amU
Aix Marseille Université

NNT : 2025AIXM0048

THÈSE DE DOCTORAT

Soutenue à AMU — Aix-Marseille Université
le 03 Mars 2025 par

Amirhossein ESMAEILI

Neural bases of others' presence:
A multi-scale probabilistic framework of social facilitation

Discipline
Doctorat en Biologie-Santé

Spécialité
Neurosciences

École doctorale
École Doctorale Sciences de la Vie et de la Santé - ED 62

Laboratoire/Partenaires de recherche
UMR1106INS - Institut de Neurosciences des Systèmes - REPAIRS Réseau européen de formation (ETN)

Composition du jury

Jerome SALLET University of Oxford	Rapporteur
Daniele MARINAZZO Ghent University	Rapporteur
Andrea BROVELLI Aix-Marseille Université	Examinateur
Athena AKRAMI University College London	Examinatrice
Julie GREZES École Normale Supérieure	Présidente du jury
Driss BOUSSAOUD Aix-Marseille Université	Directeur de thèse
Meysam HASHEMI Aix-Marseille Université	Membre invité
Viktor JIRSA Aix-Marseille Université	Membre invité

Affidavit

I, undersigned, Amirhossein Esmaeili, hereby declare that the work presented in this manuscript is my own work, carried out under the scientific supervision of Prof. Driss Boussaoud, in accordance with the principles of honesty, integrity and responsibility inherent to the research mission. The research work and the writing of this manuscript have been carried out in compliance with both the french national charter for Research Integrity and AMU charter on the fight against plagiarism.

This work has not been submitted previously either in this country or in another country in the same or in a similar version to any other examination body.

Marseille, 28-12-2024



This work is licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Public License

List of publications and conference participation

List of publications submitted in the framework of this thesis:

1. Amirhossein Esmaeili et al. “Probing other’s Presence: Probabilistic Inference Across Brain Scales Reveals Enhanced Excitatory Synaptic Efficacy”. In: *bioRxiv* (2024)

Participation to conferences and summer schools during the period of the thesis:

1. REPAIRS closing conference - Presentation - August 2024 - Groningen - Netherlands
2. FENS (Federation of European Neuroscience Societies) Forum 2024 - Presentation - June 2024 - Vienna - Austria
3. BrainHack Marseille - Project Participation - December 2023
4. 3rd Advanced Course & Symposium on Artificial Intelligence & Neuroscience (ACAIN) - Presentation - September 2023 - Grasmere - England - UK
5. Human Brain Project Summit - Presentation - March 2023 - Marseille
6. School of Cells & Circuits Camillo Golgi - Participation - November 2022 - Erice - Italy
7. CENTURI Hackathon 2022 - Project Participation - June 2022 - Marseille
8. Probabilistic Modelling Workshop by Michael Betancourt - Participation - May 2022 - Marseille

Résumé et mots clés

La perception des conspecifics constitue un élément invariable et intégral de la cognition sociale chez les espèces sociales diverses. Bien que les acquis neutrophiliologiques sous-jacents à ces stimuli restent en grande partie mystérieux, sur le plan comportemental, la simple présence d'autres individus semblent moduler la performance tâchiste en fonction de la complexité du tâche. Nous essayons de bridger ce fossé en investiguant comment la présence des autres perturbe l'efficacité synaptique à trois échelles spatiotemporelles du cerveau et comment telles perturbations pourraient conduire à des améliorations de la performance dans les macaques et participants humains. Dans les macaques accomplissant un tâche d'apprentissage d'association, la présence sociale augmente l'éfficacité synaptique excitatoire des régions cérébrales dorsolatérales préfrontales et du cortex antérieur-cingulaire. Dans les humains effectuant une tâche de interception latérale, la présence des conspecifics facilite la performance dans le groupe féminin, et cette facilitation est liée à l'amélioration de l'efficacité synaptique au sein du réseau d'attention dorsale et ventral. Nous posons que les améliorations de la performance due à la présence sont dues à un modération attentionnelle médianée par des changements dans l'efficacité synaptique excitatoire à trois échelles spatiotemporelles, à savoir microéchelle (uniques neurones), macroéchelle (colonnes corticales) et macroéchelle (cerveau entier). Nos insights du raisonnement probabiliste s'accumulent pour proposer un cadre frais et multiéchelle des bases neuronaux de la présence, ouvrant ainsi la voie à future recherches dans l'intérêt complexe entre la simple présence, le comportement et les dynamiques neuronales.

Mots clés: Effets de facilitation et d'impédiment social, Modulation attentionnelle, Inférence probabiliste, Modélisation multiscale, Estimateurs de densité neuronale profonde, Efficacité synaptique, Connexion efficace

Abstract and keywords

Perception of conspecifics is an invariant and integral prerequisite of social cognition in myriad social species. While the neurobiological underpinnings of such stimuli remain largely enigmatic, on a behavioral scale, mere presence of others reportedly modulates task performance based on task complexity. We attempt to bridge this gap by investigating how others' presence perturbs synaptic efficacy across three spatiotemporal brain scales, and how such perturbations could conduce to improvements of task performance in Rhesus monkeys and human participants. In monkeys performing an association learning task, social presence elevated the excitatory synaptic efficacy of attention-oriented regions dorsolateral prefrontal cortex and anterior cingulate cortex. In humans performing a lateral interception task, conspecific presence facilitated performance in the females subject group, and this facilitation was linked to enhancements of synaptic efficacy within the Dorsal and Ventral Attention Networks. We posit that presence-induced enhancements to performance stem from attentional modulation mediated by changes in excitatory synaptic efficacy across spatiotemporal brain scales, namely microscale (single neurons), mesoscale (cortical micro-columns) and macroscale (whole-brain). Our insights from probabilistic inference coalesce to propose a fresh, multi-scale framework on the neural bases of others' presence, paving the path for future forays into the intricate interplay between mere presence, behavior, and neural dynamics.

Keywords: Social facilitation and impairment effects, Attentional modulation, Probabilistic inference, Multiscale modeling, Deep neural density estimators, Synaptic efficacy, Effective connectivity

Acknowledgments

I find myself indebted to numerous individuals for their contribution to this project, as an interdisciplinary undertaking of this magnitude would not have been possible without each and every one of them.

First and foremost, I must acknowledge the superbly gifted supervisory pillars of this work, Driss Boussaoud, Meysam Hashemi, and Viktor Jirsa. As with all else, it is best to start with the genesis and inspiration of this entire project. Driss was there through the thick and thin of it all, from the first interview, until the last supervisory meeting, and has laid the empirical framework for this here work. Without his expertise on social cognition, his help in designing the lateral-interception task, and his ever-present support this project would not have even passed its nascency , much less come to fruition. I have been incredibly fortuitous to have such a bastion of 'facilitation', and hope that our collaboration has moved us one step closer towards our shared aim of unraveling the social brain.

I am equally at a loss on how to begin to thank Meysam for his role as a theoretical advisor; The first year of my PhD was particularly challenging, not only from a personal stance, but also with regards to my progress towards the aims of this project, as I was still conducting exploratory analysis of the data. Enter Meysam, and his brilliant suggestion to redirect my efforts towards a DCM approach. It is therefore with utmost candor that I say this project would not have reached its current potential, had it not been for his invaluable inputs and contributions. Thanks for everything dude, it was a blast, and a 'highly informative' one at that!

Viktor constitutes the last pillar, whom I must thank for his guidance and support throughout the project. As if warping the fabric of space-time itself, he was always there when we critically required his insight and expertise, despite his schedule being completely overloaded by numerous large-scale projects. His presentations and written works were always a source of inspiration which propelled me to improve upon my own. I am also immensely thankful for structuring this rare and positive 'flow' within the lab, as I found some of my best friends here.

Moving on to other colleagues and mentors, I must offer my heartfelt gratitude to my exceptional colleagues and mentors whom shouldered the brunt of the experimental design and/or data acquisition efforts of the two tasks presented in this work.

A most heartfelt thanks to Marie Demolliens, Faical Isbaine, and Pascal Huguet, for their help with the association learning task. Toiling tirelessly for years on end, Marie and Faical shouldered the entirety of stereotactic surgeries, animal training and data acquisition. In fact, the discovery of context-oriented neurons which laid the experimental foundation of this project was made possible through their perseverance and insight. I must also thank Pascal for his expertise and input which were indispensable

in the design process of our social facilitation tasks.

The lateral interception task was similiarly carried through the exemplary effort of Frank Zaal and Marjolein Viersen. The majority of the code-base for the behavioral aspect of this task came from Frank's previous work, the absence of which would have significantly diminished the prospects of this project. And almost the entirety of experimental data acquisition in this task were conducted by Marjolein, to whom I owe a big beer at some point!

A thanks of equal enormity goes to Abolfazl Ziaeemehr, who was not only instrumental for programming and optimization of the coupled Jansen-Rit model, but also made for incredibly fun and helpful discussions regarding various challenges within computational neuroscience. I am also deeply grateful to Spase Petkoski, Michael 'Marmaduke' Woodman and Andrea Brovelli as their input has been tremendously formative for optimizing the whole-brain inference process. Among my colleagues, I must not forget to thank Paul Triebkorn, Vinicius Lima, and Borana Dollamaja for having the patience to answer my innumerable inquiries with regards to the project.

My efforts in this project were entirely funded through the REPAIRS consortium, and by extension, the European Union's Horizon 2020 research and innovation programme Marie Skłodowska-Curie grant agreement No 956003, without which none of this would be possible. Throughout this project, I had the pleasure and privilege of being exposed to numerous scientific environments and learned from some of the leading experts in computational neuroscience, probabilistic inference, and social cognition, none of which would have been possible without REPAIRS. On a similar note, I want to thank the crew at CODEMART, whom I've learned a great deal from during my latest secondment; Thanks to Lia and Theodora for making time to chat with me about the code, despite their busy workloads and my limited knowledge, to Mihai for the philosophical discussions during breaks, to David for letting me invade his workspace during those two weeks, and to Catalin for headbanging with us in Rockstadt!

Then there are those that genuinely carried me through these three years; Cheers to Taso, Camille, Meysam, and Nariman, whose company is akin to having a second family. May our journeys be as awesome as ever! A second cheers goes to my newly-wed friends Giovanna and Manef for the occasional but much-needed coffee breaks!

Speaking of family, I would like to give endless thanks to my mother, whose passion for science kick-started my entire career, and I whose support I would not be here without.

A final thanks to you dear reader, for having the patience to peruse the endeavor depicted here.

Contents

Affidavit 2

List of publications and conference participation 3

Résumé et mots clés 4

Abstract and keywords 5

Acknowledgments 6

Contents 8

List of Figures 10

List of Tables 12

Glossary 13

1. Introduction 17
- 1.1. What is social facilitation? . . . 18
  - 1.1.1. Cognitive frameworks of SFI . . . 20
  - 1.1.2. Neural Correlates of mere presence . . . 21
- 1.2. A Probabilistic portrait of SFI . . . 25
  - 1.2.1. Bayesian representation of neural systems . . . 25
  - 1.2.2. Causal perspective of attentional modulation . . . 28
  - 1.2.3. Sampling from posterior distributions . . . 32
  - 1.2.4. Amortized probabilistic inference . . . 33
  - 1.2.5. Low-dimensional subspaces of neural systems . . . 35

2. Experimental evidence for facilitation 37
- 2.1. Association Learning task . . . 38
  - 2.1.1. Subjects . . . 38
  - 2.1.2. Behavioral task . . . 38
  - 2.1.3. Electrophysiological Recordings . . . 40
  - 2.1.4. Previous results . . . 41
  - 2.1.5. Neural firing rates . . . 43
  - 2.1.6. Event-related potentials . . . 45
- 2.2. Lateral interception task . . . 48
  - 2.2.1. Subjects . . . 49

2.2.2. Experimental setup . . . 49
2.2.3. Movement data . . . 51
2.2.4. Electrophysiological recordings . . . 52
2.2.5. Functional connectivity . . . 56
2.2.6. Measures of integration . . . 57

3. Data-generative processes **60**
3.1. Brunel balanced network . . . 61
3.2. Jansen-Rit neural mass model . . . 66
3.3. Whole-brain Jansen-Rit network . . . 69

4. Probabilistic inference across scales **73**
4.1. Microscale inference . . . 74
4.2. Mesoscale inference . . . 77
4.2.1. HMC validation . . . 80
4.3. Macroscale inference . . . 83

5. Discussion **87**
5.1. Summary . . . 87
5.2. Arousal versus attentional views of SFI effects . . . 89
5.3. Homeostatic framework of SE within the social brain . . . 89
5.4. Limitations and future directions . . . 90

Conclusion **92**

Bibliography **93**

Appendix **110**
A. Validation of single-neuron clustering . . . 111
B. Robustness of ERP amplitude as a feature . . . 112
C. Degeneracy of a full Jansen-Rit NMM . . . 114
D. Optimization of external input in coupled JR NMMs . . . 115
E. Computation of source signal-to-noise ratio . . . 116
F. Sensitivity analysis of generative models across brain scales . . . 117
G. Hypothesis testing of empirically-informed inference . . . 118

List of Figures

1.1. Meta-analysis of social facilitation effects in the Stroop task 19
1.2. Allocentric neural representation 22
1.3. Schematic of select discoveries in social cognition 23
1.4. Neurons as dynamical systems 25
1.5. Dynamical systems in neuroscience 26
1.6. Forward and inverse problems of DCM 30
1.7. Neural representations of conspecific social characteristics 31
1.8. General outline of a simulation-based inference workflow 35
1.9. Attractor manifolds of neural activity 36

2.1. Association-learning task outline 39
2.2. Recorded cortical sites 41
2.3. Differences of behavioral and neural patterns in presence versus absence. 42
2.4. Social and asocial neurons 43
2.5. Presence-induced firing rate patterns 44
2.6. Characteristics empirical evoked potentials across conditions 47
2.7. Lateral interception task outline 48
2.8. Task block design 50
2.9. Paddle movement data analysis 52
2.10. Eye movement independent component templates 53
2.11. Power spectra of participants 55
2.12. EEG source-localization steps 56
2.13. Functional connectivity metrics of lateral interception 57
2.14. Relationship between integration and task performance 58

3.1. Schema of common models in computational neuroscience 62
3.2. Brunel spiking network 63
3.3. Jansen-Rit neural mass model 67
3.4. Whole-brain JR network 69

4.1. Microscale summary statistics 74
4.2. Synthetic validation of microscale inference 75
4.3. Microscale inference of synaptic efficacy 76
4.4. Synthetic validation of macroscale inference 77
4.5. Posterior predictive check of mesoscale inference 78
4.6. Distribution of mesoscale synaptic efficacies 79
4.7. Validation of Hamiltonian Monte-Carlo 80
4.8. HMC validation of mesoscale SBI 81

4.9. Integration across functional networks 83
4.10. Validation of macroscale inference 84
4.11. Distribution of SE across macroscale networks 85
4.12. Integration versus synaptic efficacy heatmaps 86
4.13. Correlation of synaptic efficacy with task performance 86

A1. Validation of single-neuron clustering 111
A2. Analysis of ERP amplitude as a data feature 113
A3. Convergence and degeneracy of full-parameter HMC 114
A4. Correlation of the JR parameters 115
A5. Macroscale external current optimization 116
A6. Signal-to-noise ratio of source activity 116
A7. Sensitivity analysis of multiscale brain models 117
A8. Significance testing of empirically-informed synaptic efficacies 118

List of Tables

3.1. Parameters for LIF neurons. 65
3.2. Parameters for ERP model. 68
3.3. Parameters for whole-brain EEG model. 71

4.1. Prior range for ERP model parameters. 77
4.2. Prior distribution of synaptic efficacies in HMC. 82

Glossary

ABC Approximate Bayesian computation. A statistical method approximating a posterior distribution without direct sampling. 33

ACC Anterior cingulate cortex. 40

ASD Autism spectrum disorder. A neurodevelopmental disorder characterized by persistent deficits in social communication and social interaction. 89

BAP Ball arrival position. The horizontal position of the virtual ball on the interception axis at the end of the trial. 50

BDP Ball departure position. The horizontal position from which the virtual ball initiates its trajectory. 50

BNN Bayesian neural network. An extension of traditional neural networks that incorporates Bayesian inference to estimate model parameters and make predictions. 47

CA1 Cornu Ammonis area 1. A region within hippocampus proper. 23

conspecifics Of, or belonging to the same species. 18

DAN Dorsal attention network. 54

DCM Dynamical causal modeling. A statistical framework for estimating causal relationships between variables in complex systems. 28

dlPFC Dorsolateral prefrontal cortex. 40

DMN Default mode network. 54

DST Dynamical systems theory. The study of complex systems that evolve over time, often using mathematical models to analyze and understand their behavior. 26

DTI Diffusion tensor imaging. An MRI technique that visualizes the microstructure of tissues by measuring the diffusion of water molecules in different directions. 29, 70

E/I Excitation/Inhibition. 31

EEG Electroencephalography. A non-invasive electrophysiological technique that records the electrical activity of the brain through electrodes placed on the scalp. 24

ERP Event-related potential. A short-term electrophysiological response to a specific sensory or cognitive event. 45

FC Functional connectivity. Coherence of time series between regions/nodes. 56

FIR Finite impulse response. 53

fMRI Functional magnetic resonance imaging. A neuroimaging technique that measures changes in blood flow and oxygenation in the brain. 24

FPN Frontoparietal network. 54

GCPR G protein-coupled receptor. A type of membrane-bound receptor that responds to ligands by activating a G protein complex, initiating downstream signaling pathways involved in various cellular processes. 90

HMC Hamiltonian Monte Carlo. A variant of MCMC that uses the Hamiltonian dynamics of a system to generate proposals for posterior samples, rather than relying on random-walk. 32

ICA Independent component analysis. A statistical method for separating mixed signals into independent source components by analyzing the statistical independence of their features. 53

IN Interneuron. A type of –typically– inhibitory cell in the brain that is integral in shaping and refining information processing within cortical circuits. 67

JR Jansen-Rit. A neural mass model capable of simulation of evoked and tonic activity. 67

KS Kolmogorov-Smirnov. A non-parametric statistical measure of distance between two distributions. 79

LDS Linear dynamical system. A dynamical model characterized by linear state equations. 44

LFP Local field potential. A measurable electrical signal generated by the collective activity of neurons within a specific region of the brain, serving as an indicator of neural population dynamics and cortical function. 40

LIF Leaky integrate-and-fire. A type of neural model with a leak current, capable of simulating spiking activity. 61

LIM Limbic network. 54

MCMC Markov chain Monte Carlo. A computational algorithm for approximating Bayesian inference and sampling from probability distributions. 32

MEG Magnetoencephalography. A non-invasive neuroimaging technique that detects magnetic fields generated by the electrical activity of the brain. 24

NMM Neural mass model. A mathematical representation of a cortical region. 66

NSF Neural spline flow. A type of neural network architecture that uses splines to model complex transformations through continuous, piecewise-defined mappings. 34

NUTS No U-turn sampler. An adaptive proposal distribution used in Markov Chain Monte Carlo (MCMC) methods to propose new samples, particularly useful in high-dimensional parameter spaces or complex posterior distributions. 80

ODE Ordinary differential equation. Mathematical equations involving a dependent variable and its derivatives with respect to a single independent variable. 66

PN Pyramidal neuron. The most abundant type of neuron in the cerebral cortex, characterized by a triangular cell body. These excitatory cells are play a pivotal role in cortical information processing and transmission. 67

PSP Postsynaptic potential. A change in the membrane potential of a postsynaptic neuron in response to synaptic input from another neuron. 45

PSTH Peristimulus time histogram. Graphical representation of the distribution of neuronal responses over time following a specific stimulus presentation. 40

ReLU Rectified linear unit. An activation function commonly used in neural networks, particularly in deep learning. 34

SBI Simulation-based inference. A class of statistical methods that leverage computer simulations to make inferences about the parameters of a model when traditional likelihood-based methods are intractable. 33

SC Stellate cell. An excitatory 'interneuron' characterized by an star-shaped soma. Stellate cells are said to be typically involved in regulation of activity of other neurons. 67

SE Synaptic efficacy. The efficacy of synaptic communication between neural elements. 31

SEM Structure equation modeling. A statistical approach to testing hypotheses about the relationships among latent and observed variables. 28

SFI Social facilitation/inhibition. Modulations of task performance under mere presence of conspecifics. 18

SFM Structured flow on manifolds. A mathematical object describing the evolution of systems via flows which span the space of all state variables. 35

SMN Somatomotor network. 54

SNR Signal-to-noise ratio. A measure of the quality of a signal by comparing its amplitude to the level of background noise or interference present in the signal. 58

SSM State space modeling. A class of statistical models that describe systems as a combination of underlying state variables and observation equations. 45

StC Structural connectivity. The physical connections between different brain regions which can be mapped using techniques such as DTI. 70

TTC Trials-to-criterion. Threshold metric for learning speed. 41

VAN Ventral attention network. 54

VIS Visual network. 54

WS Wasserstein. A non-parametric statistical measure of distance between two distributions that is particularly useful in high-dimensional spaces. 81

1. Introduction

Summary

1.1. What is social facilitation? 18
1.1.1. Cognitive frameworks of SFI 20
1.1.2. Neural Correlates of mere presence 21
1.2. A Probabilistic portrait of SFI 25
1.2.1. Bayesian representation of neural systems 25
1.2.2. Causal perspective of attentional modulation 28
1.2.3. Sampling from posterior distributions 32
1.2.4. Amortized probabilistic inference 33
1.2.5. Low-dimensional subspaces of neural systems 35

1.1. What is social facilitation?

Imagine a lone researcher working through a manuscript at his office after hours when he hears a knocking noise, and interprets the source of this sensation as the presence of another person. As most of us have experienced a similar version of this heuristic before, sudden awareness of another's presence brings about a very fast –albeit potentially subtle– transition in our behavioral and cognitive states. The ubiquitous nature of this heuristic in humans –and other social species–, hints at shared neural mechanisms for registration of others' presence. Indeed, the study of how the presence of conspecifics (i.e. members of the same species) influences our behavior has been a cornerstone of social psychology for over a century [202].

A well-documented subset of these effects known as "social facilitation" describe the phenomenon where participants tend to perform better on simple or well-learned tasks in the presence of others compared to when they are alone. Despite being the topic of ongoing debates, this enhancement in performance might stem from the awareness of being observed by an audience [15]. The historical roots of social facilitation research can be traced back to the late 19th century with Norman Triplett's classic study on cyclists, who demonstrated that riders performed faster when competing against each other than when cycling alone [183]. While the findings of this study itself are now shown to be not as significant as reported [183], social facilitation effects have been found in a variety of motor-oriented tasks since then [18, 167]. Moreover, these effects reportedly extend beyond simple motor performances, and have been made across various cognitive tasks and species [44, 89]. It is crucial to reiterate that this effect is most pronounced on tasks that are already familiar and well-rehearsed [15]. Social inhibition on the other hand, involves the decrease in performance due to increased anxiety and self-consciousness, and is commonly found in complex or novel tasks [126]. Moving forward, we will use the term social facilitation/inhibition (SFI) when referring to effects which rely primarily on the registration of others' presence.

A complete understanding of mere presence effects is integral in SFI investigations. Mere presence effects encompass any change in behavior driven by others' presence, regardless of the directionality of the behavioral modulations; An undirected modulation of behavior –whether facilitative or inhibitory– due to the simple fact that someone else occupies the same space. We can therefore categorize mere presence effects as the superset of social facilitation, with the latter incorporating the subdomain where others presence leads to a tangible improvement to behavioral responses. In contrast with facilitation studies, investigations of mere presence can be of an exploratory nature, in order to break new ground regarding the gradient of performance changes as a result of conspecific presence.

