Blueprint for Task-Capable Prototype (D3.19 – SGA3)

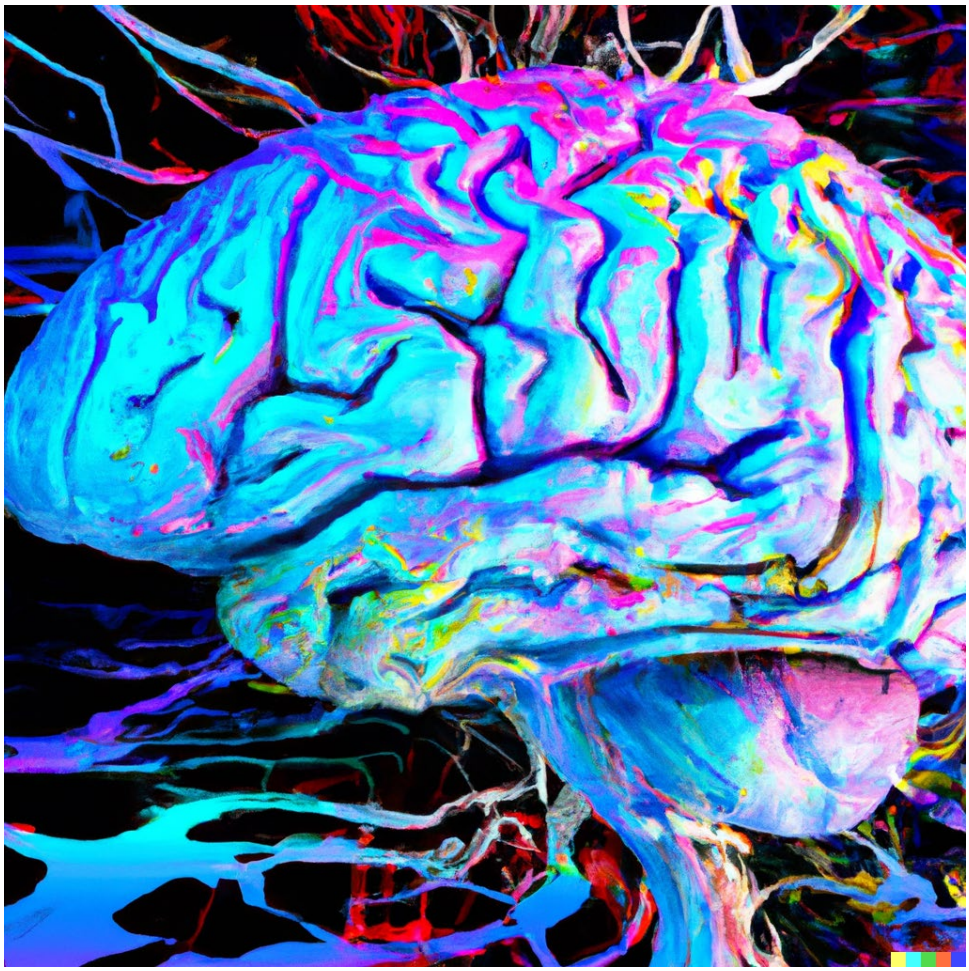

Figure 1: Artistic rendition of the envisioned task-capable whole-brain model.

Project Number:	945539	Project Title:	HBP SGA3

Document Title:	Blueprint for Task-Capable Prototype
Document Filename:	D3.19 (D119) SGA3 M40 SUBMITTED 230830.docx
Deliverable Number:	SGA3 D3.19 (D119)
Deliverable Type:	Report
Dissemination Level:	PU = Public
Planned Delivery Date:	SGA3 M40 / 31 JUL 2023
Actual Delivery Date:	SGA3 M41 / 30 AUG 2023
Author(s):	Rainer GOEBEL, UM (P117), Mario, SENDEN, UM (P117), Viktor JIRSA, AMU (P78), Mavi SANCHEZ, IDIBAPS (P93), Alain DESTEXHE, CNRS (P10)
Compiled by:	Rainer Goebel, UM (P117), Mario, SENDEN, UM (P117)
Contributor(s):	
WP QC Review:	Jeannette BOSCHMA, UM (P117)
WP Leader / Deputy Leader Sign Off:	Rainer Goebel, UM (P117)
T7.4 QC Review:	N/A
Description in GA:	Blueprint for task-capable prototype, including written document summarizing results of activities and conclusions. The document will identify lead paradigms, likely to be derived from existing cognitive tasks in WP2-WP3, discuss, interrogate, and analyse key challenges and propose solutions leading to the creation of a blueprint guiding the proposed prototype. The discussion will also include an evaluation of how models of brain states and cognition can be optimally integrated and an evaluation of the capacity of use of neuromorphic simulation. Tasks contributing to this Deliverable: T1.11, T2.1, T2.4 and T3.1
Abstract:	Roadmap for developing the first prototype of a Task-Capable Whole-Brain model. To foster innovation and exploit multiple ideas, each WP worked on specific approaches towards solving the challenging problem to integrate task-performing networks with connectome-based networks of neuronal dynamics that model brain measurements. In this blueprint, the different initial attempts, and how they can be integrated to build at least one prototype are described.
Keywords:	Whole-Brain Modelling, Neurocomputational Models, Visual Processing, Convolutional Neural Network (CNN), Representational Dissimilarity Matrices, Representational Connectivity
Target Users/Readers:	computational neuroscience community, computer scientists, neuroscientific community, neuroscientists

Table of Contents

Table of Figures

1. Introduction

In recent years, the scientific community has witnessed a significant paradigm shift towards a comprehensive, quantitative understanding of brain function. Whole-brain models serve as cornerstones of this intellectual endeavour, offering a mathematical description of the intricate dynamics and interconnectedness of various neural populations distributed across different brain regions. Traditionally, these models have been structured around three fundamental components: brain parcellation, an anatomical connectivity matrix, and local dynamics.

Brain parcellation presents a spatial resolution framework where brain dynamics occur, encompassing cortical, sub-cortical, and cerebellar regions. The anatomical connectivity matrix defines the network of interconnections between these regions, often based on the human connectome. Local dynamics signify the chosen activity within each region, dictated by interactions with other areas. These components, unified, offer a powerful tool to predict and investigate brain dynamics.

Underpinning the creation of these models is an array of diverse biophysical variables and their joint evolution, typically represented by coupled differential equations, or in some instances, discrete time step models. The equations are either derived from biophysical mechanisms governing brain activity forms or chosen phenomenologically, depending on the desired dynamic outcomes. To imbue these models with empirical relevance, neuroimaging tools, such as fMRI, EEG, MEG, DTI, and PET, play a vital role, providing in vivo estimates of anatomical connectivity networks, structural connectivity, statistical observables, metabolic information, and neuromodulator receptor density maps.

Despite the predictive power and comprehensive scope of these models, a salient limitation is their lack of consideration for the computational and informational processing inherent in the nervous system. This is a realm typically catered to by neurocomputational models, which seek to illuminate the representational and computational underpinnings of high-level biological functions such as perception and cognition. These models, however, are often more abstract, less biologically specific, and restricted to limited brain regions.

In response to this dichotomy, we seek to combine the strengths of both whole-brain and neurocomputational models (see figure 1 for an artistic rendering). Each involved WP was asked to suggest ideas to approach this challenging problem and to develop and work on a first prototype by initiated across-WP collaborations. While the provided ideas reflect the specific expertise of a WP, the roadmaps clearly indicate that approaches converge allowing to construct first prototypes of Task-Capable Whole-Brain models that simultaneously encapsulate the global dynamics and specific computational characteristics of the brain. The sections below will outline the suggested paths for developing prototypes from the viewpoint of each WP. This is then followed by a discussion of how these initial ideas to develop Task-Capable Whole-Brain Modeling can be further integrated.

