# The Role of Inhibition in fMRI Resting-State Negative Correlations

Shreyas Harita[1, 2], Davide Momi[2, 3], Zheng Wang[2], Sorenza P. Bastiaens[1, 2], John D. Griffiths* [1, 2, 4, 5]

*Affiliations:*
1 = Institute of Medical Science, University of Toronto
2 = Krembil Centre for Neuroinformatics, Centre for Addiction and Mental Health (CAMH), Toronto
3 = Unit of Neuroimaging and Neurointervention, Department of Neurological and Neurosensorial Sciences, AOUS, 53100, Siena, Italy
4 = Department of Psychiatry, University of Toronto
5 = Institute of Biomedical Engineering
* = Corresponding Author

*Highlights:*

- The reduced Wong-Wang (RWW) neural mass model was used to understand and study the physiological origin of negative correlations in rs-fMRI.
- The simulated data can partially capture the dynamics of the empirical rs-fMRI data.
- Increasing the level of inhibition leads to an increase in the number of negative correlations.
- The levels of excitation and inhibition have differential effects on model stability and negative correlations.
- Insights from this work can contribute towards fine-tuning model parameters and in the future, aim to further validate findings using empirical data from clinical studies, and explore alternate parameters influencing negative correlations in rs-fMRI FC.

*Keywords:*

Neural mass model, Negative correlation, Resting-state, Functional Connectivity, Structural Connectivity.

**Abstract**

Resting-state brain activity, as observed via functional magnetic resonance imaging (fMRI), displays non-random fluctuations whose covariance structure (aka functional connectivity; FC) is commonly parsed into spatial patterns of positive and negative correlations (PCs and NCs). Mapping NC patterns for certain key seed regions has shown considerable promise in recent years as a tool for enhancing neuro-navigated targeting and clinical outcomes of repetitive transcranial magnetic stimulation (rTMS) therapies in psychiatry. These successes bring to the fore several major outstanding questions around the neurophysiological origins of fMRI NCs, the answers to which should prove useful for future therapeutic protocol development. In this work, we studied candidate mechanisms for the emergence of fMRI NCs using connectome-based computational modeling. Simulations of fMRI data under manipulation of inhibitory parameters $W_I$ and $\lambda$, representing local and network-mediated inhibition respectively, were explored, focusing on the impact of inhibition levels on the emergence of NCs. Despite the considerable difference in time scales between GABAergic neuronal inhibition (tens of ms) and fMRI FC (dozens of seconds), a clear relationship was observed, whereby the greater levels of overall inhibition led to significantly greater magnitude and spatial extent of NCs. We show that this effect is due to a leftward shift in the FC correlation distribution, leading to a reduction in the number of PCs and a concomitant increase in the number of NCs. Relatedly, we observed that those connections available for recruitment as NCs were precisely those with the weakest corresponding structural connectivity. Relative to nominally default values for the models used, greater levels of inhibition also improved, quantitatively and qualitatively, single-subject fits of simulated to empirical FC matrices. Our results provide new insights into how individual variability in anatomical connectivity strengths and neuronal inhibition levels may determine individualized expression of NCs in fMRI data. These, in turn, may offer new directions for optimization and personalization of rTMS therapies and other clinical applications of fMRI NC patterns.

## Introduction

### *Negative correlations in resting-state functional MRI data*

The human brain's spontaneous activity at rest, without any external input, shows non-random, highly structured patterns of organization, which can be measured using resting-state functional magnetic resonance imaging (rs-fMRI). Despite the name, during rs-fMRI recordings, the brain is not 'at rest', with neurons and neural populations continuing to exert local and long-range excitatory and inhibitory influences. The most commonly used approach for studying these intrinsic neural fluctuations is by measuring correlations amongst voxel- or region-wise rs-fMRI time series, known as 'functional connectivity' (FC). Initially, rs-fMRI FC was reported in the primary motor cortex (Biswal et al., 1995). Since then, FC patterns have been studied intensely by the neuroimaging community, leading to the progressive discovery that several so-called 'functional networks', spanning the entire brain, appear as prominent common variance components in rs-fMRI data. These include the Visual (VN), Somatomotor (SMN), Dorsal Attention (DAN), Ventral Attention (VAN), Limbic (LN), Frontoparietal (FPN), and Default Mode Network (DMN) (Yeo et al. 2011; Schaefer et al. 2018). These networks are consistent across groups of individuals and with different fMRI protocols (Rosazza & Minati, 2011).

In addition to this consistent modular network structure, a prominent feature of rs-fMRI FC patterns is the presence of anti- or negative correlations (NCs), which manifest at both the local (voxel/parcel) level and the regional/network level. It is uncertain what the exact driving source is for these NCs seen in rs-fMRI. Some studies have suggested that the popular preprocessing technique of global-signal regression (GSR) artificially introduces NCs, and therefore their appearance in fMRI data must be interpreted with caution (Fox et al., 2009; Murphy & Fox, 2017). However, other studies have shown strong NCs to be present in rs-fMRI time series both with and without GSR (using a principal component-based noise reduction technique) (Chai et al., 2012). Thus, it is likely that at least some of the NCs seen in rs-fMRI FC are driven by real properties of neural dynamics.

A body of evidence suggests that a potential source of NCs in rs-fMRI data could be inhibition, which in turn leads to reduced neuronal activity. For example, in a study looking at responses to electrode stimulation from the visual cortex (V1) of anesthetized monkeys, NCs in the fMRI recordings were linked to decreases in neuronal activity when compared to spontaneous activity beyond the stimulated V1 regions (Shmuel et al., 2006). The authors of this study demonstrated that it is possible to predict the time course of one response based on the other, suggesting a link between them. While this coupling of reduced neuronal activity and NCs does not necessarily imply causation, it does suggest that NCs can potentially be driven by reduced neuronal activity (mediated via inhibition), which in turn leads to reduced local cerebral blood flow. Another comprehensive study employing neuroimaging modalities and electrophysiological recordings highlights this link between the NCs and neuronal inhibition, this time in the somatosensory cortex of anesthetized rats (Devor et al., 2007). In this work, it is concluded that neuronal inhibition and concurrent arteriolar vasoconstriction are associated with a decrease in blood oxygenation. This, in turn, could potentially underscore NCs observed in fMRI BOLD data.

### *Clinical significance of negative correlations*

While the exact physiological origins of NCs remain unknown, their clinical importance has been well established, in particular as a tool for improving neuro-navigated targeting in repetitive transcranial magnetic stimulation (rTMS) treatment for major depressive disorder (MDD) (Cash et al., 2020; Cole et al., 2020; Fox et al., 2012). Since its introduction in the early 1990s, rTMS protocols for MDD have primarily targeted the dorsolateral prefrontal cortex (dlPFC) with varying degrees of clinical improvements reported (Asgharian Asl & Vaghef, 2022; Brunoni et al., 2017; Leuchter et al., 2015; Levkovitz et al., 2015). In a seminal clinical research study, it was reported that the precise rTMS target location within dlPFC is most effective when it is maximally negatively correlated with the subgenual cingulate cortex (SGC). This was done initially in healthy controls and then tested for reproducibility in an independent MDD patient cohort with MDD (Fox et al., 2012). This retrospective finding has subsequently been replicated in prospective clinical trials, across three distinct geographical cohorts, with results staying consistent across different scanner types, preprocessing methodologies, TMS devices, and patient populations (Weigand et al., 2018; Cash et al., 2020; Cole et al., 2020).

Despite the clear importance of rs-fMRI NCs for both fundamental and clinical rs-fMRI, there has been relatively little research to date attempting to understand these phenomena in terms of their underlying neurophysiological bases. Related work has studied NCs at a more fine-grained, intra-regional (visual cortex) level, finding that stronger inhibition and weaker excitation can lead to a higher magnitude of NCs in neuronal firing rates (Chelaru & Dragoi, 2016; Goelman et al., 2014). Whether these micro-scale NCs bear any direct relationship to rs-fMRI FC NCs, which are significantly slower temporally and larger spatially, is however far from clear.

### *Present Study*

As noted above, previous research has implicated neuronal inhibition generating NCs seen in neuroimaging data (Cole et al., 2020; Goelman et al., 2014). We investigated this question using whole-brain connectome-based neural mass modelling of rs-fMRI FC data in the present study. Neural mass models consider the mean activity of an entire neuronal population (representing a brain area or parcel), instead of tens of thousands of individual neurons in that patch of tissue. Connectome-based brain network models combine the activity of several neural masses based on structural (anatomical) connectivity, typically derived from diffusion MRI tractography (Breakspear, 2017). This methodology has been used to study resting-state brain activity (Deco et al., 2011, 2013; Honey et al., 2009), brain oscillations (Breakspear et al., 2010), alpha waves (Griffiths et al., 2020; Jansen & Rit, 1995), and traveling waves (Breakspear, 2017; Heitmann et al., 2012), to name a few. Here we developed a connectome-based brain network model employing the reduced Wong-Wang (RWW) neural mass model, as implemented in The Virtual Brain (TVB) software library (Ritter et al. 2014; Sanz-Leon et al. 2015), to study the role of inhibition in generation of rs-fMRI FC NCs. The RWW has been extensively used in prior work to study resting and task-evoked brain dynamics, with the emphasis of modelling studies using RWW typically being rs-fMRI activity (this being what it was designed for; Deco et al. 2013, 2014). By studying rs-fMRI NCs as a function of inhibition levels in this neural mass model, and focusing on patterns that are consistent across the simulated and empirical rs-fMRI time series, our aim is to identify the key mechanistic drivers and organizational principles of this important feature of human brain activity.