Therefore, findings of this agnostic perspective can be used to further enhance facilitation studies. For instance, a study on human participants found that cortisol levels increased only in conditions involving explicit negative social evaluation, rather than mere social presence, which suggests that social-evaluative threats –and not just the presence of others– is responsible for stress-related physiological changes [46].

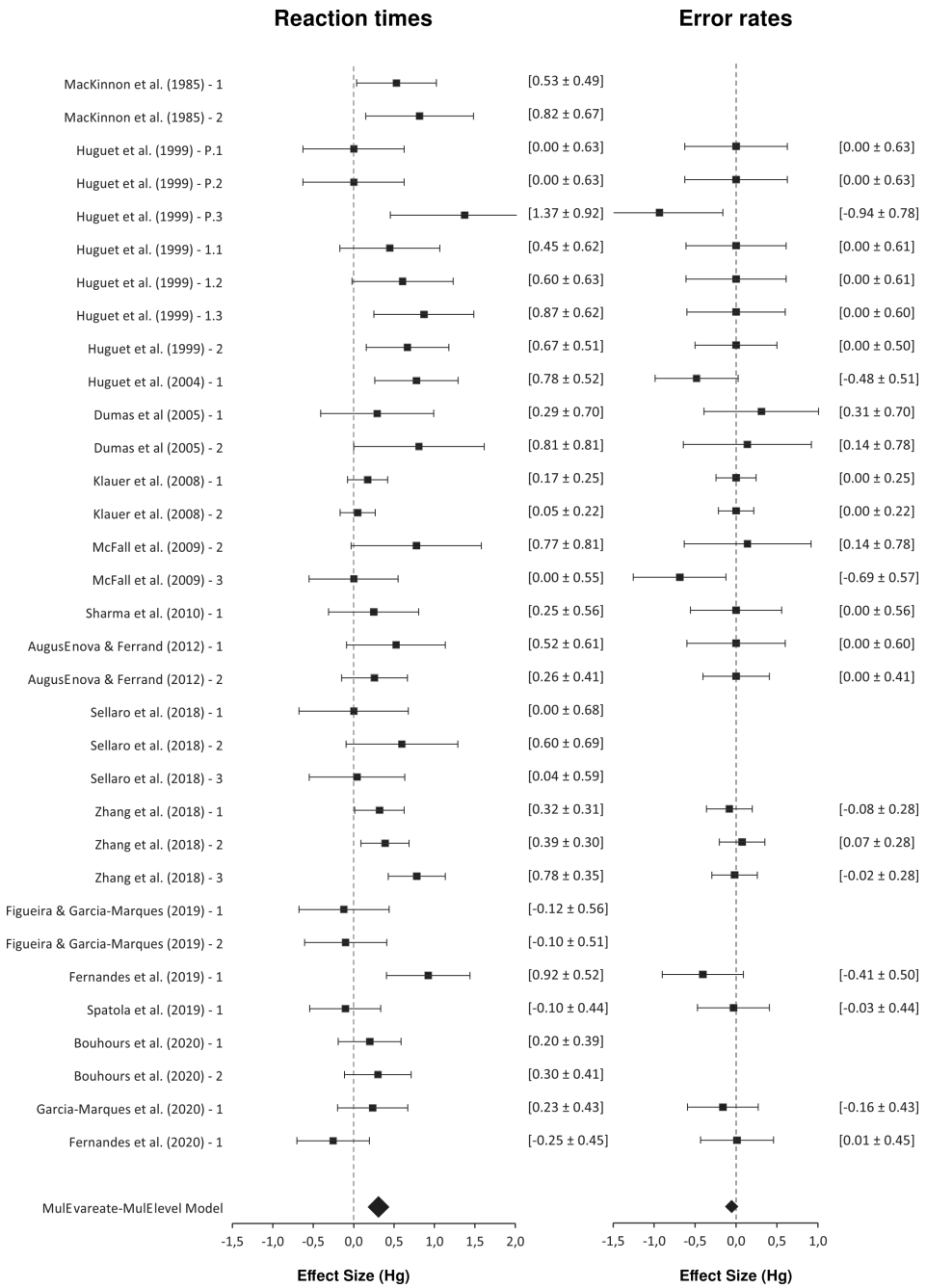

Figure 1.1.: Meta-analysis of social facilitation effects in the Stroop task. Presence of others reduces the magnitude of Stroop effects (Stroop interference), with a stable effect size across numerous studies. Figure adapted from [68].

This finding was further resolved in a recent meta-study, in which attentiveness of spectators was found to heavily influence the strength of facilitative effects; The strongest effects were observed with an attentive audience, which were reduced by mitigation of said attentiveness, and completely nullified in the presence of an

evaluative audience [68]. Insights such as these are usually used to create a more robust set of experimental conditions for social facilitation.

Measurements of mere presence effects can be influenced by myriad factors. Cultural norms can significantly impact social behavior [27], and might stem from gene-culture coevolution, which biases individuals towards traits which enhance social coordination [34]. On an individual scale, factors such as anxiety can significantly affect one's behavioral output [37]. It is therefore paramount to carefully control over both intra-individual (i.e. social) and inter-individual task parameters in an SFI experiment. While this might prove to be a difficult endeavor, rigorous investigations of SFI effects can have profound implications across a wide array of fields such as academia, sports, and industry. In the following, we will delve deeper into the complex portraits of SFI effects by contrasting its proposed cognitive mechanisms, and investigating its neural underpinnings.

1.1.1. Cognitive frameworks of SFI

Which cognitive phenomena contribute to the emergence of the aforementioned performance modulations within SFI? Two main schools of thought dominate this discourse, namely the motivational and attentional framework [15].

The motivational theories emphasize the role of social rewards in driving performance. They posit that the desire for approval, recognition, or competition can motivate individuals to exert greater effort when performing in front of others. This perspective suggests that the presence of an audience triggers a complex interplay of internal and external rewards, influencing our behavior towards achieving desired outcomes. Some have been able to find a link between motivational reactivity and social facilitation of well-learned behaviors [175]. These findings have since been subjected to heavy criticism, due to the use of skin conductance as a metric for motivation; a metric which has been strongly rebuked as a genuine measure for motivation, given our current understanding of the reward system [71, 49]. Notwithstanding the relatively sparse –and quite dated– body of work regarding the motivational framework, later investigations have also struggled to provide sufficient evidence in support of its validity. For instance, research revolving around the motivational impact of self-presentation on SFI found that it alone could not account for the entirety of observations [159]. Putting aside all of the above factors, proponents of this framework should contend with the fact that the motivational framework inherently relies on audience's evaluation; a factor that –as previously mentioned–, has been shown to be heavily involved in nullifying any facilitation effects [68].

The attentional framework posits that the presence of conspecifics induces attentional modulation in the participants, which in turn impacts task performance [89]. Imagine a musician practicing scales before an audience. In this example, the awareness of being observed potentially intensifies their concentration, subsequently leading to a narrowing of their focus that enhances their performance. The distraction-conflict theory proposed by Baron constitutes the first attentional perspective of SFI effects, which posits that the distraction induced by social presence leads to diversion

of attention from the task, thereby generating attentional conflict [9]. Attentional conflict refers to the tug-of-ward between stimuli for the foci of attention (i.e. the task or the audience in our case). Said conflict –and the potential cognitive overload– is then resolved via restriction of the organism's foci of attention [15]. Interestingly, this restriction could generate a similar set of task outcomes which would have been classically categorized via Zajonc's 'dominant response energization' framework (performance facilitation in simple tasks and inhibition in complex tasks). The ambiguity of whether to interpret the results under attentional modulation or dominant responses has since been fortuitously resolved by research revolving around poorly-learned task which involved a limited set of stimuli. In this context, the former hypothesis predicts facilitation of responses where as the latter predicts impairment [15].

The Stroop task is an excellent candidate for settling this. Participants are tasked with identifying the color of words and control cues. This task takes it's name from Stroop interference, the effect in which color identification takes longer for incongruent words (e.g. the word 'blue' printed in yellow) than the control signs (e.g. '+++' printed in any color) or congruent words (e.g. the word "blue" printed in blue) [118]. Interferences induced in this task are stems from the conflict of an intended and controlled cognitive process (i.e. color naming), with an unintended, dominant, and automatic one (i.e. word reading, [68, 15]). Zajonc's solution forecasts an elevation of Stroop interference during social presence, as word reading represents the dominant tendency in this task. On the other hand, the proposed restriction of attentional focus should reduce this interference by enabling the participants to better focus on letter-color stimuli. Huguet and colleagues indeed found that social presence significantly decreases Stroop interference [89]. Said reduction was has been successfully replicated in the mere presence of behaviorally unpredictable humanoid androids [178]. Moreover, research has found motivation to be lacking in prevention of Stroop interference, which further supports an attentional basis for SFI effects than a motivational one [88, 118]. A large body of work has since been dedicated to clarification and replication of said presence-induce interference reduction [68, 15, 4, 169, 109, 48].

The modern take of SFI effects revolves around modulations of cognitive control; Attentional modulation underlying SFI effects is said to be dependent on the timescale of responses, in which only responses with long reaction time lead to the reduction of Stroop interference [169], as opposed to the timescale-independent prediction from Baron's perspective. The human brain is a complex tapestry, woven as a result of innumerable evolutionary pressures. It should therefore be noted that mere presence effects could be indeed partially driven by factors other that attentional modulation. Nonetheless, the attentional framework lays the foundation of the endeavor outlined here, due to the vast amount of previous evidence which corroborating its involvement in social facilitation [68, 15].

1.1.2. Neural Correlates of mere presence

Cognitive frameworks are essential in conducting a thorough analysis of social facilitation. What they are unable to do by themselves, is providing a sufficiently granular

lens for capturing the potential neurobiological underpinnings of social presence by themselves, particularly on a neuronal/synaptic level. This presents a fascinating frontier for research, as deciphering how specific neuronal characteristics and connections give rise to these behavioral changes can provide invaluable insights into how our –largely social– brain operates in different contexts [52]. One can reasonably argue that social facilitation should logically –and inherently– be driven by a set of 'social' stimuli; a set which can be reasonably assumed to intertwine with other facets of social cognition, and has long been posited to recruit disparate neural subsystems than non-social stimuli. As research surrounding the neural correlates of SFI effects is rather sparse, I will provide a perspective on a selection of relevant literate that either explicitly revolve around social facilitation, or have strong implications for it.

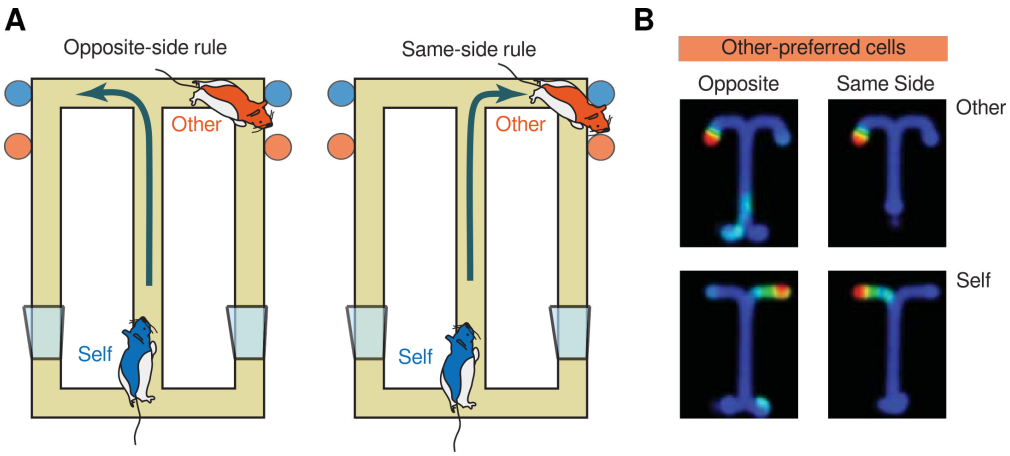

Figure 1.2.: Allocentric neural representation. **A**) Outline of the T-maze task, with two experimental conditions; A 'same-side' condition, during which the self rat (i.e. the recorded animal, in blue) had to choose the same side as the other rat (in orange), and an "opposite-side" condition, in which the self rat had to choose the opposite side as the other rat. **B**) Rate map of other-preferred cells. Note the specific firing pattern of these cells, in which the maximal firing rate is observed only when the self rat is on the same side as the other rat. Figure adapted from [40].

The majority of relevant evidence for the existence of said correlates is at the whole-brain scale (i.e. macroscale); For instance, research has shown that brain responses are increased during live action and face observations, as opposed to virtual ones [36]. On cognitive frameworks, the brain structures involved in automatic attentional orienting are reportedly different with regards to non-social versus social stimuli. In other words, subcortical mechanisms are reportedly recruited for non-social orienting, while predominantly cortical mechanisms are recruited for social orienting [79]. But which regions are better tuned towards social stimuli? Some have proposed that the activity of the attention oriented frontoparietal network is associated with presence-induced improvements in behavior. Intriguingly, this facilitation was not associated

with motivation-oriented modulations of brain activity or stress levels [132]. Some recent work has even been able to isolate 'presence' stimuli at the microscale (i.e. single neurons). For one, hippocampal area CA1 has been demonstrated to encode the spatial representation of conspecifics in rodents [40], Figure 1.2.

These spatial allocentric maps are of course, not tuned specifically to their sheer presence. And while further research into these cells might provide invaluable insight into the fundamental building blocks of social cognition, not every finding within rodents can be immediately translated to primate social cognition. In fact, there are but only few who have attempted to uncover the neural underpinnings of mere presence in higher primates. One of the only lines of microscale evidence with regards to social facilitation comes from the work of Demolliens and colleagues. In their study, the authors were able to find neural subpopulations within the brain of Rhesus monkeys, which are preferentially tuned to the presence or absence of conspecifics [44]. These context-specific neural subpopulations lay the foundation for the microscale and mesoscale components of this work, which we will thoroughly explore in the following chapter. The above studies –while disparate in certain aspects-, provide strong evidence for uniqueness of neural activity encoded by social networks against that of non-social networks [24] Figure 1.3.

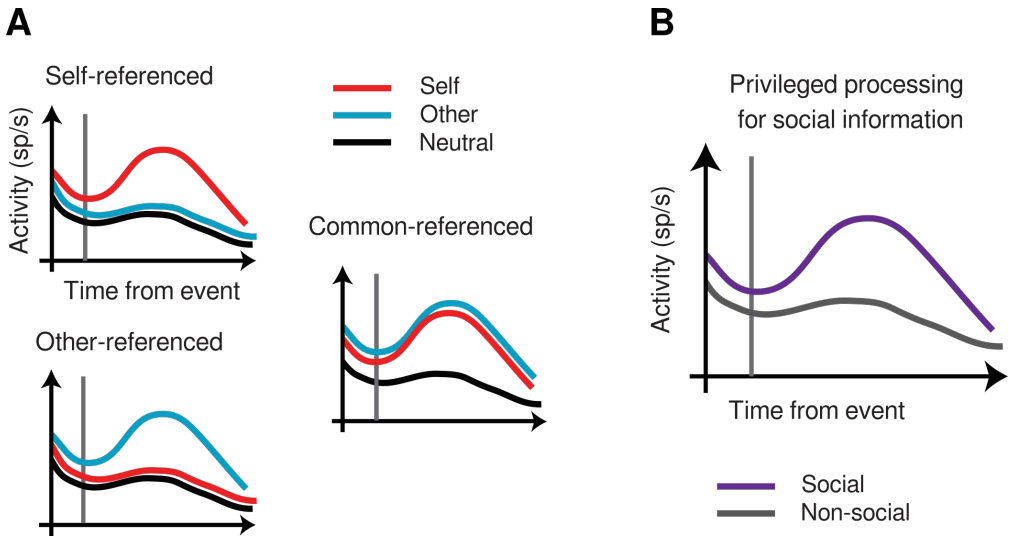

Figure 1.3.: Schematic of select neural principles obtained from primate social neurophysiology research. **A**) Three neuron tuned to disparate task events in referenced to self, others, or both (neutral). **B**) Example of the activity of a neuron which is preferentially modulated towards certain social stimuli (e.g. faces, eyes). Figure adapted from [24].

Notice that each of the above perspectives provide an invaluable but incomplete view of how social facilitation is represented in the brain: Invasive studies are capable of generating neurobiologically-grounded hypotheses, but are inherently limited in number of the explored brain regions, as conducting in-vivo experiments (e.g.

those involving non-human primates and rodents) simultaneously across the entire brain would require an impossibly large array of electrodes. Conversely, whole-brain investigations address global changes to brain dynamics –or their correlates–, but are unable to propose reasonable biological mechanisms that underpin SFI, as the spatiotemporal resolution of current neuroimaging techniques in humans, (e.g. fMRI, EEG/MEG), is not suited for inferring single-unit activity from any brain region.

These constraints elucidate the prohibitive nature of designing experiments in order to test the social brain hypothesis in its entirety. Consequently, there are two facets of neural data when it comes to social cognition. Granular, single-neuron recordings from a subset of –all potentially involved– regions, and whole-scalp non-invasive activity (which does not linearly correlate to cortical activity before rigorous source-localization). Experimental forays into social neuroscience are thus forced to rely on indirect or incomplete measures in order to make progress towards understanding the neural correlates of social facilitation and mere presence effects. As we will delve into in the following chapters however, I will outline how we can overcome said limitations and bridge the gap between the two facets of neural data via model-based approaches, in order to conduct a more comprehensive and unified investigation of social facilitation effects.

1.2. A Probabilistic portrait of SFI

1.2.1. Bayesian representation of neural systems

The human brain is an intricate tapestry of interconnected elements, with its emergent properties rarely originating from a mere sum of their constituents, rather from complex feedback loops and interactions where small changes in one element can cascade throughout the network, leading to dramatic shifts in the system's output.

This presents formidable challenges for understanding neural functions, particularly in the realm of social cognition. Moreover, the dichotomy of granular single-neuron recordings in a handful of regions, versus 'coarser' whole-brain recordings necessitates innovative methodologies that transcend the constraints of experimental paradigms and bridge this gap in our understanding. In this section, I will provide the framework through which we will attempt to surmount these constraints.



Figure 1.4.: Neurons as dynamical systems. Figure adapted from [94].

Neural systems are time-dependent in the sense that the behavior of the system in the next moment in time typically depends –among others– on the input it receives and its state in the previous moment(s) of time. We can thus mathematically represent certain aspects of them as dynamical systems, by defining a set of variables which de-

scribe the state of the neural system, and principles that outline how the dynamics of those state variables unfold (i.e. evolve) through time [94]. Dynamical systems theory (DST) is a formidable mathematical tool which allows us to address the abovementioned constraints, particularly when dealing with the immense nonlinearity of neural systems. In the nomenclature of DST, a snapshot of the system's current configuration (i.e. collection of all state variables) at a specific moment is defined as a 'state', with state variables representing a measurable aspect of a particular neural element at a given time [102]. Examples of state variables encompass various properties of a neural element, such as membrane potential, firing rate of a neuron/region, or post-synaptic potentials. The evolution of these variables through time are governed by differential equations; In the most reduced sense, one can mathematically define the state of the desired system at any given time as a vector $\vec{x}(t)$, where each element corresponds to the activity level of a specific neuron, neuronal population, or an entire cortical region. This state vector's temporal dynamics is described by a set of coupled differential equations $\frac{d\vec{x}}{dt} = F(\vec{x}, t)$ where F represents the system's dynamics, encapsulating the –nonlinear– interactions between different components and the influence of external stimuli [102].

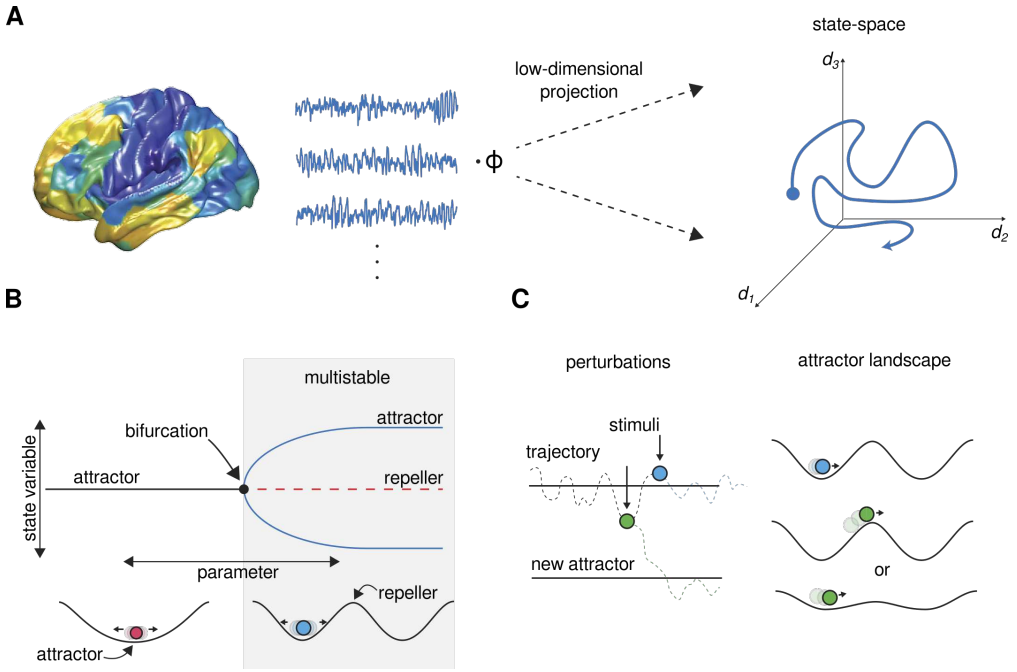

Figure 1.5.: Dynamical systems in neuroscience. **A**) Portrayal of brain as a reduced dynamical system in state-space. **B**. Schematic of a (pitch-fork) bifurcation where a single equilibrium splits into three equilibria due to a change in a system parameter, typically characterized by a symmetry-breaking instability. **C**) The influence of external stimuli on inducing state transitions. Figure adapted from [102].

DST allows us to categorize the qualitative characteristics of trajectories, such as

stability, which is a central concept in computational neuroscience [102]. This framework has been used to investigate a diverse range of topics within the field, such as decision making, social cognition and network-wide perturbations of healthy brain function in neuropathologies [102, 51]. One can delineate stability via the notion of attractors, where a region in the state space draws in nearby trajectories. Systems in attractor states can be perturbed, but subthreshold perturbations lead to the eventual convergence of their trajectories towards the stable 'basin' of attraction. Unstable states on the other hand, can be represented via repellers, in which trajectories diverge from a point in the state-space. Each dynamical system can be configured to exert a specific qualitative behavior by making changes to its parameters, where going from one such behavior to another is referred to as a 'bifurcation' [102]. A relevant example of a bifurcation is going from quiescence to spiking (e.g. as a result of external stimuli) in single-neuron models. One can even portray attentional modulation through bifurcations, such as representing focused attention via a deep attractor. In this portrayal, a strong distractor (i.e. sufficiently strong perturbation) can still push the system towards another state [102]. DST thus enables us to robustly represent myriad neural and cognitive phenomena, given sufficient knowledge of their underpinnings.

For the sake of our discussion here, let us assume that we have our chosen dynamical model of SFI effects at hand. After doing so, we would soon encounter an issue of equal importance: how does one pioneer a rigorous link between model parameters, and empirical observations?

Performing a parameter sweep (i.e. brute-forcing) is among the classical approaches of establishing such a link. In this method, the parameter ranges of interest are divided into grids, and values corresponding to grid elements are used to simulate the target model. Empirical data are then fit to a specific region of (or point in) the parameter space, by optimizing which region in the parameter space produces synthetic features that best resemble our chosen empirical ones [10]. As you can guess, not only can this method prove to be quite cumbersome, but its application verges on computational inefficiency –or even intractability– with increases in parameter dimensionality [116].

