $D_{KS} = 0.08$ for the E:I model and $D_{KS} = 0.11$ for the BEI model (Figure 2I). However, despite these small differences in model fits, we show in the following sections marked differences between models in other dynamical properties.

Ignition capacity

We now evaluate additional dynamical properties of each model that cannot be directly evaluated with respect to empirical fMRI data but which are thought to be essential features of complex neural dynamics. We first consider ignition capacity. Ignition refers to the capacity of a sensory input to trigger self-supporting, reverberating, temporary, metastable, and distributed network activity, and is thought to be a necessary condition for conscious perception of the stimulus (*4*, *34*). This process is non-linearly related to stimulus intensity, such that evoked activity is confined to local sensory areas for low levels of stimulation (which supports subliminal perception) and then leads to explosive activation of distributed cortical systems once an appropriate threshold is reached.

Figure 1F-G shows the strategy that we follow to quantify ignition capacity. For a given specific working point of the whole-brain model, we compute how the brain broadcasts information after artificially stimulating the lateral occipital cortex region of the Desikan-Killiany atlas (*35*) bilaterally, which contains the foveal representation of V1. We measure the evoked responses at the level of population firing rates rather than simulated BOLD signal changes to have direct access to the millisecond timescale. To quantify the effect of occipital stimulation on activity in each of the other 66 brain regions, we plot, for each non-stimulated region, how its population firing rate changes as a function of occipital stimulation intensity. Two quantities are relevant here: (1) the maximum firing rate achieved at the highest stimulation intensity, r_{max}; and (2) the speed with which the firing rate increases beyond a given intensity threshold, c_{max}, which we quantify as the concavity of the regional response function (i.e., the maximal second derivative). We then summarize the ignition capacity of each non-stimulated region i as $I(i) = c_{max}(i) \times r_{max}(i)$, and estimate the global ignition capacity of the brain as the mean across all regions, $I = \frac{1}{N-1} \sum I(i)$.

Figure 3A shows the response time courses of all left hemisphere regions of the E:I model to the maximum stimulus intensity. Strong responses were evoked throughout the brain, including prefrontal cortex. Figure 3B shows cortical surface renderings of regional $I(i)$ values obtained following occipital stimulation over 30 runs of the homogeneous BEI and heterogeneous E:I models. The occipital stimulation in the E:I model triggers more widespread ignition-like propagation than in the BEI model. This observation is confirmed quantitatively in Figure 3C. Specifically, the E:I model shows higher ignition capacity than all other models ($20 < z < 39$, all $p_{bonf} < 0.05$), suggesting that this model is able to rapidly propagate focal perturbations with higher fidelity than in the other models. Notably, the spatial null models showed higher ignition capacity than the PC1, T1w:T2w and BEI models ($17.4 < z < 33.5$, all $p_{bonf} < 0.05$), suggesting that generic spatial gradients with the same autocorrelation as the E:I maps can yield strong ignition-like dynamics on average, although considerable variability was observed around the median (Figure 3C). The PC1 model showed significantly greater ignition capacity than the both the T1w:T2w and BEI models ($z = 32.5, p =< .001$ and $z = 29.2, p < 0.001$, respectively), which did not differ from each other ($z = 1.22, p = 0.22$). A separate evaluation of ignition-like dynamics following stimulation of the postcentral gyrus, which contains primary somatosensory cortex, also revealed greater ignition capacity for the E:I compared to BEI model, although the differences were not as pronounced as in the visual system (Figure S3). Together, these results indicate that ignition-like propagation following focal visual visual stimulation is most potent when regional heterogeneity is constrained by the transcriptional activity of E:I receptor genes.

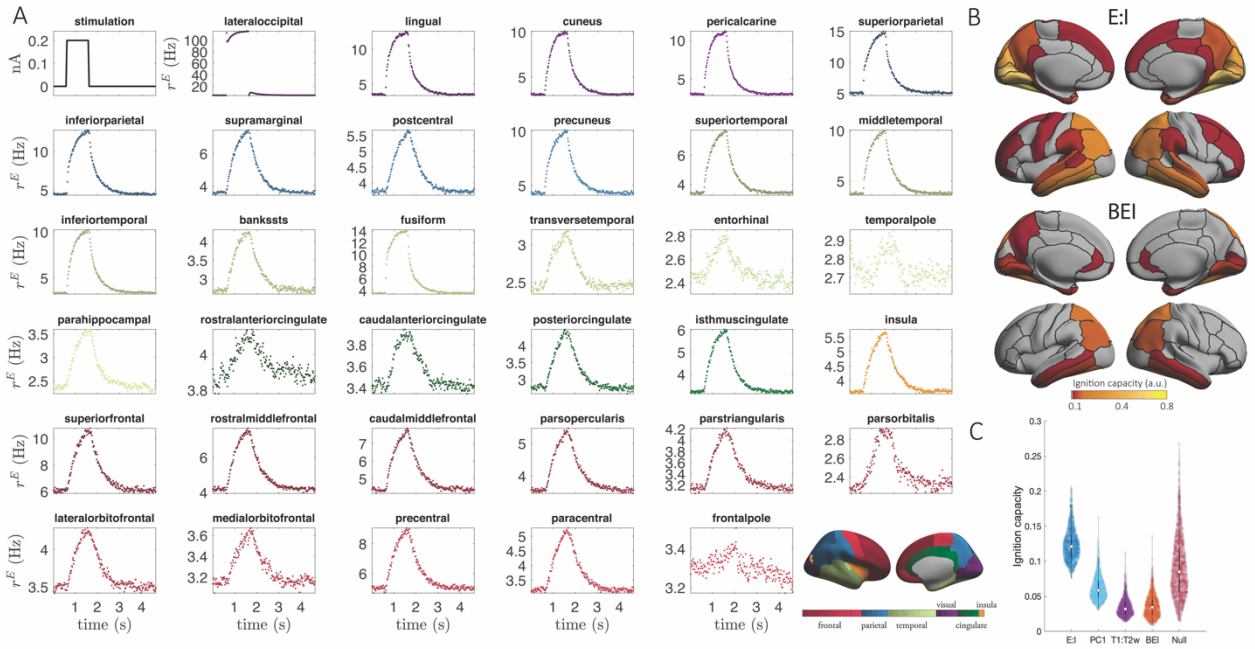

Figure 3. Evaluating the ignition capacity of each dynamical model. A: Time course of the impulse stimulation (top left) used to study ignition-like dynamics and decay timescales, followed by regional responses to the maximum stimulus intensity for all left hemisphere regions. Bottom right shows each region on a cortical surface map with the colours corresponding to the response traces and the stimulation site indicated by the yellow marker. **B:** Cortical surface renderings of average regional ignition capacity values, measured in arbitrary units (a.u.), obtained after lateral occipital cortex stimulation in the E:I model (top) and BEI model (bottom). The E:I model shows widespread ignition-like propagation across distributed areas extending to prefrontal cortex whereas the BEI model shows negligible ignition of non-visual areas. **C:** Distributions of global ignition capacity values obtained for 1000 different runs of the optimal E:I, PC1, T1w:T2w, BEI, and spatial null models. Ignition capacity is highest in the E:I model.

Regional activity timescales

We characterize temporal hierarchies in the models using two approaches, following Chaudhuri et al. (*15*). First, we examine the decay rate of each region's activity following impulse stimulation of lateral occipital cortex using the same virtual stimulation approach as in our analysis of ignition capacity (Figure 1H). Specifically, we apply a virtual impulse stimulation and quantify the temporal decay of activity in each non-stimulated target area by fitting an exponential curve to its firing rate as it returns to spontaneous conditions after termination of the maximal stimulation intensity (Figure 1H). The value of the exponent can be used to index the decay timescale of each brain region and offers one way to examine temporal hierarchies in the brain (*15*) (see Methods).

These fitted exponential curves are shown for all regions in the E:I model in Figure 4B. A clear trend is evident for rapid decay in visual regions and more prolonged responses in prefrontal areas. Supporting the validity of this putative temporal hierarchy, we find a significant positive correlation between regional decay values and an independent measure of cortical hierarchy, T1w:T2w ($\rho = 0.44$, $p_{spatial} < .001$; Figure 4C), such that higher T1w:T2w is associated with more rapid decay timescales (higher decay exponents).

We next estimate the intrinsic timescale of each region's activity using the autocorrelation of its activity time series simulated during continuous white noise stimulation of lateral occipital cortex. This approach is more directly comparable to the experimental paradigms previously used to characterize intrinsic timescales of neuronal activity (*17*, *36*, *37*). Figure 4C shows the fitted autocorrelation functions for every brain region for the E:I model (se Methods). Once again, a clear trend is evident for more rapid timescales in visual areas and slower timescales in prefrontal regions. We define the richness of regional timescales in the model as the inter-quartile range of decay rates across target areas, such that high values reflect a greater diversity of timescales across the brain, which implies a broader dynamic range for information processing. We term this measure regional decay variability. Figure 4D indicates that the PC1 model shows the highest decay variability, followed by the E:I and spatial null models, all of which show greater decay variability than homogeneous BEI model ($14.2 < z < 38.2$, all $p_{bonf} < 0.05$). The decay variability of the BEI and T1w:T2w models is comparable ($z = 0.25, p = 0.80$). Importantly, both the E:I and PC1 models show greater decay variability than the spatial null model ($z = 16.45, p < 0.001$ and $z = 38.54, p < 0.001$, respectively) indicating that empirical transcriptomic maps of heterogeneity confer a wider dynamic range than generic spatial gradients with similar autocorrelation.

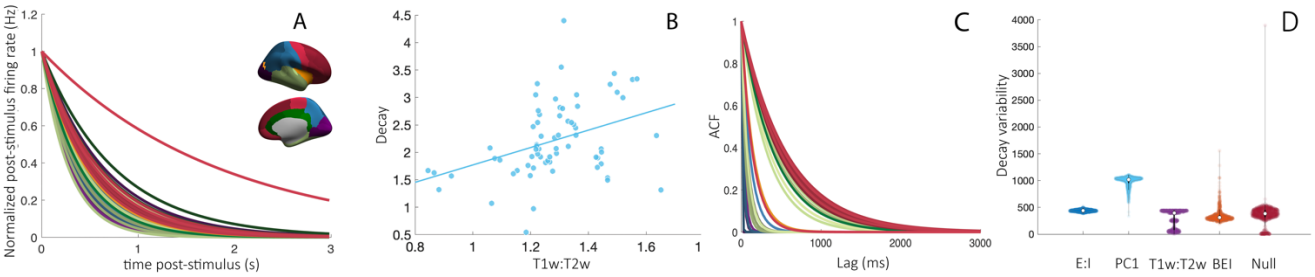

Figure 4. Timescales of regional dynamics. A: Normalized decay curves for all left hemisphere brain regions showing the rate at which regional activity returns to baseline following an impulse stimulation to lateral occipital cortex in the E:I model. Decay rates are slower in prefrontal compared to visual areas. Colours correspond to the regions shown in the inset. **B**: Association

between regional decay rates (in s^{-1}) in the E:I model and empirical T1w:T2w values. **C:** Intrinsic activity timescales (in ms), quantified as autocorrelation functions of each brain region obtained following an input of continuous white noise to lateral occipital cortex. Prefrontal areas again show prolonged timescales. **D:** Variability of intrinsic regional timescales (inter-quartile range) obtained for 1000 runs of each model.