2. Task-Capable Whole-Brain Model: WP1 Roadmap

We here take a theoretical perspective to the principal question of “how to make virtual brains think?”. Obviously, this statement is to be understood with a grain of salt, but it conveys well the principal challenge we encounter of linking task-relevant function, typically described as a goal-oriented process evolving over time, to brain activity, emergent from networks comprising the aforementioned components, that is brain parcellation, anatomical connectivity matrix, and local dynamics. Networks of the latter type are referred to as virtual brains (using anatomical connectivity matrix, often derived from diffusion weighted MRI). In this project, we reformulate the question to make it theoretically more accessible, as follows: given an appropriate representation of task-related function, what are the neural mechanisms to make the corresponding mathematical object arise from a virtual brain dynamics?

2.1 Structured Flows on Manifolds (SFMs)

Here we take the perspective of Structured Flows on Manifolds (SFMs), where the task-related function is represented as a process unfolding on a low-dimensional manifold associated with structured (aka task relevant meaningful) flow. The SFM is the mathematical object, representing the lawful behavior of the function, which in turn is hypothesised to be represented in the dynamics of the brain [1, 2]. In dynamic system theory, such lawful behavior can be captured by various fully equivalent means, including the totality of all possible behaviors as measured by the corresponding time series, or a set of ordinary differential equations (essentially capturing the task-specific rules underlying the function), or discrete maps, or integral equations. In cognitive neuroscience, predictive coding is one of the most influential contemporary theories of brain function and its internal model representing the subjects understanding of world is essentially the conceptualization of the SFM. Predictive coding is based on the intuition, that the brain operates as a Bayesian inference system, which "makes emerge" the internal model (for our purposes, here, the SFM) creating predictions about the outside world. These predictions are continuously compared against sensory input, resulting in prediction errors and updating the internal model.

We are now in a position to rephrase our initial reflection in a more formal manner, interrogating the mechanisms underlying the emergence of SFMs in virtual brain networks. As a first step, we consider the resting state of the virtual brain dynamics from the perspective of SFMs (see figure 2). We investigate the resting state manifold itself, and characterise its flow. A substantial body of work has shown the relevance of the resting state brain activity patterns to the task functional networks, or cognitive task performance [3, 4, 5], motivating the consideration of resting state dynamics together with the task. The SFMs formalism can also be applied to the resting state of the brain, allowing us to treat both conditions in the same conceptual framework. Then we discuss the necessary actions towards changes of the resting SFM to make it functionally meaningful for a particular task, which establishes the link to the other activities in WP2 and WP3. Finally, the question of validation arises, addressing the challenge of how the SFMs can be extracted from experimental brain imaging data.

2.2 SFMs arising from brain networks at rest

The SFM description of the resting state follows two veins in parallel. First builds the manifold and the flow bottom-up in the connectome-based model, providing a mechanistic insight into how the individual features structural features such as connectivity or spatial heterogeneity shape the resting state activity. Second extracts the manifold and the flow from the empirical data providing constraints on the modeling efforts. Having the SFM as a shared object between these two streams allows us to go beyond the usually applied descriptive statistics (such as functional connectivity), and, more importantly, bears the promise of being able to relate to the SFMs of particular tasks.

The working hypothesis is that the recurrent functional connectivity states of the resting state correspond to distinct subspaces on a low-dimensional manifold associated with distinct structured flows. For the bottom-up part, we construct the whole-brain model using a subject-specific connectome derived from dMRI data, and equip each of the nodes with the Montbrio-Pazo-Roxin mean field model (MPR, [6]) in the bi-stable regime supporting up- and down-states. Scaling the coupling strength appropriately sets the model to the regime exhibiting fluid dynamics as captured on the level of fMRI observables (co-activations, variable functional connectivity dynamics) and low-dimensional manifold with a characteristic flow [7]. Main features of the resting-state SFM are the transitions between high- and low-activity subspaces, where the former support the structured flow of network-coordinated cascades of activity resulting in rare co-activation events [8]. While the symmetry breaking through the connectome alone is enough to reproduce the co-activations and fluid dynamics, further refinement of the model e.g. through spatially heterogeneous parameters can be pursued to bring the simulated resting state SFM to the one derived from empirical data, which remains an ongoing effort.