Bayesian inference on the other hand, provides a principled approach for delineating how empirical data are best represented in terms of model parameters, and in a probabilistic fashion [165]. Bearing some similarity to certain principles of the scientific method, Bayesian inference provides an alternative framework centered around updating beliefs about the world based on new evidence. In contrast to the classical frequentist approach of simply calculating fixed values and repeating experiments, this framework quantifies uncertainty via distributions, and allows us to reason probabilistically about hypotheses and parameters. Indeed, Bayesian learning aligns well with the inherent nonlinearity and variability observed in biological systems like the brain. Bayes' Theorem sets the foundation for this inference framework, elegantly capturing its ever-updating essence:

$$p(\theta|y) = \frac{p(y|\theta)\,p(\theta)}{p(y)} \tag{1.1}$$

In terms of model-based statistics, $p(\theta|y)$ represents the posterior distribution (our

updated beliefs), or the probability of parameter θ being true given observed data y [165]. $p(y|\theta)$ is probability of observed data given parameter θ (i.e. likelihood distribution), $p(\theta)$ is our prior belief (i.e. domain expertise, or previous knowledge) about parameter θ before considering any observations, and $p(y)$ is the probability of empirical observations regardless of whether θ is true. $p(y)$ is usually intractable, and therefore the unnormalized posterior distribution is approximated by a constant value in many applications of Bayesian inference [165].

In short, Bayesian inference combines our previous knowledge with new observations to arrive at a revised, more informed belief about the hypothesis under consideration [165]. The impact of this framework stems not only from representing phenomena via probability distributions, but also from its iterative nature, which allows for continuous learning and refinement as more evidence becomes available [83]. By representing measures of neural activity such as synaptic strengths or neuronal firing rates not as fixed values but as probability distributions, we can quantify our uncertainty regarding their true values [165, 173]. Furthermore, we can incorporate prior knowledge gained from existing theories, experimental findings, or even computational simulations into our models. These priors provide a starting point for our beliefs, guiding our exploration of the data and helping us focus on more plausible hypotheses. Employing this framework in neuroscience translates to a dramatic shift away from phenomenological representations and towards ones that are more grounded in biological reality, via acknowledging the inherent stochasticity and variability of neural systems, and our continuous re-evaluation of them given new findings [173].

1.2.2. Causal perspective of attentional modulation

Despite the described general frameworks which will form the foundation of this work, we are still in need of a way to *formally* link empirical observations with alterations of neurobiological characteristics. For this purpose, we turn to Dynamical Causal Modeling (DCM), which is a Bayesian framework dedicated to delineating how neural data are best represented in terms of neurobiologically-grounded parameters [182, 66]. DCM enables one to infer latent state variables (unobservable neural processes) from observable neural activity (electrophysiolgical recordings), by conceptualizing the brain as a nonlinear and deterministic system subjected to continuous inputs and yielding potentially predictable outputs. These hidden states are characterized in terms of posterior estimations of neurobiologically-plausible parameters (e.g. the effective synaptic influence of neural populations on each-other) [66]. Instances of this framework also bear the same name. DCMs –in accordance with their nomenclature– are characterized by being dynamic (through the use of nonlinear differential equations), causal (e.g. how dynamics of one neural element leads to modulations of dynamics in another), driven by biologically-plausible models, and fully driven by Bayesian principles [66, 67].

DCM's unique contributions are best evident when contrasted with previously established methods such as Structural Equation Modeling (SEM), and autoregressive

models [124]. While capable of capturing complex relationships between variables, SEM often relies on assumptions of linear dependency and cannot faithfully model nonlinear interactions prevalent in neural systems. Put another way, SEM might find correlations between brain regions during a task, but it cannot necessarily determine if one region causally influences the other or if there are alternative pathways at play. In addition, autoregressive models attempt to forecast future values of neural activity based on previous ones, by essentially treats brain activity as a time series [66]. Such correlational measures do not expressly represent the underlying connections between different brain regions, and thus offer limited insights into the causal structure underlying brain dynamics.

DCM's moves beyond mere calculations of correlation by explicitly modeling direct (i.e. causal) connections between neural populations, which allows for a more nuanced understanding of how disparate brain regions influence each other [42]. Integration of anatomical information with computational models further strengthens the explanatory power of this framework; DCM's multimodal approach of incorporating prior knowledge about neuronal connectivity based on anatomical data (e.g. DTI or diffusion tensor imaging, which reveals white matter tracts connecting different brain regions) allows for a richer understanding of how brain structure and function are intertwined [182, 42].

Every application of DCM is hence comprised of two distinct but related variables; 1) Observable variables (or electrophysiological activity) which are measured through modalities such as fMRI or EEG, and are not directly informative regarding the biological configurations of the recorded neural regions or subpopulations [182]. 2) Unobservable and time-dependent properties of neural systems (e.g. synaptic strengths), which are referred to as hidden/latent state variables. Hidden states are usually inferred from their causal influence on observed signals via model inversion [182].

Therein lies the duality of every DCM application, namely, the forward and inverse models (or problems). The forward model revolves around going from latent states (at a specific time point), to observable data (at subsequent time points), bearing assumptions about the temporal dynamics of neuronal populations. These models incorporate biophysically-relevant characteristics like the strength of synaptic weights and postsynaptic potential amplitude, along with assumptions about connectivity patterns derived from anatomical data ([182], Figure 1.6). For instance, simulations of a model of two neural subpopulations with synaptic transmission delays could show that activation of one could lead to a lagged but proportionally larger response in the other.

The inverse problem, on the other hand, refers to the challenge of inferring the hidden states from observed brain activity. This involves working backward from the data to determine the most representative neurobiological variations that could have produced the observable data Figure 1.6. To tackle this, DCM utilizes sophisticated Bayesian statistical methods such as Laplace approximation [182, 62]. By comparing different model structures and their corresponding predictions for the observed data, we can identify the model that best explains the observed patterns. This process involves calculating the probability of observing the data given different models and

hidden state configurations [66].

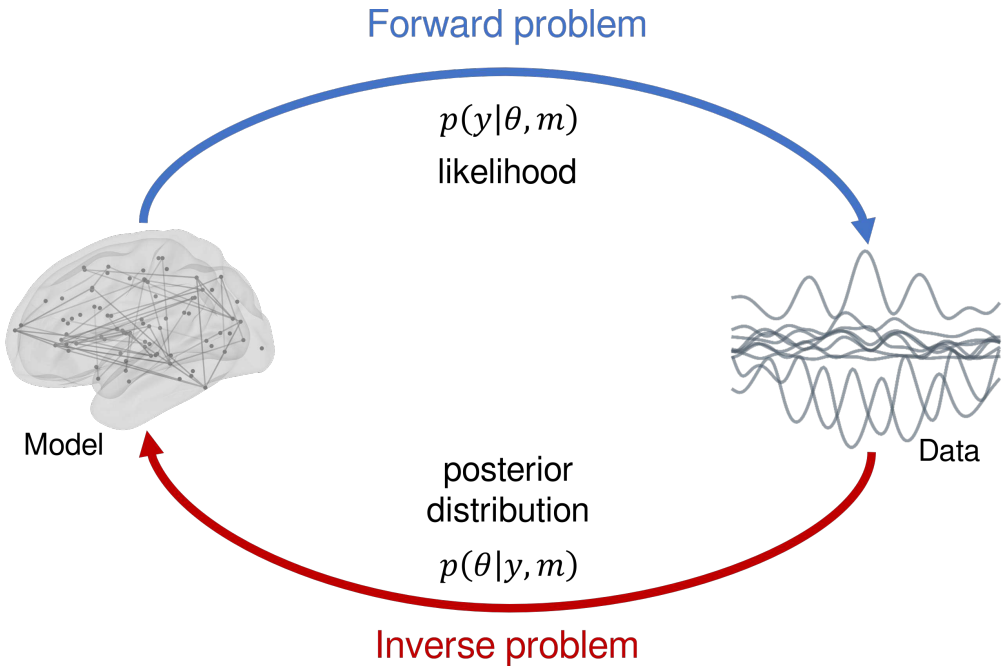

Figure 1.6.: Forward and inverse problems of DCM. Simulations of a generative –forward– model of whole-brain dynamics can be used to infer biologically-relevant parameters from empirical observations, via solving the inverse problem. The likelihood distribution $p(y|\theta)$, is defined as the probability of observed data y given model parameter θ. In contrast, the posterior distribution $p(\theta|y)$ represents the our updated beliefs or the probability of θ being true given observed data.

A key strength of DCM lies in its rigorous use of Bayesian statistics for model selection, inversion, and evaluation, in order to make mathematically sound inferences regarding the causal mechanisms underlying neural activity [182]. DCM utilizes various criteria for evaluating model fit, such as free energy minimization, and leave-one-out cross-validation [190, 144]. Free energy (i.e. upper bound on evidence or log-likelihood) measures the overall goodness of fit of a model to the data, taking into account both complexity and accuracy [144]. It reflects an approximation of the probability of observing the data given the model, considering both how well it explains the observed patterns and its inherent complexity [190]. Alternatively, leave-one-out cross-validation involves iteratively removing one data point and evaluating the model's performance on the remaining data, providing a robust measure of how well the model generalizes to unseen data [190]. This helps assess the model's ability to predict future observations and avoid overfitting to the specific training data [190].

Those that have dealt with dynamical models of biological systems are fully aware of the fact that it is computationally prohibitive to fully represent every known component of smaller biological systems, which verges upon impossibility for large-scale

simulations. The concept of effective connectivity is born from this exact realization. Simply put, effective connectivity attempts to capture the causal influence of neural elements (e.g. subpopulation, cortical micro-column etc.) on each other by reducing certain aspects of a neurobiological system [63, 64]. A substantial body of literature has corroborated the existence of a link between alterations in effective connectivity and attentional modulation [67]; For one, previous research points to correlations between posterior parietal cortex and motion-oriented area V5 in action attendance [67], premotor and prefrontal cortices in attention to visual –motion– stimuli [157], and medial temporal lobe and visual cortex [32]. Moreover, evidence seems to support a positive association between an elevation in task attentional demands, and enhanced effective connectivity of the descending attention pathway [96].

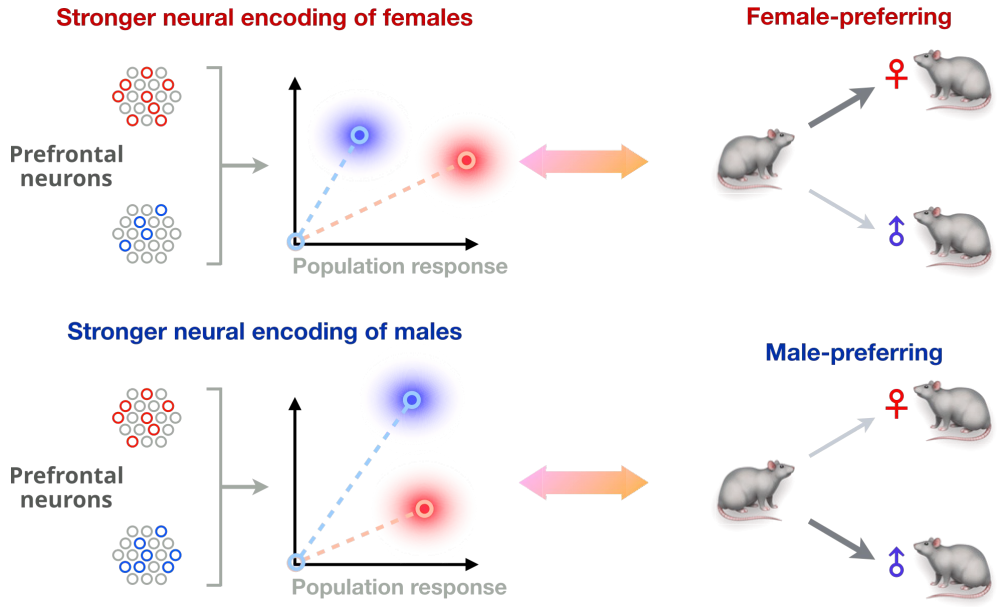

Figure 1.7.: Neural representations of conspecific social characteristics. Activity of prefrontal neurons tuned to conspecific sex (left) informs social behavior (right). Adapted from [108].

From a neurobiological perspective, this modulation is said to occur through enhancement of communications at the synaptic scale, particularly via selective adjustment of 'synaptic efficacy' (SE). SE represents a microscale subset of effective connectivity [20], the modification of which can exert drastic influences on the intricate balance between excitation and inhibition in neural ensembles, and in a degenerate manner [54]. Excitation/inhibition (E/I) balance stemming from connectivity alterations reportedly enacts a major role in efficient information processing and stabilization of network activity [200], with its disruption being implicated in numerous neuropathologies such as autism spectrum disorder [172], and its modulation rescuing social behavior [168]. Most crucially for our purposes, social behavior has been suggested to mold representations of conspecifics characteristics on cellular and

synaptic levels [108, 26, 153], which in turn have been linked to perturbations in E/I balance ([108, 26, 153], Figure 1.7).

While the majority of the above findings are based on rodent studies, it is reasonable to assume the existence of similar social representations in higher primates, albeit in a more complex manner, which gets us to the crux of the endeavor undertaken here; Informed by the shared evolutionary past with our mammalian counterparts, I posit that the observed behavioral enhancements observed during social facilitation stem from presence-induced neuromodulation of context-aware (i.e. social) neurons in attentional circuits, the overall effects of which could be faithfully measured via probabilistic inference of synaptic efficacies across brain scales. In the following parts, I will outline the exact methodologies used in order to demonstrate a well-founded link between neurobiological and SE measurements.

1.2.3. Sampling from posterior distributions

As the name suggests, sampling algorithms allow one to draw samples without knowing the underlying probability distribution [116]. This provides not only a point estimate of the most likely parameter values but also a full distribution capturing the uncertainty surrounding those estimates. This quantification of uncertainty is crucial in scientific modeling, where one should refrain from drawing definitive conclusions due to the inherent variability present in biological systems [173, 116].

As mentioned before, dynamical models are mathematical representations of time-dependent systems. Naturally, increasing the sophistication or granularity of said representations compels an elevated number of parameters, thereby compounding the difficulty of navigating the resulting high-dimensional parameter space [116]. Looking back to the Bayes rule, the integral over such a high-dimensional parameter space θ can often be analytically intractable [69]. Markov Chain Monte Carlo (MCMC)algorithms are a family of computational techniques that can address this specific challenge [136, 69]. The Markov Chain component represents the avid "navigator" that explores the vast parameter space by leveraging its local geometry and topology. With each step, our navigator makes a move from one state (i.e. set of parameters) to another, based on some transition probabilities (of moving from one state to the other). This probability distribution is typically drawn from the current state, ensuring that the explorer moves towards regions with higher likelihoods. The Monte Carlo component introduces randomness into our navigator's movement, allowing it to better navigate this space. While MCMC lays the framework for some of the state-of-the art methods for rigorous sampling of posterior distributions, gradient-free implementations of it are not suitable for many complex neural models due to its slow convergence and issues with non-normal posteriors [136, 69]. Enter Hamiltonian Monte Carlo (HMC), a gradient-based MCMC technique that addresses many limitations of its parent class [127, 16, 136]. Instead of relying solely on random walks (Monte Carlo), HMC modifies the Markov Chain by introducing a "momentum" component (via simulation of Hamiltonian dynamics). This gradient-based motion translates to more informed decisions about which states to visit next, significantly improving

the ability of bypassing local optima and avoiding random-walk [127, 16]. While this approach enables one to sample from complex posterior distributions, and asymptotically converges to the true posterior distribution (given a suitably long run time), its application is constrained by a number of factors. For one, simulation of Hamiltonian dynamics during each timestep can prove to be extremely computationally prohibitive in large-scale simulations (such as tens of thousands of cells in a spiking network). The most significant limitation however lies in its geometrical issues, as empirical data (i.e. time series) must be provided in order to obtain their posterior estimation, thereby creating two major challenges for the purposes this work [127, 16]. First, this per-observation estimation means that the researcher must re-run the algorithm in cases of new observation or changes in the preprocessing of empirical data. Second, the aforementioned approach disallows for a thoroughly 'agnostic' take on the changes in empirical data (as the estimation is based on the intricate changes in each timestep of the data, and less so on its overall informative features) [33, 187]. As I will show in section 4.2 however, efficient implementations of HMC in high-level machine learning libraries such as Numpyro allow for reliable validation of findings obtained via other probabilistic inference approaches [146].

1.2.4. Amortized probabilistic inference

Approximate Bayesian Computation (ABC) is a family of algorithms that bypass the direct computation of the likelihood in favor of its approximation [12]. ABC begins by generating numerous simulated datasets from the model under consideration, each with different parameter sets drawn from a prior distribution. These simulated datasets are then compared to the observed data using a distance metric –a measure quantifying the discrepancy between the summary statistics of the simulated and real data [35].

In essence, the likelihood is approximated by the probability that the condition $\rho(xsim, x_{obs}) < tol$ is satisfied, where ρ is some distance measure and tol is a tolerance. The accepted samples then follow an approximate version of the posterior [35]. Traditional ABC methods, however, suffer from the curse of dimensionality, in that the required number of simulations needed for obtaining a reasonable posterior estimation exponentially increases in the face of high-dimensional data. Therefore, such measures are often appropriate where the number of observations are extremely low [33].

Fortuitously, cutting-edge ABC variants such as simulation-based inference (SBI) have largely addressed the aforementioned limitations. In complex systems with many parameters or intricate interactions between variables, calculating the likelihood directly can become prohibitively expensive. SBI circumvents this challenge by focusing on simulating from the model distribution, effectively sidestepping the need for explicit likelihood calculation (Figure 1.8). Formally, posterior estimation in SBI is done via 'normalizing flows', a family of generative models which are tailored to probabilistic inference. These models exercise a series of invertible transformations capable of efficiently molding a simple initial distribution (e.g. a Gaussian or Gamma) into any

complex (or multimodal) target distribution ([140], Figure 1.8). Neural Spline Flows (NSFs; [50]) represent a sophisticated advancement within the realm of normalizing flows, specifically designed to retain perfect invertibility whilst refining the flexibility of the mapping. Capitalizing on the flexibility of splines as coupling functions, NSF's neural networks are able to acquire the optimal configurations for representing the data by controlling the parameters of these splines. In a nutshell, NSFs can efficaciously represent multimodal distributions which directly translates into improved sampling performance. With the exception of the mesoscale validation, the entirety of inferences conducted in this work will leverage NSFs.

One often begins an SBI workflow with a mechanistic model that can appropriately represent the desired phenomena (e.g. spikes, firing rates, whole-brain oscillations). The next step involves drawing random samples from the prior parameter distribution(s), which are used to generate model simulations (Figure 1.8). This crucial step heavily relies on domain expertise, as both the range of the prior distribution and the chosen parameters can exert a drastic influence on the inference process. Albeit one can use raw simulations (i.e. time series) in order to move to the next phase of the SBI pipeline, it is more prudent and computationally efficient to calculate data features (i.e. summary statistics) from each simulation. It is also paramount to leverage previous empirical findings in such calculations, as one optimally aims for features which best characterize the data. Vectors of sampled parameters, and the summary statistics computed therefrom, are subsequently given to a neural density estimator [74, 82] in order to obtain an invertible mapping between the two (Figure 1.8). With the neural network trained, similarly computed empirical data features are provided in order to obtain their respective parameter distributions [82].

More formally, a composite of N simulations are generated for training the density estimator by sampling parameters $\vec{\theta}$ from the prior $p(\vec{\theta})$ distribution of interest. The set of input used for training can be defined as $\{(\vec{\theta}_i, \vec{x}_i)\}_{i=1}^{N}$, where $\vec{\theta}_i \sim p(\vec{\theta})$ and $\vec{x}_i$ is the summary statistics from simulations obtained from parameters $\vec{\theta}_i$. Subsequent to training via normalizing flows, one can readily estimate the posterior $q_\phi(\vec{\theta} \mid \vec{x})$ with 'learnable' parameters ϕ, such that given empirical data features $\vec{x}_{obs}$: $q_\phi(\vec{\theta} \mid \vec{x}_{obs}) \simeq p(\vec{\theta} \mid \vec{x}_{obs})$. Note that the 'amortized' characteristic of SBI [74, 82] stems from the fact that learning the aforementioned mapping requires no initial input of empirical data, and the pre-trained weights could thus be infinitely reused (e.g. after re-running preprocessing steps or inclusion of additional data points) without any bias from empirical observations [187]. The complex connections between model parameters and variables, in addition to the extreme high-dimensionality of the forenamed neural data (e.g. neuron/network count) thus champions adoption of SBI [33] for the majority of the inferences conducted in this endeavor (with the exception of one particular use-case of HMC). Inferences were done via NSFs, using a neural network architecture comprising of 5 flow transforms with two residual blocks of 50 hidden units each, ReLU nonlinearity, and 10 spline bins, adhering to the previously implemented *sbi* toolbox [187]. In the following, I will delineate the quantification uncertainty of SE across all three scales (Figure 1.8).

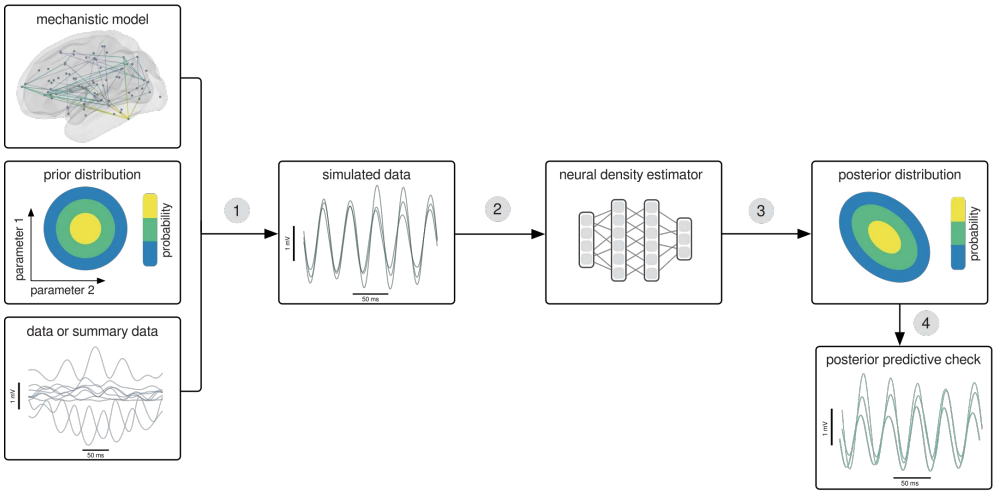

Figure 1.8.: General outline of a simulation-based inference workflow: 1) Sampled parameters from the prior distribution of the chosen mechanistic model (e.g. coupled NMMs) are used to generate fake data resembling empirical recordings. 2) Low-dimensional and informative summary statistics of them are used to train a neural density estimator on establishing an invertible mapping between model parameters and input data (features). 3) Empirical findings (or new synthetic simulations) can then be used in order to sample from the learned posterior distribution. 4) Simulations generated based on the samples from estimated posterior can be used for validation of the inference process. Figure adapted from [74].

1.2.5. Low-dimensional subspaces of neural systems

A critical component of applying simulation-based inference in empirical contexts is the use of highly informative summary statistics, which raises the question: how does one go about finding such informative (and low-dimensional) features? Nature has fortuitously paved the path for us in this case; It has been known for sometime that cognitive function emerges at the scale of populations/networks, and not single neurons. Hence, many functions of high-dimensional neural system should –theoretically– possess lower-dimensional representations. The subspaces spanned by such low-dimensional representations are often referred to as 'manifolds'.