Discussion

It has long been recognized that cortical areas vary along multiple dimensions, including cytoarchitecture, myeloarchitecture, chemoarchitecture, and inter-regional connectivity (*8–10*, *38*). The specific dynamical consequences of each of these variations remain a mystery, but they are likely to have an ultimate influence on the local ratio of excitatory to inhibitory cell activity, resulting in a variable E:I balance across different brain regions. Incorporating such heterogeneity into large-scale models of neuronal dynamics has proven challenging, but the recent availability of diverse data that can be used to index different forms of cellular, molecular, and anatomical variability across brain areas has provided a new means for imposing local variations in such models, thus offering an opportunity to understand in detail the dynamical effects of different forms of regional heterogeneity.

To this end, we compared models in which regional heterogeneity was constrained by transcriptional measures of E:I receptor gene expression, the dominant mode of brain-specific gene expression, and T1w:T2w, which is sensitive to variations in myeloarchitecture and which closely tracks regional position in the cortical hierarchy (*28*). We show that the transcriptional E:I model more faithfully reproduces empirical properties of static and dynamic FC while also showing the highest ignition capacity and a broad range of regional activity timescales. Our findings thus highlight a central role for regional variations in E:I balance in generating ignition-like dynamics and timescale hierarchies, and suggest that transcriptional atlas data provide a viable means for constraining heterogeneity in large-scale models of brain activity.

Evaluating different forms of regional heterogeneity

The starting point for our analysis was an evaluation of how each model captured diverse empirical properties of FC. Early work focused on reproducing the correlation between model and empirical static FC matrices at the edge-level (*11*, *31*). Subsequent studies introduced additional constraints, such as evaluating the fit to dynamic FC properties (*25*) or regional variations in average FC (*22*). To our knowledge, no prior study has evaluated all three properties

simultaneously. Our analysis shows that edge- and node-level measures of static FC offer loose constraints for model optimization, showing comparably high fit statistics across a broad range of values of the global coupling parameter, G. In contrast, fits to FCD show a clear optimum (at $G = 2.1$), mirroring similar results reported previously (*25*). Fitting models to both static and dynamic properties of FC is thus important for identifying an appropriate working point for each model.

Across all three FC properties, heterogeneous models provide a better match to the data than the nominally homogeneous BEI model. This result suggests that regional heterogeneity plays a role in shaping the empirical FC architecture of spontaneous BOLD-dynamics. Critically, the E:I model was the best-performing model across all three benchmarks, indicating that constraining regional heterogeneity by transcriptional markers of E:I balance yields a more faithful replication of empirical FC. However, the differences in fit statistics between models are small, suggesting that these statistics have only a limited capacity to tease apart dynamical differences between the models.

The model-based evaluations of ignition capacity and decay variability provide a more complete picture. In both cases, differences between the homogeneous BEI model and heterogeneous models, particularly those constrained by transcriptomic measures, were apparent. The median global ignition capacity across runs was 112% larger for the E:I model compared to the BEI model, while the median intrinsic timescale variability was 34% larger for the E:I model. The median timescale variability of the PC1 model was particularly high, being 81% larger than the E:I model, which was the second most variable. These results indicate that both ignition-like dynamics and timescale hierarchies depend on some degree of regional heterogeneity, and that biologically plausible forms of transcriptional heterogeneity increase these properties beyond the expectations of a generic spatial gradient (as represented by the spatially constrained null ensemble).

The E:I model showed the greatest ignition capacity, indicating that regional differences in E:I receptor gene expression may shape local dynamics in a way that promotes rapid and high-fidelity communication within the visual system. The E:I model also showed a wide range of intrinsic timescales, although the PC1 model showed the highest variability, implying that the dominant mode of gene expression in the brain, which correlates with cortical hierarchical position (*28*), may facilitate a broad dynamic range of information processing. However, whether increased timescale variability in a model provides a more faithful representation of neural dynamics

remains unclear. Further work could aim to validate model results against either invasive recordings, which allow precise estimation of decay timescales to focal perturbations, or through specifically designed functional MRI experiments (*39*, *40*). Critically, we evaluated the dynamical properties of each model at their empirically optimal working points, meaning that the models show these properties in a parameter regime that closely matches actual data. This is an important constraint, given that prior modelling work evaluating ignition dynamics and timescale hierarchies in Macaque did not fit empirical FC data (*3*, *15*). It is thus difficult to ascertain the biological plausibility of the dynamical regime that was studied. The E:I model considered here not only showed the best fit to empirical data, but it also showed the highest ignition capacity and a broad range of decay timescales, suggesting that it offers a parsimonious account of these different dynamical phenomena.

How does heterogeneity shape neuronal dynamics?

In our models, we introduced heterogeneity by modifying the regional excitability of local population activity, as determined by each region's gain response function. Different approaches have been used to incorporate regional heterogeneity in past work (*15*, *22*, *23*, *41*, *42*). For example, Wang et al. (*23*) used a "bottom-up" approach in which large-scale model inversion was used to infer regional variations in recurrent excitation strengths. The inferred parameters suggested that recurrent excitation is stronger in sensory and weaker in association cortices. This result contradicts the "top-down" approaches used by Chaudhuri et al. (*15*) and Demirtaş et al. (*22*), where regional heterogeneity was constrained using an independent biological property of hypothetical relevance. More specifically, Chaudhuri et al. (*15*) used a hierarchical ordering of cortical areas in macaque based on laminar patterning of inter-regional connectivity to linearly scale the excitatory input strength of each population. Demirtaş et al. used T1w:T2w to linearly scale the strengths of local excitatory-to-excitatory and excitatory-to-inhibitory populations when modelling cortical BOLD FC in humans, given evidence that T1w:T2w tracks hierarchical position and is negatively correlated with pyramidal spine density (*15*, *28*, *43*). In both cases, recurrent excitation was assumed to increase from sensory to association areas. Our analysis also used a top-down approach, relying on transcriptomic estimates of inhibitory and excitatory receptor expression to modify regional gain functions. Our resulting E:I ratio map showed only a moderate positive correlation with T1w:T2w ($\rho = 0.40$), indicating that it quantifies a distinct aspect of regional microcircuit heterogeneity. Our evaluation of model fits to empirical FC properties indicates that the specific aspect of heterogeneity captured by the transcriptomic E:I

ratio plays a more central role in driving realistic brain-wide dynamics when compared to the global mode of gene expression and T1w:T2w.

A comprehensive comparison between these different approaches for incorporating heterogeneity is complicated by the different methods used for defining heterogeneity (top-down vs bottom-up), different constraints on heterogeneity (T1w:T2w vs transcriptomic data), and different ways of modelling heterogeneity (modifying recurrent connection strengths vs regional gain functions). We focused on modulating population gain response functions since local variations in E:I balance will affect the net excitability of the population, which is captured by the gain function parameter, M_i. We thus assume that changes in regional gain are the final common effect of different ways in which heterogeneity might influence a specific population or interaction between populations. Critically, regional variations of M_i are modulated by just two terms—the bias, B, and the scaling factor, Z—which leads to a marked improvement in computational efficiency when compared to more highly parameterized models (e.g., (*22*, *23*)) that require the use of heuristic optimization algorithms. Such algorithms can shift some parameter values from empirically-constrained ranges and do not guarantee a globally optimum solution. The further refinement of optimization algorithms capable of robustly and effectively searching high-dimensional parameter spaces will allow a more comprehensive comparison of different heterogeneous models.

Implications for understanding ignition-like dynamics

Ignition-like dynamics are thought to be a necessary condition for conscious experience (*4*, *34*). As the intensity of a focal sensory stimulus increases, the ignition model predicts that a threshold is reached beyond which widespread activation of other areas is rapidly triggered. Stimuli that do not reach this threshold may only be perceived subconsciously.

Our results indicate that regional heterogeneity of the E:I ratio plays an important role in driving ignition-like dynamics. To the extent that our transcriptional E:I measures can be taken as a proxy for receptor abundance (see below), this result suggests that regional variations specifically in E:I receptor activity may play an especially critical role in supporting rapid, high-fidelity broadcasting of sensory stimuli throughout the network.

Recent work using a dynamical model simulated on the Macaque connectome also showed evidence of ignition-like dynamics (*3*). Critically, the authors showed that removal of feedback

projections and randomization of the connectivity weights of the SC matrix were each sufficient to abolish this dynamical behaviour, suggesting that feedback connectivity and heterogeneity in SC connection strength give rise to ignition-like activity. In our analysis, the SC matrix is derived from diffusion MRI data and thus cannot distinguish feedforward from feedback connections, although we did incorporate empirical variability in connection strengths. Thus, the fact that we observe ignition-like dynamics with heterogeneous but not homogeneous models suggests that feedback connectivity and heterogeneous edge weights alone are not sufficient for ignition. Instead, heterogeneity of regional population dynamics, and of E:I balance in particular, may represent a minimal additional requirement for the emergence of ignition-like propagation throughout the network.

While this propagation was most pronounced following visual stimulation, we also found some evidence for greater ignition capacity in the E:I model compared to the homogeneous BEI model following stimulation of somatosensory cortex. This result supports the generality of our findings, although the differences between the models were less pronounced than those observed in the visual system. This effect could be driven by variations in the connectivity of the stimulation sites, the specificity with which these sites are parcellated, intrinsic differences in the ignition capacity of visual and somatosensory systems, or the possibility that regional E:I variations may be well-placed to support ignition-like propagation through the visual processing hierarchy, which is consistent with anterior-to-posterior gradients in excitatory and inhibitory cell density (*44*, *45*). The phenomenon of ignition has been most extensively described in the visual system (e.g., (*46*)). Studying the properties of ignition-like propagation in different sensory systems will be an important avenue of further work (*47*). Moreover, while our approach offers one way to quantify ignition-like propagation throughout the brain, it will be important to develop more refined quantitative estimates that can be directly related to physiologically validated markers of ignition-driven changes in regional activity.