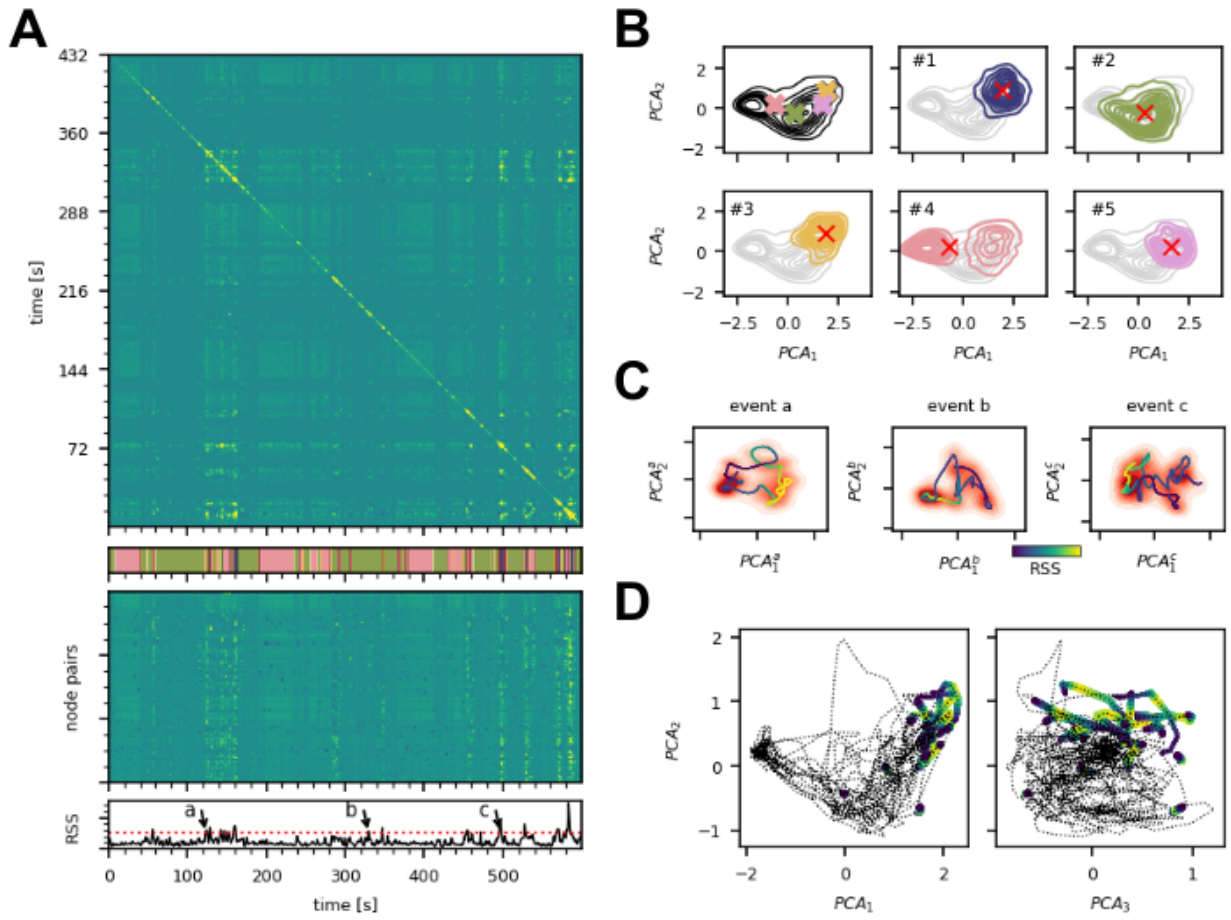

Figure 2: Resting state manifold subspaces and characteristic dynamics

(A) The edge-based dynamic functional connectivity *DFCe* of a simulation of the model in the working point (top) shows the off-diagonal structure of similarity of the system's activity across time. The edge time series (middle) shows the time evolution of the functional connectivity of the simulated BOLD signal between each of the node pairs, and exhibits the characteristic co-activation events defined as time points with the root sum squared (RSS, bottom) crossing the threshold of 95th percentile. Dividing the edge-time series into 5 clusters (k-means, shown in the colorbar under the dFC) has separated the event and non-event time points, and also differentiated the events based on their respective similarity. (B) The time intervals in *r(t)* corresponding to the 5 clusters were selected; in the first panel the centroids of the time points of the individual clusters are marked with a cross in the projection to the first two principal components of the whole time series, following panels show the projection of the *r(t)* intervals of particular clusters. Cluster #2 captures the high-activity subspace, cluster #4 corresponds to the low-activity state, and the clusters #1, #3, and #5 capture the co-activation events. (C) Local trajectories in the manifold subspaces: the time series of the three example events (a,b,c, marked in the panel A bottom) was projected to the first two components of PCA applied to each time segment individually. The smoothened trajectory marks the advance of the system through the event and out of it, and is colored by the value of RSS (yellow at the peak of the event, blue at the beginning and the end). (D) The event trajectories on the manifold. The trajectory of the simulated BOLD signal is projected in the space defined by the first three PCA components with the events colored by the RSS value (yellow at the peak of the event).

In addition to these observations, having full access to the model inner workings allows us to relate the properties of the resting state SfM to the network and node features. For example, the contours of the manifold can be sampled from the fixed-points of the system given by up- and down-state combinations, and the propensity to participation of a node in the cascades related to the stability of the node in a particular fixed point.

For validation in brain imaging data, we will focus on the extraction of the SFM from empirical resting state data, and then expand the approach of 1) model construction, 2) simulation of the emergence of task-related SFMs from connectome based brain networks, and 3) prediction of data features to be observed in brain imaging data. This approach is task-agnostic but requires the characterization of dynamic data features of the behavior, then to be translated into an SFM.

In summary, the joint resting state and task representation as Structured Flows on Manifolds has a potential to provide a powerful framework to integrate diverse and so far separate efforts across the three WPs.

References

[1] Pillai, A. S., & Jirsa, V. K. (2017). Symmetry Breaking in Space-Time Hierarchies Shapes Brain Dynamics and Behavior. Neuron, 94(5), 1010–1026. https://doi.org/10.1016/j.neuron.2017.05.013 (P1221)

[2] Jirsa, V. (2020). Structured Flows on Manifolds as guiding concepts in brain science. In K. Viol, H. Schöller, & W. Aichhorn (Eds.), Selbstorganisation – ein Paradigma für die Humanwissenschaften: Zu Ehren von Günter Schiepek und seiner Forschung zu Komplexität und Dynamik in der Psychologie (pp. 89–102). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-29906-4_6

[3] Gusnard, D. A., Raichle, M. E., & Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. Nature Reviews. Neuroscience, 2(10), 685–694. https://doi.org/10.1038/35094500

[4] Damoiseaux, J. S., Rombouts, S. A. R. B., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., & Beckmann, C. F. (2006). Consistent resting-state networks across healthy subjects. Proceedings of the National Academy of Sciences of the United States of America, 103(37), 13848–13853. https://doi.org/10.1073/pnas.0601417103

[5] Shine, J. M., Bissett, P. G., Bell, P. T., Koyejo, O., Balsters, J. H., Gorgolewski, K. J., Moodie, C. A., & Poldrack, R. A. (2016). The Dynamics of Functional Brain Networks: Integrated Network States during Cognitive Task Performance. Neuron, 92(2), 544–554. https://doi.org/10.1016/j.neuron.2016.09.018

[6] Montbrió, E., Pazó, D., & Roxin, A. (2015). Macroscopic Description for Networks of Spiking Neurons. Physical Review X, 5(2), 021028. https://doi.org/10.1103/PhysRevX.5.021028

[7] Fousek, J., Rabuffo, G., Gudibanda, K., Sheheitli, H., Jirsa, V., & Petkoski, S. (2023). Symmetry breaking organizes the brain's resting state manifold. In bioRxiv (p. 2022.01.03.474841). https://doi.org/10.1101/2022.01.03.474841

[8] Rabuffo, G., Fousek, J., Bernard, C., & Jirsa, V. (2021). Neuronal Cascades Shape Whole-Brain Functional Dynamics at Rest. eNeuro, 8(5), 2020.12.25.424385. https://doi.org/10.1523/ENEURO.0283-21.2021

[9] Zerlaut, Y., Chemla, S., Chavane, F., & Destexhe, A. (2018). Modeling mesoscopic cortical dynamics using a mean-field model of conductance-based networks of adaptive exponential integrate-and-fire neurons. Journal of Computational Neuroscience, 44(1), 45–61. https://doi.org/10.1007/s10827-017-0668-2

3. Task-Capable Whole-Brain Model: WP2 Roadmap 1

Neurons are endowed with a diverse range of ionic channels and neurotransmitter receptors with complex biophysical properties [1]. Neuromodulators act by altering both the properties of synaptic conductance and the intrinsic membrane properties of individual neurons and are involved in the modulation of brain states, for instance, the transition between sleep and awake states [2,3]. These two different brain states, slow wave sleep, and wakefulness, are rather different from a network dynamics perspective: highly synchronous, characterized by the spontaneous emergence of slow oscillations, and desynchronized, with a suppression of slow frequency content, respectively. Neuromodulators like acetylcholine (Ach) and norepinephrine (NE) have been shown to induce state transition in cortical networks [4]. Moreover, Dalla Porta et al [5], have shown that one single ionic current, the muscarinic-sensitive potassium current (M-current), is able to significantly modulate network dynamics during slow oscillations, generating a pattern of activity analogous to a transition towards wakefulness [6]. These studies highlight the importance of intrinsic neuronal properties and their impact on brain state dynamics. However, at the whole-brain level, the impact of single cells properties on the network dynamics is still not clearly understood. Some authors have addressed this question by linking synaptic dynamics to the whole-brain model, by means of regional heterogeneity, i.e., simulating each cortical region according to a receptor gradient. For instance, using a whole-brain model, Deco et al. [7] showed that regional heterogeneity, by means of excitatory and inhibitory receptors balance, closely reproduces the large-scale human brain activity, when compared to a homogeneous model. In another whole-brain model study, Luppi et al [8] introduced heterogeneity in the model through the cortical distribution of GABA receptors, reproducing then MRI activity dynamics during anaesthesia with propofol. More recently, Feng et al [9] have introduced a whole-brain scale model built with NMDA receptor density allowing for the emergence of cognitive function, in their case, distributed working memory. Both studies highlight the importance of local heterogeneity in whole-brain modelling. However, despite efforts in understanding the advantages of local heterogeneity in whole-brain modelling, these studies have only focused on the synaptic heterogeneity level. Thus, it remains to be explored the impact of ionic channel neuromodulators, such as Ach, at the whole-brain level.