In line with prior research, we hypothesize that elevated inhibition levels would lead to an augmented occurrence of NCs in simulated rs-fMRI FC data. Additionally, we explored candidate explanations for the observed spatial patterns of these NCs. Despite the fundamental and clinical significance of rs-fMRI NCs, limited research has delved into comprehending their neurophysiological foundations. A thorough understanding of these phenomena not only holds the potential to enhance clinical applications but also provides insights into the broader dynamics of brain function.

## Methods

*Overview*

Resting-state FC is typically used to characterize the spatiotemporal patterns of neural dynamics in the human brain. We use theoretical models to help understand the link between underlying structural connectivity (SC) and FC. The first step in building these models is to determine the density of white matter tracts connecting various brain regions (each represented by a neural mass), depicted by the SC matrix to construct the connectome. This is typically done using diffusion-weighted MRI (dw-MRI) tractography. In the majority of previous work, SC matrices used have had relatively low anatomical resolution of 50-100 brain regions (parcels) (Deco et al., 2013, 2014; Hagmann et al., 2008; Saenger et al., 2018; Zheng et al., 2020). In this study, we use a higher resolution 200 x 200 SC matrix, based on the Schaefer parcellation atlases (Schaefer et al., 2018) with dw-MRI data from the Human Connectome Project (HCP) (Van Essen et al., 2013). Figure 1 shows a schematic of our approach. The model equations and neuroimaging data analysis processing details are described below.

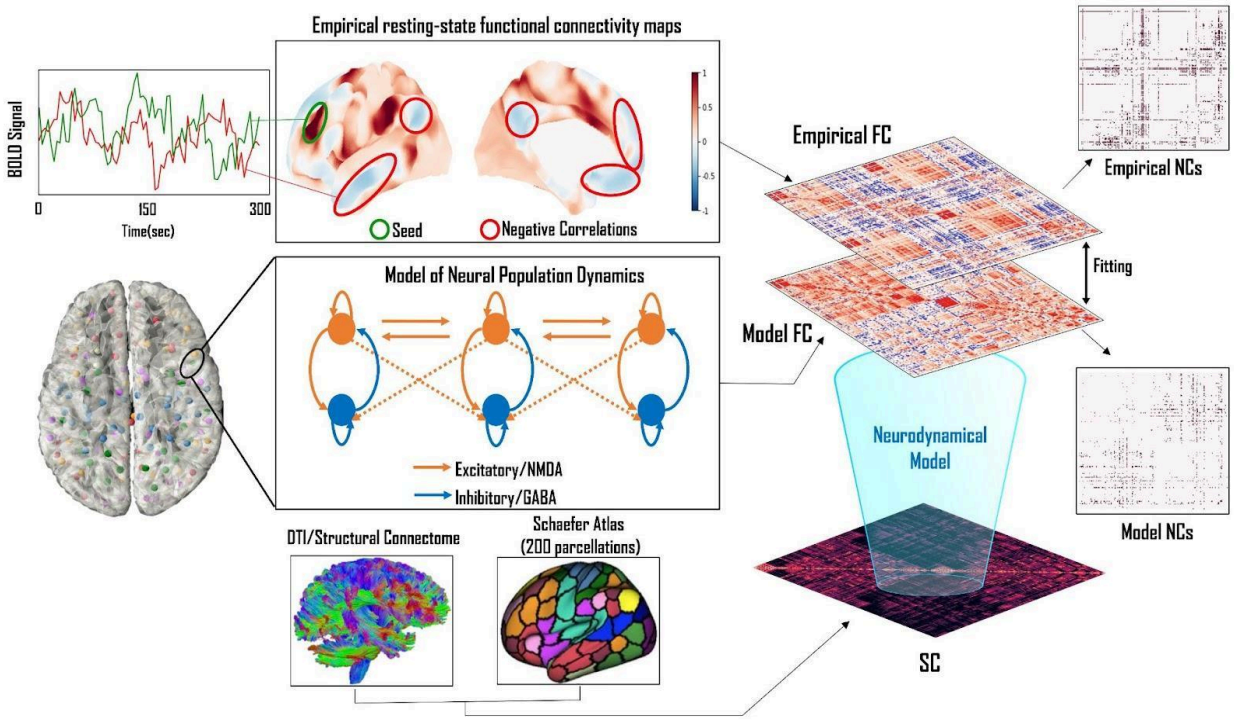

**Figure 1:** ***Schematic of the methodological approach***. We extend the whole-brain modelling methodology from previous studies to focus on the role of inhibition in negative correlations in resting-state fMRI data. The reduced Wong-Wang (RWW) model, governed by a set of stochastic differential equations coupled according to the SC matrix, was used to simulate resting-state BOLD fMRI time series data. To determine the validity of the RWW model, we compared the model-derived functional connectivity to the ones observed in empirical resting-state fMRI BOLD data. We then manipulated specific inhibitory parameters in the two models to identify potential sources of negative correlations. FC = functional connectivity; SC = structural connectivity; NCs = negative correlations.

*Neuroimaging data analyses*

DW-MRI data were extracted from 24 randomly selected subjects from the HCP WU-Minn cohort database (Van Essen et al., 2013). Data preprocessing pipeline was run on Ubuntu 18.04 LTS and included tools in FMRIB Software Library (FSL 5.0.3; www.fmrib.ox.ac.uk/fsl) (Jenkinson et al., 2012), MRtrix3 (www.MRtrix.readthedocs.io) (Tournier et al., 2012) and FreeSurfer (Fischl et al., 2004). All DW-MRI data downloaded were already corrected for motion via FSL's EDDY (Andersson and Sotiropoulos, 2016) as part of the HCP minimal preprocessing pipeline (Glasser et al., 2016). The multi-shell multi-tissue response function (Christiaens et al., 2015) was estimated using a constrained spherical deconvolution algorithm (Jeurissen et al., 2014). Simultaneously, the T1w images, which were already coregistered to the b0 volume, were segmented using the FAST algorithm (Zhang et al., 2001). Following this, anatomically constrained tractography was employed to generate the initial tractogram with 10 million streamlines (Amico et al., 2017) using second-order integration over fiber orientation distributions (Tournier et al., 2010). Then, the Spherical-deconvolution Informed Filtering of Tractograms (SIFT2) methodology (Smith et al., 2015) was applied to provide more biologically accurate measures of fiber connectivity.

Following tractography reconstructions, an anatomical parcellation of the grey/white matter interface was used to segment the whole-brain streamline sets into pairwise bundles, resulting in an anatomical connectivity matrix for each subject. The parcellation used was the 200-node variant of the atlas developed by (Schaefer et al., 2018), which incorporates the assignment of each parcel into one of the 7 canonical functional networks identified by Yeo et al. (2011) - VN, SMN, DAN, VAN, LN, FPN, and DMN. Parcels were defined at the level of surface vertices, mapped to each individual's FreeSurfer reconstruction using spherical registration (Fischl et al., 1999), and converted to native-space image volumes in register with the dwMRI tractography streamlines. Entries in the resultant 200x200 SC matrices extracted for each subject represent the number of white matter streamlines connecting each pair of ROIs.

*Mathematical modeling of neural population dynamics*

We modelled activity at each node in the 200x200 network defined by the above tractography analysis using the framework developed by Deco and colleagues (Deco et al., 2013, 2014), which employs a mean-field approximation of a leaky-integrate-and-fire spiking network based on the adiabatic assumption that the system is dominated by slow dynamics. As this formulation was first proposed by Wang and colleagues (Brunel and Wang, 2001; Wong and Wang, 2006),

we follow others (Sanz-Leon et al., 2015) in referring to the resulting equations as the RWW model. The RWW model was first applied to whole-brain fMRI connectivity dynamics by Deco et al. (2013), which focused solely on networks of slow excitatory (NMDA) neurons, explicitly omitting fast excitatory (AMPA) and inhibitory (GABA-A) neurons, due to their having significantly shorter decay time constants than NMDA receptors (10 ms vs. 100 ms, respectively). This was then extended in Deco et al. (2014) to re-introduce the GABA-A inhibitory neural population, yielding the following two-dimensional system of coupled nonlinear differential equations, representing input currents and population firing rates

$$I_i^{(E)} = W_E I_0 + w_+ J_{NMDA} S_i^{(E)} + G J_{NMDA} \sum_j C_{ij} S_j^{(E)} - J_i S_i^{(I)} + I_{external} \quad \textbf{(1)}$$

$$r_i^{(E)} = H^{(E)}(I_i^{(E)}) = \frac{a_E I_i^{(E)} - b_E}{1 - exp(-d_E(a_E I_i^{(E)} - b_E))} \quad \textbf{(2)}$$

$$\frac{dS_i^{(E)}(t)}{dt} = -\frac{S_i^{(E)}}{\tau_E} + (1 - S_i^{(E)})\gamma r_i^{(E)} + \sigma v_i(t) \quad \textbf{(3)}$$

$$I_i^{(I)} = W_I I_0 + J_{NMDA} S_i^{(E)} - S_i^{(I)} + \lambda G J_{NMDA} \sum_j C_{ij} S_j^{(E)} \quad \textbf{(4)}$$

$$r_i^{(I)} = H^{(I)}(I_i^{(I)}) = \frac{a_I I_i^{(I)} - b_I}{1 - exp(-d_I(a_I I_i^{(I)} - b_I))} \quad \textbf{(5)}$$

$$\frac{dS_i^{(I)}(t)}{dt} = -\frac{S_i^{(I)}}{\tau_I} + r_i^{(I)} + \sigma v_i(t) \quad \textbf{(6)}$$