The framework of structural flows on manifolds (SFM) is central to forming a robust understanding the formation of such subspaces [100, 147]. SFM posits that complex biological systems often exhibit low-dimensional structure underlying their behavior. This means that the system's state can be effectively captured by a significantly smaller set of variables, which evolve along a specific trajectory defined as a manifold embedded within the larger, more complex space. Consider a flock of birds: individually they might exhibit simple behaviors, but collectively their coordinated movements form intricate patterns [100, 147]. SFMs can help us identify the low-dimensional

'grammar' governing these collective behaviors, capturing how individual interactions come together to produce emergent behavior at a larger scale.

In the context of neuroscience, SFMs have proven particularly insightful in capturing both brain and behavioral dynamics. Through rigorous application of SFM principles, one can identify certain stable subspaces that encapsulate essential aspects of neural activity, such as cognitive states or motor patterns. For instance, research suggests that specific manifolds in the brain may underly a wide range of cognitive phenomena including motor control [106, 90] and resting state dynamics [61].

The discovery of 'attractor manifolds' is another line of evidence supporting the existence of such embeddings. Attractor manifolds are defined as low-dimensional stable patterns (i.e. stable states) emerging from the aggregate activity of neural populations [75, 197]. These stable representations can be robustly preserved through sparse connectivity and redundant/distributed encoding, irrespective of the time-dependent nature of neural activity [47, 192]. Attractor networks reportedly maintain persistent activity for myriad neural processes, such as social cue representation, error correction and working memory ([192, 107, 25], Figure 1.9). The reforming of attractor manifolds under electrical stimulation has even led some to suggest their formation during plasticity in general [13]. A prominent subtype of such manifolds in biological systems are line attractors where all trajectories –of a neural system– converge to an infinite number of fixed points [192]. Line attractors are posited to emerge due to the dominance of slow timescales in dictating the system's behavior [192, 197].

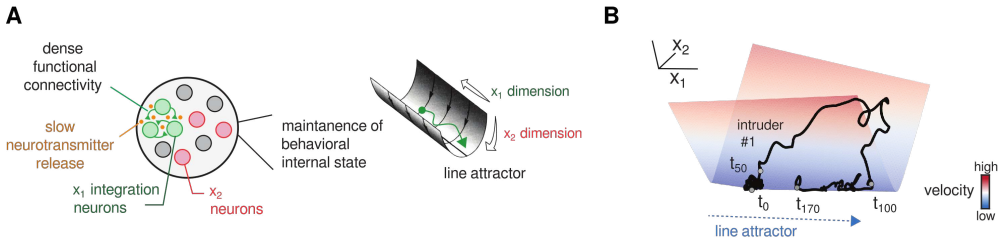

Figure 1.9.: Attractors are posited as inherent characteristics of biological neural circuits which stem from interactions between system components. **A**) Schematic of proposed implementation of a line attractor encoding an internal state. **B**) Line attractor of neural activity constrained on aggression observation. Figure adapted from [192].

This framework elegantly bridges the gap between high-dimensional observable reality and our capacity to comprehend it, by dissecting the qualitative behavior of complex systems into more manageable components. To answer our question then, any of myriad dimensionality reduction algorithms could enable us to reveal emergent data features (i.e. low-dimensional properties) which are most conducive towards a rigorous application of simulation-based inference. In the subsequent chapter, I will fully describe two facilitation tasks and their experimental data, on which the above frameworks will be applied.

2. Experimental evidence for facilitation

Summary

2.1. Association Learning task . . . 38
2.1.1. Subjects . . . 38
2.1.2. Behavioral task . . . 38
2.1.3. Electrophysiological Recordings . . . 40
2.1.4. Previous results . . . 41
2.1.5. Neural firing rates . . . 43
2.1.6. Event-related potentials . . . 45
2.2. Lateral interception task . . . 48
2.2.1. Subjects . . . 49
2.2.2. Experimental setup . . . 49
2.2.3. Movement data . . . 51
2.2.4. Electrophysiological recordings . . . 52
2.2.5. Functional connectivity . . . 56
2.2.6. Measures of integration . . . 57

Social facilitation reportedly works across subject groups, tasks, and even species. In the following sections, I will fully delineate the behavioral and neural data obtained from two experimental paradigms. 1) An association learning task in Rhesus monkeys designed by Dr.s Driss Boussaoud, Pascal Huguet, Marie Demolliens and Faical Isbaine, as part of Marie's thesis. The behavioral training (spanning several months) and data acquisition (spanning 3 years) were entirely conducted by Marie and Faical [44, 93]. 2) A lateral-interception task in humans designed by Driss, Dr. Frank Zaal, Marjolein Viersen and myself. The pilot sessions were conducted by myself, and the data acquisition of experimental sessions conducted by Marjolein.

My central aim in this chapter is to delineate low-dimensional embeddings of empirical recordings across three spatiotemporal brain scales that are highly correlated with behavior, which will subsequently be employed as highly-informative data features in chapter 4.

2.1. Association Learning task

2.1.1. Subjects

The subjects of this study were two adult male Rhesus macaques (Macaca mulatta) housed together since the age of three years. Rhesus monkeys are well-established models in neuroscience research due to their cognitive abilities, complex social structures, and relatively large brain size, facilitating electrophysiological recordings and detailed analysis of neuronal activity [11]. Subjects were maintained on a standard dry diet supplemented with fresh fruit and vegetables. Their liquid consumption and body weight were meticulously monitored daily to ensure optimal health and well-being throughout the study. At the time of the experiments, both monkeys weighed between 8 and 12 kg. Animal care and housing strictly adhered to the European Directive (2010/63/UE) on the use of non-human primates in scientific research, guaranteeing their welfare and ethical treatment in accordance with internationally recognized standards [141]. While the relatively small sample size might appear unconventional, several factors justify its use for research of SFI effects. Firstly, the focus was on longitudinal observations within each individual monkey, tracking changes in their neural activity and behavior over time as they learned an association between a sensory cue and a specific outcome. Longitudinal studies often benefit from smaller sample sizes when the emphasis lies on individual trajectories rather than broad population generalizations. This allows for more detailed analysis of individual learning patterns, identifying subtle variations in how each subject adapts to the experimental paradigm. This long-term cohabitation allowed for the development of a well-established social dynamic, characterized by consistent interactions and a clear dominance hierarchy. This approach allows for an in-depth understanding of the dynamic processes underlying associative learning at a finer granularity than would be possible with larger, less focused samples. Lastly, using a homogeneous sample consisting solely of male monkeys minimizes potential confounding variables related to sex differences in learning and memory [130]. With regards to social interactions, monkey A consistently exhibited dominant behavior, as evidenced by preferential access to food and water resources within their shared living space. This dominance structure is crucial in Rhesus macaque societies, influencing not only resource allocation but also social interactions and even physiological responses [58, 122]. The inherent social complexity of these primates makes them ideal subjects for studying associative learning, particularly within a familiar and stable social context.

2.1.2. Behavioral task

The aim of this behavioral task was to investigate how social presence influences performance and neural activity during cognitive processing. The test subjects were trained to associate abstract images with specific targets on the touchscreen under the absence or presence of a conspecific (Figure 2.1**A**). Under social presence, both monkeys were seated in primate chairs facing each other, with their heads immobi-

lized using custom head restraints. This setup allowed for direct visual contact while minimizing extraneous movement during the task. The designated "actor" monkey had access to the touchscreen and received rewards for correct responses, while the "spectator" was not rewarded or incentivized to perform in any specific way. To prevent observational learning, the spectator never performed the task on the same day as the actor, and a new set of stimuli was used when tested. This protocol ensured that the observed behavioral effects were driven by social influence rather than mere imitation or learned responses from observing the other monkey. In contrast to the social presence condition, the isolation testing involved one monkey being alone in the experimental room while the other remained securely housed at a distance, effectively preventing any visual, auditory, or olfactory communication between them. This stringent isolation procedure minimized the potential for even subtle cues from the absent (and distant) conspecific, allowing for a more direct comparison between the effects of social presence and absence on task performance.

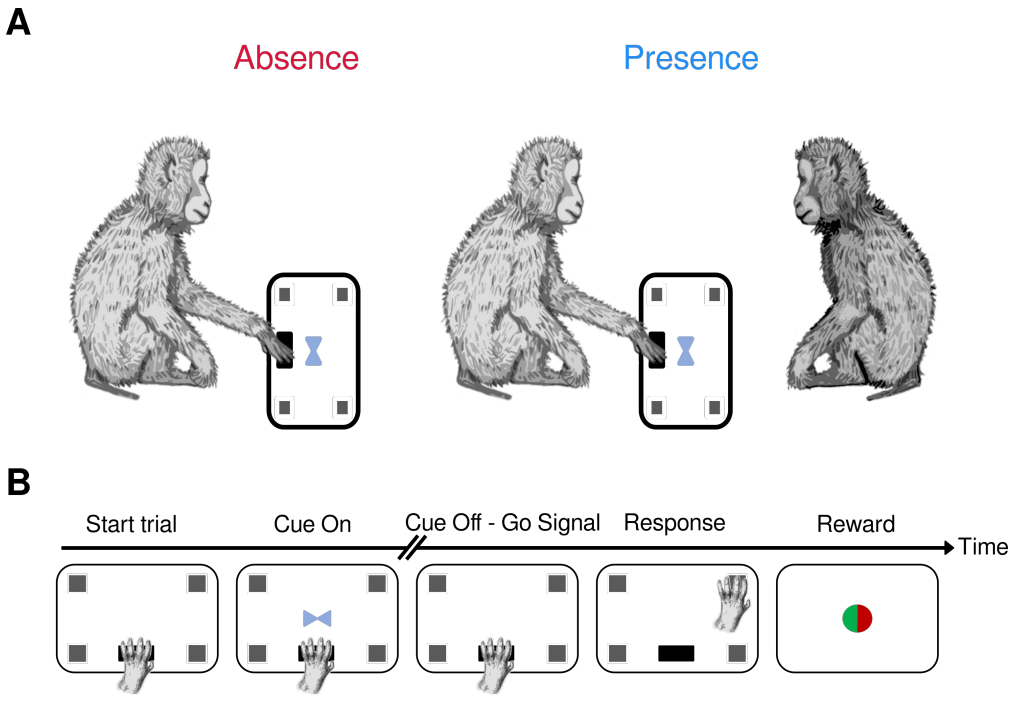

Figure 2.1.: Outline of the association learning task. **A**) Schematic of the experimental conditions of presence (in blue), and absence (in red). **B**) Task timeline during which subjects had to associate which of the four corners of the touchpad was associated with the cue in order to obtain a food reward. Images of subjects in this figure were adapted from [44].

Each trial was initiated by the actor touching a white rectangle at the center of the touchscreen. This action triggered the presentation of a cue image at the center of the screen, followed by a variable delay (500-700 ms). The cue image disappeared after this delay, signaling the "go" signal for the monkey to choose the appropriate target

among four white squares positioned around the screen. Feedback was provided instantaneously following the animal's response. If correct, a green circle appeared on the screen indicating success, and a reward (fruit juice) was delivered after a 1 second delay. Incorrect responses were signaled by a red circle, with no reward delivered (Figure 2.1**B**). Each session consisted of two sets of 80 trials – one under isolation learning followed by another set under social presence conditions. This fixed order was implemented for two primary reasons: First, initial tests showed that the monkeys were more reluctant to perform the task in isolation after experiencing social presence, suggesting a potential decrease in motivation or an increase in stress when separated from their familiar companion. This finding aligns with prior research indicating that the presence of conspecifics can reduce stress levels in primates [194]. Second, studies have demonstrated that the presence of others may still exert influence even during proceeding isolation trials [121]. This potential carryover effect necessitates a control condition where each monkey performs under isolation twice, allowing for a more accurate assessment of social influence on learning and performance.

2.1.3. Electrophysiological Recordings

Extracellular neuronal activity was recorded from dorsolateral prefrontal cortex (dlPFC) and the anterior cingulate cortex (ACC) in order to investigate the neural underpinnings of social presence during the association learning task (Figure 2.2). These regions were chosen due to their pivotal role in cognitive control, attentional modulation, decision-making, and reward processing [5, 201]. Each monkey underwent implantation of a recording chamber over the hemisphere contralateral to their dominant hand, ensuring access to both the dlPFC and ACC. A single chamber allowed for simultaneous recording of multiple neurons in addition to local-field potentials (LFP) within these two crucial brain areas. The recordings were conducted via a multichannel system (AlphaLab, Alpha Omega) equipped with single tungsten electrodes (FHC Instrument, 0.8–1.2 MX impedance), which enabled the isolation and recording of electrical signals generated by individual neurons within the dlPFC and ACC.

In order to obtain single-neuron activity, individual spike waveforms were separated from background noise using a semi-automatic algorithm. This process ensures that each spike corresponds to a single neuron's action potential, enabling accurate quantification of firing patterns. Subsequently, spike counts –for each neuron– in each time bin were baseline-normalized on a trial-by-trial basis, creating standardized representations of neuronal firing patterns across trials. Peri-stimulus time histograms (PSTHs) were then generated by plotting the normalized spike counts for each neuron over time, aligned with specific events within the task (e.g., cue presentation, feedback delivery). Finally, a Mann-Whitney U-test ($p < 0.05$) was used to compare neuronal activity between the isolation and social presence conditions for each defined epoch, highlighting neurons which exhibited significantly altered firing patterns in response to the presence of another monkey during the task (Figure 2.3**A**).

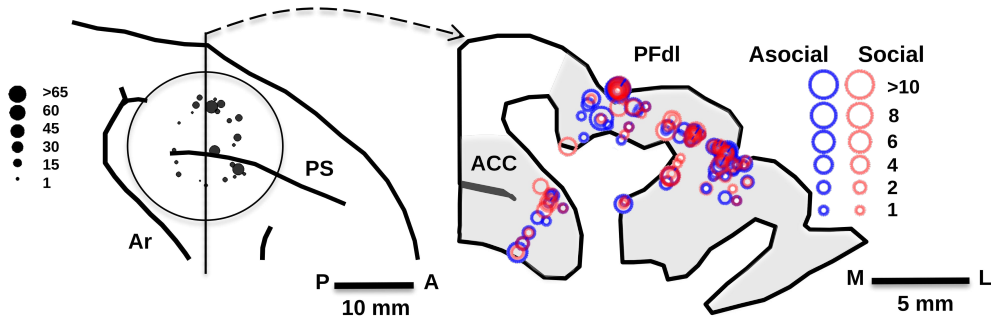

Figure 2.2.: Anatomical sites recorded. Left: Recording chamber location (black circle) and the sites shown on the lateral view of the right frontal portion of the brain. Right: Recorded sub-regions of asocial (blue) and social (pink) neurons acquired from a coronal section from the mid point of the recording chamber (dashed vertical line in the left). The radii of the recording sites represents the number of analyzed neural activations. Abbreviations: Ar, Arcuate sulcus; PS, principal sulcus. Figure adapted from [44].

2.1.4. Previous results

Since the raw microscale/mesoscale data in this work originate from the previously published investigation by our colleages [44], I will first briefly summarize their findings.

The behavioral data consisted of a total of 181 sessions for monkey *A*. In order to assess the influence of mere presence on learning speed –i.e., trials-to-criterion (TTC)–, sessions were set as the unit of analysis. In the context of this task, TTC is defined as the third of five consecutive correct trials. Analysis of correct trial sequences revealed a reduction of number of TTCs during the presence condition compared to the absence condition (from 11.02 to 9.26 across sessions in monkey *A*, $F(1,180) = 10.84$; $p < 0.001$, $\eta_p^2 = 0.06$). To further strengthen this conclusion, test-retest data were analyzed, which showed that the improvement in performance under social presence could not be attributed to mere task repetition or familiarity. This finding reinforces the conclusion that social presence directly influences learning speed and is not merely a consequence of prior experience or task order, which was further validated using the non-parametric Mann-Whitney U test ($ps < 0.05$).

The presence of a spectator also significantly influenced the activity of a substantial majority of prefrontal neurons, whether they errors (85%) or rewards (82%). An analysis of this phenomenon revealed two principal categories of neurons, distinguished by their unique responses to social presence. These categorizations were confirmed through statistical tests, including Mann-Whitney U tests that compared PSTHs across conditions. Social neurons, comprising 41% and 38% of the populations responding to negative and positive feedback, respectively, exhibited increased firing rates in response to social presence compared to solitary conditions. Conversely, asocial

neurons, representing 41% and 47% of the same neural populations, displayed higher firing rates during solitary conditions relative to social presence (Figure 2.4). To validate this categorization, unsupervised hierarchical clustering analysis was performed on the data, yielding results that corroborated the existence of two primary neuronal categories with less than 3% overlap between them. Additionally, finer sub-clustering was observed within each main category. Social and asocial neurons were distributed similarly across both of the examined prefrontal regions, with histological reconstructions being unable to reveal any apparent spatial segregation of these neuronal categories within either cortical area.

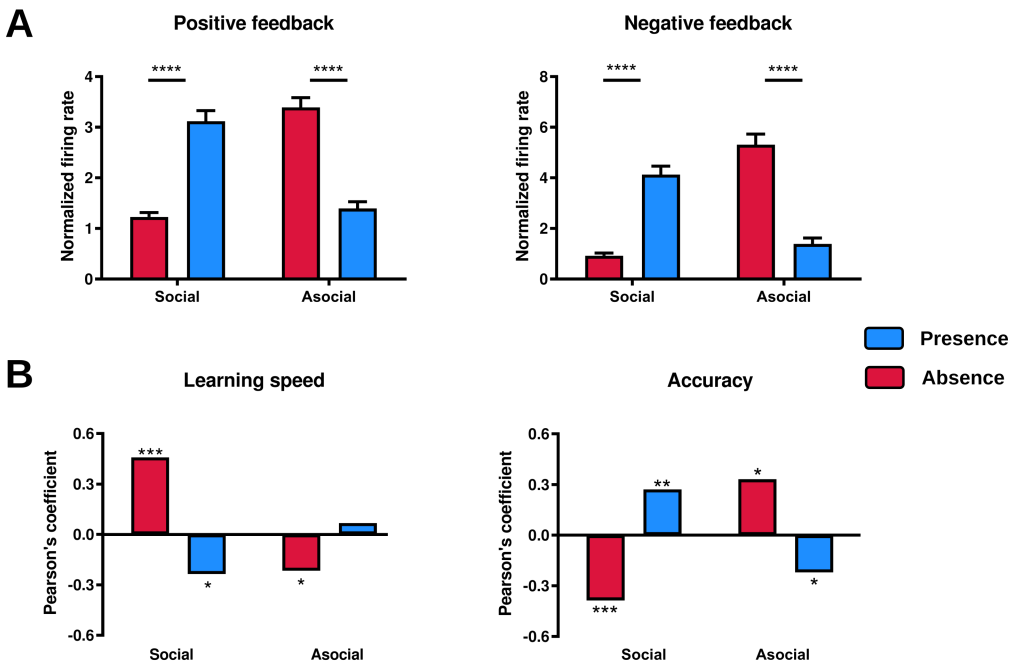

Figure 2.3.: **A**) Average neural activity across the two experimental conditions. **B**) Correlation between task performance and the firing rate of context-oriented neurons. Figure adapted from [44].

The intricate relationship between behavioral performance and neural activity under different social contexts was further explored through correlation analyses. A prominent finding was the significant negative correlation between the firing rate of social neurons associated with negative feedback and learning speed under social presence – the higher the firing rate, the faster the learning. Conversely, asocial neurons showed a negative correlation between firing rate and learning speed during isolation. A similar trend was observed for accuracy as a metric; Social neurons' activity was highly correlated with task performance during presence, with asocial neurons mirroring this pattern but for their preferred context (i.e. absence) (Figure 2.5**A**).

Notably, these correlations systematically reversed for both neuron types when their

respective preferred context was altered, highlighting a strong dependency on social context for both neuronal coding and behavioral performance. Overall, these results suggest that the activity of these neurons might contribute to the observed social facilitation effect.

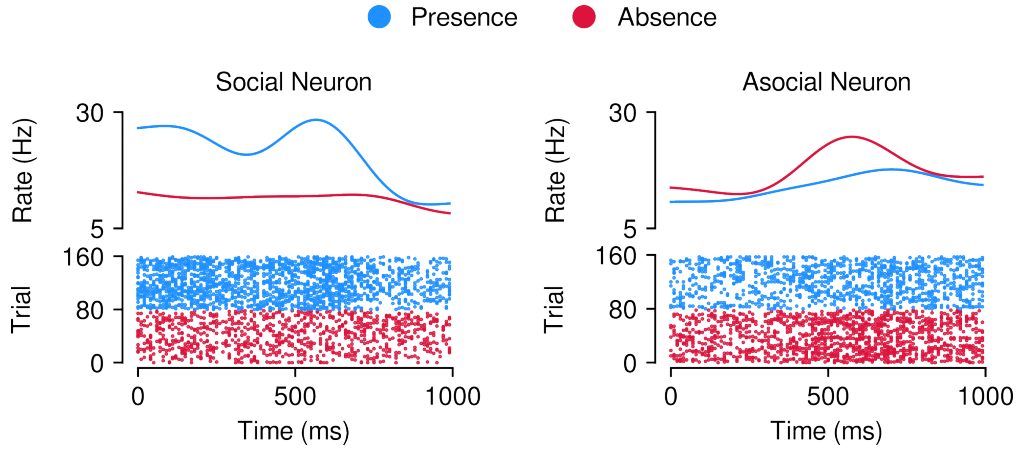

Figure 2.4.: Samples from the categorized Social (left) and asocial (right) neural subpopulations. Top: Average firing rate during presence (blue) or absence (red). Bottom: Raster plot of the entire recording session for each neuron.

2.1.5. Neural firing rates

Notwithstanding of the significant finding of context-specific neural subpopulations and the correlation of their activity with task performance, the above work did not address how social presence influenced the overall change of the spiking activity within the recorded regions. To that end, I set out to compute and analyze the overall (i.e. average) firing rate of the previously discovered context-dependent subpopulations.

The neural data were pooled together into a single neuronal sample due to the similarity in the characteristics of event-oriented activity in dlPFC and ACC. The most significant correlation between behavioral performance and single-neuron firing rates was found in negative feedback events. Therefore, a subset of all neurons in monkey A ($n = 92$) which were tuned to this event, and were explicitly categorized as social or asocial (i.e. no 'neutral' neuron was chosen). A window of 500 ms preceding and following the negative feedback event was chosen, in lieu of the structure of trials used in the previously conducted analyses.

Neurobiological data inherently exhibits a large degree of variability across acquisition dimensions (e.g. trials, sessions, or subjects). With the aim of conducting a robust comparison of firing rates between presence and absence across experimental sessions and neurons, single-neuron firing rates in each condition were divided by the absolute maximum firing rate of each neuron during that session, yielding normalized values between zero and one. Subsequent to this normalization, the peak firing rate of

single neurons were found to be significantly increased during the presence condition ($p < 0.001$; Figure 2.5**B**, **C**).

In order to have an agnostic validation of this between-condition difference, the firing rate of the entire neural dataset was fitted to a linear dynamical system (LDS) to infer the hidden dynamics in the two experimental conditions. Linear dynamical systems [17], characterized by linear evolution functions over time, are frequently employed to model data due to their mathematical tractability for learning and inference. In an LDS framework, at each time step t, a high-dimensional emission $y_t \in \mathbb{R}^N$ is observed, driven by low-dimensional hidden states $x_t \in \mathbb{R}^D$. The state dynamics are governed by the equation: $x_t = Ax_{t-1} + Vu_t + b + w_t$, where A represents the dynamics matrix, V is the input-to-state control matrix, u_t denotes a control input, b is an offset vector, and $w \sim \mathcal{N}(0, \Sigma_w)$ represents process noise.