Here, we seek to build a whole-brain model which explicitly incorporate neuronal intrinsic properties to simulate realistic whole-brain dynamics across different brain states. We aimed at the cholinergic and noradrenergic neuromodulation, which are both main neuromodulators involved in the arousal system, and in the excitatory NMDA, responsible for cognitive functions such as working memory. We depart from a model recently proposed by Goldman et al [10], which is open-access and executable into The Virtual Brain (TVB) simulator of EBRAINS. The model accounts for the mean-field dynamics of excitatory-inhibitory circuits together with neuromodulation effects. Structural connectivity, synaptic weights and synaptic delays are defined by human tractography data and human diffusion tensor imaging data, respectively [12, 13]. The model is set to reproduce two rather different dynamic brain states corresponding to wakefulness and slow wave sleep dynamics.

3.1 Transcriptional data

The project relies on the Allen Human Brain Atlas (AHBA) dataset, which comprises microarray data quantifying transcriptional activity of > 60,000 genes in > 4000 different tissue samples distributed throughout the brain, taken from six post-mortem samples [14]. From this dataset genes coding for muscarinic, norepinephrine and NMDA receptors were used to constrain regional heterogeneity in the model (Figure 3). More specifically, the genes selected for muscarinic receptors were CHRM1, and CHRM2; the genes for norepinephrine receptors were ADRA1A, ADRA1B, ADRA1D, ADRA2A, and ADRA2C; and the genes for NMDA receptors were GRIN1, GRIN2a, GRIN2b, and GRIN2C. Gene expression within a given brain area were normalized using a scaled robust sigmoid normalization. Gene expression data were retained for 68 cortical regions according to Desikan-Killiany (DK) parcellation map [11] through the abagen toolbox [15].

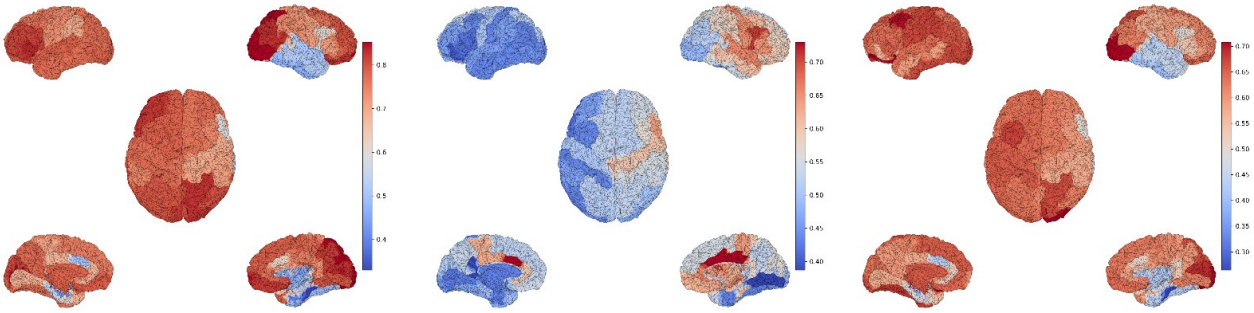

Figure 3: Gene expression spatial map for muscarinic (left), norepinephrine (middle), and NMDA (right).

3.2 Whole-brain model

We depart from the human whole-brain model recently proposed by Goldman et al [10], which is open-access and executable into The Virtual Brain (TVB) simulator of EBRAINS. The model consists of 68 cortical regions according to Desikan-Killiany (DK) parcellation [11]. Structural connectivity sets the local and long-range connections, and it was defined by human tractography data (https://zenodo.org/record/4263723, Berlin subjects/QL_20120814) from the Berlin empirical data processing pipeline [12]. The weights between each connection and the synaptic delays were estimated in human diffusion tensor imaging data (DTI) [13]. Each node, that represents a small section of the brain's cortical surface, was modelled as an excitatory-inhibitory circuit, taking into account spike-frequency adaptation for populations. The dynamics of each node was simulated by the TVB-AdEx model, a mean-field (MF) equation derived from Adaptive Exponential (AdEx) spiking-neuron network simulations which accounts for regular spiking (RS; excitatory), and fast spiking (FS; inhibitory) neurons.

Heterogeneity in the model was introduced by regional expression of cholinergic, noradrenergic and NMDA receptors as follows:

- Adaptation (parameter b in the AdEx-MF model) was modulated by the muscarinic receptors spatial maps by an amount of approximately 20% (e.g., if b = 60pA, 50pA are fixed and 10pA follow the receptor maps);
- Depolarization (parameter EL) was modulated by the norepinephrine receptors spatial maps also by an amount of approximately 20% (corresponding to approximately 10 mV).
- Excitation (parameter τE) was modulated by the NMDA receptors spatial maps by changing the time-decay scale by an amount of approximately 10% (corresponding to approximately 0.5 ms).

To reproduce the transition from slow wave sleep dynamics towards wakefulness, we parametrically decreased the parameter *b* and *EL* in the model, which corresponds to the strength of adaptation and leak reversal potential, respectively. Both variables are modulated by acetylcholine and norepinephrine, respectively.

3.3 Brain states dynamics: wakefulness and slow wave sleep

Using the whole-brain model governed by AdEx Mean-Field and constrained by gene expression data, we reproduced the dynamics characteristic of awake-like (Fig 4-5) and sleep-like (Fig 6-).

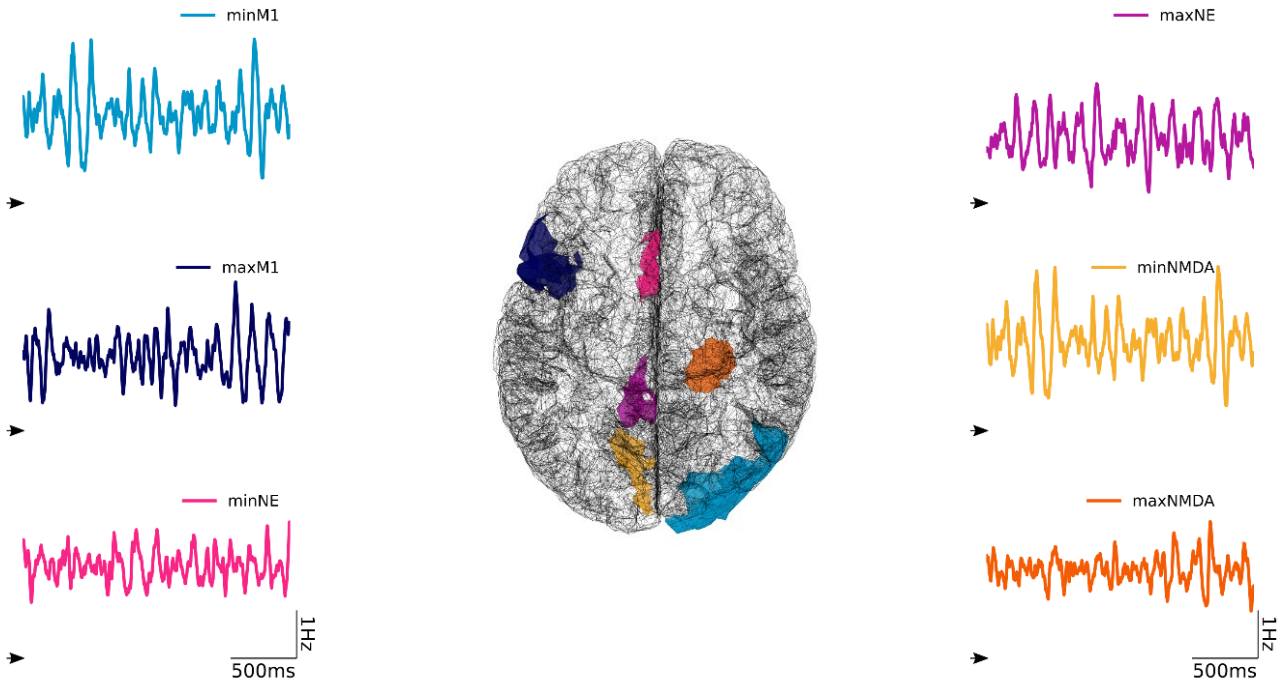

Figure 4: Awake-like activity in different cortical regions across the whole-brain.