Here, the population firing rate of the excitatory (*E*) and inhibitory (*I*) neuronal populations in a given brain region *i* is given by $r_i^{(E,\ I)}$. $S_i^{(E,\ I)}$ represents the level of synaptic gating for the *E* and *I* populations, respectively. Synaptic gating refers to the regulatory control or modulation of information flow between neurons at synapses. It involves mechanisms that influence whether a synaptic signal is effectively transmitted or inhibited. These gating processes play a crucial role in shaping neural communication and contribute to the overall dynamics of neuronal networks. $I_i^{(E,\ I)}$ are the excitatory/inhibitory input currents to the population *i*. $I_{external}$ is the external stimulation into the ensemble, which is always zero in the present study because we consider only the resting state. A key concept within this model and the present study is that of feedforward inhibition, which refers to the driving of inhibitory population activity by long-range excitatory connections from other nodes in the network. This is represented above by the final terms in equation 4, and its contribution is scaled parametrically by $\lambda \in [0,1]$.

The global scaling factor (G) modulates the strength of inter-areal connections in the neural population model, which are determined by the neuroanatomical matrix (SC), denoted by $C_{ij}$. The SC matrix specifies the density of fibers between brain regions. *G* offers control over the overall connectivity strength within the global cortical system. In our study, a systematic G

sweep explored various values, revealing how $G$ influences the dynamics of the model. This exploration sheds light on $G$'s impact on simulated brain dynamics, including synchronization, information flow, and network connectivity. The $G$ sweep is crucial for navigating the parameter space and identifying the optimal $G$ value that aligns the simulated neural population dynamics with real rs-fMRI data. This iterative parameter optimization enhances the model's fidelity in replicating empirical observations, providing a more accurate representation of underlying neural processes.

The input-output functions of the excitatory and inhibitory populations are $H^{(E)}$ and $H^{(I)}$, respectively. These are mathematical representations of how the activity of neuronal populations responds to incoming inputs. They account for the collective synaptic input received by neurons, incorporating both excitatory *(Eq. 2)* and inhibitory signals (*Eq. 5*). The excitatory ($a_E$, $b_E$, and $d_E$) and inhibitory ($a_I$, $b_I$, and $d_I$) gating variables influence the input currents to the excitatory ($I_i^{(E)}$) and inhibitory ($I_i^{(I)}$) populations, respectively. This integrated input is then transformed into a firing rate, which represents the rate at which neurons within the population generate action potentials. The exponential form for the firing rate equation allows for an accurate representation of the non-linear response profiles of neural tissue, with greater input currents leading to higher firing rates past a certain inflection point. The kinetic parameter is $\gamma = 0.641/1000$. $I_0 = 0.382$ (nA) is the overall effective external input, which is scaled by $W_E$ and $W_I$ (the scaling input weights) for the excitatory and inhibitory populations, respectively. The noise amplitude is $\sigma = 0.01$ (nA), with $v_i$ being standard Gaussian noise. The local excitatory recurrence is $w_+ = 1.4$. The excitatory synaptic coupling $J_{NMDA}$ and the local feedback inhibitory synaptic coupling $J_i$ were determined manually using Equations 1-6, to achieve a firing rate of 5 Hz. The corresponding $J_{NMDA}$ and $J_i$ values were 0.18 and 1.03, respectively. We note that this is slightly above the reported 3.55 Hz but we did not see much difference in the final results (Deco et al., 2014).

*Haemodynamic observation model*

The Balloon-Windkessel (BW) model translates simulated neural activity into a simulated BOLD time series, elucidating the hemodynamic processes underlying the neurovascular coupling. The model begins with the neural activity time series as its input, representing the dynamic behavior of neural populations. The hemodynamic response function is a central component, capturing the intricate relationship between changes in neural activity and subsequent alterations in blood flow, volume, and oxygenation within the brain. The "Balloon" component mirrors the compliance of blood vessels, considering the expansion and contraction of vessels in response to cerebral blood flow changes, with parameters such as time constants defining inflow and outflow (Buxton et al., 1998; Buxton & Frank, 1997). Simultaneously, the "Windkessel" component represents the damping effect of the vasculature on blood flow, incorporating parameters like arterial transit time (Mandeville et al., 1999). Through a set of differential equations, the model computes the BOLD signal, reflecting fluctuations in blood oxygenation levels across different brain regions.

The signal that induces blood flow is generated by neuronal activity $u(t)$, where $u(t)$ is $S_i^{(E)}/S_i^{(I)}$ for excitatory and inhibitory neuronal activity, respectively.

$$\dot{s} = \varepsilon u(t) - s/\tau_s - (f_{in} - 1)/\tau_f \quad \textbf{(7)}$$

where ε, $\tau_s$, and $\tau_f$ have typical values associated with them.

The inflow of blood $f_{in}$ is also related to the signal generated by $u(t)$:

$$f'_{in} = s \quad \textbf{(8)}$$

With the inflow of blood, there is the extraction of oxygen from the hemoglobin and the outflow of blood $f_{out}$ with the change in deoxyhemoglobin ($q/q'$).

$$\tau_0 q' = f_{in} \frac{E(f_{in}, E_0)}{E_0} - f_{out}(v) q/v \quad \textbf{(9)}$$

where $E(f_{in}, E_0)$ is the fraction of the oxygen extracted from the inflowing blood, $\tau_0$ is the time constant, and $v$ is the volume of venous blood. The rate of change of $v$ is given by

$$\tau_0 v' = f_{in} - f_{out}(v) \quad \textbf{(10)}$$

Finally, the resultant BOLD signal $y(t)$ is a static nonlinear function of normalized $v$, normalized total $q$, and the resting net oxygen extraction fraction ($E_0$), given by the equation

$$y(t) = V_0(k_1(1 - q) + k_2(1 - q/v) + k_3(1 - v) \quad \textbf{(11)}$$

where $k_1 = 7E_0$, $k_2 = 2$, $k_3 = 2E_0 - 0.2$, and $V_0$ is the resting-blood volume fraction.

To generate a simulated BOLD time series for comparison with empirical data, TVB incorporates an alternative to the differential equations of the BW model described above that is mathematically and computationally simpler without loss of precision. Instead of directly implementing the BW model, TVB utilizes a Volterra Kernel that approximates the behavior of the BW model (Friston et al., 2000). The Volterra Kernel serves as an efficient and computationally tractable method to represent the hemodynamic processes involved in converting neural activity to BOLD signals within the context of the RWW model. This methodological choice not only streamlines the simulation process but also maintains the fidelity of the simulated BOLD time series, facilitating a comprehensive comparison with empirical rs-fMRI data.

*Exploring the effects of varying the inhibition parameters*

Numerical simulation of the RWW equations was carried out using the Virtual Brain (TVB) Python library (Sanz Leon et al., 2013), which is designed to simulate and model large-scale brain networks. Equations 1-6 were used to simulate resting-state BOLD time series within the TVB environment (Sanz-Leon et al., 2015). The focus of this study was to observe the effects of

inhibition on NCs in our simulated BOLD time series data. To this effect, with the exception of the control analysis, we exclusively manipulated the inhibitory parameters in the RWW model. The first was the inhibitory input scaling weight - $W_I$, which was increased from the default value of 0.7 to a maximum of 1.0 in increments of 0.05. We also increased the feed-forward inhibition parameter - $\lambda$ from the default value of 0.0 to a maximum of 1.0 in increments of 0.2. All other parameters in the model were left at their respective default values (see above). The inhibitory parameters $W_I$ and $\lambda$ were increased separately to study how increasing each parameter affects the model dynamics and its impact on NCs. In the control analysis for the RWW model, we manipulated $W_E$, which is the excitatory analog of $W_I$. To study how local dynamics affect NCs, we altered the excitatory and inhibitory couplings - $J_{NMDA}$ and $J_i$, respectively. Both the couplings were 'cut' i.e., set to a value of zero while we increased either $W_E$ or $W_I$. This was done to obtain a more comprehensive insight into how these excitatory and inhibitory parameters exert their effects.

*Code and Data Availability*

All analyses reported in this paper were conducted on CentOS Linux compute servers running Python 3.7.3, using the standard scientific computing stack and several open-source neuroimaging software tools - principally TVB (simulated BOLD simulations; (Sanz Leon et al., 2013), Nibabel (neuroimaging data I/O; (Brett et al., 2020) and Nilearn (neuroimaging data visualizations; (Abraham et al., 2014). All code and analysis results are openly available at github.com/GriffithsLab/HaritaEtAl2024_tvb-negative-correlations.