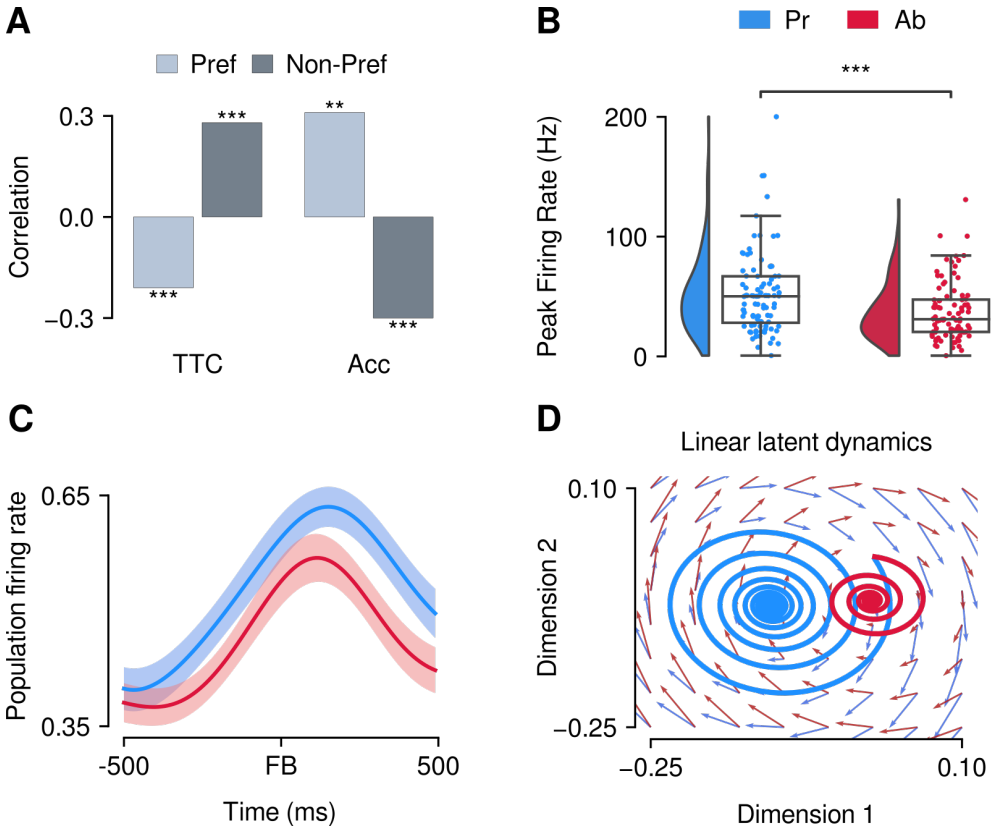

Figure 2.5.: **A**) Correlation of learning speed (i.e. TTC) and accuracy with the firing rate of social/asocial neurons. **B**) Peak firing rate of all categorized neurons in the two experimental conditions ($p < 0.001$). **C**) Representative –average– firing rate of all selected neurons as in panel B. **D**) Latent linear dynamics obtained from single-neuron firing rates in presence vs absence.

The emissions y_t are generated from the hidden states x_t according to a simple linear-Gaussian emission model $y_t = Cx_t + Fu_t + d + v_t$, where C represents the mea-

surement matrix, F is the feed-through (input-to-emission) matrix, d represents an offset or bias term, and $v \sim \mathcal{N}(0, \Sigma_v)$ denotes observation noise. Parameter estimation in LDS was performed using the open-source state-space modeling package (SSM; [114]), which streamlines the estimation of hidden states and model parameters from observed time series data (employing the Expectation-Maximization algorithm). An LDS was constructed and fitted to the firing rate time series in each condition ($n = 92$), assuming Gaussian dynamics and emission, to learn the vector fields driven by the measurement matrix A for each condition. To ensure stability, all eigenvalues of A must lie strictly within the unit circle in the complex plane, implying their magnitudes must be less than 1. The fitted systems were subsequently sampled ($n = 5000$) to forecast the system dynamics for each condition. Sampling from the learned dynamics revealed that variations in firing rate corresponded to changes in the decay rate (i.e., relaxation of population activity) [191, 134]. Decay in this context reflects the convergence towards the stable fixed point of the modeled linear system (Figure 2.5**D**). Functionally, this framework enabled the estimation of the rate at which the linear system (i.e., neural firing rate) returns to equilibrium after a perturbation in its measured output. A slower decay rate was observed in the presence condition compared to the absence condition, suggesting the presence of conspecifics could potentially exert a neuromodulatory influence on neural activity.

2.1.6. Event-related potentials

While the insights obtained from single-neuron activities are indeed promising, major facets of the data from the association learning task were not explored in the previous investigations [44]. One significant question is how the overall activities of dlPFC and ACC are modulated under the presence of conspecifics. To that end, I set out to process and analyze event-related potentials (ERP) from previously recorded local-field potentials in the association learning task.

ERPs –or evoked potentials– manifest as transient fluctuations in cortical excitability following the presentation of specific stimuli or task events and are derived by the averaging of event-locked electrophysiological trials of electroencephalography (EEG) or magnetoencephalography (MEG) signals recorded during a task [42]. Evoked potentials are said to reflect the summed activity of postsynaptic potentials (PSPs) due to the synchronization of pyramidal (sub)populations within the cortex [42]. Indeed, ERP signals and their respective features (e.g. amplitude, latency, and topography) have been frequently employed in the investigation of numerous cognitive functions such as attentional modulation, sensory processing, and working memory [174].

The LFPs (used to calculated the ERPs) comprised of four channels, and were processed using time-locking (or epoching), a technique employed to isolate the neural activity corresponding to specific events or stimuli. The signals were then epoched to a 350 ms window preceding feedback timestamps, allowing for the isolation of the neural responses elicited by the feedback event. A high-pass filter with a cutoff frequency of 2 Hz was applied to the LFPs prior to epoching, as this is typical for removing low-frequency noise. Additionally, electrodes exhibiting flat signals were excluded

from analysis due to their reduced (and often infinitesimal) signal-to-noise ratio, which often leads to inaccurate calculation of evoked potentials. Furthermore, only electrodes meeting excellent and good signal quality standards (defined using criteria such as electrode impedance or visual inspection) were included in the calculation of event-related potentials.

The raw ERPs ($n = 117$) for both positive and negative feedback conditions were averaged, yielding a comprehensive dataset for each task condition. Following the application of a low-pass Butterworth filter with a 20 Hz cutoff frequency, the evoked potentials underwent upsampling to a temporal resolution of 1000 time-steps (Figure 2.6**A**). This increased sampling rate is typically used to enhance the precision and accuracy of ERP estimates. To ensure comparability across sessions, these evoked potentials were normalized with respect to their maximum amplitude. Similar to the findings from single neurons, the ERPs in the presence condition exhibited a significantly higher peak amplitude compared to the absence condition ($p < 0.001$; Figure 2.6**B**). This close resemblance to the previous spatiotemporal scale however, might not translate to meaningful correlations with task performance. Moreover, the calculation of this correlation requires a robust metric by which to measure between-condition changes in neural, behavioral and synthetic variables. In this light, and from this point forward, let us define between-condition differences in behavioral and neural metrics X, as the percent change of variable X between the presence P and absence A conditions: $X_{Ratio} = \frac{X_P}{X_P + X_A} \times 100$.

With our definition in hand, behavioral performance ratio (accuracy per session) was found to strongly correlate with the ratio of peak ERP between conditions ($\rho_{pearson} = 0.33$; Figure 2.6**C**), which suggest that the neuromodulatory role of others' presence persists not only in single-neurons or subpopulations, but also across different brain regions or cortical micro-columns. As with the spiking dataset, agnostic validation of the significance of between-condition differences in ERPs was done via a state-of-the-art machine learning framework. In short, normalized ERP time series were fed –as input– into a machine learning framework utilizing CEBRA [164], which generated low-dimensional embeddings (i.e. representations) of the data through an unsupervised learning process. CEBRA employs contrastive learning to map high-dimensional data onto a lower-dimensional space. In this context, triplets of data points in lower-dimensional (sub)space are subdivided into positive pairs (similar data points) negative pairs (dissimilar data points). CEBRA then trains a neural network to maximize the similarity between positive pairs and minimize the similarity between negative ones, ensuring that the resulting embeddings capture the underlying structure of the original data while reducing noise and improving interpretability [164].

Visual inspection of the embeddings can indeed validate the existence of a significant amplitude difference between presence and absence. However, such embeddings only provide a static description of the dataset (i.e. which time series belonged to which experimental condition). Going a step further would be to use the output of CEBRA in order to construct a smooth probability grid delineating the likelihood of an embedding stemming from presence or absence trials. Bayesian neural networks

(BNN) enable one to do just that via incorporation of prior knowledge and uncertainty into models. In order to construct such a probabilistic description, we used the PyMC [142, 1], a probabilistic programming package in Python which aims to simplify the construction and fitting of Bayesian models. ERP embeddings were fed to a BNN with variational inference and a perceptron network architecture (comprising two hidden layers, each with five nodes). The resulting probability grid (Figure 2.6**D**) overlayed with the embeddings verifies previous findings with regards to between-condition differences in ERP amplitude.

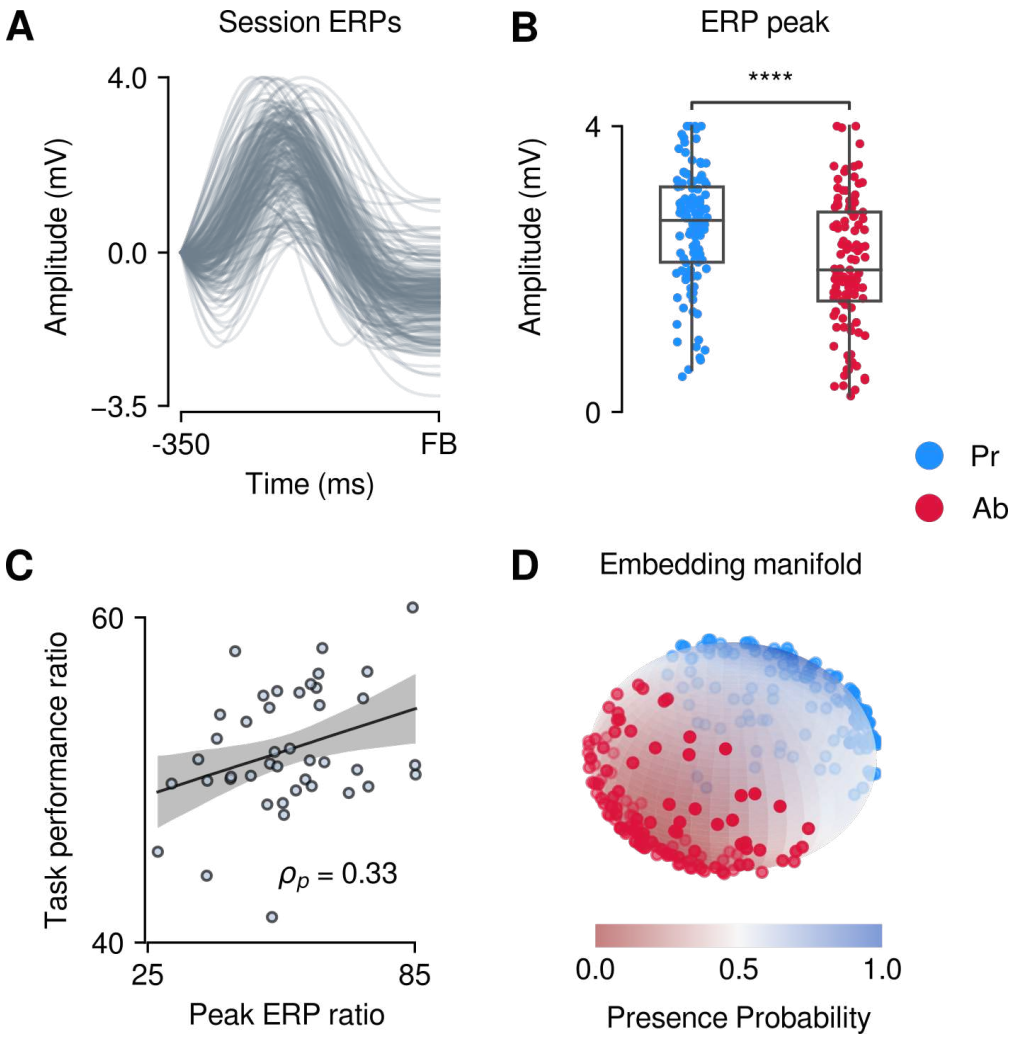

Figure 2.6.: Characteristics empirical evoked potentials across conditions. **A**) Subset of empirical ERPs locked to the feedback event. **B**) Peak amplitude of task ERPs during presence vs absence ($p < 0.001$). **C**) Correlation of –between condition ratio– of peak ERP amplitude with behavioral performance. **D**) Probability map of an ERP belonging to the presence condition, based on the embedding manifold of peak activity.

2.2. Lateral interception task

The inception of the lateral interception task came about when thinking of the empirical findings and limitations of the association learning task. Despite the importance of these findings for attaining a better grasp on the neural and behavioral underpinnings of social facilitation, modulations of neural activity at micro/mesoscale –across a handful of regions– cannot be directly used to infer how others' presence would influence brain/network dynamics.

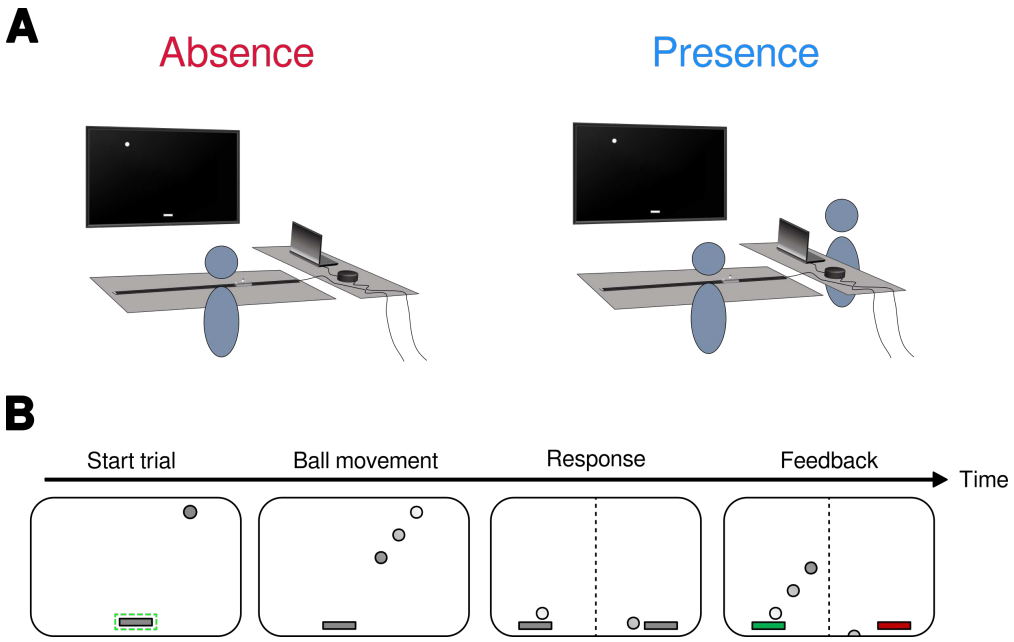

Figure 2.7.: Outline of the lateral interception task. **A**) Schematic of the experimental conditions of presence (in blue), and absence (in red). **B**) Task timeline during which subjects had to intercept a downward-moving ball with a handheld paddle.

One can argue that presence of others could disparately modulate the functional dynamics of various networks within the brain. Moreover, social facilitation effects are reportedly not limited to a specific task design [131, 148]. Hence, the next logical step would be to devise an experiment which attempts to investigate the influence of social presence on whole-brain dynamics, but also addresses the limitation of the previous dataset, such as the number of subjects and sex.

We designed a lateral-interception task in collaboration with –and conducted in– University Medical Center Groningen (UMCG), with the goal of addressing said limitations. Participants were tasked with intercepting a downward-moving ball, in the presence or absence of a conspecific. During this 'pong-like' virtual game, players controlled an on-screen virtual paddle via a handheld slider which allowed them to attempt to intercept the ball (Figure 2.7**A**).

2.2.1. Subjects

A total of 28 healthy right-handed participants ((n = 27, n_{female} = 14, n_{male} = 14), with an average age of 23.6 ± 2.1 years) were recruited for the study. All participants provided written consent forms prior to the initiation of the experiment. One male participant was removed from the participant pool due to an error in recording of the EEG data (n_{male} = 13). All procedures conducted in this study were approved by the Central Ethics Board of the University Medical Center Groningen (UMCG) and according to the Declaration of Helsinki (World Medical Association).

2.2.2. Experimental setup

The presence condition involved a seated spectator positioned 1.5 meters from the participant in order to be within boundaries of the participants' social space [80]. To minimize potential confounding factors, the spectator's age and sex were closely matched to those of the participant. The spectator was positioned to be within the peripheral vision of the actor, instructed to avoid looking at the screen and instead focus approximately 70% of their attention on observing the participant's hands in order to induce a subtle sense of being observed. In order to obfuscate the spectator's true aim to the participants, they were told a cover story prior to the experiment, in which the spectator presence was required for monitoring the EEG equipment. Following each interception attempt, participants received visual feedback: their paddle color briefly changed to green for successful trials and red for unsuccessful ones (Figure 2.7**B**). To further reduce evaluation-induced anxiety in the presence condition, the observer (i.e. spectator) was strategically placed within the participant's peripheral vision. This observer occasionally monitored the participant's hand movements but refrained from looking at the screen, echoing a similar strategy employed by Culot et al. ([37], Figure 2.7).

Participants were informed that the observer was there to monitor equipment, effectively maintaining their unawareness of the true purpose behind this presence. Overall, participants demonstrated a rapid grasp of the task and maintained consistent pursuit of the ball throughout the trials (Figure 2.7). Both kinematic (paddle movement) and electrophysiological data (EEG activity) were collected during the experiment. Participants engaged in the experiment at a table situated 2 meters from a large 55-inch Samsung LED television display (1920 x 1080 pixels, 120 Hz refresh rate). Their task involved controlling a virtual paddle displayed on the screen using a handheld knob connected to an in-house designed slider. The objective was to intercept a descending virtual ball as it crossed the screen. Continuous recording of slider positions was done via a National Instruments data acquisition card, capturing kinematic data on both the paddle and ball movements, along with success or failure in intercepting the ball. The entirety of the experiment was designed and controlled using the PsychoPy platform [143].

Electroencephalographic (EEG) signals were recorded in parallel with the behavioral task, using a TMSi SAGA data recorder and a 64-electrode TMSi infinity gel headcap

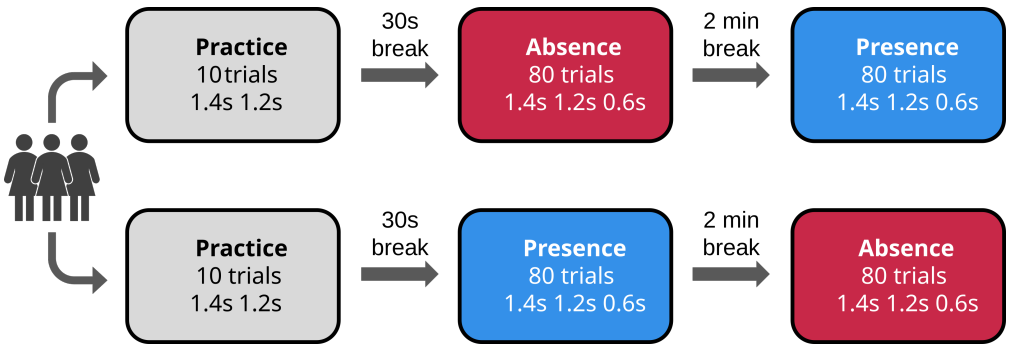

Figure 2.8.: Timeline of the training and condition (experimental) blocks within the lateral interception task. During each session, participants were assigned to one of the two counter-balanced block sequences starting with practice trials, which was followed by their pre-assigned sequence (presence-absence, or absence-presence).

(TMSi, Enschede, The Netherlands). The ground electrode was placed on the left wrist of the participant, and an online average reference was applied. The following steps were followed for each participant in order to retain consistency across different recording sessions: Skull measurements were taken, the Cz electrode was placed at the intersection of the two midlines going from inion and nasion, and from the tragus of each ear. The sampling frequency of the EEG signals were set at 2048 Hz, impedance of electrodes was controlled to be under 10 $k\Omega$, and shielded cables (TMSi) were employed in order to reduce cable movement artifacts as much as possible. Triggers were sent to the docking station using a custom-developed system, with each trigger marking the onset of an event such as the moment when the ball crossed the interception axis (in the case of failed trials) or made contact with the paddle (i.e. successful trials).

Each participant was randomly designated to one of the counter-balanced groups. As mentioned before, the lateral interception task involved intercepting a downward-moving virtual ball (radius = 16 px) on the screen using a virtual paddle (width = 48 px, height = 13 px). Participants needed to place their paddle within a red rectangle situated at the bottom-middle of the for the duration of 100 frames in order to initiate each trial. Upon doing so, the box would turn green, indicating the beginning of that trial and the subsequent appearance of the ball at the top of the screen.

With the aim of keeping the task relatively simple to undertake, the ball velocities and trajectories were manipulated such that the interception rate of participants would range between 80 and 85 percent. The total duration of each trial depended on that trial's ball velocity, which were set such that trials would last 1.4 seconds, 1.2 seconds, or 0.6 seconds. The ball departure positions (BDP) at the onset of each trial had a fixed Y-coordinate of 501, and five possible positions along the X-axis: -672, -336, 0, 336, 672. Likewise, ball arrival positions (BAP) had a fixed Y-coordinate of 501, and five possible arrival positions along the X-axis: -672, -336, 0, 336, 672. A random

X-offset was added to the BDP and BAP of each trial to preclude over-familiarization of participants to ball trajectories. Every experimental session started with a 10-trial practice block in order to familiarize participants with the general structure of the task, followed by two 80-trial experimental blocks (presence-absence, or absence-presence). Participants encountered only interceptable trajectories of 1.4 seconds and 1.2 seconds in the practice block, and 36 trajectories of 1.4 seconds and 1.2 seconds (each), as well as 8 fast balls (0.6 seconds) that were often un-interceptable within the experimental blocks. These fast balls were added in order to maintain and maximize interest and attention levels of participants during the task (Figure 2.8).

Prior to the start of the experiment, the EEG caps were fitted to each participant's individual head size. During the process of applying conductive gel to the electrodes and securing the EEG cap, the experimenter began introducing a cover story, mentioning potential slider defects and the possibility of someone entering the room to 'read the slider resistance'. Before the actual experiment commenced, the experimenter informed the participant that she would leave the room 'to look for technical support to test the equipment'. Participants were informed that they should continue the experiment unremitted, as it would be entirely conducted without interventions from the experimenter. In the instances where the first block was the 'presence' condition (followed by an absence block), a 30-second break was programmed between practice block and the first experimental block to allow the spectator to enter the room in a timely manner. In addition, an automatic two minute break between the experimental blocks enabled the spectator to enter/leave the room. The absence/presence condition blocks were counter-balanced, which meant the spectator would enter the room during the first block for half of the participant pool, and during the second block for the other half (Figure 2.8). For both female and male subject groups, a pre-determined –and unchanging– spectator of the same sex was assigned. Participants were instructed to never communicate or interact with the spectator during the experiment, since it would introduce unwanted noise to the recorded EEG data.