Implications for understanding timescale hierarchies

We assessed timescale hierarchies by examining regional dynamics following both impulse and continuous stimulation of visual cortex. These analyses converged to identify prolonged activity timescales in prefrontal and association areas compared to primary regions. Comparison of regional intrinsic timescale variability across models indicated that the E:I, PC1, and spatial null models showed a wider dynamic range than the T1w:T2w and homogeneous BEI models, with the E:I and PC1 models both showing greater variability than the spatial null. Thus, spatial maps

with an autocorrelation that matches the gene expression data can enhance the dynamic range of the biophysical model, but the largest gains come when using empirical maps of transcriptional heterogeneity. The relatively wide range of the PC1 model implies that the dominant mode of gene expression, which correlates with T1w:T2w and is related to cortical hierarchy (*28*), may play a particularly important role in shaping timescale differences between regions.

Past work has shown that several other properties correlate with various measures of regional activity timescales. In a comprehensive survey of 6390 different regional time series properties estimated for fMRI recordings in the mouse brain, Sethi et al. (*19*) found that timescale-related measures such as relative high-frequency signal power, correlate with structural measures of in-degree, such that the time series of regions with many incoming connections showed greater low-frequency power. Similar results were later reported in human (*18*). Other work in humans has also shown that regional variations in autocorrelation and other time series properties track the cortical hierarchy (*20*, *48*). These additional measures could be used to provide additional constraints on model fitting, although the temporal resolution of functional fMRI may be too limited to provide a detailed understanding of regional variations in activity timescales.

Limitations

We relied on a single coarse scale parcellation of the brain to allow exhaustive parameter sweeps over several different models. This choice was essential to reduce computational burden, and it ensures that our results are directly comparable with other studies using the same parcellation (e.g., (*11*)). Nonetheless, parcellation type and scale can have an effect on empirical connectome properties (*49*, *50*), and further evaluation of the degree to which our results generalize to other parcellations is required.

Our E:I model used regional measures of excitatory and inhibitory receptor gene expression as a proxy for the relative abundance of each receptor type in each region, which in turn is assumed to be related to the relative activity levels of inhibitory and excitatory cells. Gene expression in the AHBA was assayed using microarray, which quantifies messenger RNA (mRNA) abundance. The correspondence between mRNA levels and protein abundance can be complex and may vary across different tissues or brain regions (*51*). Recent work has used autoradiographic data to derive maps of regional E:I (*52*), which provides a more direct estimate of protein abundance. Qualitatively, the maps resulting from these two methods show some convergence, although some discrepancies are also apparent. Direct comparison between these maps is difficult due to the

different parcellations used, the incomplete anatomical coverage and laminar segregation of the autoradiographic data, and different methods for estimating the E:I ratio. More specifically, where Goulas et al. (*52*) used data from 15 different receptors, including those for neuromodulatory systems such as dopamine, serotonin, and acetylcholine (derived from *53*), we relied on a more conservative list of genes directly related to receptors with a clear role in excitation and inhibition. Future work can use the analysis framework developed here to directly compare the efficacy of different estimates of receptor abundance and E:I ratios in improving model performance. Such efforts will also benefit from the increasing availability of single-cell RNAseq data, which can provide more precise estimates of regional cell-type abundances (*54*). Critically, any imprecision in our estimates should limit the ability of the E:I model to yield physiologically-plausible dynamics. Our results may therefore represent a lower bound on the accuracy that such a model may achieve.

In summary, we show that regional heterogeneity improves model fits to empirical FC properties, particularly when models are constrained by transcriptomic estimates of E:I receptor gene expression. Critically, at the same working point, this model also shows the highest ignition capacity and a broad range of activity timescales, suggesting that regional variations in the activity levels of excitatory and inhibitory receptors may play an important role in the emergence of these dynamical properties. Our analysis indicates that recently constructed transcriptional atlas data provide a fruitful and biologically principled means for imposing regional heterogeneity in large-scale dynamical models.

Methods

Empirical MRI data

Image acquisition

Eyes-closed resting-state fMRI data were acquired in 389 healthy individuals (170 males; mean age = 24 years; SD = 4.2 years) using a Siemens Skyra 3 Tesla scanner with a 32-channel head coil located Monash Biomedical Imaging in Clayton, Victoria, Australia using the following parameters: time-to-repetition (TR), 754 ms; time-to-echo (TE), 21 ms; flip angle 50°; multiband acceleration factor, 3; field-of-view, 190 mm; voxel size, 3 mm^3 isotropic. A total of 620 volumes were acquired. T1-weighted structural images were also acquired using 1mm^3 isotropic voxels; TR = 2300 ms; TE = 2.07 ms; FOV=256 x 256 mm. Diffusion MRI data were available for a subset of 289 individuals, acquired using the following parameters: 2.5mm3 voxel size, TR = 8800 ms, TE = 110 ms, FOV of 240 x 240 mm, 60 directions with b = 3000 s/mm2 and seven b =

0 volumes. A single b = 0 s/mm2 volume was obtained with the reverse-phase encoding for use in distortion correction. Further details can be found in (*55*, *56*). All participants were recuited and data required in accordance with local ethics committee gulidenes.

DWI processing and Connectome construction

The diffusion data were processed using MRtrix3 version 3.0 (https://www.mrtrix.org/) and FSL version 5.0.11 (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki), as detailed in Oldham et al. (*56*). Briefly, the diffusion images for each individual were corrected for eddy-induced current distortions, susceptibility-induced distortions, inter-volume head motion, outliers in the diffusion signal (*57*), within-volume motion (*58*), and B1 field inhomogeneities (*59*). Tractography was conducted using the Fibre Orientation Distributions (iFOD2) algorithm, as implemented in MRtrix3 (*60*), which utilises fibre orientation distributions estimated for each voxel using constrained spherical deconvolution, which can improve the reconstruction of tracts in highly curved and crossing fibre regions (*61*, *62*). Streamline seeds were preferentially selected from areas where streamline density was under-estimated with respect to fibre density estimates from the diffusion model (*63*). We used Anatomically Constrained Tractography to further improve the biological accuracy of streamlines (*64*). To create a structural connectivity matrix, streamlines were assigned to each of the closest regions in the parcellation within a 5mm radius of the streamline endpoints (*65*), yielding an undirected 82 × 82 connectivity matrix. A group-representative connectome was generated by selecting 30% of edges showing the lowest coefficient of variation (based on the streamline count) across participants (*66*). Threshold selection is a challenging problem in human connectomics (*67*) and the correct connection density depends on numerous factors, such as the sensitivity/specificity of the tractography method used, the spatial resolution of the parcellation, and the size of the human brain (relative to other species) (*68*). An important goal for future work will be to understand how different parcellation and thresholding strategies affect model fits.

fMRI processing and FC estimation

The functional MRI data were processed as detailed in (*55*). Briefly, the functional images were initially preprocessed in FSL FEAT following a basic pipeline, which included removal of the first four volumes, rigid-body head motion correction, 3mm spatial smoothing to improve signal-to-noise ratio, and high pass temporal filter of 75s to remove slow drifts (*69*). The data were then denoised using FSL-FIX, an independent component analysis (ICA)-based denoising method that uses an automated classifier to identify noise-related components for removal from the data (*70*).

The algorithm was trained on a manually labelled held-out set of 25 individuals scanned with identical imaging parameters. Time courses for noise-labelled components, along with 24 head motion parameters (6 rigid-body parameters, their backwards derivatives, and squared values of the 12 regressors), were then removed from the voxelwise functional MRI time series using ordinary least squares regression.

Denoised functional data were spatially normalized to the International Consortium for Brain Mapping 152 template in Montreal Neurological Institute (MNI) space using ANTs (version 2.2.0) (*71*), via a three-step method: 1) registration of the mean realigned functional scan to the skull-stripped high resolution anatomical scan via rigid-body registration; 2) spatial normalization of the anatomical scan to the MNI template via a nonlinear registration; and 3) normalization of functional scan to the MNI template using a single transformation matrix that concatenates the transforms generated in steps 1 and 2. Mean time series for each parcellated region were then extracted and Pearson correlations between each pair of regional time series were used to estimate inter-regional FC matrices. Functional connectivity dynamics (FCD) were also calculated for the empirical data, as outlined below.

Cross-validation sample

To test the generalizability of our model fit estimates, we examined the out-of-sample performance of our models in fMRI data from an independent sample of 1003 individuals drawn from the Human Connectome Project (HCP). The HCP website (http://www.humanconnectome.org/) provides the full details of the participants, acquisition protocol, and preprocessing of the resting state fMRI data, which were processed using the minimal preprocessing pipeline and ICA-FIX for denoising (see (*72*)). Regional time series and functional connectivity estimates were generated using the same parcellation.

The Balanced Excitation-Inhibition (BEI) Dynamic Mean Field Model

We simulated spontaneous neuronal activity using a whole-brain model that couples the local population dynamics of different brain areas through an anatomical inter-regional connectivity matrix derived from diffusion-weighted MRI data. The local population dynamics of each area are described by a dynamic mean field model, as proposed by (*11*). The model expresses the activity of large ensembles of interconnected excitatory and inhibitory spiking neurons as a reduced set of dynamical equations describing the population rate activity of coupled excitatory (E) and inhibitory (I) pools of neurons, following the original derivation of Wong and Wang (*73*).

At the local neuronal population level, the excitatory synaptic currents, $I^{(E)}$, are mediated by NMDA receptors and the inhibitory currents, $I^{(I)}$, are mediated by $GABA_A$ receptors. Within each brain area, the E and I neuronal pools are mutually connected. The inter-area coupling between local neuronal population dynamics corresponding to two different brain areas *i* and *j* is mediated only at the E-to-E level and defined by the scaled structural connectivity matrix *C*. The elements of this matrix, $C_{ij,}$ define the inter-regional anatomical connectivity with connection weights estimated using tractography of diffusion-weighted MRI data, as described above.