## Results

Using the above-outlined whole-brain computational modelling framework, we explored the role of inhibition, local dynamics, and SC in the generation of negative correlations in the BOLD rs-fMRI time series. In the following, we describe our results using the RWW model. Equivalent simulations were done using the Jansen-Rit (JR) model. This additional analysis was done to validate our results from the RWW model. Our results from the JR model uniformly returned the same conclusions as RWW, which are detailed in the Supplementary Information section at the end of this study.

### ***Model simulations fit data well and reproduce the primary characteristics of empirical FC matrices.***

For each of the 24 subjects, RWW-simulated BOLD FC was generated for a range of values of *G* factor (*Eq. 1 and 4*), and the fit to empirical data (Pearson correlation between simulated and empirical FC matrix upper triangles) was computed. *G* values yielding maximal simulated-empirical FC fits in this brute-force 1-dimensional parameter sweep for each subject were selected for subsequent model runs. This process returned an average fit of R = 0.24 (Std. Dev = 0.06), ranging from R = 0.10 to R = 0.41 (Figure 2A).

The simulated FC from the RWW model can partially capture features of the empirical FC matrix. The model can replicate the FC within some functional networks better than others. For example, the model is better at replicating the FC within the VN when compared to the DMN within the left and right hemispheres. Additionally, the intra-hemispheric simulated FC shows a higher correlation to the intra-hemispheric empirical FC than does the cross-hemispheric simulated FC. Within the left hemisphere, the average fit of the simulated to the empirical data was R = 0.32 (Std. Dev = 0.07), ranging from R = 0.20 to R = 0.48. Within the right hemisphere, the corresponding values were R = 0.30 (Std. Dev = 0.09), ranging from R = 0.10 to R = 0.46. However, the cross-hemispheric fit of the simulated to the empirical was significantly lower than the intra-hemispheric fit. Between the left and right hemispheres, the average fit of the simulated to the empirical data was R = 0.1 (Std. Dev = 0.1), ranging from R = 0.10 to R = 0.38. Hence, the model exhibits a partial replication of the empirical data. When looking at the actual number of NCs, we note that the model does not inherently have many NCs to begin with. The number of NCs increases when the level of inhibition is increased as noted below.

***Increasing inhibition increases negative correlations in simulated data***

The purpose of selectively manipulating inhibitory parameters in the RWW model was to understand the role inhibition plays in NCs. In the RWW model, we systematically increased the values of $W_I$ and $\lambda$. Both parameters, when increased, lead to a significant ($p < 0.01$) increase in the overall number of NCs. When $W_I$ was increased from 0.7 (default value) to 1.0, the mean number of NC increased from 50.08 ± 92.77 to 13111.25 ± 5313.50. Similarly, when $\lambda$ was increased from 0 (default) to 1, the mean NC increased from 50.08 ± 92.77 to 2161.00 ± 1859.02 (Figure 2B). The overall fit to empirical data, at the maximum inhibition values, was reduced to an average of R = 0.18 ± 0.05 (at a $W_I$ value of 1.0; Figure 4 B [top]), and remained relatively unchanged at an average of R = 0.24 ± 0.04 at a $\lambda$ value of 1, across all subjects.

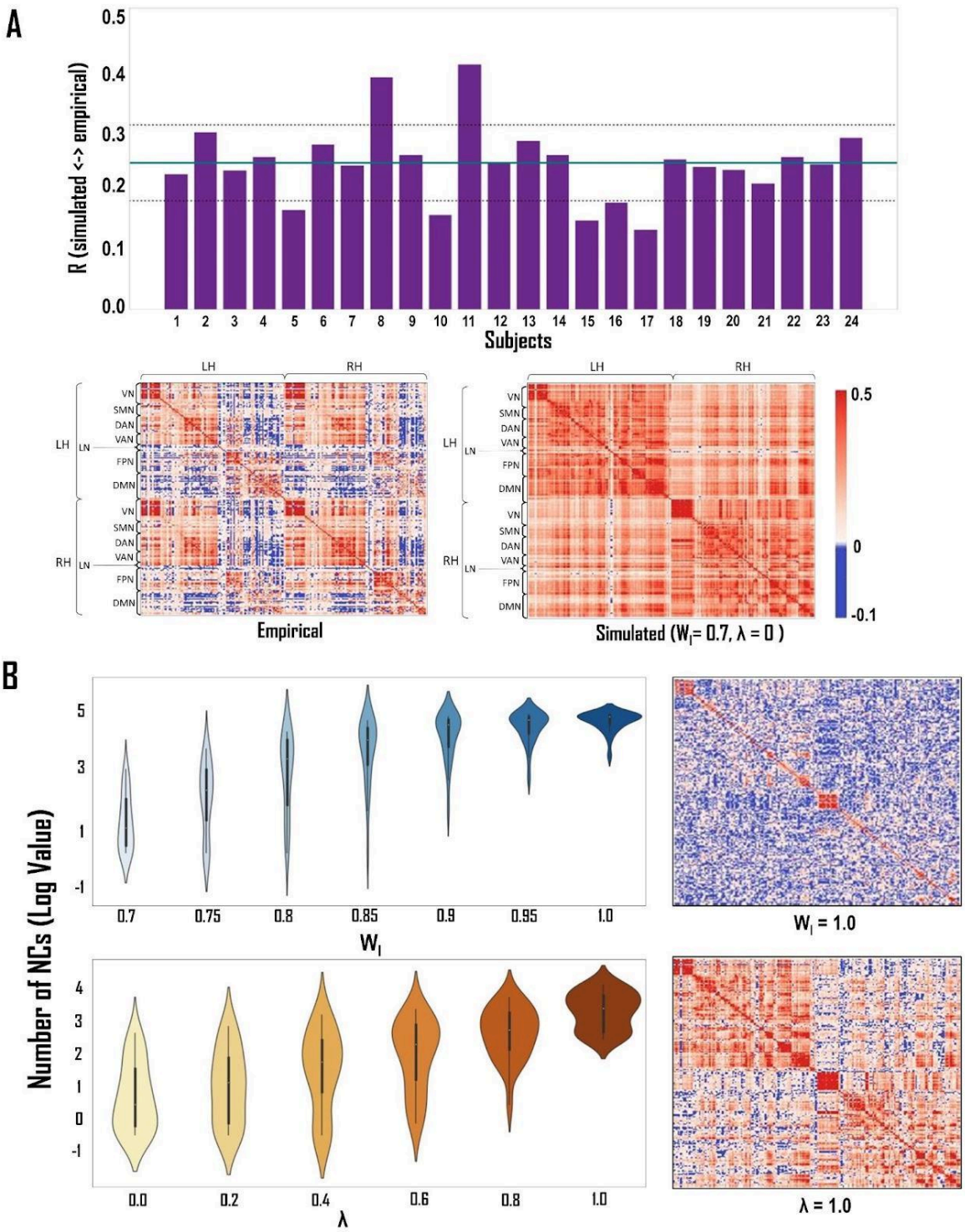


**Figure 2:** ***Model Fit to Empirical Data and Inhibition-Induced Changes in Negative Correlations.*** **A** - *Goodness of fit of the RWW model.* Top: Across our subject cohort, the RWW model showed an average fit of 0.24 ± 0.06 to the empirical data (max. = 0.41). Bottom: The heatmap plots for the empirical and simulated data of a randomly selected subject. The various functional networks are outlined on the left of each FC matrix. **B -** *The effect of increasing the level of inhibition on the number of negative correlations.* Violin plots show the effect of increasing the level of inhibition on negative correlations. - *Top:* $\mathbf{W_I}$; *Bottom:* $\boldsymbol{\lambda}$. The y-axis in these plots is represented in the

log scale. The heatmap plots on the right showcase another way of visualizing this observation. The increase in the number of NCs is seen when compared to the simulated heatmap in part A.

The increase in NCs was accompanied by a reduction in positive correlations (PCs) in our simulated data when the inhibition parameters were increased. To further investigate this, we studied kernel density estimate (KDE) plots for the overall simulated functional connectivity distribution when inhibition is increased. We observed that increasing inhibition leads to a reduction in the skew of the right tail of the distribution. Additionally, there was a narrowing of the distribution and a corresponding shift to the left, at higher inhibition values (Figure 3).