2.2.3. Movement data

Movement trials were generated by time-locking to the feedback events, comprising of a 700 ms time window preceding said events. This window was chosen as it was sufficiently long to cover the most significant sections of paddle movements (up until the interception threshold), but precluded the initial flat section of paddle dynamics from the beginning of the trial until the initiation of ball movement. Movement trials were low-pass filtered (5 Hz of $order = 2$), normalized with respect to ball arrival positions and subsequently converted to absolute values to ensure uniform computation of paddle speeds. Finally, artefactual trials with mismatched starting position were not included in the subsequent analyses.

Previous studies on social facilitation have outlined the importance of using kinematic measures of performance in motor tasks, such as movement speed/duration –and features thereof– as opposed to outcome-based metrics such as accuracy or hit-rate [129].

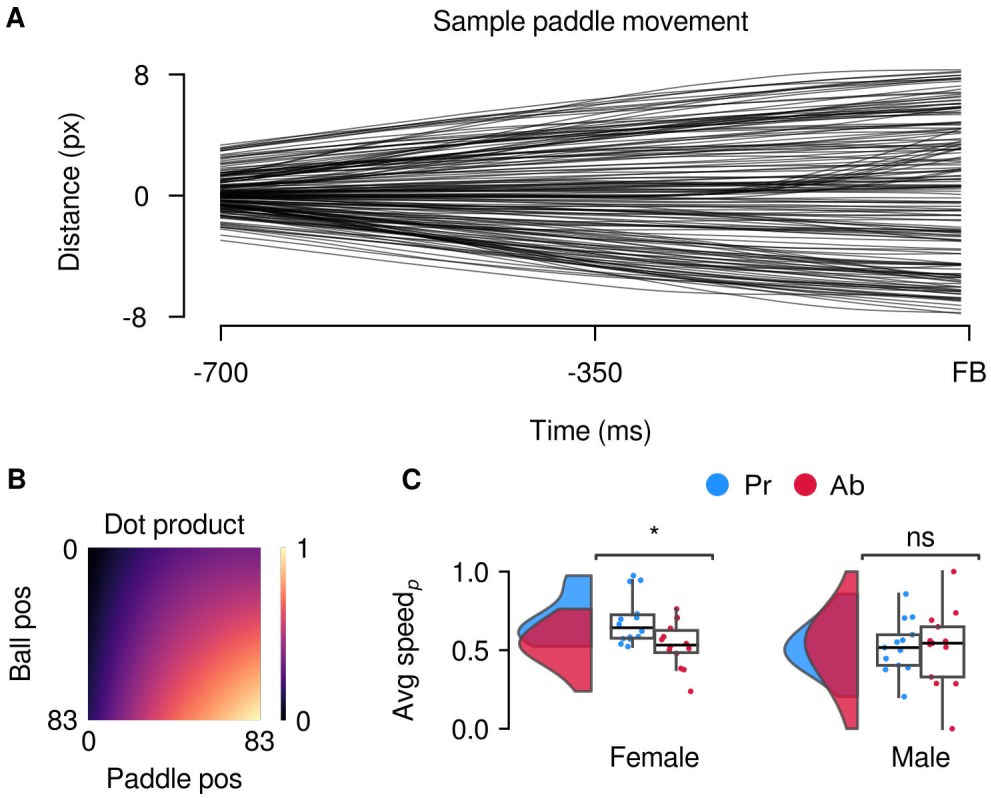

Figure 2.9.: Characteristics of the movement data. **A**) Paddle movement trials of one participant after time-locking to feedback events. **B**) Dot product (i.e. covariance) of ball and horizontal paddle trajectories across all trials. Note the maximization of covariance between the position of the ball and the paddle (i.e. player) towards the ends of each trial. **C**) Average paddle speed across conditions and subject groups. Presence of others significantly increases the performance in the female subject group ($p < 0.05$), compared to the male subject group ($p > 0.05$).

Therefore, we analyzed average and peak paddle velocity within a 700 ms window (half-width of the longest trials) before the feedback time (Figure 2.9**A**). Participants learned the overall aim of the task rapidly, and were subsequently very consistent in pursuing the ball within the experimental blocks (Figure 2.9**B**). We found that average paddle speed was significantly higher in the female subject group in the presence condition compared to absence ($p < 0.05$; Figure 2.9**C**, left panel), whereas there was no significant difference between the experimental conditions in the male subject group ($p > 0.05$; Figure 2.9**C**, right panel), indicating a social facilitation effect for one subject group only.

2.2.4. Electrophysiological recordings

The results of manual preprocessing of electrophysiological recordings can be strongly influenced by the subjective choices or biases of its conductors [43, 154]. Thus, we

devised an agnostic and automatic pipeline in order to minimize such biases. Preprocessing was done entirely via the MNE-python package [77, 78, 113].

In each subject, bad EEG channels were first identified and interpolated via the Ransac module of the AutoReject package [98], with input epochs from -500 ms to 1000 ms after the movement onset (with a baseline window of -500 ms to -200 ms). Average referencing was subsequently used in order to re-reference the channels. Independent component analysis (ICA; $method$ ='fastica') was done in order to project out the blinks and saccades components from the recordings; data were first band-passed from 1 to 46 Hz via a FIR filter and uniformly epoched. A computed rejection threshold in addition to these equally-sized epochs were used as inputs to the ICA method. We chose the first participant ($p01$) to compute the common templates for saccade/blink components. Subsequently, template-matching was ran over all participants during which these components were projected out of the data ($threshold = 0.9$).

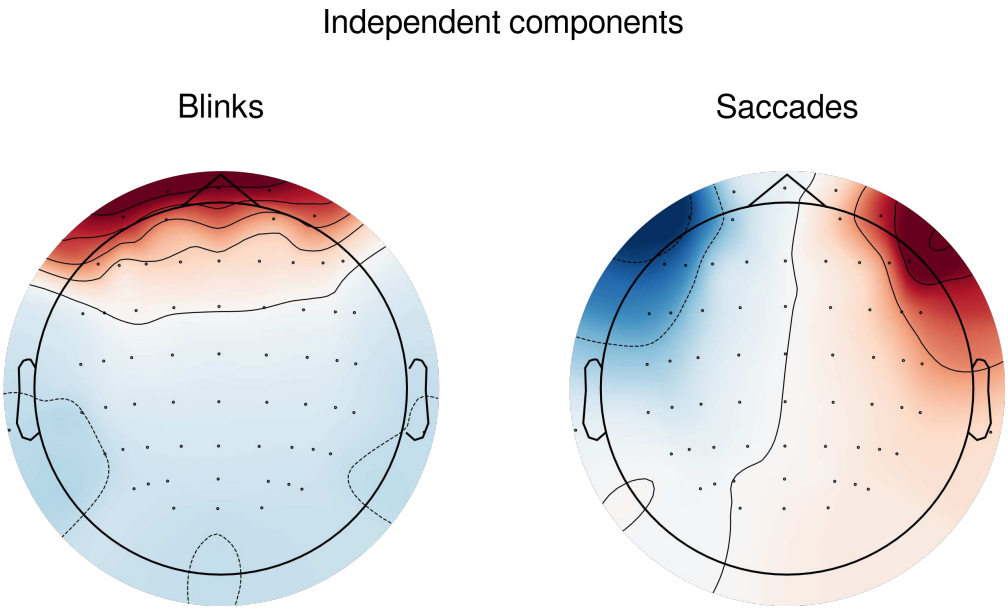

Figure 2.10.: Eye movements artifacts during the recording of EEG data introduce strong artifacts which can severely reduce the reliability of source estimation. ICA alleviates this issue by decomposing the signal into independent time-dependent components, a subset of which can be chosen to be projected out of the data (i.e. removed). For our task, precomputed independent component templates obtained from a single subject were used to remove saccade and eye-blink artifacts from the EEG data of all participants.

Inferring the activity of functional brain networks involves source-localization, as EEG simply measures voltages across the scalp [171]. In the absence of per-participant structural scans, source-localization was conducted via the freesurfer [45] template 'fsaverage'. It should be noted that even with per-participant structural imaging, between-condition or between-group comparisons of source activity should opti-

mally be conducted after morphing of the structural data of each participant to a representative 'template' [181]. Calculation of the forward solution ($mindist = 5mm$) was done with the help of the boundary-element model ('fsaverage-5120-5120-5120-bem-sol') and the source space model 'fsaverage-ico-5-src'). The noise covariance was computed via broadband epochs similar to those in the identification of bad channels (only with a $tmax = -0.2$ ms). With the forward solution and the noise covariance in hand, the inverse operator was generated which allowed for the extraction of source time-courses (regularization parameter $loose = 0.2$).

Oscillatory rhythms have been famously linked to different aspects of cognition. In this study, the analysis of EEG data were restricted to the alpha band, as not only the majority of the participants exhibited a pronounced alpha peak during the task, but also due to its reportedly heavy involvement in attentional modulation (Figure 2.11). Alpha-band synchronization is indeed widely recognized as a crucial mechanism in attentional modulation, particularly within visuospatial attention. Studies have demonstrated that alpha oscillations are modulated by the orientation of attention and associated with reduced reaction times to attended stimuli [39, 115]. Research also indicates that attention drives synchronization of alpha rhythm in regions such as the right inferior frontal cortex and primary sensory neocortex [158]. Finally, studies have shown that alpha-band phase can modulate bottom-up feature processing [99, 119].

Hence, raw data were band-passed to the alpha band (8-12 Hz, multitaper), and epoched to a 500 ms time window preceding the ball movement (with a baseline from -200 to 0 which was cropped before saving the epochs). This specific window was chosen since attentional modulation should occur closer to the visual cue, and in advance of any behavioral modulations (Figure 2.12**A**).

The precomputed inverse operator was then employed (via the MNE algorithm) to extract source activity in vertex format. Cortical activity in this format is extremely high-dimensional (tens of thousands of vertices), and hence not efficient or easy to analyze functional activation of various brain network. Therefore, the Schaefer 2018 atlas (400 parcels) was chosen for extracting source time-courses from seven functional brain networks [199, 161]. These networks comprise dorsal attention network (DAN), ventral attention network (VAN), somatomotor network (SMN), visual network (VIS), frontoparietal network (FPN), default-mode network (DMN), and limbic network (LIM) (Figure 2.12**B**). The activity from the limbic network was included in the computations purely for thoroughness, and was not included in the interpretation of any analyses, as the definition and function of this network is still highly controversial [72].

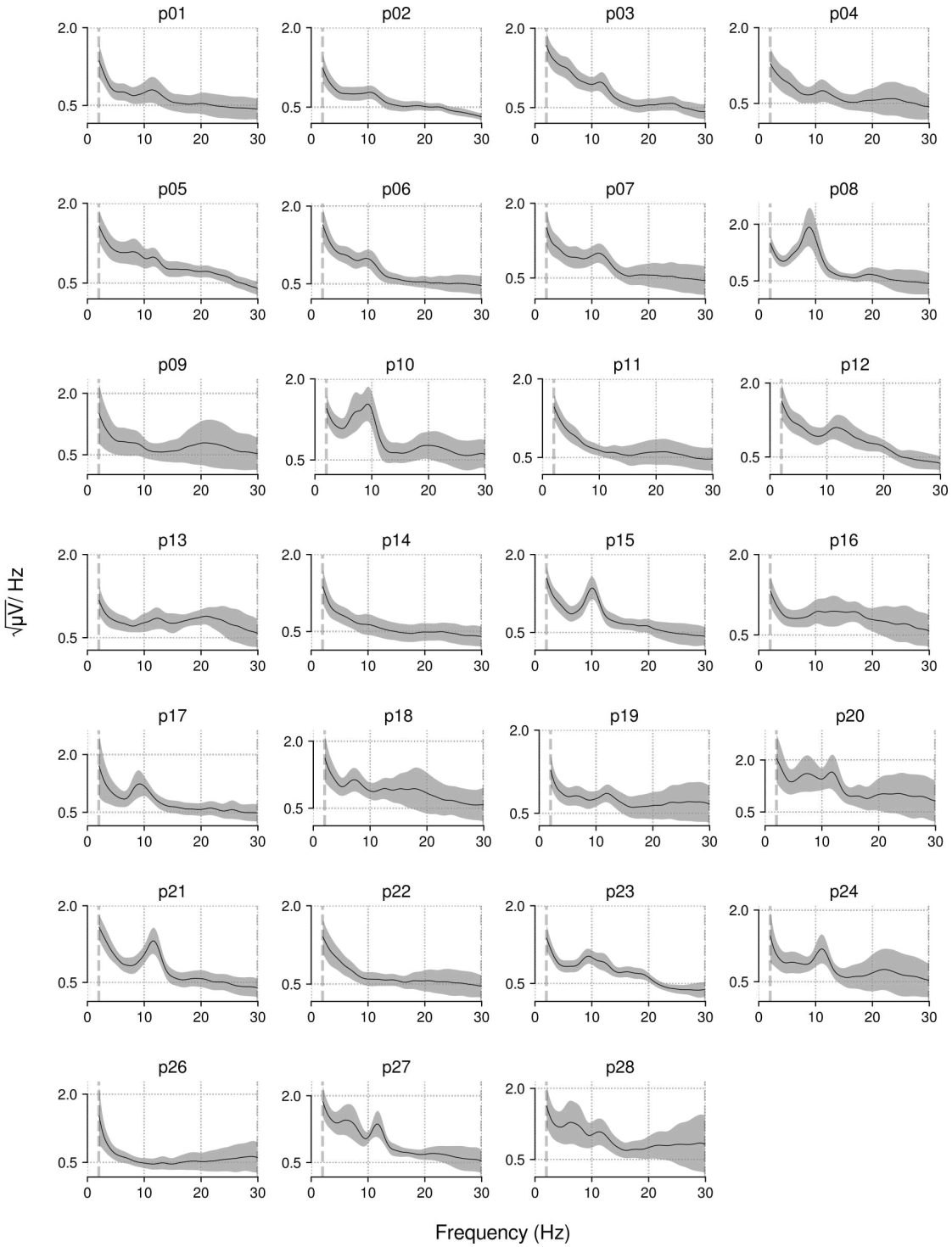

Figure 2.11.: Power spectra of the analyzed participants in the lateral interception task. The black lines and shaded regions depict the average and standard deviation of power spectra across electrodes, respectively. The pronounced alpha band (10 Hz) peak present in the majority of subjects suggests the importance of these oscillations for task performance.

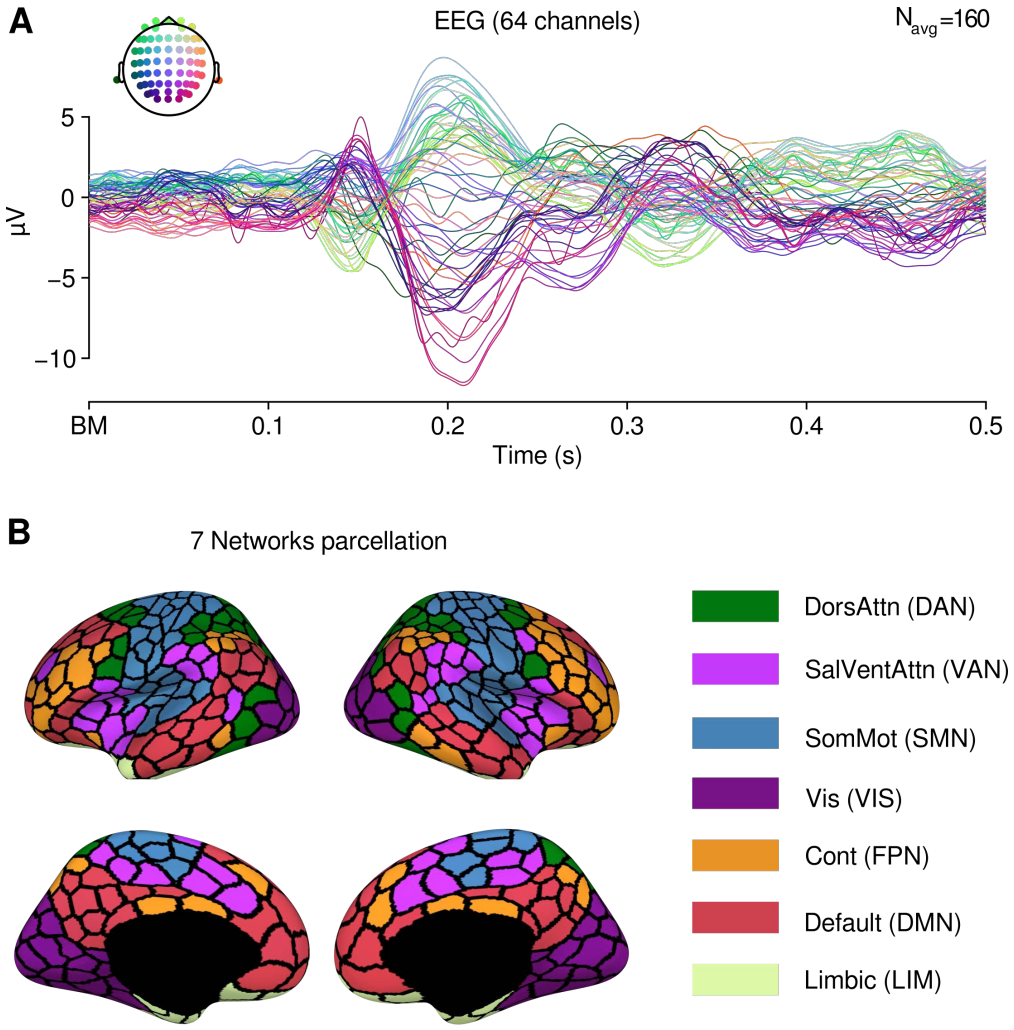

Figure 2.12.: **A**) Evoked activity of one participant, acquired by time-locking to a 500 ms window starting from the ball movement. **B**) Cortical parcellation (400 parcels) used in order to extract the source time-courses of the 7 functional networks based on the Schaefer atlas.

2.2.5. Functional connectivity

There exist myriad measure of computing coherence across neural regions. Among such measures, functional connectivity (FC) has been shown to be a reliable proxy for synchronization not only in resting-state activity, but also task-oriented ones [64]. Functional connectivity revolves around estimating the coordination of activity between the nodes of a network, which reportedly serves as a unique identifier of focused attention [128]. An apt analogy would be to consider an orchestral performance where one metric of the performance could be the 'volume' of all instruments, as measured

by power analysis. However, coherence of the instruments (i.e. correlation of activity) would allow one to understand if the instruments are in tune (and in time) with each other, which is pivotal for the production of "cohesive" orchestral sound.

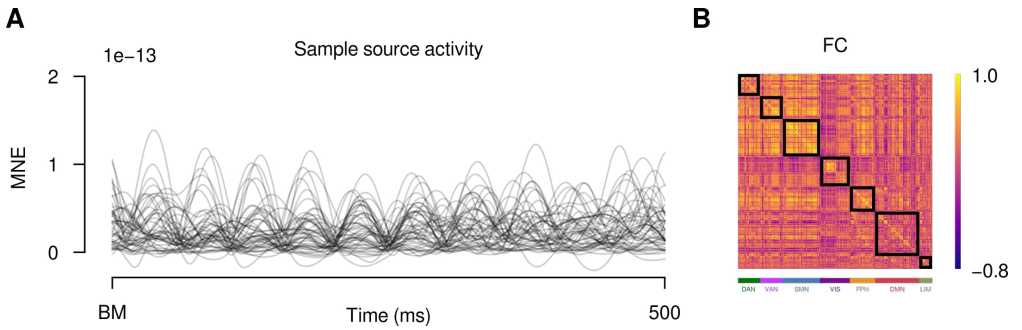

Figure 2.13.: **A**) Source time-courses within a 500 ms window starting from the ball movement. **B**) Sample source functional connectivity of a participant obtained via the Schaefer atlas.

FC can be computed as the –Pearson– correlation across the elements of time series arrays resulting in an n_{elem} by n_{elem} square matrix (with n_{elem} representing the number of nodes/regions in the time series array). Formally, in the correlation matrix FC, the element FC_{ij} depicts the statistical dependency between two regions/elements i and j, which is evaluated as $FC_{ij} = \frac{C_{ij}}{\sqrt{C_{ii} * C_{jj}}}$, where C_{ij} is represents the covariance between said labels, and C_{ii} represents the variance of i [63]. The elements of this matrix are then subsequently sorted and isolated into masks in accordance with the chosen functional networks. Elements of the left and right hemispheres –of each network– are concatenated together in order to obtain the overall activity of each network (Figure 2.13).

2.2.6. Measures of integration

As previously mentioned, functional connectivity can provide us with an invaluable metric for approximating the coherence of neural elements. The pairwise correlation of neural elements however leads to each FC matrix comprising hundreds of elements, even in the smallest of parcellations. Raw matrices of functional connectivity, therefore, do not provide one with a convenient or directly informative summary statistic for comparing across conditions/subjects. One widely used FC feature for summarizing network-wide influences is integration/segregation ratio. Integration in this context is defined as the efficacy of the communication of nodes inside a network with each other, and segregation as the efficacy of communication of the same network nodes with the rest of the nodes inside the system (i.e. brain) [29].

With the aim of comparing the strength of communications across *elements* of functional networks (instead of coherence across networks), we can discard the second half of the above computations in order to arrive a measure for –network– integration.

This coherence measure is supported by previous findings. For instance, the intrinsic connectivity within attentional networks is said to alter the strength of distractor suppression [149]. Integration can be formally defined as the sum of the upper triangular mask of each network: $Integration = \sum FC_{network}$. Despite integration being an empirically-driven measure, it could serve a suitable proxy for alterations of effective connectivity within a network [179]. It would therefore be reasonable to posit increased activity coherence of attentional networks (i.e. dorsal attention network) during the presence condition, should attention-oriented brain regions indeed play a significant role in social facilitation [44].

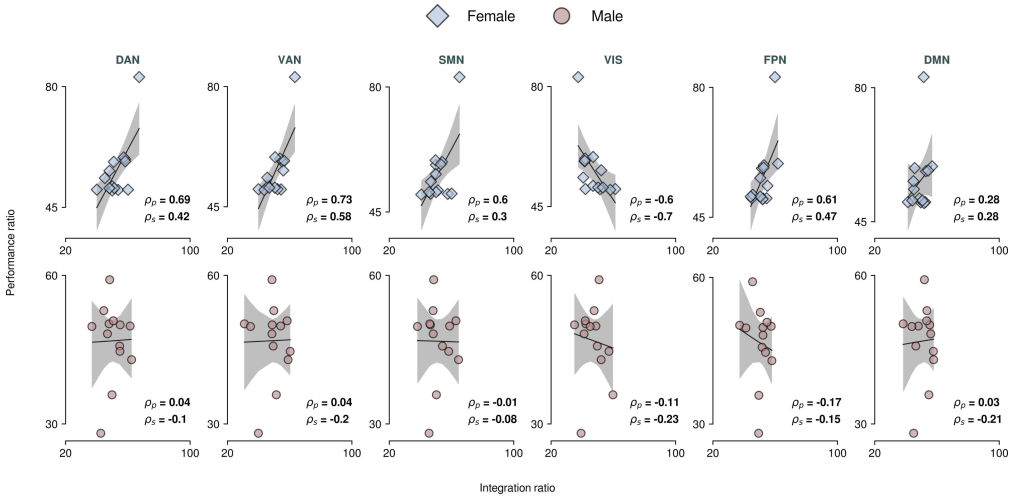

Figure 2.14.: Correlation of between-condition ratios of integration values with task performance in both subject groups. The uncertainty cloud stems from linearly bootstrapping the data-points, with the black line representing the mean of this deviation (i.e. roughly corresponding to the Pearson correlation values).