The balanced dynamic mean-field whole-brain model is mathematically described by the following system of coupled differential equations:

$$I_i^{(\mathrm{E})} = I_i^{ext} + W_E I_0 + w_+ J_{NMDA} S_i^{(E)} + G J_{NMDA} \sum_j C_{ij} S_j^{(E)} - J_i S_i^{(I)}, \tag{1}$$

$$I_i^{(\mathrm{I})} = W_I I_0 + J_{NMDA} S_i^{(E)} - S_i^{(I)}, \tag{2}$$

$$r_i^{(E)} = H^{(E)}(I_i^{(E)}) = \frac{a_E I_i^{(E)} - b_E}{1 - \exp(-d_E(a_E I_i^{(E)} - b_E))}, \tag{3}$$

$$r_i^{(\mathrm{I})} = H^{(\mathrm{I})}(I_i^{(\mathrm{I})}) = \frac{a_I I_i^{(\mathrm{I})} - b_I}{1 - \exp(-d_I(a_I I_i^{(\mathrm{I})} - b_I))}, \tag{4}$$

$$\frac{dS_i^{(E)}(t)}{dt} = -\frac{S_i^{(E)}}{\tau_E} + (1 - S_i^{(E)})\gamma r_i^{(E)} + \sigma \upsilon_i(t), \tag{5}$$

$$\frac{dS_i^{(\mathrm{I})}(t)}{dt} = -\frac{S_i^{(\mathrm{I})}}{\tau_I} + r_i^{(\mathrm{I})} + \sigma \upsilon_i(t), \tag{6}$$

In the last equations, for each excitatory *(E)* or inhibitory *(I)* pool of neurons in each brain area *i,* $r_i^{(E,I)}$ (in Hz) indicates the firing rate, $I_i^{(E,I)}$ (in nA) denotes the total input current, and $S_i^{(E,I)}$ represents the synaptic gating variable. The total input current received by the E and I pools is transformed into firing rates, $r_i^{(E,I)}$, by $H^{(E,I)}$, which corresponds to the sigmoidal neuronal response functions defined in (*74*). In the excitatory *(E)* and inhibitory *(I)* neuronal response functions, $a_E = 310$ nC^{-1} and $a_I = 615$ nC^{-1} define the gain factors determining the slope of *H*, $b_E =$ 0.403 nA, and $b_I = 0.288$ nA are the threshold currents above which the firing rates increase linearly with the input currents, and $d_E = 0.16$ and $d_I = 0.087$ are constants determining the shape of the curvature of *H* around *b*, respectively. The synaptic gating variable of excitatory pools, $S_i^{(E)}$, is

mediated by NMDA receptors with a decay time constant $\tau_{NMDA} = 0.1$ s and $\gamma = 0.641$, whereas the average synaptic gating in inhibitory pools is mediated by GABA receptors with a decay time constant $\tau_{GABA} = 0.01$s. The local external input impinging in each population is $I_0 = 0.382$ nA and is weighted with $W_E = 1$ for the excitatory pools and with $W_I = 0.7$ for the inhibitory pools. The parameter I_i^{ext} captures external stimulation to the excitatory population. In all cases, we set $I_i^{ext} = 0$ for all regions, except when assessing the ignition capacity and response decay profiles of the model, in which case we set $I_i^{ext} > 0$, as detailed below. The local recurrent excitation in each excitatory pool is weighted by $w_+ = 1.4$. All excitatory synaptic couplings are weighted by $J_{NMDA} = 0.15$ nA. In equations (5) and (6) υ_n denotes uncorrelated standard Gaussian noise with an amplitude of $\sigma = 0.01$ nA. All parameters were set as in (*11*). We have shown previously that the synaptic gating variable of NMDA receptors has a much longer decay time constant (100 ms) than the AMPA receptors, which means that the dynamics of the NMDA gating variable dominate the time evolution of the system, while the AMPA synaptic variable instantaneously reaches its steady-state. Our research has shown that we can leave out the contribution of the AMPA receptors and that this greatly reduced approximation does not compromise performance in terms of fitting the empirical data (*31*). Note that the balanced model maintains the spontaneous rate activity of the local population at the same low level across the whole brain, irrespective of the anatomical coupling. This makes the model even more biophysically realistic, as shown in detail in (*11*), and improves the model's dynamical properties and fits to empirical data and improves the encoding capabilities.

One method for introducing heterogeneity to this model is to define the parameters such that each isolated brain area shows asynchronous spontaneous activity with low firing rate ($r^{(E)}$~3Hz and $r^{(I)}$~9Hz), as shown in in electrophysiological experiments (*12–14*). In this regime, the whole-brain model is balanced. The balance is obtained by adjusting each local feedback inhibition weight, J_i, for each brain area *i* such that the firing rate of the excitatory pools $r_i^{(E)}$ remains clamped at 3Hz even when receiving excitatory input from connected areas. This regulation for achieving balance is known as Feedback Inhibition Control (FIC) and is obtained numerically by the method described in (*11*). Indeed, this work demonstrated that FIC improves the fitting of empirical resting-state FC and more crucially, achieves more realistic evoked activity. We refer to this model as the balanced excitation-inhibition (BEI) model. Despite the regional heterogeneity introduced by local tuning of the FIC parameter, we use this model as our homogeneous

benchmark for comparing to other models, since the FIC adjustments ensure uniform firing rates across all regions and all regional gain parameters are fixed (as detailed below).

The BEI model is fitted to empirical fMRI data by finding the optimal global coupling factor *G,* which uniformly scales all inter-area E-to-E connections. The parameter *G* is the only control parameter that is adjusted to move the homogeneous model to its optimal working point, where the simulated activity maximally fits the empirically-observed resting-state FC architecture. In particular, we fit static edge-level FC, node-level FC, and FCD, by choosing the coupling factor *G* such that the correlation between the empirical and simulated FC is high and the distance between the data and model FCD distributions is minimal (i.e., the minimal D_{KS} between empirical and simulated FCD; see below).

Simulations were run for a range of G between 0 and 3 with an increment of 0.01 and with a time step of 0.1 ms. For each value of G, we averaged 389 simulations of 464.464 seconds each, which were then put through a BOLD forward model (see below) in order to emulate the empirical resting-state data, which comprised haemodynamic recordings in 616 volumes acquired in each of 389 human subjects. We optimized the coupling factor G before introducing regional heterogeneity to the model using the gain and scale parameters (i.e., B=0 and *Z*=0 during this fitting procedure; see below*)*.

Simulating BOLD signals

Regional BOLD signals are simulated in the model using the generalized hemodynamic model of (*75*). We calculate the BOLD signal in each brain area *i* from the simulated firing rate of the excitatory pools $r_i^{(E)}$. In this hemodynamic model, an increase in the firing rate causes an increase in a vasodilatory signal, s_i, that is subject to auto-regulatory feedback. Blood inflow, f_i, responds in proportion to this signal, inducing changes in blood volume, v_i, and deoxyhemoglobin content, q_i. The differential equations describing the hemodynamic model coupling these biophysical variables are:

$$\frac{ds_i(t)}{dt} = 0.5r_{iE} + 3 - k_i s_i - \gamma_i(f_i - 1), \tag{7}$$

$$\frac{df_i(t)}{dt} = s_i, \tag{8}$$

$$\tau_i \frac{dv_i(t)}{dt} = f_i - v_i^{1/\alpha}, \tag{9}$$

$$\tau_i \frac{dq_i(t)}{dt} = \frac{f_i(1-(1-\rho_i)^{1/f_i})}{\rho_i} - \frac{q_i v_i^{1/\alpha}}{v_i}. \tag{10}$$

In these equations, ρ denotes the resting oxygen extraction fraction, τ is a time constant and α represents the resistance of the veins. In order to compute in each area i, the BOLD signal, B_i, we calculate a volume-weighted sum of extra- and intravascular signals, which comprises a static nonlinear function of volume, v_i, and deoxyhemoglobin content, qi, and is expressed as follows:

$$B_n = V_0\left[(k_1(1-q_i) + k_2\left(1 - {}^{q_i}/_{v_i}\right) + k_3(1-v_i)\right] \tag{11}$$

We adopted the biophysical parameters recommended by (*75*). To focus on the most functionally relevant frequency range in resting-state conditions, we bandpass filtered both empirical and simulated BOLD signals (0.08 > f > 0.008 Hz).

Introducing regional heterogeneity

One of the key drivers of local population activity in the dynamical model is the balance between excitatory and inhibitory activity. E:I balances across brain regions are thought to play a critical role in synchronized dynamics and their disruptions have been implicated in diverse diseases (*76*, *77*). In the absence of accurate density estimates for each cell type in different regions of the human brain, regional variations in E:I balance could, in principle, be parameterized in many different ways. Here, we use data on regional expression levels of excitatory and inhibitory receptors to constrain our model of E:I heterogeneity across the brain. To this end, we use data from the Allen Institute Human Brain Atlas (AHBA), which comprises microarray data quantifying the transcriptional activity of >20,000 genes in >4,000 different tissue samples distributed throughout the brain, taken from six post-mortem samples (*26*, *30*).

The AHBA data were processed according to the pipeline developed in Arnatkevičiūtė et al. (*29*). Briefly, probe-to-gene annotations were updated using the Re-Annotator toolbox (*78*), resulting in the selection of 45,821 probes corresponding to a total of 20,232 genes. Tissue samples derived from brainstem and cerebellum were removed. Then, probes that did not exceed background noise in more than 50% of remining samples were removed, excluding 13,844 probes corresponding to

4,486 genes. A representative probe for each gene was selected based on the highest mean intensity across regions (*79*). To increase the anatomical accuracy of sample-to-region matching, samples were first divided into four separate groups based on their location: hemisphere (left/right) and structure assignment (cortex/subcortex) and then assigned to the corresponding parcellation-defined regions of interest separately for each of the four groups. Donor-specific grey matter parcellations were generated and samples located within 2 mm of the parcellation voxels were mapped to the closest voxel assigning almost 90% of all tissue samples. We retained only data for the left cortex, as anatomical coverage in the AHBA is more complete for this hemisphere. Finally, gene-expression measures within a given brain were normalized using a scaled robust sigmoid normalization for every gene across samples. Normalized expression measures in samples assigned to the same region were averaged and aggregated into a region × gene matrix consisting of expression measures for 15,745 genes over 34 left hemisphere regions. Given the lack of notable inter-hemispheric differences in regional gene expression identified in the AHBA (*26*), the left hemisphere values were reflected on to the right hemisphere to enable whole-brain model simulations.

To capture regional heterogeneity in E:I balance, we extracted expression measures specifically for genes coding for the excitatory AMPA and NMDA receptors and inhibitory GABA-A receptor isoforms and subunits. We excluded genes encoding GluN3A and GluN3B subunits due to their low expression levels in adult cortex (*80*, *81*). More specifically, the genes selected for AMPA were GRIA1, GRIA2, GRIA3, GRIA4; the genes for NMDA were GRIN1, GRIN2A, GRIN2B, GRIN2C; and the genes for GABA-A were GABRA1, GABRA2, GABRA3, GABRA4, GABRA5, GABRB1, GABRB2, GABRB3, GABRG1, GABRG2, GABRG3. In order to increase the accuracy of gene expression measures, probes representing the expression of individual genes were selected based on the highest correlation to RNA sequencing data in two of the six donor brains (*79*).