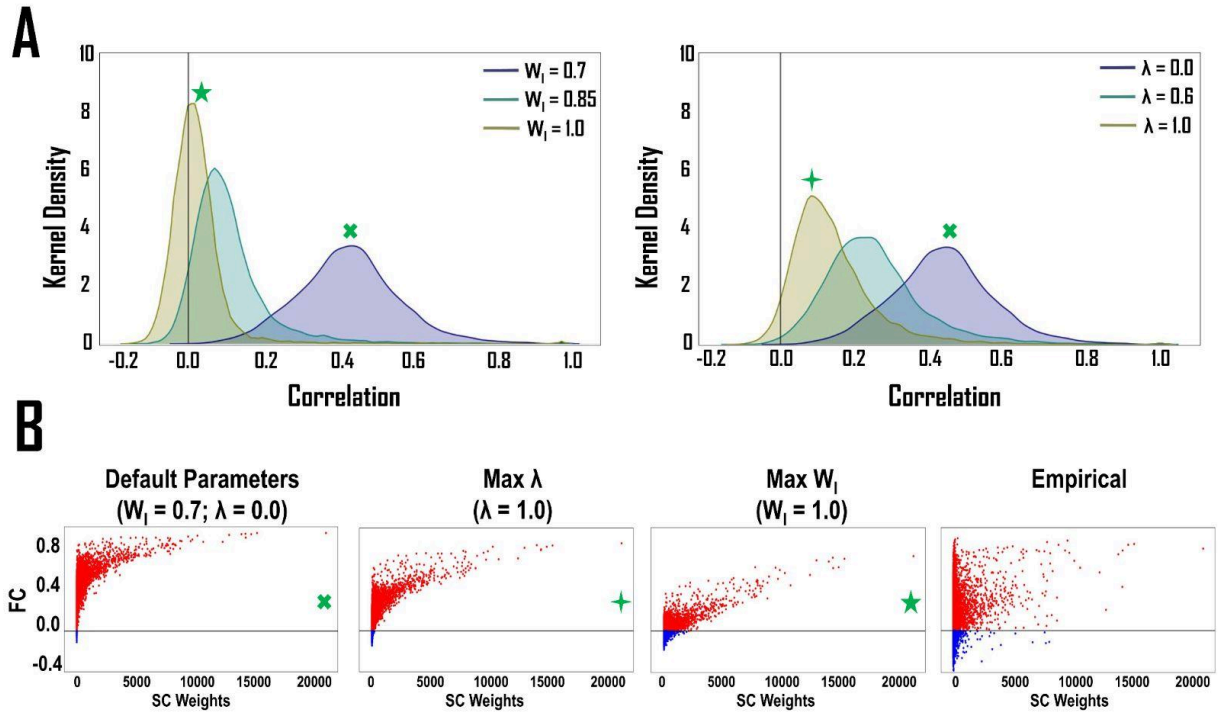


**Figure 3:** ***An increase in NCs leads to a decrease in PCs when the level of inhibition is increased.*** **A -** Kernel density estimation (KDE) plots show that increasing inhibition levels lead to an increase in negative correlations by reducing the number of positive correlations. This is seen as a reduction in the skew of the right tail of the KDE plot. Furthermore, the KDE plots narrow at higher inhibition parameter values and shift to the left. *Left*: Increasing $\mathbf{W_I}$; *Right*: Increasing λ. Both plots in this figure represent a single, randomly selected subject. **B -** Scatter plots showcase the relationship between the FC correlations and SC weights. Negative correlations

### *The levels of excitation and inhibition have differential effects on model stability and negative correlations*

The results above demonstrate the effects of increasing inhibitory parameters on NCs. To further understand how inhibition affects simulated data, we conducted control analyses with the RWW model. Control analyses involve modifying an alternative parameter to the inhibitory parameter of interest, to see if variation in NC levels is exclusively determined by inhibitory neuronal processes in the model. We selected $W_I$'s excitatory analog, the external input scaling weight - $W_E$.

When $W_E$ was increased from 1.0 (default) to 2.0, there was an increase in the number of NCs from 58.67 ± 113.24 to 18359.50 ± 556.62, respectively (Figure 4 A [top]). The model's overall fit was significantly ($p < 0.01$) reduced from 0.24 ± 0.06 when $W_E = 1$ (default value) to 0.08 ± 0.03 at $W_E = 2.0$. Note that when both $W_E$ and $W_I$ were increased, there was an increase in the number of NCs. To further understand how $W_E$ and $W_I$ exert their effects to increase NCs, the excitatory and inhibitory synaptic couplings were manipulated by removing the respective connections (i.e., setting the corresponding synaptic coupling - $J_{NMDA}$ (excitatory) or $J_i$ (inhibitory) to 0). In addition to this manipulation, $W_E$ and $W_I$ were also increased to see how NCs were affected. This led to four different conditions - 1) Increase $W_E$, $J_{NMDA} = 0$; 2) Increase $W_I$, $J_{NMDA} = 0$; 3) Increase $W_E$, $J_i = 0$; 4) Increase $W_I$, $J_i = 0$. All other parameters in these conditions were set to their respective default values.

Our results from this analysis can be broadly classified into two parts. One where $J_{NMDA}$ was set to 0 and the other where $J_i$ was set to 0. When $J_{NMDA}$ was removed, we noticed that the model fit to the empirical data was 0.00. Furthermore, the negative correlations in these simulations were random, not structured, and accounted for approximately half the values in the correlation matrix. This observation was the same when either $W_E$ or $W_I$ was increased. For $W_E$, the mean number of NCs ranged from 19851.50 ± 166.82 at $W_E = 1.0$, to 19866.00 ± 146.48 at $W_E = 2.0$. For $W_I$, the mean number of NCs ranged from 19928.50 ± 195.74 at $W_I = 0.7$, to 19929.83 ± 138.33 at $W_I = 1.0$. The values reported are averaged values across our 24-subject cohort.

When $J_i$ was removed, we observed an increase in the number of negative correlations when $W_E$ was increased. When the $W_E$ value was increased from 1.0 (default) to 2.0, we observed an average increase in the number of NCs from 9865.50 ± 3433.89 to 18952.83 ± 421.82, across all subjects, respectively. The trade-off with this increase in the number of NCs was a significantly ($p < 0.01$) lower model fit to the empirical data. The model fit to the empirical data decreased from 0.22 ± 0.03 at $W_E = 1.0$ (default), to 0.06 ± 0.02 at $W_E = 2.0$, averaged across all subjects (Figure 4 A [bottom]). When the $W_I$ value was increased from 0.7 (default) to 1.0, the number of NCs increased slightly from 9966.33 ± 3292.50 to 10289.83 ± 3410.60. However, this increase was not significant. The model fit to the empirical data was stable and did not change from 0.22 ± 0.03 when the $W_I$ value was increased (Figure 4 B).

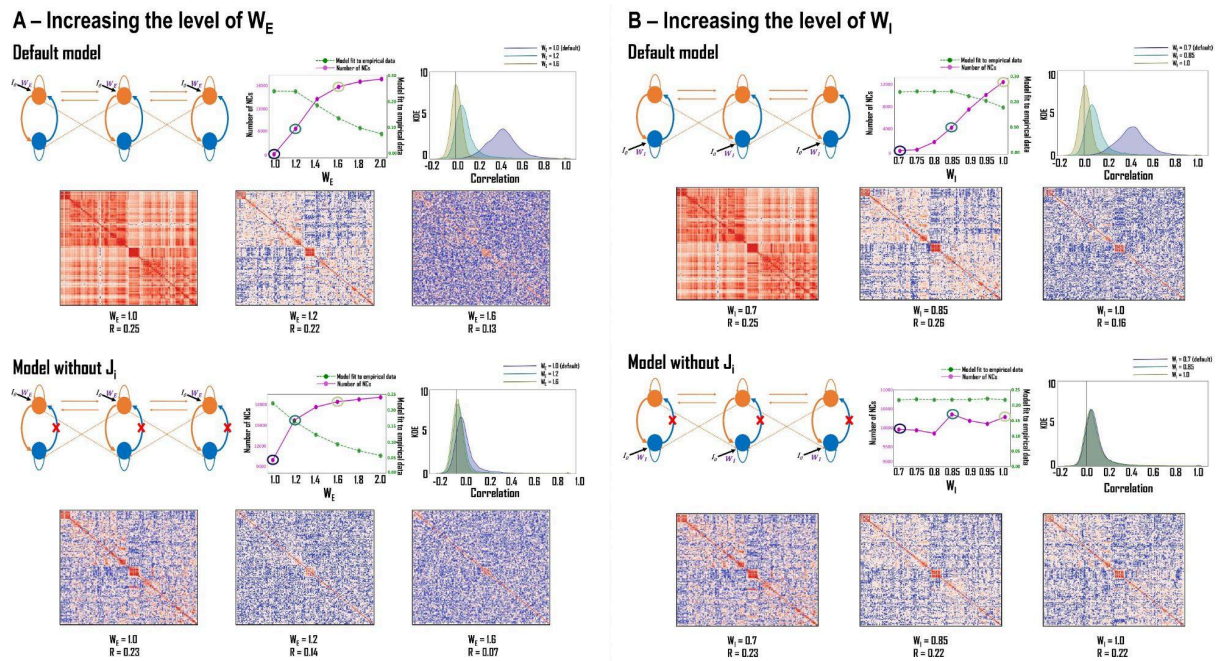


**Figure 4:** ***The levels of excitation and inhibition have differential effects on model stability and negative correlations.*** **A** - *The effect of increasing the $W_E$ value on model stability and negative correlations.* ***Top***: When the $W_E$ value is increased in the default model, there is an increase in the number of NCs. However, the model stability (**R**) is drastically reduced leading to random correlations in the resulting simulated FC matrix. ***Bottom***: When $J_i$, is set to zero and the $W_E$ value is increased, the model starts by being stable and having an increased number of NCs (at the default $W_E$ value), however, this stability deteriorates at higher $W_E$ values with random correlations in the simulated FC matrix. **B** - *The effect of increasing the $W_I$ value on model stability and negative correlations.* ***Top***: When the $W_I$ value is increased in the default model, initially, there are very few NCs, but there is a significant increase in NCs at higher $W_I$ values but the model stability is reduced, but not as much as when $W_E$ values are increased. ***Bottom***: In the model without $J_i$, there are a higher number of NCs to begin with (at the default $W_I$ value), however, increasing the $W_I$ value does little to affect the overall number of NCs with only a modest increase at higher $W_I$ values.

| Parameter | Value | Default Model | Model without $J_i$ |
|---|---|---|---|
| **$W_E$** (Scales $I_0$ for excitatory neurons) | Default (1.0) | FC ✓ NCs ⇔ | FC ✓ NCs ⇧ |
| | Max (2.0) | FC ✗ NCs ⇧⇧ | FC ✗ NCs ⇧⇧ |
| **$W_I$** (Scales $I_0$ for inhibitory neurons) | Default (0.7) | FC ✓ NCs ⇔ | FC ✓ NCs ⇧ |
| | Max (1.0) | FC ✓ NCs ⇧ | FC ✓ NCs ⇧ |

Table 1: ***The effect of increasing the $W_E$ and $W_I$ value on model stability and negative correlations***. This table complements the information provided in ***Figure 4***. Model stability is represented by FC ✓ (structured correlations) or FC ✗ (random correlations). The changes in the number of NCs are compared to the default model, represented by ⇔. An increase in NCs is depicted by ⇧. An excessive increase in NCs (random correlations) is represented with ⇧⇧.