We therefore measured the between-condition ratios of integration and task performance across all networks. Since the small sample size might influence statistical measures due to outlier effects, all linear –Pearson– correlations were corroborated via the Spearman ranked test. The results of this analysis also revealed a group-based pattern; In the female subject group, integration was shown to be strongly correlated with task performance in attentional networks such as such as DAN, VAN, and FPN, compared to other networks ($\rho_{Pearson} = 0.69$, $\rho_{Spearman} = 0.42$ for DAN, $\rho_{Pearson} = 0.73$, $\rho_{Spearman} = 0.58$ for VAN, and $\rho_{Pearson} = 0.61$, $\rho_{Spearman} = 0.47$ for FPN Figure 2.14 **top**). The SMN also showed a high but linear correlation with behavioral performance, which was significantly diminished via the Spearman test ($\rho_{Pearson} = 0.6$, $\rho_{Spearman} = 0.3$). Conversely, a similar correlation analysis within the male subject group revealed little to no relation between integration values and performance (Figure 2.14 **bottom**).

In order to ensure the above results did not simply originate due to noise, we measured the signal-to-noise ratio (SNR) across all functional networks. As evident

from (Figure A6), SNR is consistent across networks and conditions, thereby verifying the authenticity of integration measures. We have therefore derived our sought-after informative summary statistics from the empirical data across all three scales, and can turn to data generative processes constituting the forward model in the following chapter (chapter 3).

3. Data-generative processes

Summary

3.1. Brunel balanced network 61
3.2. Jansen-Rit neural mass model 66
3.3. Whole-brain Jansen-Rit network 69

The neurophysiological data presented in the previous chapter comprise three distinct spatiotemporal scales, namely, microscale (single-neurons), mesoscale (cortical micro-columns), and macroscale (whole-brain). As previously mentioned, our ultimate aim in this endeavor revolves around examining how others' presence influences synaptic efficacy. Extra vigilance must therefore be retained in selecting the data-generative processes (i.e. theoretical models) that faithfully represent said scales with regards to modulations of effective connectivity. In the following, I will depict the chosen models for said three scales in order to lay the groundwork for conducting the probabilistic inference of empirical data in the next chapter (Figure 3.1).

3.1. Brunel balanced network

Single-neurons comprise the first spatiotemporal scale of our empirical data, which are typically modeled via myriad implementations of spiking models. Spiking neuron models are a class of dynamical systems employed for simulating the behavior of individual neurons. In other words, they explicitly capture the temporal dynamics of action potentials – brief electrical impulses that serve as the fundamental unit of communication in the nervous system [94]. Spiking models can be broadly categorized into phenomenological and biophysical models. Biophysical models aim for –reasonable– realism by explicitly incorporating the complex interplay of ion channels responsible for generating and shaping action potentials. While biophysical models offer a more detailed representation of neuronal function, they often come with increased complexity and computational demands. Indeed, even the most sophisticated of spiking models are simplifications of reality, as they capture essential features of neuronal activity but cannot fully replicate the complexity of biological neurons, which possess myriad ion channels, intensely complex arborization, and pivotal glial communications. Phenomenological models, on the other hand, focus on capturing essential features of neuronal firing behavior without delving into the intricate details of ionic currents underlying action potentials. These models provide a valuable alternative by capturing key dynamics with relative simplicity, making them suitable for large-scale simulations and exploring emergent network properties [94]. Although the choice between these approaches depends on the specific research question and the level of detail required, spiking neuron models remain invaluable tools for understanding the fundamental principles governing brain function and provide a crucial foundation for building more complex network simulations.

DCM of spiking networks typically involve –at the minimum– tens of thousands of simulations with each one possessing neuron counts of a similar scale. This precludes the possibility of more biophysically-oriented implementations such as Hodgkin-Huxley neurons in favor of ones that are more computationally efficient, such as the Leaky Integrate-and-Fire (LIF) neuron. Introduced more than a century ago, LIF neurons derive their name from their fundamental mechanism, in which the membrane potential of a neuron continuously integrates incoming synaptic inputs, while simultaneously 'leaking' away potential through a passive conductance [112, 186]. When the integrated potential surpasses a predetermined threshold, the neuron is said to fire an 'imaginary' action potential, resetting the membrane potential and triggering a brief refractory period during which it is less excitable, effectively limiting the frequency at which it can fire. This period plays a crucial role in shaping the temporal patterns of neuronal firing and preventing continuous spiking. It is paramount to note that while LIF neurons are typically bundled with other spiking models (within the context of this work and in the literature), they do not possess genuine spike generation dynamics [186, 94]. In truth, the rise and fall phases of the spike are instantaneous, and are not explicitly modeled via differential equations, hence the previous use of the term 'imaginary'.

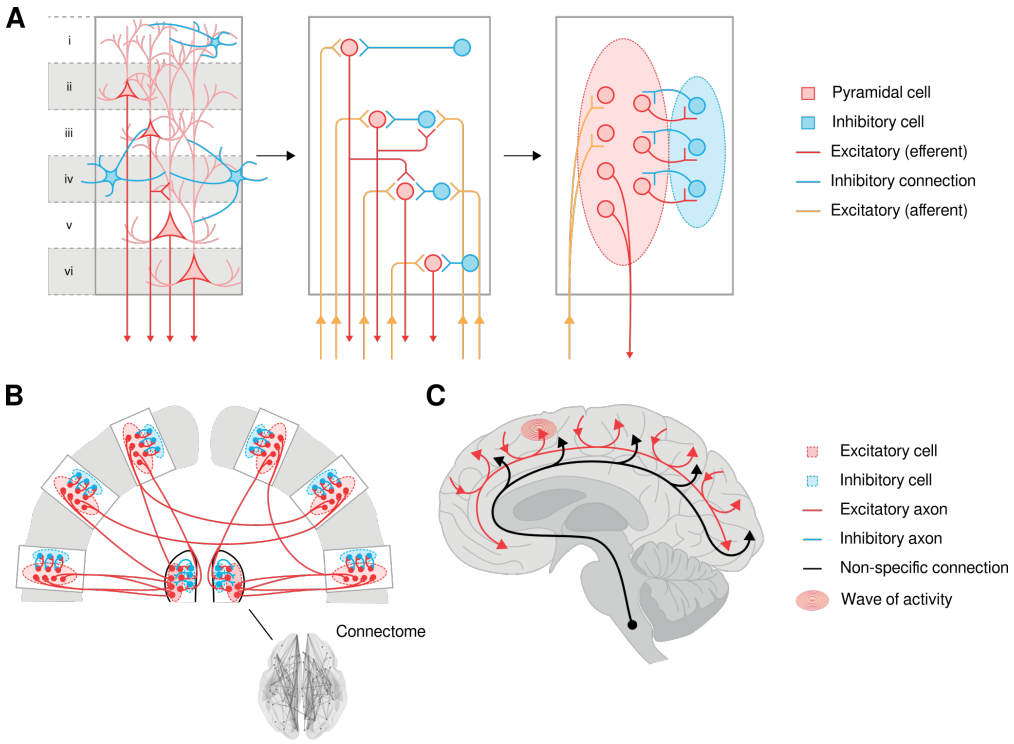

Figure 3.1.: Schema of common models in computational neuroscience. **A**) Left panel: Anatomical depiction of cell types within a cortical micro-column. Middle panel: Spiking networks are constructed by coupling numerous instances of single-cell models (e.g. Pyramidal neurons, interneurons, and excitatory interneurons) together. Right panel: During the process of deriving neural mass models (via mean-field approximations), the majority of neurobiological characteristics of a biological neural network (e.g. cell type, internal/external connections, and anatomical organization) are mathematically reduced (i.e. abstracted). **B**) Whole-brain model of coupled neural mass models connected via structural connectivity. **C**) Neural field models are another variant of whole-brain models, and represent cortex as a sheet of coupled vertices, thereby allowing for smooth propagation of neural activity. Figure adapted from [19].

The LIF neuron's dynamics centers around a single first-order differential equation:

$$\frac{dV_{\mathrm{m}}}{dt} = -\frac{V_{\mathrm{m}} - E_{\mathrm{L}}}{\tau_{\mathrm{m}}} + \frac{I_{\mathrm{total}}}{C_{\mathrm{m}}} \tag{3.1}$$

Where V represents the membrane potential, t signifies time, τ_m is the membrane time constant reflecting the neuron's characteristic integration timescale, C_m represents the membrane capacitance governing the passive leak current, and I_{total} encapsulates the sum of all inputs received by the neuron. This term typically comprises excitatory postsynaptic currents (EPSPs) arising from connections with presynaptic neurons releasing excitatory neurotransmitters, and inhibitory postsynaptic currents

(IPSPs) resulting from connections with presynaptic neurons releasing inhibitory neurotransmitters. The precise balance between EPSPs and IPSPs significantly influences the membrane potential dynamics and ultimately dictates the firing behavior of the LIF neuron. The computational efficiency of the LIF allows one to investigate the influence of synaptic input convergence on neuronal firing via highly scalable neural network simulations.

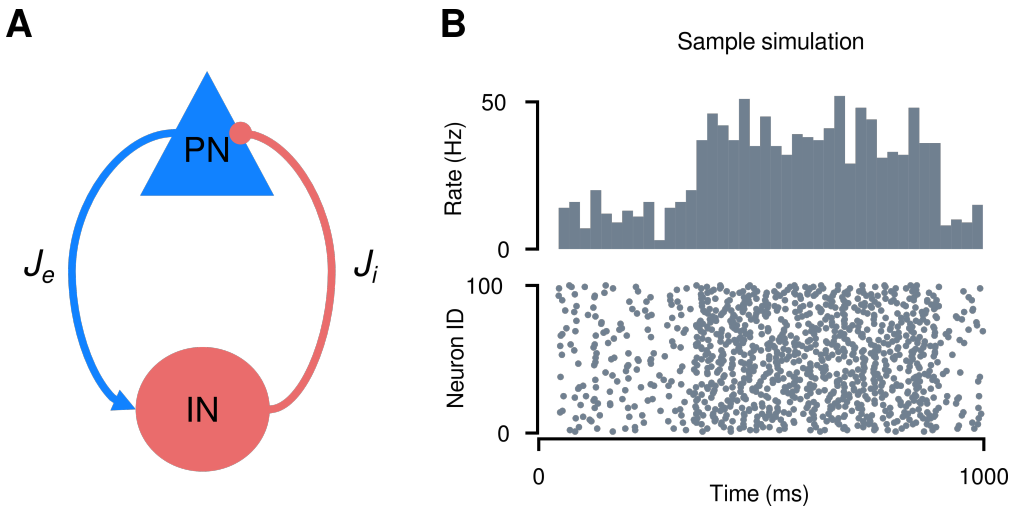

Figure 3.2.: **A**) Schema of the Brunel balanced spiking network. Excitatory and inhibitory neurons are portrayed as a single node for simplicity. **B**) A sample simulation raster plot (bottom) and spike histogram (top) of the Brunel network.

Having a single-neuron model is only the first step of construction of a biologically-tuned neural network. For our purposes, the balanced spiking network of Brunel stands out the most due to its particular connectivity architecture (Figure 3.2**A**, [166]). The concept of balanced networks stems from the observation that biological neural populations heavily maintain an equilibrium between excitation and inhibition, despite experiencing significant stochasticity of inputs and and dynamics [21], which suggests the existence of a fundamental mechanism for retaining coherent neural communications even in the presence of variability. Indeed, excitatory/inhibitory (E/I) balance reportedly bears a pivotal role in shaping cortical network properties [185]. In addition, balanced networks exhibit –and mimic– the inherent noise-resilience of biological systems, which is crucial for reliable signal propagation and perception in real-world environments with highly variable sensory inputs [166].

Neurons in the Brunel neural network are categorized into two subpopulations: excitatory and inhibitory leaky integrate-and-fire (LIF) neurons, which are randomly and sparsely coupled [21]. With regards to synaptic transmission, I have chosen to employ Alpha synapses which are not instantaneous but instead exhibit more biologically plausible temporal dynamics. Alpha synapses exhibit specific time constants for the rise and decay of their postsynaptic potentials (PSPs), which directly influences the temporal integration of synaptic inputs: Faster rise time constants allow for quicker

responses to presynaptic spikes, enhancing the impact of recent inputs on postsynaptic activity. Longer decay time constants prolong the duration of synaptic influence, allowing for a more cumulative effect from past inputs and contributing to persistent firing patterns. Each network simulations comprises one excitatory ($n = 10000$) and one inhibitory ($n = 2500$) subpopulation, with a constant 4:1 cell count ratio for biological plausibility [21]. Figure 3.2 **B** illustrates an example simulation of this network.

Connection strength among said subpopulations is denoted by $-g$, where the parameter g represents the ratio of inhibitory to excitatory weights across the entire network; An increase in g results in reduced overall activity of the network, which means it could be used as a measure of synaptic efficacy. This simple but effective implementation allows us to examine how changes in synaptic efficacy within a neural subpopulation can contribute to modulations of firing rate.

$$\begin{aligned}
&\frac{dV_\mathrm{m}}{dt} = -\frac{V_\mathrm{m} - E_\mathrm{L}}{\tau_\mathrm{m}} + \frac{I_\mathrm{syn} + I_\mathrm{ext} + I_\mathrm{stim}}{C_\mathrm{m}} \\
&I_{syn}(t) = I_{syn,e}(t) + I_{syn,i}(t) \\
&I_{syn,x}(t) = \sum_n J_n \sum_k i_{syn,x}(t - t_j^k - d_j) \\
&J_i = -gJ_e \\
&i_{syn,x}(t) = \frac{e}{\tau_{syn,x}} t e^{-\frac{t}{\tau_{syn,x}}} \Theta(t).
\end{aligned} \tag{3.2}$$

If $V_\mathrm{m}(t_k) < V_{th}$ and $V_\mathrm{m}(t_{k+1}) \geq V_\mathrm{th}$, then a spike is said to be discharged at time-step $t^* = t_{k+1}$, followed by resetting the membrane potential to $V_\mathrm{m}(t) = V_\mathrm{reset}$ for $t^* \leq t < t^* + t_\mathrm{ref}$.

Understanding the function of these parameters is important in order to create an empirically-tuned instance of the Brunel network. However, as the sole aim here is to focus on the effect of synaptic efficacy on excitability, we fix every parameter with the exception of g, but will provide a brief explanation of some: The firing threshold V_{th} determines the membrane potential value a neuron must reach to 'fire' an action potential; higher thresholds necessitate stronger depolarization before firing and thus reduce the overall excitability. E_L represents resting membrane potential when no input is received, meaning that a more negative E_L can make neurons less excitable which necessitates stronger synaptic inputs to reach V_{th}. The membrane time constant τ_m governs how quickly the membrane potential returns to its resting state after perturbation by synaptic inputs and is equally important. Smaller values of this parameter translate to faster decay and make neurons more responsive to rapid changes in input [21, 166]. Synaptic time constants are equally pivotal as they shape the temporal integration of synaptic inputs received by a neuron [21, 166].

On the formulation of synaptic currents $I_{syn,x}$, subscript x depicts either inhibitory (i) or excitatory (e) synapses. Both synapse variants share a single time constant τ_{syn}, and the neurons within either subpopulation are indexed with n. Considering a

neuron with index n, symbols d_j and k represent the delay from said neuron, and its spike time indices, respectively. Table 3.1 exhibits the rest of the parameters and their values:

Parameter	Representation	Value
τ_{syn}	Synaptic time constant	0.5 ms
τ_m	Membrane time constant	20.0 ms
τ_{ref}	Duration of refractory period	2.0 ms
C_m	Membrane capacitance	250.0 pF
V_{th}	Spike threshold	20.0 mV
E_L	Resting membrane potential	0.0 mV
V_r	Reset potential of the membrane	10.0 mV
V_m	Membrane potential	0.0 mV

Table 3.1.: Parameters for LIF neurons.

The external current I_{ext} was set in such a way as to approximate the median baseline activity of neurons in the empirical data, and was thus modeled as a Poisson generator with a rate of p_{rate} = 13341.8 Hz. Finally, elevation of the network's firing rate due to feedback anticipation was crucial in rigorously reconstructing empirical patterns. A step-current I_{stim} with a mean of 150 pA and standard deviation of 1 pA was applied to the network of neurons from 350 ms to 900 ms. Each neuron received a different variation of the current based on its standard deviation: $I_{stim}(t) = \mu + \sigma N_w$. For time-steps within the activation interval of the current w, N_w is a sample drawn from the zero-mean unit-variance normal distribution.

3.2. Jansen-Rit neural mass model

Evoked potentials from prefrontal and anterior cingulate cortices comprise the second spatiotemporal scale of the empirical data. ERPs measure the overall synaptic activity of entire neural populations, rendering them insufficient for establishing a direct link to the effectiveness of synaptic communications between different neural subpopulations. Our choice of dynamical model of evoked activity should therefore possess reasonably realistic synaptic dynamics for establishing a link between changes in effective connectivity and modulations of ERP amplitude. Despite the recent advents in high-performance computing however, simulations of cortical regions are still a far cry from being comprised of hundreds of millions of interconnected and biophysical spiking neurons. In addition, our understanding of the neurobiological underpinnings of social cognition in humans barely reaches a fraction of our understanding of similar phenomena in primates and rodent models [132, 15]. Thus, it would stand to reason that current simulations of cortical activity maintain a reasonable balance between computational and biological plausibility.

Enter neural mass models (NMM), which represent a simplified approach to understanding the complex dynamics of neuronal populations within cortical regions [52, 133]. These models conceptualize a cortical region as a "mass" of neurons, characterized by a set of average firing rates that evolve over time. Forgoing the explicit modeling of individual neurons and their intricate synaptic connections, neural mass models focus on capturing the emergent properties arising from the collective behavior of these neuronal populations, which in turn allows for computationally efficient simulations of large-scale regional and network dynamics. The derivation of NMMs relies on mean-field theory, a mathematical framework used to describe the collective behavior of large systems with interacting components [52, 133].

Typically, mean-field derivations of NMMs involve defining a set of state variables which represent the average firing rate and average membrane potential of each neuronal population within the model, and constructing a system of ordinary differential equations (ODEs) that describe how these state variables evolve over time. These ODEs incorporate terms representing synaptic inputs, as well as intrinsic properties of neurons such as their firing thresholds and time constants. In sum, the mean-field approximation comes into play when we replace the complex interactions between individual neurons with an average influence. By solving these equations numerically, we obtain a dynamic representation of the activity patterns within the modeled cortical region [52]. This simplified approach to modeling allows researchers to focus on how the macroscopic properties of the neuronal population such as synaptic strengths or intrinsic neuronal properties, influence the overall behavior of the system, without getting bogged down in the unique characteristics of each neuron or synapse.

A suitable neural mass model can indeed offer an empirically-grounded method for delineating the synaptic characteristics within a cortical micro-column from evoked potentials. There exist a –figurative– ocean of previously implemented NMMs which might seemingly complicate the task of choosing an appropriate model [133]. While a certain degree of reduction is –by definition– present within all NMM implementa-

tions, many are reduced beyond biological recognition such as the Kuramoto model [2]. That is, the majority of neural mass models do not possess the granularity of synaptic representation necessary for the purpose of this work.

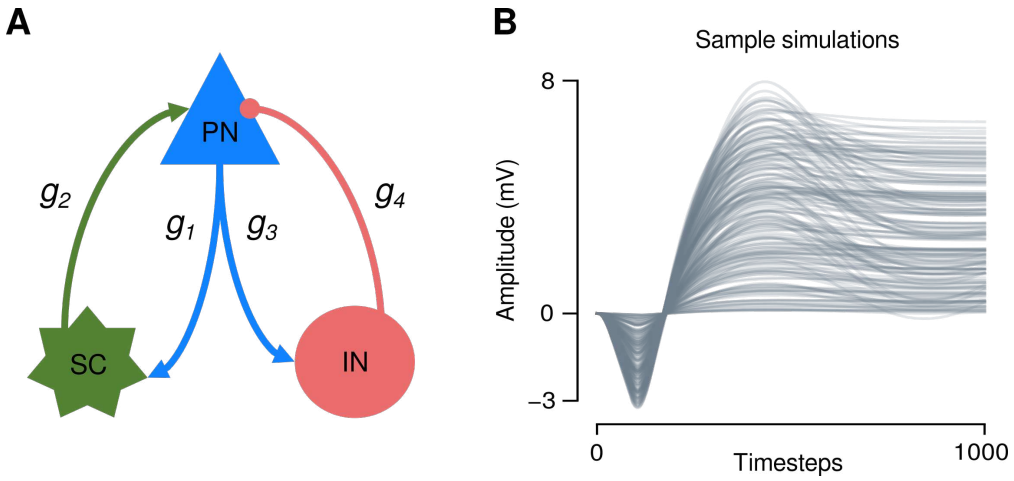

Figure 3.3.: **A**) Schematic illustration of the Jansen-Rit neural mass model. **B**) Simulations of evoked potentials generated from the Jansen-Rit neural mass model.

Our model of choice is therefore a modified version of the Jansen-Rit (JR) neural mass model ([41, 42, 97], Figure 3.3). A tremendous body of work has illustrated the ability of this model in reproduction of the majority of dynamics observed in empirical EEG data [41, 42, 97]. Moreover, the derivation of David and colleagues is centered around rigorous generation of evoked potentials (in addition to tonic oscillations). The Jansen-Rit model represents a cortical micro-column as three coupled neural subpopulations: pyramidal neurons (PNs), interneurons (INs), and stellate cells (SCs), striking a reasonable balance between biological plausibility and computational efficiency (Figure 3.3).

The model comprises ten parameters as $g_{1,2,3,4}$ (synaptic efficacies), $\tau_{e/i}$ (membrane rate-constants), $h_{e/i}$ (post-synaptic potential maximum amplitude), δ (intrinsic delay), and u (external input). The latent states $x(t)$ within this NMM are based on nine ordinary differential equations which are first-order approximations of delay-differential equations (via $x(t-\delta) = x(t) - \delta\dot{x}(t)$). The most relevant of these states are the ones governing the dynamics of the pyramidal neurons x_9, spiny-stellate cells x_1, and inhibitory interneurons x_7 (Equation 3.3).

However, this adherence to plausibility comes with its own set of challenges, as increasingly granular implementations often allows one to generate virtually identical (or extremely similar) outputs by varying disparate –albeit interacting– system components. Formally referred to as degeneracy, this phenomena stems from the innumeracy of interactions or the redundancy of components (e.g. amino-acid codons) present within complex systems [54]. As evident from model equations, the increased biological plausibility of the Jansen-Rit model also leads to model degeneracy (Equation 3.3, Figure A3)

$$
\begin{aligned}
\frac{dx_1(t)}{dt} &= x_4(t) \\
\frac{dx_2(t)}{dt} &= x_5(t) \\
\frac{dx_3(t)}{dt} &= x_6(t) \\
\frac{dx_7(t)}{dt} &= x_8(t) \\
\frac{dx_4(t)}{dt} &= \frac{h_e(g_1(\frac{1}{e^{-0.56x_9(t-\delta)}+1} - 0.5) + u)}{\tau_e} - \frac{x_1(t)}{\tau_e^2} - \frac{2x_4(t)}{\tau_e} \\
\frac{dx_5(t)}{dt} &= \frac{h_e g_2(\frac{1}{e^{-0.56x_1(t-\delta)}+1} - 0.5)}{\tau_e} - \frac{x_2(t)}{\tau_e^2} - \frac{2x_5(t)}{\tau_e} \\
\frac{dx_6(t)}{dt} &= \frac{h_i g_4(\frac{1}{e^{-0.56x_7(t-\delta)}+1} - 0.5)}{\tau_i} - \frac{x_3(t)}{\tau_i^2} - \frac{2x_6^t}{\tau_i} \\
\frac{dx_8(t)}{dt} &= \frac{h_e g_3(\frac{1}{e^{-0.56x_9(t-\delta)}+1} - 0.5)}{\tau_e} - \frac{x_7(t)}{\tau_e^2} - \frac{2x_8(t)}{\tau_e} \\
\frac{dx_9(t)}{dt} &= x_5(t) - x_6(t)
\end{aligned}
\tag{3.3}
$$

In a nutshell, a single observable outcome, like an evoked potential, might result from diverse underlying mechanisms (e.g. by changing either h_e, or g_1 due to their multiplication within the differential equations), leading to multiple parameter sets fitting the data equally well. How can one ascertain the true causal drivers of any empirical observation in the face of biologically-plausible yet highly degenerate models? One of the most rigorous methods is to simply constrain the degeneracy through empirically-grounded premises (also referred to as domain expertise in the terminology of Bayesian inference). As my hypothesis wholly revolves around synaptic efficacy, I will fix the remainder of the parameters (such as PSP amplitudes and timescales), which allows for establishing a more reasonable link between alterations in ERP amplitude and SE. Table 3.2 exhibits the fixed parameters in the model.