In our implementation, we quantify regional E:I balance simply by summing all expression values for AMPA and NMDA, summing all values for GABA, dividing the former by the latter, and then normalizing the values to the unit interval. Specifically, for each brain region, *i,* we estimate $E{:}I_i' = \frac{E{:}I_i}{E{:}I - E{:}I_{min}} - \frac{E{:}I_{max}}{E{:}I_{max} - E{:}I_{min}} + 1$, where $E{:}I_i = (\sum NMDA + \sum AMPA)/\sum GABA$ and $E{:}I_{max}$ and $E{:}I_{min}$ the respective maximum and minimum E:I ratios observed across regions. We use the same normalisation for all other sources of heterogeneity. We note that

estimating E:I balance from gene expression data is challenging given that each receptor is a heteromer, consisting of differing subunit compositions (AMPAR: heterotetrameric assembly from 4 different subunits; NMDAR: heterotetrameric assembly from 7 different subunits; GABA-AR: homo- and heteropentameric assembly from 19 subunits)(*80*, *82–84*). The exact combination of these subunits can alter receptor properties, including biophysical, pharmacological and signalling attributes. For example, the GluN2 subunit is essential for determining the gating properties of NMDA receptors, with different subunit compositions impacting excitatory postsynaptic decay times (*84*). Here we have adopted a pragmatic approach and chosen genes encoding the major subunit types present in the human adult cortex for each receptor, assuming the relative expression levels provide an adequate estimate of the E:I ratio within a given cortical region.

We used the transcriptomically-constrained local E:I ratio R_i to modulate the gain of the local neuronal response functions of the corresponding excitatory and inhibitory pools of each brain region, *i*, following a number of experimental and theoretical studies (*85*, *86*). Specifically, we consider that R_i modulates the gain of the neuronal response function $H^{(E,I)}$ in each brain area according to the modified equations (3) and (4):

$$r_i^{(E)} = H^{(E)}\left(I_i^E\right) = \frac{M_i(a_E I_i^{(E)} - b_E)}{1 - exp(-d_E M_i\left(a_E I_i^{(E)} - b_E\right))}, \quad (12)$$

$$r_i^{(I)} = H^{(I)}\left(I_i^I\right) = \frac{M_i(a_I I_i^{(I)} - b_I)}{1 - exp(-d_I M_i\left(a_I I_i^{(I)} - b_I\right))}. \quad (13)$$

Similar to (*22*), we assume that the heterogeneity factors (in our case, the transcriptomically-defined local excitation/inhibition ratios R_i), are linearly transformed by two unknown free parameters corresponding to a bias *B* and a scaling factor *Z*. Thus, in equations (12) and (13), $M_i = 1 + B + ZR_i$. In other words, the effect of the transcriptomically-constrained E:I ratio R_i is to modulate locally the slope of the corresponding excitatory and inhibitory neuronal response functions, *H*. Note that the slope of H is determined by M using the input-output function of Abbott and Chance (*74*). We studied exhaustively the two free parameters *B* and *Z* regulating the influence of transcriptional heterogeneity. Homogeneous conditions correspond to cases where *Z=0*. Note that the balancing FIC procedure is applied after fixing the gain scaling, i.e. for a specific bias *B* and a specific scaling factor *Z*. Also note that although we constrained the model

using a transcriptional index of E:I balance, we used this constraint to modify the gain of each population model. We did this under the assumption that the population gain function, capturing net regional excitability, will be sensitive to regional variations in E:I balance, and in order to limit the number of fitted parameters in the model. More specifically, we fitted only two parameters—B and Z—to model regional heterogeneity whereas other approaches have fitted four or more (*22*, *23*). We adopted this strategy to enable a more comprehensive search of parameter space, which is important to identify a unified working point for all outcome properties of the model. A more direct modulation of E:I weights might yield different results.

We additionally evaluated two other forms of biologically-constrained heterogeneity: the PC1 and T1w:T2w models. The PC1 model was also informed by AHBA data. In this case, we identified the dominant mode of spatial variation in gene expression as the first principal component (PC1) of 1,926 genes that overlapped between our final quality-controlled gene list and those previously deemed to be brain-specific in (*28*), where it was shown that this mode correlates with T1w:T2w. Regional loadings on this PC were mapped to a positive range before they were incorporated into the model.

The T1w:T2w model incorporated heterogeneity by directly relying on empirical estimates of the T1w:T2w ratio, which is sensitive to myelin content (*27*). Prior work has shown that this measure can also be used to map cortical hierarchies (*28*), and that large-scale heterogeneous biophysical models constrained by T1w:T2w more accurately reproduce empirical FC properties than homogeneous models (*22*). Regional T1w:T2w values were obtained from the Human Connectome Project dataset, averaged over 1200 subjects following 2 mm FHWM Gaussian smoothing in grayordinates space (https://balsa.wustl.edu/study/show/RVVG), and then parcellated using the Desikan-Killany atlas (*35*).

Finally, to further ensure that any results obtained for the E:I model cannot be explained by a generic spatial gradient, we also evaluated model performance after permuting the assignment of E:I values to regions. A total of 10,000 permutations were performed by spatially rotating the expression values using the approach described in (*87*). Critically, this approach preserves the spatial autocorrelation of the expression values, which is essential for ensuring that any resulting effects are not driven by low-order spatial gradients (see also (*32*)). Thus, this method allows us to distinguish the effects of actual E:I heterogeneity from generic spatial gradients with the same

autocorrelation structure (see also (*32*)). In all cases, we normalized the corresponding heterogeneity factor to the [0,1] interval, and treat it similarly to the ratio R_i.

Evaluating model performance

Models were fitted to three empirical properties of the FC data, as detailed in the following. To compare model performance, we simulated 1000 runs of each model at its own optimal parameters to obtain a distribution of fit statistics that captures the inherent stochasticity of the model. To evaluate out-of-sample performance, we repeated the same analyses using the same optimal model working points in an independent sample of 1003 individuals drawn from the Human Connectome Project (Figure S2)

Static edge-level FC

The static edge-level FC is defined as the $N \times N$ matrix of BOLD signal correlations between brain areas computed over the entire recording period (see Figure 1B). We computed the empirical FC for each human participant and for each simulated trial (the total number of trials matched the number of participants). The group-averaged empirical and simulated FC matrices were compared by computing the Pearson correlation between their upper triangular elements (given that the FC matrices are symmetric).

Static node-level FC

Node-level FC, or node strength, is another static spatial measure that characterizes the average FC strength for each area (see Figure 1E) (*68*). It has also been called global brain connectivity in previous work (*22*). Thus, node-level functional connectivity strength is defined as

$$FCS_i = \frac{1}{N}\sum_{j=1}^{N} FC_{ij}. \qquad (14)$$

We computed the empirical strength vectors for each human participant and for each simulated trial (total number of trials matched the number of participants). The node FC fit is quantified by computing the Pearson correlation between the group averaged empirical and simulated strength vectors.

Functional Connectivity Dynamics (FCD)

The spatiotemporal dynamics of resting-state activity are captured by the statistics of how FC computed over sliding windows evolves across time (see Figure 1E). Here, we computed the FC over a sliding window of 80 TRs (corresponding approximately to 1 minute) with incremental shifts of 18 TRs. In order to quantify the spatiotemporal statistics of time-evolving FC, we calculated a time-versus-time matrix, called Functional Connectivity Dynamics (*FCD*) (*25*). This FCD matrix is defined so that each entry, $FCD(t_x, t_y)$, corresponds to the correlation between the *FC* centred at times t_x and the *FC* centred at t_y. The statistical distribution of *FCD* values captures the spatiotemporal architecture of the recording session. In order to compare quantitatively the spatiotemporal dynamical characteristics between empirical data and model simulations, we generate the distributions of the upper triangular elements of the FCD matrices over all participants as well as of the FCD matrices obtained from all simulated trials for a given parameter setting. The similarity between the empirical and model FCD distributions is then compared using the Kolmogorov-Smirnov distance, D_{KS}, allowing for a meaningful evaluation of model performance in predicting the changes observed in dynamic resting-state FC. Note that FCD represents a more stringent benchmark for model performance than static FC, as it requires the model to capture the statistics of how FC varies at each edge over time.

Ignition capacity

In the literature, ignition has been described using multiple complementary perspectives as comprising three key properties: (1) a rapid, non-linear amplification of activity once stimulus intensity has surpassed a certain threshold; (2) widespread activation of spatially distributed systems, particularly frontoparietal areas; and (3) sustained reverberatory activity persisting after termination of the stimulus (*4*, *88*). Our model provides insights into the first two properties but has reduced sensitivity for capturing reverberatory activity due to its focus on noisy fluctuations and the incorporation of local feedback inhibition. In our analysis, we therefore aimed to quantify ignition-like dynamics in our model using a metric that captures elements of both non-linear amplification and widespread activation. More specifically, we quantify the ignition capacity of each model by examining how regional dynamics respond to a simulated focal perturbation. First, for a given working point, we run the whole-brain model for 3000 ms under spontaneous conditions in order to eliminate transients. Second, from 3000 to 4000 ms, we apply a stimulation of strength I_i^{ext} to the bilateral lateral occipital cortex region in the Desikan-Killiany atlas (Figure 3A), which contains the foveal representation of V1 and the confluence of other visual areas. Third, from 4000 ms to 7000 ms we reset the external stimulation to zero again, and let the

system relax again to the spontaneous state. For each condition, i.e., for each parameter setting, and each particular stimulation I_i^{ext}, we run 30 trials and compute the average, across trials, of the elicited activity as a function of time on all brain areas, in order to smooth the resulting neuronal whole-brain activity.

To compare ignition capacity across the whole cortex, we fix the location *i* where the stimulation is applied (and also stimulate the region's contralateral homologue). Then, we simulate the averaged neuronal response elicited in a given brain area as a consequence of the applied artificial stimulation and as a function of the strength I_i^{ext} (we modify I_i^{ext} from 0 to 0.2 in steps of 0.001). For each brain region, we perform a sigmoidal fitting of the elicited averaged rate activity in the period from 3500 to 4000 ms as a function of the applied strength I_i^{ext} in the stimulated specific site *i*. The ignition elicited for a given non-stimulated target brain area after the application of the artificial stimulation in the specific region *i* is defined by the product between the level of concavity of the neuronal response function elicited on that target area and the maximal rate elicited when the maximal strength of stimulation is applied (as fitted by the sigmoidal function). The level of concavity is defined by the maximal value of the second derivative of the fitted sigmoidal function. Figure 1F-G shows graphically the key concept. The global ignition capacity is the average ignition capacity across all brain areas. The ignition measure is basically describing two effects, namely the fidelity (captured by the maximal elicited rate in the target area) and explosiveness (captured by the level of concavity) of ignition-like activity. In secondary analyses, we used the same approach to simulate the effects of somatosensory stimulation by applying the stimulation to the postcentral gyrus region of the Desikan-Killiany atlas (Figure S3).

Activity timescales

We follow Chaudhuri et al. (*15*) and examine activity timescales in two ways. First, we examine the rate at which regional activity decays following the same impulse stimulation paradigm used in the analysis of ignition capacity. More specifically, we compute the temporal decay of each target brain region following simulation of the lateral occipital region in both hemispheres. We calculate the decay rate by fitting an exponential curve to the regional firing rates post-stimulation, averaged over 1000 simulations (the length of total simulation as described above). The functional form of the exponential curve is $r_i^E = A(\exp(-Dt) + B)$, where r_i^E is the firing rate (Hz) post-stimulation, t is the time post-stimulation (in seconds) and the free parameters of decay rate, D, scaling constant, A, and offset, B, are fitted using nonlinear least squares.