Max and Min refer to the maximum and minimum receptors expression, respectively.

Arrows indicate the reference baseline of 4Hz. M1, NE and NMDA stand for Muscarinic M1, Norepinephrine and NMDA, respectively.

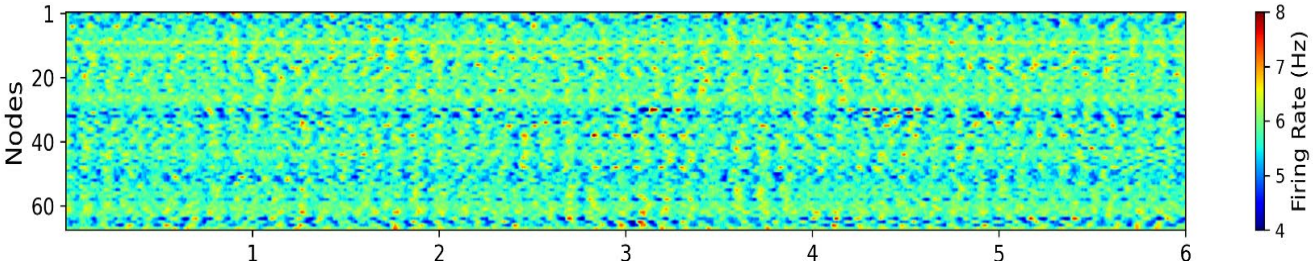

Figure 5: Time course of each node in the whole-brain model during awake-like activity.

The X-axis of the displayed ~10 Hz oscillations indicates time in seconds.

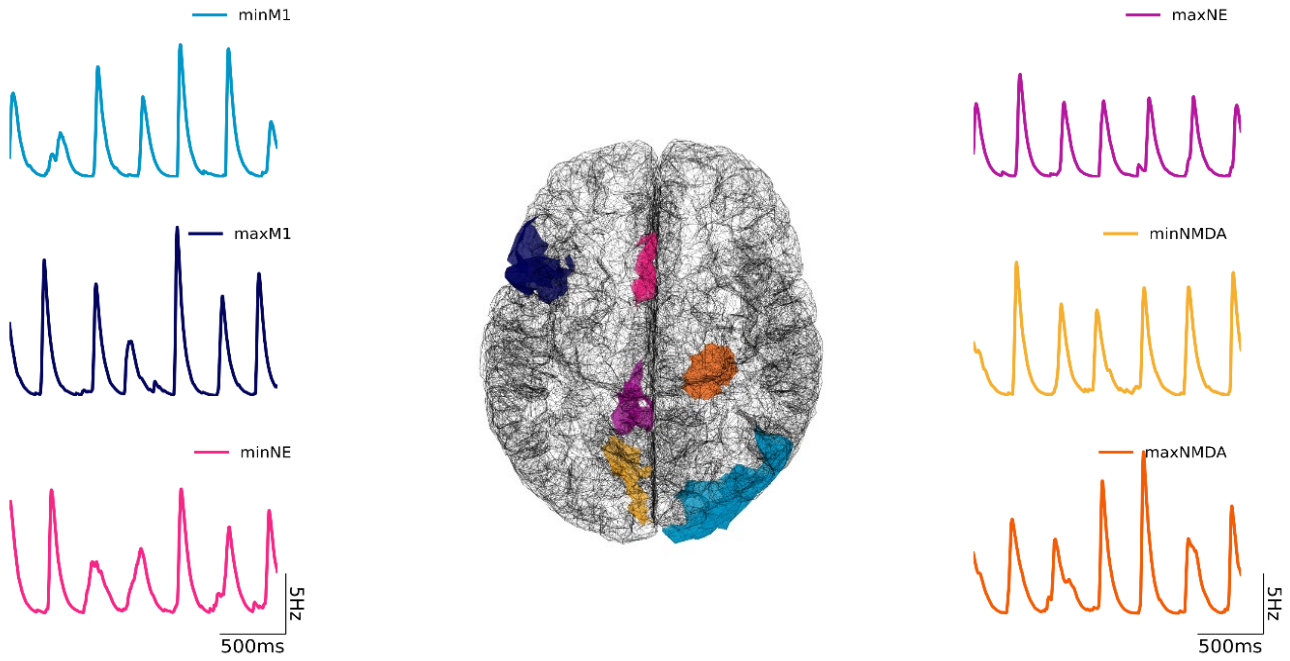

Figure 6: Sleep-like activity in different cortical regions across the whole-brain.

Mx and Min of the displayed slow oscillations refer to the maximum and minimum receptors expression, respectively. M1, NE and NMDA stand for Muscarinic M1, Norepinephrine and NMDA, respectively.

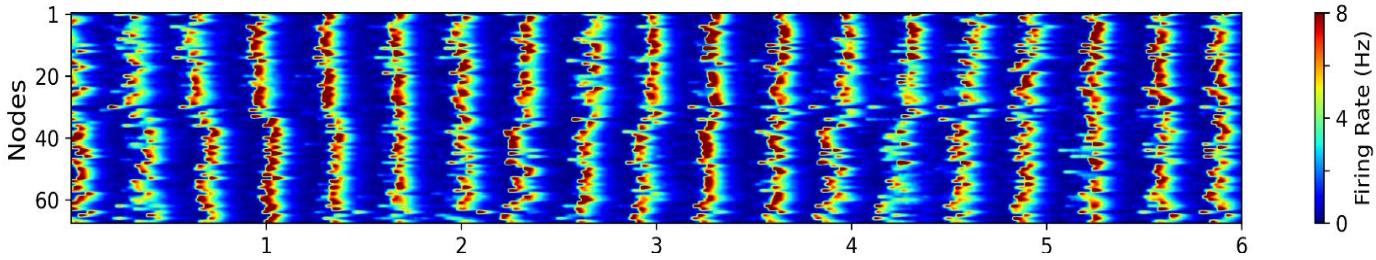

Figure 7: Time course of each node in the whole-brain model during sleep-like activity.

X-axis of the displayed ~1 Hz oscillations indicates time in seconds.

A recent work from Deco et al. [7], has shown the importance of regional heterogeneity. By implementing variations in the expression of excitatory (AMPA and NMDA) and inhibitory (GABA-A) receptors in a whole-brain model, they showed that it improves the spatiotemporal similarity between model and fMRI dynamics in humans. Moreover, they showed that heterogeneity is a key aspect for ignition-like dynamics and for hierarchical time scales.

Our model goes one step further, as it is based in the AdEx Mean-Field, it allows us to implement other types of heterogeneity observed in the human brain, such as those modulated by acetylcholine and norepinephrine, two neurotransmitters involved in the sleep-wake cycle. Furthermore, our approach allows us to easily implement different receptors densities mapped in the Allen Human Brain Atlas, that were not considered in the above presented simulations. Thus, depending on the task to be performed, different receptors can be implemented in the model to improve model-data similarity.

References:

[1] Tobin, A. E., Cruz-Bermúdez, N. D., Marder, E., & Schulz, D. J. (2009). Correlations in ion channel mRNA in rhythmically active neurons. PloS one, 4(8), e6742.

[2] Marder, E., & Thirumalai, V. (2002). Cellular, synaptic and network effects of neuromodulation. Neural Networks, 15(4-6), 479-493.

[3] McCormick, D. A. (1992). Neurotransmitter actions in the thalamus and cerebral cortex. Journal of clinical neurophysiology: official publication of the American Electroencephalographic Society, 9(2), 212-223.