## Discussion

In this study, we used the RWW brain network model to study the potential mechanistic origins of NCs in rs-fMRI data. Our results point to two cortical microcircuit parameters that influence the number of NCs in simulated fMRI data when varied parametrically. More generally, they provide a proof-of-principle that rs-fMRI NCs can be explored within a whole-brain modelling framework, despite their being an infraslow time series property that is several (temporal) orders of magnitude removed from the underlying neural dynamics. Interesting as these observations are, there are also several important caveats and clarifications in their interpretation, which we discuss below.

### *The level of inhibition and its impact on negative correlations*

By increasing the inhibitory parameter in our model and studying its effects on the model-simulated BOLD data, we aimed to understand the specific role inhibition plays in determining the number of NCs. Interesting observations emerged as we manipulated the levels of inhibition. Firstly, greater levels of inhibition led to a larger overall number of NCs, but this was accompanied by lower overall model goodness-of-fit. The average number of NCs was seen to increase at higher levels of either of the inhibitory parameters - $W_I$ (external input scaling weight to the inhibitory population) and $\lambda$ (feed-forward inhibition parameter). Moreover,

varying the levels of specific inhibitory parameters has different effects on the model fit to empirical data. In the case of $W_I$, there is a reduction in the model fit to the empirical data across all subjects, but there is a greater increase in the average number of NCs. On the other hand, when the levels of λ are increased, the model fits to the empirical data remains relatively unchanged across our subject cohort, but there is a smaller increase in the average number of NCs. Interestingly, we observed notable patterns in the relationship between inhibitory parameters and the generation of NCs. Examining the empirical data, where the average number of NCs is 3310.83 ± 2404.90 across our subject cohort, we found that the number of NCs at the maximum λ value aligns with this empirical average (2161.00 ± 1859.02), compared to the number of NCs at the maximum $W_I$ value significantly exceeds it (13111.25 ± 5313.50). This observation is consistent with our findings that an increase in λ exhibits better fits, more favorable NC patterns, and overall improved model stability compared to an increase in $W_I$. An insight into the mechanistic differences between these parameters can be found in *Eq. 4*, highlighting that λ directly influences SC ($C_{ij}$) between neuronal populations, whereas $W_I$ is associated with $I_0$, representing neural activity that contributes to the network input. Notably, changes in λ impact network connections in a node-wise manner after input summation, influencing the inputs to each node after summation has occurred. Conversely, $W_I$ represents a node-wise input that does not affect the connections between nodes. This distinction likely contributes to the differential effects of $W_I$ and λ on the number of NCs, NC patterns, and overall model stability. Thus, the nuanced interplay between these inhibitory parameters underscores their unique roles in shaping resting-state brain dynamics and the emergence of NCs.

Second, our results show that an increase in the number of NCs (due to higher levels of inhibition), is accompanied by a decrease in the number and magnitude of PCs in the simulated data. This is consistent with findings from previous computational modelling research, showing that inhibition is potentially responsible for NCs in electrophysiological studies. For example, increased inhibition and reduced excitation in a spiking network model of the visual cortex led to stronger NCs in modeled time series (Chelaru & Dragoi, 2016). As for the reduction in PCs, this may be occurring as a consequence of enhanced inhibitory feedback, which in turn causes weaker excitatory feedback within the neuronal populations of the model as the amount of inhibition is increased. This amplified inhibitory feedback may thus suppress excitatory neurons, leading to an overall decrease in PCs. Thus, the lower number of PCs observed when NCs are higher may simply result from a lower overall level of interaction and correlation between network nodes. Alternatively, an increase in inhibition could lead to a decrease in PCs due to alterations in local dynamics (Deco et al., 2014). The increased level of inhibition can influence the model's behavior, causing a reconfiguration of connectivity patterns between neuronal populations, ultimately impacting their correlation. For instance, strengthened inhibitory connections may result in increased NCs, while weakened excitatory connections may lead to reduced PCs. Increased inhibition, particularly in the context of the nonlinearity inherent in neural mass models, can give rise to complex spatiotemporal patterns, including oscillations, chaos, and multistability. These intricate patterns may nonlinearly affect correlations between neuronal populations (Breakspear, 2017; Jansen & Rit, 1995).

***Excitation and inhibition exert varying effects on model stability and negative correlations***

So far, we have seen how the level of inhibition affects the model's behavior and the number of NCs. To gain a comprehensive understanding of how these effects are mediated, we removed the excitatory and inhibitory synaptic couplings ($J_{NMDA}$ and $J_i$, respectively) from our model. Synaptic coupling refers to the interaction or connectivity between different neuronal populations. When considering the RWW model, which consists of a pair of neuronal populations with both excitatory and inhibitory neurons at each of the 200 parcellations, synaptic coupling involves the exchange of signals or information through synapses. Excitatory connections (mediated by NMDA in our model) tend to increase the likelihood of an action potential in the target neuron, while inhibitory connections (mediated by GABA in our model) decrease this likelihood. The synaptic coupling between excitatory and inhibitory neurons is a crucial aspect of neural network dynamics. It plays a significant role in regulating the overall activity and balance within the network, influencing factors such as excitability, synchronization, and the emergence of complex patterns of activity. In this study, we deliberately manipulated the synaptic coupling values to investigate the impact of these changes on the overall behavior of the model. By removing the presence of synaptic couplings, and simultaneously increasing the amount of either excitation or inhibition, we sought to elucidate how these factors shape the dynamics, connectivity patterns, and emergent phenomena within the simulated neural network. This approach allows for a comprehensive exploration of how synaptic couplings contribute to the observed behavior and functionality of the model under different conditions.

In Figure 4A, we see how the model behaves when the $W_E$ value is increased in two separate scenarios. The first is when the model is unchanged, and the second is when $J_i$ is set to 0. When the model is unchanged, increasing $W_E$ leads to an increase in the number of NCs. However, in this first scenario, the model's simulated BOLD fits to empirical data were also significantly poorer at higher $W_E$ values. We interpret this result as reflecting a process of runaway excitation, whereby higher values of excitation multiply, causing an imbalance in the model dynamics, loss of synchronous interactions between nodes, and poorer fit to empirical data. In the second case, when Ji was set to 0, and $W_E$ was increased, we observed an increase in NCs. However, model stability was quickly seen to deteriorate with increasing $W_E$ values. This is because without Ji, there is an uncontrolled activity of excitatory neurons and there is no dampening effect of the inhibitory neurons, unlike the unchanged model. As the $W_E$ values are increased, the population activity gets more and more desynchronized resulting in poorer fits compared to the default model. We also looked at the effect of setting $J_{NMDA}$ to 0 (see Supplementary Information). With the link from the excitatory population to the inhibitory population removed, the model becomes a variant of the 1-dimensional excitatory-only system studied by Deco et al. (2013), and Hansen et al. (2015). In this case, regardless of any increase in $W_E$, the model fit to the empirical data was zero, and there was no discernable spatial pattern in the FC matrices, let alone in the NCs. Without $J_{NMDA}$, there is no input from the excitatory neuronal population to the inhibitory ones, and the overall level of excitation within the neuronal population is increased. The resultant decrease in correlation between the excitatory and inhibitory populations (Breakspear et al., 2003) leads to a more random, asynchronous activity pattern, as observed in the simulated FC patterns.

In Figure 4B, we see how the model behaves when $W_I$ is increased in two analogous scenarios. First, the model is unchanged, and second, when $J_i$ is set to 0. When the model is unchanged, increasing the level of $W_I$ leads to a significant increase in the number of NCs. However, this increase in the number of NCs is not as high as when $W_E$ is increased. However, the model's simulated BOLD fits to the empirical data remains relatively consistent. When $J_i$ was set to zero, and $W_I$ was increased, we observed only a slight increase in NCs, but this was nowhere near as potent as the unchanged model. A potential reason for this is that the increase in inhibition does nothing to affect the activity of the excitatory neurons as $J_i$ is absent. As mentioned above, when $J_{NMDA}$ is set to zero (see Supplementary Information), as in the case with $W_E$, and for the same reasons (Breakspear et al., 2003; Deco et al., 2013; Hansen et al., 2015), when the level of $W_I$ is increased, the model fails to exhibit the expected behavior or ceases to function as intended, with essentially random correlation patterns in the simulated FC matrices. The results from Figure 4 showcase the importance of the interaction between the models' excitatory and inhibitory populations. Furthermore, these results shed light on the role of the dynamics present in the model and their relationship to NCs.