Second, we examine intrinsic timescales by estimating each region's autocorrelation function (ACF) measured during an input of continuous white noise to the same stimulation target (lateral occipital cortex). This approach more closely mimics empirical characterizations of intrinsic timescales (e.g., (*17*, *36*, *37*)). To this end, we provide white noise input currents with mean 0.356 nA and standard deviation 0.05 nA, as in Chaudhuri et al. (*15*). We set the noise levels of all other regions to be low, where the noise level in each model was chosen as to maximise the range of timescales with no stimuli present. This resulted in noise levels of $10^{-8}, 10^{-5}, 10^{-8}, 10^{-4}$ and 10^{-8} in the E:I, PCA, T1w:T2w, BEI and spatial null models respectively. For each model, we calculated the average ACF across 30 simulations. To calculate the timescale estimates from these average ACFs, we truncated the ACF to be above 0.05 and fitted the curve with either a single exponential, $ACF(t = lag) = \exp(-t/D)$, or a double exponential, $AC(t = lag) = A[\exp(-t/D1)] + B[\exp(-t/D2)]$ to account for the longer tails in some ACFs. In these equations, t is the lag time (ms), D, $D1$, and $D2$ are decay constants (ms^-1) and A and B are weightings of the exponential functions. Following Chaudhuri et al. (*15*), if the sum of squared errors in the double exponential fit was 8 times larger than in the single exponential, we used the sum of the two time constants (1/D1,1/D2) from the double exponential fit, with each weighted by its amplitude (A and B respectively); otherwise the timescale was calculated as 1/D. These simulations and ACF calculations were repeated 1000 times for each model to build up distributions of the timescale estimates. We then characterized time hierarchy variability as the inter-quartile range of intrinsic ACF timescales across brain areas. A larger inter-quartile range implies greater variability in intrinsic timescales across brain regions, which is consistent with a broader dynamic range of activity and a more pronounced hierarchical ordering of brain areas.

Statistical analyses

Differences in model fits to empirical properties, as well as in ignition capacity and decay variability, were assessed using pair-wise Wilcoxon ranksum tests, Bonferroni-corrected for 10 possible comparisons (corrected p-values are denoted p_{bonf}). Spatial correlations between regional maps were assessed with reference to empirical null distributions generated by rotating one of the spatial maps using the same procedure as used when generating the spatial null models (*89*). The corresponding p-values are denoted $p_{spatial}$.

Acknowledgements

G.D. is supported by the Spanish Research Project COBRAS PSI2016-75688-P (AEI/FEDER, EU), by the European Union's Horizon 2020 Research and Innovation Programme under grant agreements n. 720270 (HBP SGA1) and n. 785907 (HBP SGA2), and by the Catalan AGAUR Programme 2017 SGR 1545. AF was supported by the Sylvia and Charles Viertel Charitable Foundation and National Health and Medical Research Council (ID: 3251549).

Competing interests

The authors declare that they have no competing interests.

Code and Data availability

All data and code needed to evaluate the conclusions in this paper are present in the paper, Supplementary Materials, and available at https://github.com/KevinAquino/HNM.

Author contributions

GD, KMA, and AF conceived the project and wrote the manuscript; GD, KMA, MK, AA, SO, NR, and KS processed and analysed the data and provided feedback on the manuscript.

References

1. E. Bullmore, O. Sporns, Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*. **10**, 186–198 (2009).

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

3. M. R. Joglekar, J. F. Mejias, G. R. Yang, X.-J. Wang, Inter-areal balanced amplification enhances signal propagation in a large-scale circuit model of the primate cortex. *Neuron*. **98**, 222-234.e8 (2018).

4. G. A. Mashour, P. Roelfsema, J.-P. Changeux, S. Dehaene, Conscious Processing and the Global Neuronal Workspace Hypothesis. *Neuron*. **105**, 776–798 (2020).

5. G. Deco, V. Jirsa, A. R. McIntosh, O. Sporns, R. Kotter, Key role of coupling, delay, and noise in resting brain fluctuations. *Proceedings of the National Academy of Sciences*. **106**, 10302–10307 (2009).

6. C. J. Honey, R. Kötter, M. Breakspear, O. Sporns, Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proc Natl Acad Sci U S A*. **104**, 10240–10245 (2007).

7. G. Deco, V. K. Jirsa, Ongoing Cortical Activity at Rest: Criticality, Multistability, and Ghost Attractors. *Journal of Neuroscience*. **32**, 3366–3375 (2012).

8. K. Brodmann, *Vergleichende Lokalisationslehre der Grosshirnrinde* (Johann Ambrosius Barth, Leipzig, 1909).

9. X.-J. Wang, Macroscopic gradients of synaptic excitation and inhibition in the neocortex. *Nature Reviews. Neuroscience*, 1–10 (2020).

10. J. M. Huntenburg, P.-L. Bazin, D. S. Margulies, Large-scale gradients in human cortical organization. *Trends in Cognitive Sciences*. **22**, 21–31 (2018).

11. G. Deco, A. Ponce-Alvarez, P. Hagmann, G. L. Romani, D. Mantini, M. Corbetta, How Local Excitation-Inhibition Ratio Impacts the Whole Brain Dynamics. *Journal of Neuroscience*. **34**, 7886–7898 (2014).

12. B. D. Burns, A. C. Webb, The spontaneous activity of neurones in the cat's cerebral cortex. *Proc R Soc Lond B Biol Sci*. **194**, 211–223 (1976).

13. K. W. Koch, J. M. Fuster, Unit activity in monkey parietal cortex related to haptic perception and temporary memory. *Exp Brain Res*. **76**, 292–306 (1989).

14. W. R. Softky, C. Koch, The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J Neurosci*. **13**, 334–350 (1993).

15. R. Chaudhuri, K. Knoblauch, M.-A. Gariel, H. Kennedy, X.-J. Wang, A Large-Scale Circuit Mechanism for Hierarchical Dynamical Processing in the Primate Cortex. *Neuron*. **88**, 419–431 (2015).

16. P. Barone, A. Batardiere, K. Knoblauch, H. Kennedy, Laminar Distribution of Neurons in Extrastriate Areas Projecting to Visual Areas V1 and V4 Correlates with the Hierarchical Rank and Indicates the Operation of a Distance Rule. *J. Neurosci.* **20**, 3263–3281 (2000).

17. J. D. Murray, A. Bernacchia, D. J. Freedman, R. Romo, J. D. Wallis, X. Cai, C. Padoa-Schioppa, T. Pasternak, H. Seo, D. Lee, X.-J. Wang, A hierarchy of intrinsic timescales across primate cortex. *Nat Neurosci*. **17**, 1661–1663 (2014).

18. J. Fallon, P. G. D. Ward, L. Parkes, S. Oldham, A. Arnatkevičiūtė, A. Fornito, B. D. Fulcher, Timescales of spontaneous fMRI fluctuations relate to structural connectivity in the brain. *Network Neuroscience*. **4**, 788–806 (2020).

19. S. S. Sethi, V. Zerbi, N. Wenderoth, A. Fornito, B. D. Fulcher, Structural connectome topology relates to regional BOLD signal dynamics in the mouse brain. *Chaos*. **27**, 047405 (2017).

20. R. V. Raut, A. Z. Snyder, M. E. Raichle, Hierarchical dynamics as a macroscopic organizing principle of the human brain. *Proc Natl Acad Sci USA*, 202003383 (2020).

21. G. J. Yang, J. D. Murray, X.-J. Wang, D. C. Glahn, G. D. Pearlson, G. Repovs, J. H. Krystal, A. Anticevic, Functional hierarchy underlies preferential connectivity disturbances in schizophrenia. *Proc Natl Acad Sci USA*. **113**, E219–E228 (2016).

22. M. Demirtaş, J. B. Burt, M. Helmer, J. L. Ji, B. D. Adkinson, M. F. Glasser, D. C. Van Essen, S. N. Sotiropoulos, A. Anticevic, J. D. Murray, Hierarchical heterogeneity across human cortex shapes large-scale neural dynamics. *Neuron*, 1–28 (2019).

23. P. Wang, R. Kong, X. Kong, R. Liegeois, C. Orban, G. Deco, M. P. van den Heuvel, B. T. Thomas Yeo, Inversion of a large-scale circuit model reveals a cortical hierarchy in the dynamic resting human brain. *Science advances*. **5**, eaat7854 (2019).

24. D. S. Margulies, S. S. Ghosh, A. Goulas, M. Falkiewicz, J. M. Huntenburg, G. Langs, G. Bezgin, S. B. Eickhoff, F. X. Castellanos, M. Petrides, E. Jefferies, J. Smallwood, Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proceedings of the National Academy of Sciences*. **113**, 12574–12579 (2016).

25. G. Deco, M. L. Kringelbach, V. K. Jirsa, P. Ritter, The dynamics of resting fluctuations in the brain: metastability and its dynamical cortical core. *Sci Rep*. **7**, 3095 (2017).

26. M. J. Hawrylycz, E. S. Lein, A. L. Guillozet-Bongaarts, E. H. Shen, L. Ng, J. A. Miller, L. N. van de Lagemaat, K. A. Smith, A. Ebbert, Z. L. Riley, C. Abajian, C. F. Beckmann, A. Bernard, D. Bertagnolli, A. F. Boe, P. M. Cartagena, M. M. Chakravarty, M. Chapin, J. Chong, R. A. Dalley, B. D. Daly, C. Dang, S. Datta, N. Dee, T. A. Dolbeare, V. Faber, D. Feng, D. R. Fowler, J. Goldy, B. W. Gregor, Z. Haradon, D. R. Haynor, J. G. Hohmann, S. Horvath, R. E. Howard, A. Jeromin, J. M. Jochim, M. Kinnunen, C. Lau, E. T. Lazarz, C. Lee, T. A. Lemon, L. Li, Y. Li, J. A. Morris, C. C. Overly, P. D. Parker, S. E. Parry, M. Reding, J. J. Royall, J. Schulkin, P. A. Sequeira, C. R. Slaughterbeck, S. C. Smith, A. J. Sodt, S. M. Sunkin, B. E. Swanson, M. P. Vawter, D. Williams, P. Wohnoutka, H. R. Zielke, D. H. Geschwind, P. R. Hof, S. M. Smith, C. Koch, S. G. N. Grant, A. R. Jones, An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*. **489**, 391–399 (2012).

27. M. F. Glasser, D. C. Van Essen, Mapping human cortical areas in vivo based on myelin content as revealed by t1- and t2-weighted MRI. *Journal of Neuroscience*. **31**, 11597–11616 (2011).