[4] Barbero-Castillo, A., Mateos-Aparicio, P., Dalla Porta, L., Camassa, A., Perez-Mendez, L., & Sanchez-Vives, M. V. (2021). Impact of GABAA and GABAB inhibition on cortical dynamics and perturbational complexity during synchronous and desynchronized states. Journal of Neuroscience, 41(23), 5029-5044.

[5] Dalla Porta, L., Barbero-Castillo, A., Sanchez-Sanchez, J. M., & Sanchez-Vives, M. V. (2023). M-current modulation of cortical slow oscillations: Network dynamics and computational modeling. PLOS Computational Biology, 19(7), e1011246.

[6] Tort-Colet, N., Capone, C., Sanchez-Vives, M. V., & Mattia, M. (2021). Attractor competition enriches cortical dynamics during awakening from anesthesia. Cell Reports, 35(12).

[7] Deco, G., Kringelbach, M. L., Arnatkeviciute, A., Oldham, S., Sabaroedin, K., Rogasch, N. C.& Fornito, A. (2021). Dynamical consequences of regional heterogeneity in the brain's transcriptional landscape. Science Advances, 7(29), eabf4752. (P3132)

[8] Luppi, A. I., Mediano, P. A., Rosas, F. E., Allanson, J., Pickard, J. D., Williams, G. B., & Stamatakis, E. A. (2022). Whole-brain modelling identifies distinct but convergent paths to unconsciousness in anaesthesia and disorders of consciousness. Communications biology, 5(1), 384.

[9] Feng, M., Bandyopadhyay, A., & Mejias, J. F. (2023). Emergence of distributed working memory in a human brain network model. bioRxiv, 2023-01.

[10] Goldman, J. S., Kusch, L., Aquilue, D., Yalçınkaya, B. H., Depannemaecker, D., Ancourt, K., & Destexhe, A. (2023). A comprehensive neural simulation of slow-wave sleep and highly responsive wakefulness dynamics. Frontiers in Computational Neuroscience, 16, 1058957. (P3023)

[11] Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage, 31(3), 968-980.

[12] Schirner, M., Rothmeier, S., Jirsa, V. K., McIntosh, A. R., & Ritter, P. (2015). An automated pipeline for constructing personalized virtual brains from multimodal neuroimaging data. NeuroImage, 117, 343-357.

[13] 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.

[14] Hawrylycz, M. J., Lein, E. S., Guillozet-Bongaarts, A. L., Shen, E. H., Ng, L., Miller, J. A., & Jones, A. R. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. Nature, 489(7416), 391-399.

[15] Arnatkevičiūtė, A., Fulcher, B. D., & Fornito, A. (2019). A practical guide to linking brain-wide gene expression and neuroimaging data. NeuroImage, 189, 353-367.

4. Task-Capable Whole-Brain Model: WP2 Roadmap 2

The project aims to simulate the neuronal dynamics of two auditory discrimination paradigms designed for the macaque cortex. The project considers the dynamics at the whole-brain scale via simulations based on The Virtual Brain (TVB). TVB is a computational framework in which each node models the neuronal populations of a brain region, a connectivity matrix models the connections between these nodes, and neuronal pipelines model the information routing between the nodes, i.e., the brain regions. We use an Adaptive Exponential (AdEx) neuronal population model to describe each node and openly accessible CoCoMac Connectivity Dataset for the connections between the nodes. The information routing is based on a pipeline which starts from the primary auditory cortex (A1), passes through the prefrontal cortex (PFC), and ends in the primary motor cortex (M1).

4.1 Experimental setup

We focus on two tasks. In the first task, the participant is provided sequentially two pitches with two frequencies which can be different or the same. In the second task, these pitches are provided simultaneously. The participant is provided a button. The participant is instructed to push the button to indicate that it perceived the stimuli as different. Otherwise, the participant does not push the button, indicating that it perceived the stimuli as the same; see Figure 8: Scheme of Task 1 and Figure 9: Scheme of Task 2.

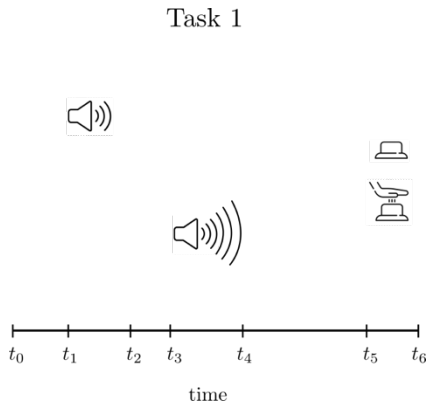

Figure 8: Scheme of Task 1

Stimuli are provided at two different instants. Then, the participant pushes the button or not, depending on if he perceived the stimuli as different or the same.

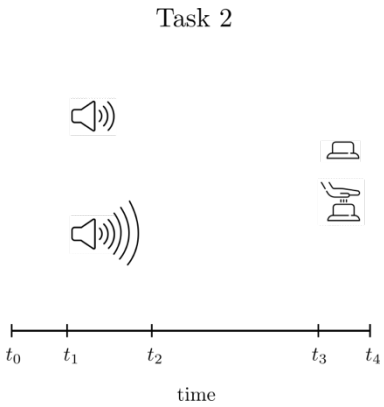

Figure 9: Scheme of Task 2

The task is similar to Task 1, with only difference of that the stimuli are provided simultaneously..

4.2 Whole-brain model

The brain model is designed on TVB. TVB allows us to approximate the neuronal population dynamics via mean-field models. The mean-field model of a neuronal network provides the averaged dynamics of the network at the coarse-grained continuum limit. This decreases the complexity of micro-level mechanisms found in the population, allowing for macro-level simulations of the brain regions. TVB provides such simplification for macro-level simulations by merging the anatomical data from brain imaging data (diffusion MRI) with the mean-field models of neuronal populations. This TVB framework is based on three categories: nodes, connections, information routing.

4.2.1 Nodes

In our TVB framework, each brain region is represented by a node, which is a set of stochastic differential equations describing the dynamics of excitatory and inhibitory neuronal populations, as well as their spike-frequency adaptation. Each node is described by an AdEx mean-field model. The mean-field model approximates the neuronal population behavior modeled by the AdEx network. In the case of the cerebral cortex, AdEx networks are used to model two cell types: Regular Spiking (RS) neurons, displaying spike-frequency adaptation as observed in the pyramidal (excitatory) neurons, and Fast Spiking (FS) neurons, with no adaptation, as observed in the interneurons (inhibitory). Hence, the AdEx networks are biophysically plausible. The AdEx mean-field is low dimensional, simpler and easier to analyze compared to the AdEx networks, yet it approximates closely the network dynamics, motivating our choice of model for each node considered in Task 1 given in Figure 8: Scheme of Task 1.

We extended the AdEx mean-field model used in Task 1 to decision-making related to a visual discrimination task applied on both human and macaque participants. It was based on two AdEx mean-field models, each representing one cortical column of PFC. In that setting, each column votes in favor of one of the two visual alternatives, where the winning vote is determined throughout a competition between the two columns. The competition is induced by intercolumnar excitation which is conflicted by intracolumnar inhibition. The excitation provokes the competition between the two columns. The winning column makes the decision. This competition is biased according to a function with two attractors. Each attractor promotes one of the two decision alternatives. This extended AdEx mean-field model is used to describe the nodes of PFC in Task 2 given in Figure 9: Scheme of Task 2.