***Implications for the use of negative correlations in clinical neuroscience***

In MDD, the predominant and most effective modern neuromodulatory therapies involve rTMS of the left dlPFC and deep brain stimulation (DBS) of the SGC (Fox et al., 2014). Previous investigations have revealed that the FC between these targets is characterized by a relatively large number of NCs, suggesting that stimulation of these regions may influence a shared, disease-specific neural network. For example, multiple studies have shown that the NC between the dlPFC and SGC can predict the outcomes of rTMS, with a more negative relationship between these regions leading to a better outcome in patients following rTMS treatment (Cash et al., 2020; Fox et al., 2012; Siddiqi & Fox, 2024). Alternatively, brain stimulation has also been shown to induce NCs between brain regions. For example, a study looking at the effects of noninvasive brain stimulation on cortico-thalamic-striatal FC, by applying rTMS to the dorsal-medial prefrontal cortex (dmPFC), showed that MDD patients who showed the most improvement following the treatment course had an increase in NCs between the dmPFC and the insula. These subjects also developed an increased NC between the dmPFC and SGC (Salomons et al., 2014). In addition, research has demonstrated that various disorders exhibit therapeutic benefits when stimulating sites that manifest NCs (i.e., negatively correlated network nodes). Excitatory stimulation of one site or inhibitory stimulation of the other, both result in symptomatic improvement. Beyond MDD, stimulation of negatively correlated sites has therapeutic effects in Parkinson's disease, addiction, Alzheimer's disease, anorexia, gait dysfunction, and pain (Fox et al., 2014). Consequently, NCs are significant as they assist in identifying stimulation sites with the potential to induce beneficial effects across various neurological and psychiatric disorders.

Based on the results of this study, the precise picture in terms of how these clinical improvements relate to the neural circuit mechanisms influencing the expression of NCs that we have focused on appears complicated. This is primarily because the topographic distribution of NCs seen in our simulations does not align very well with those in empirical data. While there is an increase in the number of NCs when the level of inhibition is increased, we are not yet able to decipher the specific spatial patterns of NCs or their variations in the population. Future studies can focus on determining the spatial specificity of NCs in model-based simulations and this can

in turn shed more light on how NCs contribute to clinical improvements in MDD and other disorders. NCs represent a pervasive characteristic of human brain networks and organization. They are observed throughout the cortex, irrespective of brain state, highlighting their biological importance. Thus, NCs may hold a particular relevance in clinical contexts.

***Limitations***

While the results of this study provide some important insights into the potential mechanisms underlying NCs in resting-state brain connectivity, there are some important limitations to highlight.

*Model Fit to Empirical Data*: Concerning model fitting, it is noteworthy that our average model fit to empirical rs-fMRI FC data is 0.24 ± 0.06 (max 0.41) in the RWW model, which is comparatively lower than the previously reported model fits using the RWW model of ~0.5 (Deco et al., 2014). The primary reason for this difference lies in our deliberate choice to employ a higher-resolution brain parcellation in our models, resulting in higher-resolution anatomical and simulated FC matrices. Most studies modeling rs-fMRI with the RWW model have traditionally utilized parcellations with ~66 brain regions (Deco et al., 2013, 2014). In our work, we opted for the 200-region Schaefer-Yeo parcellation (Schaefer et al., 2018; Yeo et al., 2011), which we believe strikes a better balance between spatial granularity and computational feasibility than 66 regions. It is crucial to note that our choice of a 200-region brain parcellation in this study, in contrast to the standard whole-brain dynamics studies with fewer regions, introduces a considerable increase in both regions and connections. While a higher 'R' value in model fits might suggest a better representation of empirical data, it does not necessarily translate to accurate data prediction (Honey et al., 2009; Messé et al., 2015; Roberts et al., 2019). An unavoidable tradeoff of transitioning to larger parcellations is, however, a reduction in the simulated data fits. The lower fits can be attributed to the following reasons: 1) in the empirical rs-fMRI data, parcel time series are noisier due to averaging over fewer voxels, and 2) in the simulated rs-fMRI data, these additional nodes introduce their own RWW equation noise terms, thereby increasing the total level of stochastic variability in the simulated network. Previous research studies have reported biophysical model fits to empirical data where ~R = 0.1 to 0.5 (Messé et al., 2015; Roberts et al., 2019). In the context of our investigation into the nature of NCs in rs-fMRI data, our focus is not on achieving the best-in-class model fits but rather on obtaining results that are reasonably adequate and qualitatively resemble the empirical data, which our findings demonstrate.

*Simplified Models*: Our study uses the RWW model, which is a simplified mathematical model of brain dynamics. This model does not capture the full complexity and diversity of real brain networks, which display a plethora of features such as oscillations, chaoticity, and spatiotemporal pattern generation. These simplifications and associated assumptions in the models may limit the generalizability of our results to real-world brain functioning. We do note however that we also used the JR model and got similar results (*See Supplementary Information*).

*Limited Parameters*: Our study focuses on the effects of increasing the level of inhibition in the RWW model on the number of NCs in simulated rs-fMRI data. However, other factors and parameters of this model that can potentially affect simulated rs-fMRI data, such as excitation, synaptic connectivity, and network topology, are not fully explored in the study. For example, we did not fully explore continuous variations of $J_{NMDA}$ and $J_i$, but rather we set them to zero, as our focus was on the effects of the inhibitory parameters, $W_I$ and $\lambda$. It is possible that these and other parameters may interact non-linearly in ways that are significant for the generation of fMRI correlation patterns. Further research on this question should consider a broader range of parameters and their interactions, which may provide a more comprehensive understanding of the physiological origins of NCs.

*Interpretation of Results*: We acknowledge the observed increase in NCs and decrease in PCs with heightened inhibition in the models and we recognize the importance of further investigation to enhance our understanding. The interpretation of the results is grounded in observed correlations within the simulated data and existing literature. To strengthen the reliability of our proposed mechanisms and interpretations, additional research and validation are essential.

## *Conclusions*

This study introduces a novel computational modeling approach to study the effects of varying inhibition and how it affects rs-fMRI dynamics (specifically, NCs) at the population level in brain network models. Our work adds to the existing literature on brain network dynamics and provides new insights into the potential mechanisms underlying NCs in brain connectivity. We employ rigorous mathematical modelling techniques, specifically connectome-based whole-brain modelling using the RWW model, to simulate neural dynamics and investigate the relationship between varying levels of inhibition and NCs in resting-state brain connectivity.

Our results show that increasing the amount of inhibition in the RWW model leads to a robust increase in the number of NCs in the simulated data. Different inhibitory parameters in the model had varying effects on the model fit to empirical data and the number of NCs. Previous research has suggested that inhibition could directly or indirectly play a role in determining the number and magnitude of NCs, but it is still unclear why an increase in NCs leads to a reduction in PCs. From a clinical perspective, our findings may have practical relevance and could potentially contribute to the development of novel therapeutic interventions for mental health disorders, such as rTMS therapy for MDD. Potential directions for future research include further validation of the findings using empirical data from clinical studies, exploration of other parameters and factors that may influence resting-state brain connectivity and clarification of the underlying mechanisms and interactions between inhibitory and excitatory dynamics in the brain.

## Acknowledgments

We are grateful to the Krembil Foundation, CAMH Discovery Fund, and Labatt Family Network for the generous funding support that has made this research possible.

*Author contributions*: SH: Conceptualization, Methodology, Formal analysis, Writing - Original Draft, Writing - Review & Editing, Visualization; DM: Methodology, Writing - Review & Editing; ZW: Methodology; SPB: Writing - Review & Editing; JDG: Conceptualization, Methodology, Writing - Review & Editing, Supervision, Funding acquisition.

**References**

Anne Weigand, Andreas Horn, Ruth Caballero, Danielle Cooke, Adam P. Stern, Stephan F. Taylor, Daniel Press, Alvaro Pascual-Leone, and Michael D. Fox. (2018). Prospective Validation That Subgenual Connectivity Predicts Antidepressant Efficacy of Transcranial Magnetic Stimulation Sites. *Biological Psychiatry*, *84*(1), 28–37.

Asgharian Asl, F., & Vaghef, L. (2022). The effectiveness of high-frequency left DLPFC-rTMS on depression, response inhibition, and cognitive flexibility in female subjects with major depressive disorder. *Journal of Psychiatric Research*, *149*, 287–292.

Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine: Official Journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, *34*(4), 537–541.

Breakspear, M. (2017). Dynamic models of large-scale brain activity. *Nature Neuroscience*, *20*(3), 340–352.

Breakspear, M., Heitmann, S., & Daffertshofer, A. (2010). Generative models of cortical oscillations: neurobiological implications of the kuramoto model. *Frontiers in Human Neuroscience*, *4*, 190.

Breakspear, M., Terry, J. R., & Friston, K. J. (2003). Modulation of excitatory synaptic coupling facilitates synchronization and complex dynamics in a biophysical model of neuronal dynamics. *Network*, *14*(4), 703–732.