28. J. B. Burt, M. Demirtaş, W. J. Eckner, N. M. Navejar, J. L. Ji, W. J. Martin, A. Bernacchia, A. Anticevic, J. D. Murray, Hierarchy of transcriptomic specialization across human cortex captured by structural neuroimaging topography. *Nature Neuroscience*. **21**, 1251–1259 (2018).

29. A. Arnatkeviciute, B. D. Fulcher, A. Fornito, A practical guide to linking brain-wide gene expression and neuroimaging data. *Neuroimage*. **189**, 353–367 (2019).

30. A. Fornito, A. Arnatkevičiūtė, B. D. Fulcher, Bridging the Gap between Connectome and Transcriptome. *Trends Cogn Sci*. **23**, 34–50 (2019).

31. G. Deco, A. Ponce-Alvarez, D. Mantini, G. L. Romani, P. Hagmann, M. Corbetta, Resting-State Functional Connectivity Emerges from Structurally and Dynamically Shaped Slow Linear Fluctuations. *Journal of Neuroscience*. **33**, 11239–11252 (2013).

32. A. F. Alexander-Bloch, H. Shou, S. Liu, T. D. Satterthwaite, D. C. Glahn, R. T. Shinohara, S. N. Vandekar, A. Raznahan, On testing for spatial correspondence between maps of human brain structure and function. *Neuroimage*. **178**, 540–551 (2018).

33. F. Váša, J. Seidlitz, R. Romero-Garcia, K. J. Whitaker, G. Rosenthal, P. E. Vértes, M. Shinn, A. Alexander-Bloch, P. Fonagy, R. J. Dolan, P. B. Jones, I. M. Goodyer, NSPN consortium, O. Sporns, E. T. Bullmore, Adolescent Tuning of Association Cortex in Human Structural Brain Networks. *Cereb Cortex*. **28**, 281–294 (2018).

34. S. Dehaene, J.-P. Changeux, Experimental and theoretical approaches to conscious processing. *Neuron*. **70**, 200–227 (2011).

35. R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman, M. S. Albert, R. J. Killiany, An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. **31**, 968–980 (2006).

36. S. E. Cavanagh, J. P. Towers, J. D. Wallis, L. T. Hunt, S. W. Kennerley, Reconciling persistent and dynamic hypotheses of working memory coding in prefrontal cortex. *Nat Commun*. **9**, 3498 (2018).

37. D. F. Wasmuht, E. Spaak, T. J. Buschman, E. K. Miller, M. G. Stokes, Intrinsic neuronal dynamics predict distinct functional roles during working memory. *Nat Commun*. **9**, 3499 (2018).

38. R. E. Passingham, K. E. Stephan, R. Kötter, The anatomical basis of functional localization in the cortex. *Nat Rev Neurosci*. **3**, 606–616 (2002).

39. H.-Y. S. Chien, C. J. Honey, Constructing and Forgetting Temporal Context in the Human Cerebral Cortex. *Neuron*. **106**, 675-686.e11 (2020).

40. Y. Yeshurun, M. Nguyen, U. Hasson, Amplification of local changes along the timescale processing hierarchy. *Proc Natl Acad Sci U S A*. **114**, 9475–9480 (2017).

41. J. F. Mejias, X.-J. Wang, Mechanisms of distributed working memory in a large-scale model of the macaque neocortex. *bioRxiv*. **10.1101/760231** (2019), doi:10.1101/760231.

42. S. Froudist-Walsh, D. P. Bliss, X. Ding, L. Jankovic-Rapan, M. Niu, K. Knoblauch, K. Zilles, H. Kennedy, N. Palomero-Gallagher, X.-J. Wang, A dopamine gradient controls access to distributed working memory in monkey cortex. *bioRxiv*. **2020.09.07.286500** (2020), doi:10.1101/2020.09.07.286500.

43. G. N. Elston, Pyramidal cells of the frontal lobe: all the more spinous to think with. *J Neurosci*. **20**, RC95 (2000).

44. D. J. Cahalane, C. J. Charvet, B. L. Finlay, Systematic, balancing gradients in neuron density and number across the primate isocortex. *Front Neuroanat*. **6**, 28 (2012).

45. B. D. Fulcher, J. D. Murray, V. Zerbi, X.-J. Wang, Multimodal gradients across mouse cortex. *Proc Natl Acad Sci USA*. **116**, 4689–4695 (2019).

46. B. van Vugt, B. Dagnino, D. Vartak, H. Safaai, S. Panzeri, S. Dehaene, P. R. Roelfsema, The threshold for conscious report: Signal loss and response bias in visual and frontal cortex. *Science*. **360**, 537–542 (2018).

47. R. Romo, R. Rossi-Pool, Towards a conscious model of consciousness. *Cognitive Neuropsychology*. **37**, 220–223 (2020).

48. G. Shafiei, R. D. Markello, R. V. de Wael, B. C. Bernhardt, B. D. Fulcher, B. Mišić, Topographic gradients of intrinsic dynamics across neocortex. **178**, 540–17 (2020).

49. A. Fornito, A. Zalesky, E. T. Bullmore, Network scaling effects in graph analytic studies of human resting-state FMRI data. *Front Syst Neurosci*. **4**, 22 (2010).

50. A. Zalesky, A. Fornito, I. H. Harding, L. Cocchi, M. Yücel, C. Pantelis, E. T. Bullmore, Whole-brain anatomical networks: does the choice of nodes matter? *Neuroimage*. **50**, 970–983 (2010).

51. Y. Liu, A. Beyer, R. Aebersold, On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell*. **165**, 535–550 (2016).

52. A. Goulas, J.-P. Changeux, K. Wagstyl, K. Amunts, N. Palomero-Gallagher, C. Hilgetag, The natural axis of transmitter receptor distribution in the human cerebral cortex. *Proc Natl Acad Sci USA*. **118**, e202574118 (2021).

53. K. Zilles, N. Palomero-Gallagher, Multiple transmitter receptors in regions and layers of the human cerebral cortex. *Frontiers in Neuroanatomy*. **11**, e1000489-26 (2017).

54. R. D. Hodge, T. E. Bakken, J. A. Miller, K. A. Smith, E. R. Barkan, L. T. Graybuck, J. L. Close, B. Long, N. Johansen, O. Penn, Z. Yao, J. Eggermont, T. H. x000F6 llt, B. P. Levi, S. I. Shehata, B. Aevermann, A. Beller, D. Bertagnolli, K. Brouner, T. Casper, C. Cobbs, R. Dalley, N. Dee, S.-L. Ding, R. G. Ellenbogen, O. Fong, E. Garren, J. Goldy, R. P. Gwinn, D. Hirschstein, C. D. Keene, M. Keshk, A. L. Ko, K. Lathia, A. Mahfouz, Z. Maltzer, M. McGraw, T. N. Nguyen, J. Nyhus, J. G. Ojemann, A. Oldre, S. Parry, S. Reynolds, C. Rimorin, N. V. Shapovalova, S. Somasundaram, A. Szafer, E. R. Thomsen, M. Tieu, G. Quon, R. H. Scheuermann, R. Yuste, S. M. Sunkin, B. Lelieveldt, D. Feng, L. Ng, A. Bernard, M.

Hawrylycz, J. W. Phillips, B. Tasic, H. Zeng, A. R. Jones, C. Koch, E. S. Lein, Conserved cell types with divergent features in human versus mouse cortex. *Nature*, 1–38 (2019).

55. K. Sabaroedin, J. Tiego, L. Parkes, F. Sforazzini, A. Finlay, B. Johnson, A. Pinar, V. Cropley, B. J. Harrison, A. Zalesky, C. Pantelis, M. Bellgrove, A. Fornito, Functional Connectivity of Corticostriatal Circuitry and Psychosis-like Experiences in the General Community. *Biol Psychiatry*. **86**, 16–24 (2019).

56. S. Oldham, A. Arnatkevic Iūtė, R. E. Smith, J. Tiego, M. A. Bellgrove, A. Fornito, The efficacy of different preprocessing steps in reducing motion-related confounds in diffusion MRI connectomics. *Neuroimage*. **222**, 117252 (2020).

57. J. L. R. Andersson, M. S. Graham, E. Zsoldos, S. N. Sotiropoulos, Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. *NeuroImage*. **141**, 556–572 (2016).

58. J. L. R. Andersson, M. S. Graham, I. Drobnjak, H. Zhang, N. Filippini, M. Bastiani, Towards a comprehensive framework for movement and distortion correction of diffusion MR images: Within volume movement. *NeuroImage*. **152**, 450–466 (2017).

59. Y. Zhang, M. Brady, S. M. Smith, Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*. **20**, 45–57 (2001).

60. J. D. Tournier, R. Smith, D. Raffelt, R. Tabbara, T. Dhollander, M. Pietsch, D. Christiaens, B. Jeurissen, C.-H. Yeh, A. Connelly, MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage*, 116137 (2019).

61. J. D. Tournier, F. Calamante, A. Connelly, Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*. **35**, 1459–1472 (2007).

62. J.-D. Tournier, F. Calamante, D. G. Gadian, A. Connelly, Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *Neuroimage*. **23**, 1176–1185 (2004).

63. R. E. Smith, J.-D. Tournier, F. Calamante, A. Connelly, SIFT2: Enabling dense quantitative assessment of brain white matter connectivity using streamlines tractography. *Neuroimage*. **119**, 338–351 (2015).

64. R. E. Smith, J. D. Tournier, F. Calamante, A. Connelly, Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information. *NeuroImage*. **62**, 1924–1938 (2012).

65. R. E. Smith, J. D. Tournier, F. Calamante, A. Connelly, The effects of SIFT on the reproducibility and biological accuracy of the structural connectome. *NeuroImage*. **104**, 253–265 (2015).

66. J. A. Roberts, A. Perry, G. Roberts, P. B. Mitchell, M. Breakspear, Consistency-based thresholding of the human connectome. *Neuroimage*. **145**, 118–129 (2017).

67. B. C. M. van Wijk, C. J. Stam, A. Daffertshofer, Comparing Brain Networks of Different Size and Connectivity Density Using Graph Theory. *PLoS ONE*. **5**, e13701 (2010).

68. A. Fornito, *Fundamentals of Brain Network Analysis* (Academic Press, Inc, London, 2016).

69. S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. J. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak, D. E. Flitney, R. K. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. De Stefano, J. M. Brady, P. M. Matthews, Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*. **23**, S208–S219 (2004).

70. L. Griffanti, G. Salimi-Khorshidi, C. F. Beckmann, E. J. Auerbach, G. Douaud, C. E. Sexton, E. Zsoldos, K. P. Ebmeier, N. Filippini, C. E. Mackay, S. Moeller, J. Xu, E. Yacoub, G. Baselli, K. Ugurbil, K. L. Miller, S. M. Smith, ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *Neuroimage*. **95**, 232–247 (2014).