4.2.2 Connections

In Task 1, we use the connectome of CoCoMac, which provides the structural connectivity obtained for 84 regions of the macaque brain via diffusion MRI. The connectome provides the long-range excitatory connections between regions. Each brain region is represented by one AdEx mean-field model and the connectivity within each region is intrinsically provided in the AdEx model architecture. In Task 2, we represent the dorsolateral prefrontal cortex (PFCdl) in terms of two AdEx mean-field models, which are connected to each other via long-range excitatory connections. This doubling is required to represent the cortical columns in PFCdl, where each column corresponds to one of the two decisions. Moreover, A1 is organized in terms of columns, similarly to the other sensory cortices. The tonotopic organization of A1 allows for pitch discrimination of the auditory stimuli. The segments of the input pattern are distributed via a gating module to separate the parts of PFCdl in an ordered manner. This results in a possible comparison of these segments with other segments coming from a second pattern, as suggested by functional MRI recordings of auditory recognition. We model this gating by distributing the output of A1 selectively to the two different PFCdl areas, which we name PFCdla and PFCdlb. This allows our model to initiate the discrimination task at A1 level in Task 2, as observed in some auditory discrimination experiments based on a spatial selective detection of auditory stimuli.

4.2.3 Information routing

The main difference of our TVB-AdEx implementation compared to the previous ones is that we introduce a cognitive context in the implementation by considering the whole sensory-motor pathway and by including the modulations of the neuronal dynamics as they propagate from one region to another. This pathway and modulatory mechanisms determine how the information is processed and transmitted from the sensory regions towards the motor regions. We provide the information routing which we used in Task 1 and Task 2 in Figure 10: Information routing for Task 1 and Figure 9: Information routing for Task 2, respectively. We assume hemispherical symmetry and consider only the right hemisphere, which is highlighted by “R” subscript in Figure 10: Information routing for Task 1 and Figure 9: Information routing for Task 2.

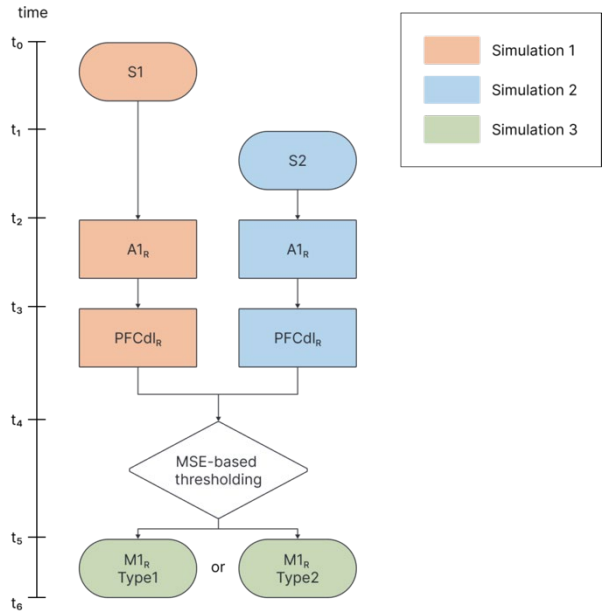

Figure 10: Information routing for Task 1

We stimulate A1-Ra in the first simulation, and A1-Rb in the second. We compare the outputs of these simulations obtained from PFCdl-Ra and PFCdl-Rb via mean square error (MSE). We perform the third simulation, which consists of the stimulation of M1 with a stimulus generated according to the result of the comparison of the PFCdl outputs of the previous step.

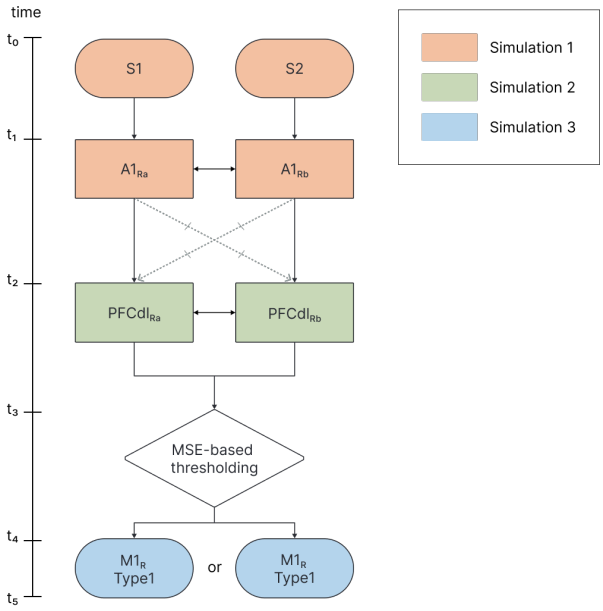

Figure 11: Information routing for Task 2

We stimulate A1 regions with two stimuli simultaneously in the first simulation. We check if the outputs of A1 regions exceed the prefixed threshold or not. In the second simulation, we stimulate the PFCdl regions with strong and/or weak stimuli, depending on the thresholding result regarding the A1 outputs of the first simulation. The outputs of PFC regions are compared via MSE. In the third simulation, we stimulate M1 with strong or weak stimulus, depending on the MSE result.

4.3 Simulation procedure

The simulation procedure for Task 1 can be summarized as follows: 1) Two stimuli are provided to A1 sequentially and independently as shown in Figure 8: Information routing for Task 1. 2) The outputs of A1 are provided as inputs to PFCdl sequentially. 3) The difference between the outputs of PFCdl is measured based on mean squared error (MSE). 4) Depending on whether the MSE value is above or below a prefixed threshold, one of the two actions is chosen. Each action is modeled as a signal corresponding to a different signal strength. 5) The chosen action is transmitted to M1 as the signal with the corresponding signal strength to the chosen action. The simulation procedure for Task 2 can be summarized as follows: 1) Two stimuli are provided to the duplicated A1 simultaneously as shown in Figure 11: Information routing for Task 2. 2) The outputs of the duplicated regions, A1-Ra and A1-Rb are checked if they exceed a prefixed threshold. If any of the output exceeds the threshold, then we provide a strong stimulus as input to the corresponding PFCdl region, otherwise a weak stimulus. 3)-5) are the same as in Task 1.

5. Task-Capable Whole-Brain Model: WP3 Roadmap

The project relies on the Natural Scenes Dataset (NSD), a resource rich in visual stimuli and corresponding brain response data. The project involves building a whole-brain model at a detailed level with each node representing a vertex on an inflated cortex mesh. The model incorporates excitatory-inhibitory circuits to simulate local dynamics and structural connectivity data from diffusion imaging to establish connectivity. A Convolutional Neural Network (CNN) replicates the role of the retina and the lateral geniculate nucleus, extracting salient features from natural images. The task capability of the model is assessed by comparing its internal representations with the empirical data using representational dissimilarity matrices and representational connectivity.

5.1 Empirical data

The project utilizes the **Natural Scenes Dataset (NSD)**, a rich, publicly available fMRI dataset based on natural scene viewing. It contains brain response data from multiple subjects who were presented with thousands of distinct natural images during fMRI scanning sessions. One of the primary features of the NSD dataset is its breadth and depth of visual stimuli, which spans a broad range of natural scenes and objects, offering extensive variability in visual content.

The NSD dataset also provides additional valuable information, such as the exact sequence of images presented and the corresponding timing, which allows for accurate replication of the stimulus presentation protocol in model simulations. The dataset also includes anatomical scans and diffusion imaging and resting-state scans.

The NSD dataset is particularly suitable for this project for several reasons:

Relevance to Visual Processing: The project aims to develop a prototype whole-brain model with a focus on visual processing. The NSD dataset, with its emphasis on natural scene viewing, provides a rich source of visual stimuli and corresponding brain response data, making it highly relevant for this project.

Detailed Connectivity Data: The NSD dataset includes diffusion imaging data that allows the estimation of structural connectivity within the brain. This information is critical for our project as it forms the basis of the connectivity profile in the whole-brain model.