Brunoni, A. R., Chaimani, A., Moffa, A. H., Razza, L. B., Gattaz, W. F., Daskalakis, Z. J., & Carvalho, A. F. (2017). Repetitive Transcranial Magnetic Stimulation for the Acute Treatment of Major Depressive Episodes: A Systematic Review With Network Meta-analysis. *JAMA Psychiatry* , *74*(2), 143–152.

Buxton, R. B., & Frank, L. R. (1997). A model for the coupling between cerebral blood flow and oxygen metabolism during neural stimulation. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, *17*(1), 64–72.

Buxton, R. B., Wong, E. C., & Frank, L. R. (1998). Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Magnetic Resonance in Medicine: Official Journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, *39*(6), 855–864.

Cash, R. F. H., Weigand, A., Zalesky, A., Siddiqi, S. H., Downar, J., Fitzgerald, P. B., & Fox, M. D. (2020). Using Brain Imaging to Improve Spatial Targeting of Transcranial Magnetic Stimulation for Depression. *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2020.05.033

Chai, X. J., Castañón, A. N., Ongür, D., & Whitfield-Gabrieli, S. (2012). Anticorrelations in resting state networks without global signal regression. *NeuroImage*, *59*(2), 1420–1428.

Chelaru, M. I., & Dragoi, V. (2016). Negative Correlations in Visual Cortical Networks. *Cerebral Cortex* , *26*(1), 246–256.

Cole, E. J., Stimpson, K. H., Bentzley, B. S., Gulser, M., Cherian, K., Tischler, C., Nejad, R., Pankow, H., Choi, E., Aaron, H., Espil, F. M., Pannu, J., Xiao, X., Duvio, D., Solvason, H. B., Hawkins, J., Guerra, A., Jo, B., Raj, K. S., … Williams, N. R. (2020). Stanford Accelerated Intelligent Neuromodulation Therapy for Treatment-Resistant Depression. *The American Journal of Psychiatry*, appiajp201919070720.

Deco, G., Jirsa, V. K., & McIntosh, A. R. (2011). Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nature Reviews. Neuroscience*, *12*(1),

43–56.

Deco, G., Ponce-Alvarez, A., Hagmann, P., Romani, G. L., Mantini, D., & Corbetta, M. (2014). How local excitation-inhibition ratio impacts the whole brain dynamics. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(23), 7886–7898.

Deco, G., Ponce-Alvarez, A., Mantini, D., Romani, G. L., Hagmann, P., & Corbetta, M. (2013). Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(27), 11239–11252.

Devor, A., Tian, P., Nishimura, N., Teng, I. C., Hillman, E. M. C., Narayanan, S. N., Ulbert, I., Boas, D. A., Kleinfeld, D., & Dale, A. M. (2007). Suppressed neuronal activity and concurrent arteriolar vasoconstriction may explain negative blood oxygenation level-dependent signal. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(16), 4452–4459.

Fox, M. D., Buckner, R. L., Liu, H., Chakravarty, M. M., Lozano, A. M., & Pascual-Leone, A. (2014). Resting-state networks link invasive and noninvasive brain stimulation across diverse psychiatric and neurological diseases. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(41), E4367–E4375.

Fox, M. D., Buckner, R. L., White, M. P., Greicius, M. D., & Pascual-Leone, A. (2012). Efficacy of transcranial magnetic stimulation targets for depression is related to intrinsic functional connectivity with the subgenual cingulate. *Biological Psychiatry*, *72*(7), 595–603.

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

Friston, K. J., Mechelli, A., Turner, R., & Price, C. J. (2000). Nonlinear responses in fMRI: the

Balloon model, Volterra kernels, and other hemodynamics. *NeuroImage*, *12*(4), 466–477.

Goelman, G., Gordon, N., & Bonne, O. (2014). Maximizing negative correlations in resting-state functional connectivity MRI by time-lag. *PloS One*, *9*(11), e111554.

Griffiths, J. D., McIntosh, A. R., & Lefebvre, J. (2020). A Connectome-Based, Corticothalamic Model of State- and Stimulation-Dependent Modulation of Rhythmic Neural Activity and Connectivity. *Frontiers in Computational Neuroscience*, *14*, 575143.

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., & Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS Biology*, *6*(7), e159.

Hansen, E. C. A., Battaglia, D., Spiegler, A., Deco, G., & Jirsa, V. K. (2015). Functional connectivity dynamics: modeling the switching behavior of the resting state. *NeuroImage*, *105*, 525–535.

Heitmann, S., Gong, P., & Breakspear, M. (2012). A computational role for bistability and traveling waves in motor cortex. *Frontiers in Computational Neuroscience*, *6*, 67.

Honey, C. J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J. P., Meuli, R., & Hagmann, P. (2009). Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(6), 2035–2040.

Jansen, B. H., & Rit, V. G. (1995). Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biological Cybernetics*, *73*(4), 357–366.

Leuchter, A. F., Cook, I. A., Feifel, D., Goethe, J. W., Husain, M., Carpenter, L. L., Thase, M. E., Krystal, A. D., Philip, N. S., Bhati, M. T., Burke, W. J., Howland, R. H., Sheline, Y. I., Aaronson, S. T., Iosifescu, D. V., O'Reardon, J. P., Gilmer, W. S., Jain, R., Burgoyne, K. S.,

… George, M. S. (2015). Efficacy and Safety of Low-field Synchronized Transcranial Magnetic Stimulation (sTMS) for Treatment of Major Depression. *Brain Stimulation*, *8*(4), 787–794.

Levkovitz, Y., Isserles, M., Padberg, F., Lisanby, S. H., Bystritsky, A., Xia, G., Tendler, A., Daskalakis, Z. J., Winston, J. L., Dannon, P., Hafez, H. M., Reti, I. M., Morales, O. G., Schlaepfer, T. E., Hollander, E., Berman, J. A., Husain, M. M., Sofer, U., Stein, A., … Zangen, A. (2015). Efficacy and safety of deep transcranial magnetic stimulation for major depression: a prospective multicenter randomized controlled trial. *World Psychiatry: Official Journal of the World Psychiatric Association* , *14*(1), 64–73.

Mandeville, J. B., Marota, J. J., Ayata, C., Zaharchuk, G., Moskowitz, M. A., Rosen, B. R., & Weisskoff, R. M. (1999). Evidence of a cerebrovascular postarteriole windkessel with delayed compliance. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, *19*(6), 679–689.

Messé, A., Rudrauf, D., Giron, A., & Marrelec, G. (2015). Predicting functional connectivity from structural connectivity via computational models using MRI: an extensive comparison study. *NeuroImage*, *111*, 65–75.

Murphy, K., & Fox, M. D. (2017). Towards a consensus regarding global signal regression for resting state functional connectivity MRI. *NeuroImage*, *154*, 169–173.

Roberts, J. A., Gollo, L. L., Abeysuriya, R. G., Roberts, G., Mitchell, P. B., Woolrich, M. W., & Breakspear, M. (2019). Metastable brain waves. *Nature Communications*, *10*(1), 1056.

Rosazza, C., & Minati, L. (2011). Resting-state brain networks: literature review and clinical applications. *Neurological Sciences: Official Journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*, *32*(5), 773–785.

Saenger, V. M., Ponce-Alvarez, A., Adhikari, M., Hagmann, P., Deco, G., & Corbetta, M. (2018). Linking Entropy at Rest with the Underlying Structural Connectivity in the Healthy and Lesioned Brain. *Cerebral Cortex* , *28*(8), 2948–2958.

Salomons, T. V., Dunlop, K., Kennedy, S. H., Flint, A., Geraci, J., Giacobbe, P., & Downar, J. (2014). Resting-state cortico-thalamic-striatal connectivity predicts response to dorsomedial prefrontal rTMS in major depressive disorder. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *39*(2), 488–498.

Sanz-Leon, P., Knock, S. A., Spiegler, A., & Jirsa, V. K. (2015). Mathematical framework for large-scale brain network modeling in The Virtual Brain. *NeuroImage*, *111*, 385–430.

Sanz Leon, P., Knock, S. A., Woodman, M. M., Domide, L., Mersmann, J., McIntosh, A. R., & Jirsa, V. (2013). The Virtual Brain: a simulator of primate brain network dynamics. *Frontiers in Neuroinformatics*, *7*, 10.

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., Eickhoff, S. B., & Yeo, B. T. T. (2018). Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cerebral Cortex* , *28*(9), 3095–3114.

Shmuel, A., Augath, M., Oeltermann, A., & Logothetis, N. K. (2006). Negative functional MRI response correlates with decreases in neuronal activity in monkey visual area V1. *Nature Neuroscience*, *9*(4), 569–577.

Siddiqi, S. H., & Fox, M. D. (2024). Targeting Symptom-Specific Networks With Transcranial Magnetic Stimulation. *Biological Psychiatry*, *95*(6), 502–509.

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., Ugurbil, K., & WU-Minn HCP Consortium. (2013). The WU-Minn Human Connectome Project: an overview. *NeuroImage*, *80*, 62–79.

Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(3), 1125–1165.

Zheng, M., Allard, A., Hagmann, P., Alemán-Gómez, Y., & Serrano, M. Á. (2020). Geometric renormalization unravels self-similarity of the multiscale human connectome. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(33), 20244–20253.