71. B. B. Avants, N. J. Tustison, G. Song, P. A. Cook, A. Klein, J. C. Gee, A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage*. **54**, 2033–2044 (2011).

72. M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, D. C. Van Essen, M. Jenkinson, WU-Minn HCP Consortium, The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage*. **80**, 105–124 (2013).

73. K.-F. Wong, X.-J. Wang, A recurrent network mechanism of time integration in perceptual decisions. *J Neurosci*. **26**, 1314–1328 (2006).

74. L. F. Abbott, F. S. Chance, Drivers and modulators from push-pull and balanced synaptic input. *Prog Brain Res*. **149**, 147–155 (2005).

75. K. E. Stephan, N. Weiskopf, P. M. Drysdale, P. A. Robinson, K. J. Friston, Comparing hemodynamic models with DCM. *Neuroimage*. **38**, 387–401 (2007).

76. V. S. Sohal, J. L. R. Rubenstein, Excitation-inhibition balance as a framework for investigating mechanisms in neuropsychiatric disorders. *Mol Psychiatry*. **24**, 1248–1257 (2019).

77. J. H. Krystal, A. Anticevic, G. J. Yang, G. Dragoi, N. R. Driesen, X.-J. Wang, J. D. Murray, Impaired Tuning of Neural Ensembles and the Pathophysiology of Schizophrenia: A Translational and Computational Neuroscience Perspective. *Biol Psychiatry*. **81**, 874–885 (2017).

78. J. Arloth, D. M. Bader, S. Röh, A. Altmann, Re-Annotator: Annotation Pipeline for Microarray Probe Sequences. *PLoS One*. **10**, e0139516 (2015).

79. J. A. Miller, V. Menon, J. Goldy, A. Kaykas, C.-K. Lee, K. A. Smith, E. H. Shen, J. W. Phillips, E. S. Lein, M. J. Hawrylycz, Improving reliability and absolute quantification of human brain microarray data by filtering and scaling probes using RNA-Seq. *BMC Genomics*. **15**, 154 (2014).

80. P. Paoletti, C. Bellone, Q. Zhou, NMDA receptor subunit diversity: impact on receptor properties, synaptic plasticity and disease. *Nat Rev Neurosci*. **14**, 383–400 (2013).

81. I. Pérez-Otaño, R. S. Larsen, J. F. Wesseling, Emerging roles of GluN3-containing NMDA receptors in the CNS. *Nat Rev Neurosci*. **17**, 623–635 (2016).

82. J. M. Henley, K. A. Wilkinson, Synaptic AMPA receptor composition in development, plasticity and disease. *Nat Rev Neurosci*. **17**, 337–350 (2016).

83. E. Sigel, M. E. Steinmann, Structure, function, and modulation of GABA(A) receptors. *J Biol Chem*. **287**, 40224–40231 (2012).

84. S. Vicini, J. F. Wang, J. H. Li, W. J. Zhu, Y. H. Wang, J. H. Luo, B. B. Wolfe, D. R. Grayson, Functional and pharmacological differences between recombinant N-methyl-D-aspartate receptors. *J Neurophysiol*. **79**, 555–566 (1998).

85. G. Deco, J. Cruzat, J. Cabral, G. M. Knudsen, R. L. Carhart-Harris, P. C. Whybrow, N. K. Logothetis, M. L. Kringelbach, Whole-Brain Multimodal Neuroimaging Model Using Serotonin Receptor Maps Explains Non-linear Functional Effects of LSD. *Curr Biol*. **28**, 3065-3074.e6 (2018).

86. M. L. Kringelbach, J. Cruzat, J. Cabral, G. M. Knudsen, R. Carhart-Harris, P. C. Whybrow, N. K. Logothetis, G. Deco, Dynamic coupling of whole-brain neuronal and neurotransmitter systems. *Proceedings of the National Academy of Sciences*. **117**, 9566–9576 (2020).

87. F. Váša, R. Romero-Garcia, M. G. Kitzbichler, J. Seidlitz, K. J. Whitaker, M. M. Vaghi, P. Kundu, A. X. Patel, P. Fonagy, R. J. Dolan, P. B. Jones, I. M. Goodyer, the NSPN Consortium, P. E. Vértes, E. T. Bullmore, Conservative and disruptive modes of adolescent change in brain functional connectivity. *bioRxiv*, 1–8 (2019).

88. S. Dehaene, C. Sergent, J.-P. Changeux, A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proc Natl Acad Sci U S A*. **100**, 8520–8525 (2003).

89. F. Váša, R. Romero-Garcia, M. G. Kitzbichler, J. Seidlitz, K. J. Whitaker, M. M. Vaghi, P. Kundu, A. X. Patel, P. Fonagy, R. J. Dolan, P. B. Jones, I. M. Goodyer, NSPN Consortium, P. E. Vértes, E. T. Bullmore, Conservative and disruptive modes of adolescent change in human brain functional connectivity. *Proceedings of the National Academy of Sciences*. **117**, 3248–3253 (2020).

Supplementary Materials for

Dynamical consequences of regional heterogeneity in the brain's transcriptional landscape

Gustavo Deco[†], Morten L. Kringelbach[†], Aurina Arnatkevičiūtė, Stuart Oldham, Kristina Sabaroedin, Nigel C. Rogasch, Kevin Aquino[#], and Alex Fornito[#*]

[†] equal author contributions
equal author contributions

*Corresponding author. Email: alex.fornito@monash.edu

Fig. S1.

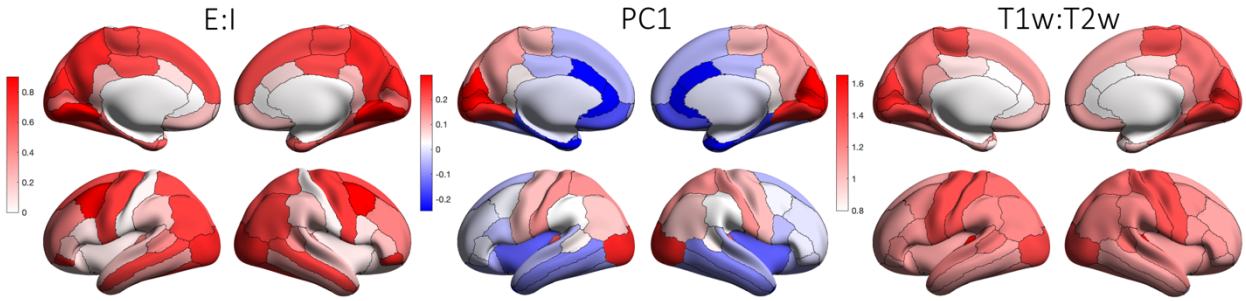

Figure S1. Different maps of regional heterogeneity. This figure shows the different maps used to incorporate regional heterogeneity in the E:I, PC1, and T1w:T2w models. E:I is the ratio of excitatory to inhibitory receptor gene expression. PC1 is regional loadings on the first principal component of brain-specific gene expression. T1w:T2w is the ratio of T1-weighted to T2-weighted MRI signal. See Methods for further details.

Fig. S2.

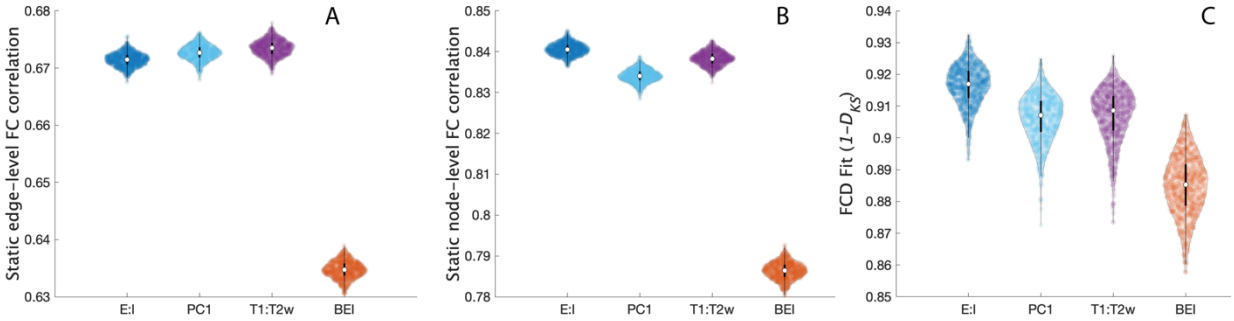

Figure S2. Out-of-sample model performance. Panels show out-of-sample model fits to (**A**) static edge-level functional connectivity (FC), (**B**) static node-level FC, and (**C**) functional connectivity dynamics (FCD). Fits were evaluated in an independent sample of dataset of 1003 individuals using the same working point as identified in our initial analysis. Heterogeneous models fit the data better than the homogeneous model in all cases. Model fits are comparable for the static edge-level FC but are superior for the E:I model when considering static node-level FC and FCD.

Fig. S3.

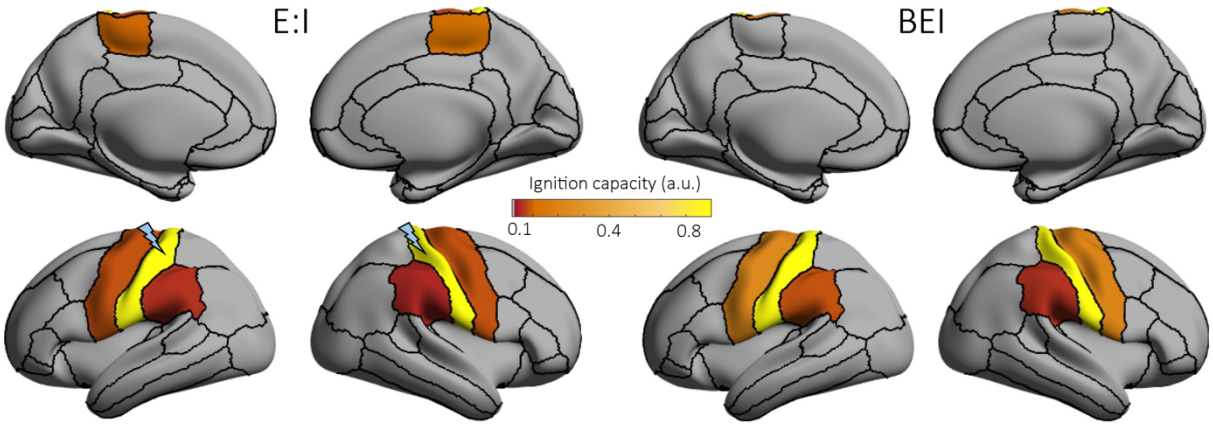

Figure S3. Regional ignition-like propagation following stimulation of somatosensory cortex. Ignition capacity for the E:I (left) and BEI (right) models, measured in arbitrary units. Blue marker in bottom left shows stimulation site.