Basis for Model Validation: The dataset provides a basis for comparing and validating the results of the whole-brain model. For instance, the model's simulated BOLD signals can be compared with the empirical BOLD signals in the NSD dataset.

Resource for Optimizing Parameters: Resting-state measurements from the NSD dataset can be used to optimize model parameters to reproduce empirically observed functional connectivity, dynamic functional connectivity, and metastability.

Note that for the prototype, we will limit ourselves to data from a single participant of the NSD dataset.

5.2 Whole-Brain Model

The whole-brain model in this project operates at a detailed level where each node represents a vertex on an inflated cortex mesh, which is a significant enhancement of detail over traditional models that represent regions of interest (ROIs) as single nodes. Each vertex in the model thus represents a small section of the brain's cortical surface. The choice of resolution is crucial for the project because it allows the model to incorporate computations realized through interactions between vertices (neuronal populations) within and between ROIs.

The local dynamics of each vertex, modeled as an excitatory-inhibitory (E-I) circuit, reflect the underlying biological mechanisms that govern neural activity in the brain. By employing a dynamic mean field model, the project leverages a balance between computational tractability and biological realism. The choice of such dynamics is suitable for the project as it allows for the simulation of both resting state and stimulus-evoked brain dynamics.

The connectivity of the model, which includes both intra-ROI and inter-ROI connections, realizes computations in the model. In the long term, the model's connectivity can be optimized (e.g., through gradient-descent) to represent a wide range of brain functions. Initially, the model will focus on visual processing with fixed connectivity. Intra-ROI connections are determined by a combination of structural connectivity data from diffusion imaging and a distance-dependent decay rule. The structural connectivity data provides an empirical basis for the connections, indicating where the most probable pathways exist based on actual brain imaging data. The distance-dependent decay rule then models the natural tendency for connections to be stronger between closer nodes and weaker between more distant ones, reflecting the biological principle of "locality" in neural connectivity. Inter-ROI connections, on the other hand, describe the relationships or

connections between different ROIs. These connections will be entirely based on structural connectivity estimated from diffusion imaging.

Nodes within early visual cortex ROIs of the whole-brain model will additionally receive retinotopically organized connections from the lateral geniculate nucleus (LGN) that will provide features extracted from natural images (see section 5).

5.3 Parameter Optimization

The process of parameter optimization in the whole-brain model revolves around two critical parameters: the global coupling scaling parameter and local feedback inhibition.

1) Global Coupling Scaling Parameter: This parameter scales the structural connectivity. It is tuned to find the working point for which the model reproduces global dynamics exhibited by the cortex. We characterize global cortex dynamics in terms of functional connectivity (FC), dynamic functional connectivity (DFC), and metastability (MS). Functional connectivity is essentially the statistical dependence or correlation between fluctuations in the BOLD signal of different brain regions. Dynamic functional connectivity reflects the dynamics of statistical dependencies as short-lived global network states that may dissolve and re-emerge at different moments in time. Metastability reflects the overall variability of network states of the system; i.e., the extent to which the system exhibits transient synchronization dynamics. The optimization of the global coupling scaling parameter is done to reproduce these characteristics observed in empirical data.
2) Local Feedback Inhibition: This parameter is central to controlling the balance of excitatory and inhibitory dynamics within each node in the model and plays a crucial role in determining the dynamic behavior of the system. To fine-tune this parameter, we employ the Feedback Inhibition Control (FIC) procedure. This involves repeatedly simulating the whole-brain model and adjusting the local feedback inhibition until the input to the excitatory pool is slightly dominated by inhibition. Deviations from the desired input level will be incrementally adjusted using a Hill climbing approach.

The data used for optimization comes from the resting state measurements of the NSD dataset.

5.4 Retina+LGN Model

A shallow convolutional neural network (CNN) emulates the role of the retina and the lateral geniculate nucleus (LGN) in the visual processing pathway. This CNN is designed to extract salient features from natural images and thus replicate the primary stages of visual perception in the human brain.

To train the CNN, a dataset of natural images will be used. The CNN will be pre-trained on classifying these images, allowing it to learn relevant features and patterns inherent in the visual data. The aim here is to approximate the image processing that occurs in the human visual system, from the retina to the LGN. The model employs a retinal sampling layer (RSL), which resamples images according to retinal ganglion cell density, mimicking the non-uniform resolution across the visual field in human vision.

Once trained, the CNN's output will be connected to the whole-brain model through pooling field mapping. This process maps the feature maps produced by the final layer of the CNN to the early visual cortex regions (V1 ROIs) of the whole-brain model. Pooling fields are an extension of population receptive fields. They fit a 2D Gaussian to the CNN's feature maps, effectively determining the spatial extent of the pooling field. In other words, each region in the V1 ROI will be responsive to certain areas of the visual input, similar to how neurons in the early visual cortex respond to stimuli in their receptive field.

The preference for particular features (such as edges or textures) within each region of the V1 will be determined using a ridge regression. This regression is based on responses to natural images presented to the CNN and the participant. By associating network units (CNN) to V1 vertices (whole-

brain model), we effectively instill the model with the ability to represent complex visual features across its nodes.

The use of a CNN in this context is advantageous for several reasons. First, CNNs have proven extraordinarily successful in image classification tasks, capturing complex visual features through their hierarchical structure. This mirrors the hierarchical processing that occurs in the human visual system. Second, by training the CNN on natural images, the model becomes grounded in realistic, ecological stimuli, enhancing the generalizability of the model's visual processing capabilities. Lastly, the coupling of the CNN with the whole-brain model bridges the gap between localized feature extraction and whole-brain dynamics, which is essential for infusing task-capability into whole-brain models.

5.5 Task Capability (Information Representation)

To assess the task-capability of the whole-brain model, we analyze how it represents natural images, which we feed into the model via the Convolutional Neural Network (CNN) retina+LGN model. The key to this assessment lies in comparing the internal representations in the model with the empirical data from the NSD dataset.

We begin by conducting a General Linear Model (GLM), with each image from the NSD dataset serving as a predictor. This model is applied separately for each vertex in the model. The output from this analysis is a set of regression weights (betas) for each voxel and image, providing an image-specific representational signature of responses across vertices.

Following this, we examine how images are represented within each ROI by computing representational dissimilarity matrices (RDMs) for each ROI. These matrices quantify the difference between the representations (beta distributions across vertices) of each pair of images within an ROI, giving us a measure of how dissimilar the images are according to the model's internal representation. We can then compare these RDMs from the model with the empirical RDMs from the NSD dataset.

In addition, we also compute representational connectivity (RC), which refers to the pairwise correlation between RDMs for all pairs of ROIs. This metric provides a measure of the similarity in representational structure across different brain regions. We will compare model and empirical RC matrices.

This detailed analysis of the internal representations within the model, based on empirical fMRI data, allows us to assess the degree to which the model can process and represent natural scenes, which is a significant component of many cognitive tasks. This makes it an appropriate method to evaluate the task-capability of the whole-brain model. It also serves as a stepping stone for further optimizations and improvements, with the ultimate goal of creating a whole-brain model that can faithfully reproduce a wide range of cognitive functions.

6. Brief Discussion and Looking Forward

The presented roadmaps provide solutions to create first prototypes of task-capable whole-brain connectome-based models integrating expertise from WPs 1, 2 and 3. Importantly, work is on target to provide actual implementations of the described prototypes. While all proposed solutions are based on a similar approach to create task-capable whole-brain connectome-based models, concreate realizations are based on WP-related background knowledge and specific aims as detailed in the original task proposal. Interactions between work packages, including a physical meeting in Barcelona, have been instrumental to support each other and to discuss how to integrate the different specific roadmaps further to present at least one running prototype as a joint effort across WPs by the end of September.