Parameter	Representation	Value
δ	Intrinsic delay	8.41
τ_i	Inhibitory membrane rate constant	7.77
h_i	Inhibitory maximum PSP amplitude	27.87
τ_e	Excitatory membrane rate constant	5.77
h_e	Excitatory maximum PSP amplitude	1.63
u	External input strength	3.94

Table 3.2.: Parameters for ERP model.

3.3. Whole-brain Jansen-Rit network

The third and final spatiotemporal scale revolves around whole-brain dynamics. Whole-brain models move beyond the confines of individual neurons or localized cortical regions, strive to represent the intricate tapestry of interconnected neuronal populations spanning the entire brain, and provide us with a robust framework for understanding how large-scale brain function emerges from the nonlinear interactions of vast neural communities.

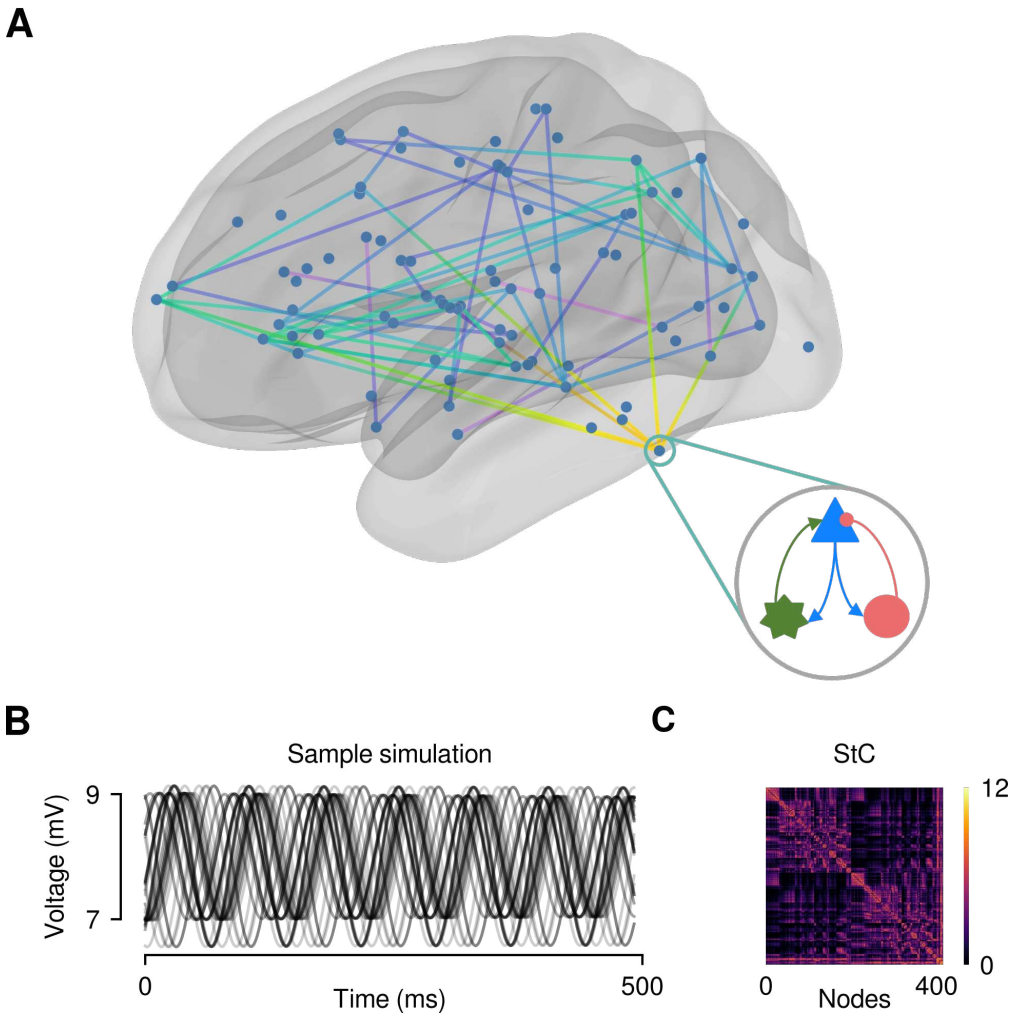

Figure 3.4.: **A**) Schematic of the macroscale model comprising Jansen-Rit NMMs coupled via structural connectivity. **B**) A simulation of the Jansen-Rit whole-brain network, which is used to compute synthetic measures of functional connectivity. **C**) Structural connectivity matrix used for macroscale simulations obtained from the Schaefer atlas [161].

The way by which these models are implemented however, can vary. Neural field models for instance, treat the brain as a continuous expanse of interconnected vertices represented by mathematical functions (Figure 3.1**C**). This approach captures oscillatory activity as a distributed field across the cortex, allowing for more nuanced representations of spatial patterns and propagation of information [145]. However, the increased spatial continuity of these models comes at the cost of significantly reduced biological plausibility [30].

A more prevalent approach is to employ neural masses as building blocks in the construction of the whole-brain model. This approach not only allows for inclusion of more plausible representations of brain regions, but it also enables researchers to simulate the interactions between different cortical and subcortical areas, exploring how information flows and patterns of activity emerge across the entire brain (and not just the cortices, in cases where subcortical representation is essential) (Figure 3.1**B**). Coupled NMMs have indeed been used in order to investigate how information flows between regions, how synchronization patterns emerge, and how these dynamics contribute to various cognitive processes such as attention and perception [105]. Most crucially, the concept of coupling modeling whole-brain dynamics via coupled oscillators is instrumental for establishing a reasonable link between neural scales.

The Jansen-Rit neural mass provides a valuable framework for understanding 'local' neuronal dynamics, though it falls short when attempting to capture the complex interplay between diverse brain regions that underpins higher cognitive functions. The granular architecture of coupled NMMs enables us to incorporate any potential finding from the mesoscale into the macroscale implementation. The nodes comprising whole-brain models can be coupled via a structural connectivity (StC) matrix, which is often obtained by incorporating anatomical information derived from diffusion tensor imaging (DTI) data. DTI allows one to map the white matter tracts connecting various brain areas, and provides a quantitative measure of the density and directionality of axonal fibers. Put another way, the weight assigned to each connection –between regions– reflects the strength of the anatomical pathway, thereby influencing the degree to which these regions influence each other's activity. Inspired by the connectome-oriented and network-oriented perspectives [160, 87, 101] found in previous research, I constructed a whole-brain JR network by coupling JR NMMs together via a representative structural connectivity matrix (Equation 3.4).

A similar atlas was used for generation of the structural connectivity matrix in order to match the parcellation used in order to obtain the functional connectivity matrices in the empirical data ([161], Figure 3.3**A**). Formally, the StC matrix was generated via tractography algorithms [163], and parcellations from the Schaefer Atlas (400 regions). Finally, the average/representative connectivity was computed by averaging per-subject StCs, the result of which was normalized via $StC_{norm} = \log(StC + 1)$ (Figure 3.3**B**). Note that the structure of each neural mass is the same as in mesoscale simulations, with the caveat that the stellate subpopulation now receives input from other regions as well. Here, the sigmoidal function is used to model the nonlinear relationship between the input and output of subpopulations.

The $Coupling$ term defines the influence of other regions on the dynamics of a

given region, in which the strength of said coupling is modulated by the structural connectivity matrix and the activity differential between the interneural and stellate subpopulation of each input region. It should be noted that G serves as a global scaling parameter of the coupling of the all nodes (and not as a proxy for effective connectivity). $P(t)$ serves as the external current, and was modeled as a Gaussian random noise ($mean = 0.295$ and $SD = 0$). Euler integration method was used so as to conduct all simulations ($dt = 0.05\ sec$). Table 3.3 outlines the fixed parameters for the model.

$$
\begin{aligned}
\frac{dy_{0,i}(t)}{dt} &= y_{3,i}(t) \\
\frac{dy_{1,i}(t)}{dt} &= y_{4,i}(t) \\
\frac{dy_{2,i}(t)}{dt} &= y_{5,i}(t) \\
\frac{dy_{3,i}(t)}{dt} &= Aa\,\mathrm{Sig}(y_{1,i}(t) - y_{2,i}(t)) - 2ay_{3,i}(t) - a^2 y_{0,i}(t) \\
\frac{dy_{4,i}(t)}{dt} &= Aa\Big(P(t) + g_2\,\mathrm{Sig}(g_1 y_{0,i}(t)) + G \times \mathrm{Coupling}_i\Big) - 2ay_{4,i}(t) - a^2 y_{1,i}(t) \\
\frac{dy_{5,i}(t)}{dt} &= Bb\Big(g_4\,\mathrm{Sig}(g_3 y_{0,i}(t))\Big) - 2by_{5,i}(t) - b^2 y_{2,i}(t) \\
\mathrm{Sig}(v) &= \frac{v_{max}}{1 + \exp(r(v_0 - v))} \\
\mathrm{Coupling}_i &= \sum_j \mathrm{STC}_{ij}\mathrm{Sig}(y_{1,j} - y_{2,j})
\end{aligned}
\tag{3.4}
$$

Parameter	Representation	Value
A	Maximum amplitude of EPSP	3.25 mV
B	Maximum amplitude of IPSP	22.0 mV
1/a	Rate constant of EPSP	0.01 s
1/b	Rate constant of IPSP	0.05 s
G	Global coupling	1
g_1	Average number of synapses from PNs to SCs	135.0
g_3	Average number of synapses from PNs to INs	33.75
g_4	Average number of synapses from INs to PNs	33.75
v_{max}	Maximum firing rate	6.0 Hz
v_0	Potential at half-maximum of firing rate	6.0 mV
r	Sigmoid function slope at v_0	0.56 mV^{-1}

Table 3.3.: Parameters for whole-brain EEG model.

We are now primed to embark on the journey of inferring synaptic efficacies, with the help of our empirical and synthetic data generative processes (i.e. task and models;

chapter 4).

4. Probabilistic inference across scales

Summary

4.1. Microscale inference 74
4.2. Mesoscale inference 77
4.2.1. HMC validation 80
4.3. Macroscale inference 83

Now that we have suitable generative models across all three brain scales, we can begin the process of inferring synaptic efficacies from the informative empirical data features identified in chapter 2.

4.1. Microscale inference

Synaptic efficacy of the Brunel model is represented via the inhibitory-to-excitatory weight ratio parameter g. SEs were randomly sampled from a truncated uniform prior distribution with $g \in [5,8]$, which were used in order to simulate the model ($n = 10000$) with each simulation lasting for 1000 ms ($dt = 0.1\ ms$). The remaining parameters (e.g. synaptic/leak time constants, external input) were fixed (Table 3.1).

A correct application of SBI calls for the employment of identical data features in the entirety of the inference pipeline. Hence, the empirical summary statistics identified in the first chapter will lay the basis for computation of similar synthetic metrics in each scale (in order to train the deep neural density estimators). If you recall, peak firing rate of the clustered neurons was found to be significantly different between presence and absence conditions. Recorded rasters of the excitatory population within the Brunel model were thus used in order to compute the spike-time histograms ($nbins = 100$). The histograms were subsequently smoothed using a low-pass filter (Butterworth, $order = 5$) with a critical frequency of 20 Hz (based on the original sampling frequency) as to obtain the firing rate time series.

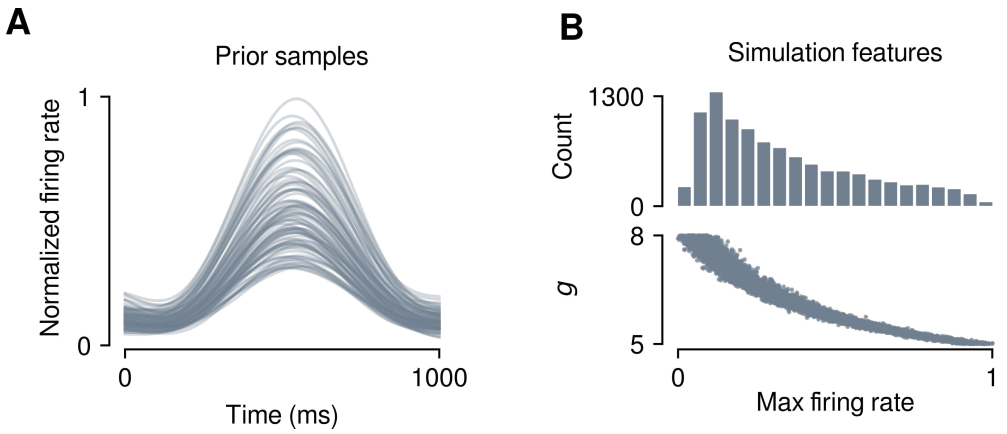

Figure 4.1.: **A**) Firing rate time-series obtained from simulations of the Brunel model, which are used in feature calculation. **B**) The characteristics of the chosen data feature for training the neural density estimator (i.e. peak firing rate). Top: Histogram of the features. Bottom: The inverse relationship between synaptic efficacy of the Brunel model and maximum firing rate.

Those that have dealt with both synthetic and empirical data before can quickly recognize that similarity in data features does not equate to proper alignment between the two datasets. Biological data are prone to noise and myriad other confounding factors, while the majority of synthetic simulations can be subject to stochasticity (e.g. noise-driven models). Though one can get away with inputting raw (i.e. unscaled) data features in some cases, simulation-based inference –as with other applications of deep learning– greatly benefits from data normalization [180]. Therefore, synthetic firing rates were normalized in the range of 0 to 1, such that the absolute minima of

the entire dataset would lie at 0, and the absolute maxima would lie at 1 (Figure 4.1**A**). The peak value within each time series was then saved as the final and corresponding data feature for the value of g that was used for its simulation. Vectors of normalized firing rates and sampled SEs were then fed to the neural density estimator for training (Figure 4.1**B**).

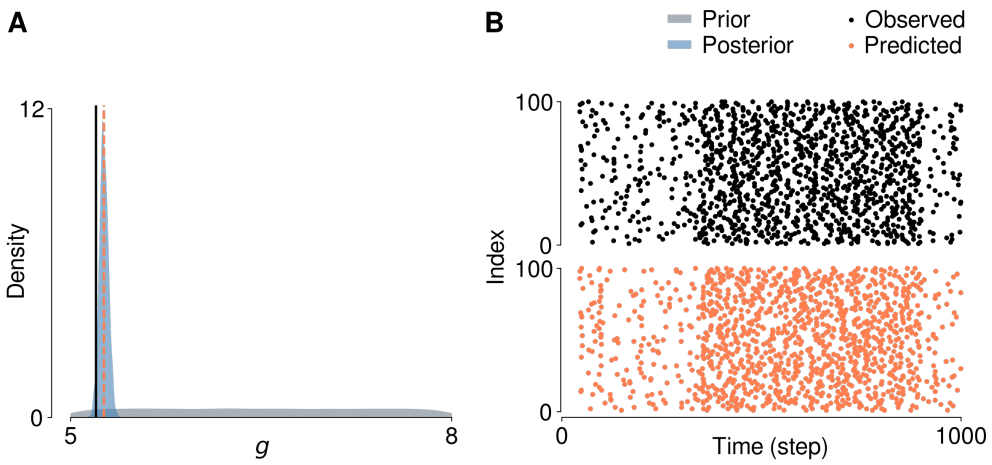

Figure 4.2.: **A**) Alignment of observed and predicted values of synaptic efficacy. **B**) Raster plot obtain from the observed (i.e. ground-truth) value of SE (top), versus the one obtained from the predicted value of SE (bottom).

Before introducing the empirical data, we need to ensure the robustness of the training procedure. One way to implement such a validation is by inputting random sampled parameters as to generate in-silico observations and subsequently sampling from the learned posterior. The maximum/median of the sampled posterior (chosen based on its degree of multimodality) can be checked against the initial parameter values in order to measure how close can the algorithm get to the ground-truth values. In this light, we validated the NSF model's training using randomly sampled parameters, which provides strong evidence for the robustness of the inference conducted at microscale (Figure 4.2**A**, **B**).

Previously outlined empirical peak firing rates were subsequently employed so as to obtain per-neuron posterior distributions of synaptic efficacy in each experimental condition. Looking closer into the sampled posteriors, we find that g significantly decreases during presence both across all –pooled– neurons ($p < 0.0001$, Figure 4.3**C**), and when comparing condition averages ($p < 0.0001$, Figure 4.3**D**). For the sake of correctly interpreting the inference results, it is vital to remember that the strength of connections between the two subpopulations is scaled by $-g$, which means that an elevated g results in an overall reduced network activity. Posterior predictive checks from neuron-averaged activity further outlines the goodness-of-fit between the simulated and empirical time series (Figure 4.3**A**, **B**). Taken together, the above findings illustrate that others' presence can induce an elevation in the synaptic efficacy of neural subpopulations of frontal areas involved in attentional modulation such as

dlPFC and ACC.

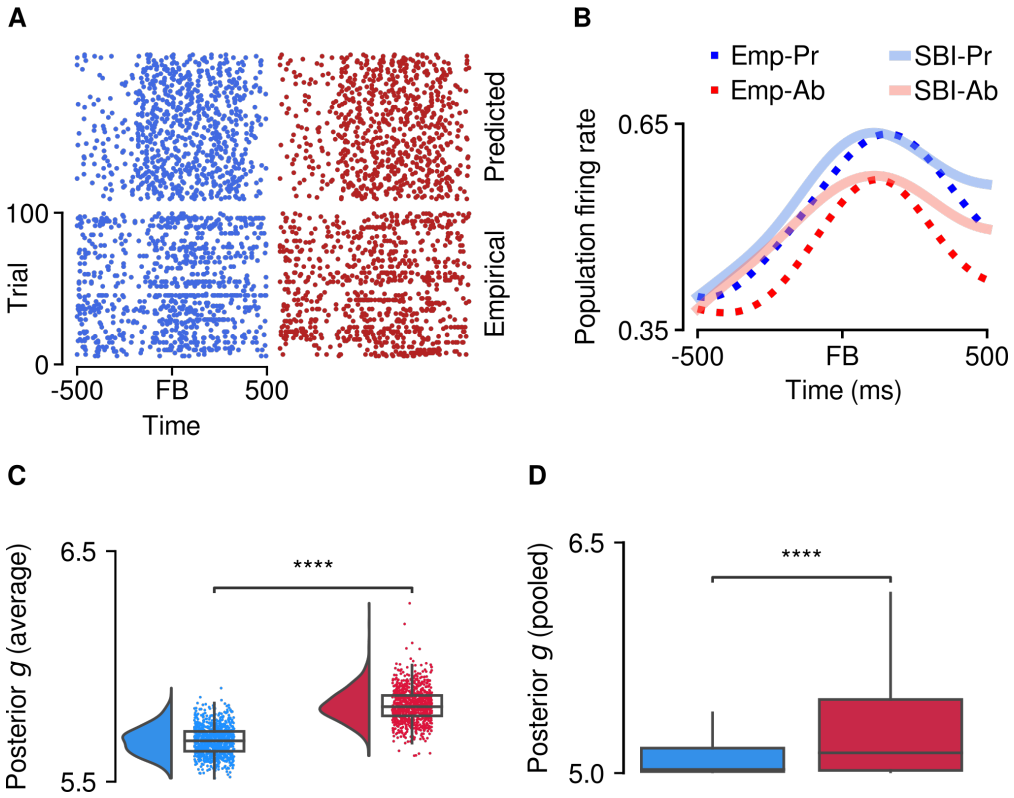

Figure 4.3.: **A**) Empirical versus predicted spiking activity for each condition. **B**) Empirical versus predicted (SBI) average firing rates across experimental conditions. **C**) Pooled distribution of single-neuron SEs. **D**) Distribution of SE for the average neural activity across conditions.

4.2. Mesoscale inference

Average number of synapses across the three subpopulations constitutes the four synaptic efficacies of the Jansen-Rit NMM g_{1-4}. Similar to the previous scale, SEs were randomly sampled from uniform prior distributions whose range is given in Table 4.1. Since task ERPs were already properly scaled, the range of prior distributions was set in such a way to cover the absolute minima and maxima of empirical data, bypassing the need for further normalization of synthetic data. All other parameters (e.g. axonal delay, input strength, etc.) were fixed due to the biological plausibility (i.e. degeneracy) of the Jansen-Rit model along with the previously-outlined hypothesis of this work. SEs were randomly drawn (i.e. sampled) in order to generate synthetic evoked time series ($n = 10000$, $dt = 0.1$). Maximum amplitude of synthetic ERPs was used as the training data features, per the analyses of evoked data in the first chapter. Neural density estimators were then trained on the set of sampled parameters and peak ERP amplitudes so as to estimate the joint posterior distribution of synaptic efficacies. In-silico validation was conducted prior to sampling from the learned posterior, which again confirmed the reliability of the training process (Figure 4.4).

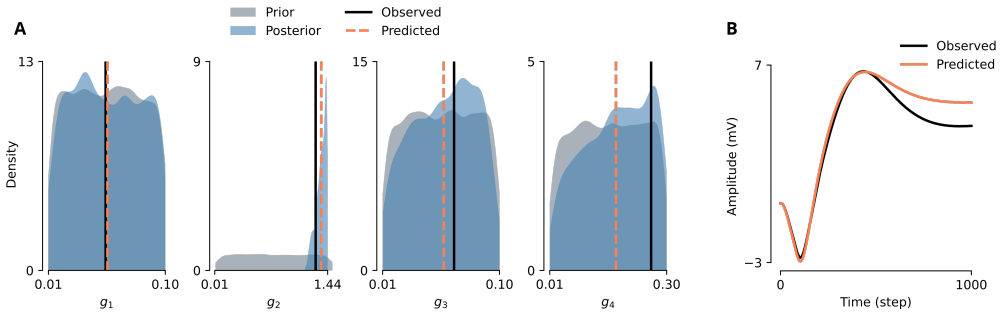

Figure 4.4.: **A**) Alignment of observed and predicted values of synaptic efficacy with regards to parameter g_2. **B**) Evoked potentials obtained from the observed values of SE (black) and the predicted value of SE (orange). Notice the similarity of amplitude between observed and predicted time series.

Parameter	Representation	Min	Max
g_1	Average number of synapses from PNs to SCs	0.01	0.1
g_2	Average number of synapses from SCs to PNs	0.02	1.5
g_3	Average number of synapses from PNs to INs	0.01	0.1
g_4	Average number of synapses from INs to PNs	0.01	0.3

Table 4.1.: Prior range for ERP model parameters.

The previously computed peak evoked activities were given to the trained NSF model in order to sample from the learned posterior distributions for each condition, experimental session, and region. Visual inspection of the nonlinear hidden dynamics

(latent states) generated from the sampled SEs of presence-dominant sessions points to similar trajectory characteristics as that of the task data (Figure 4.5**A**, **B**).

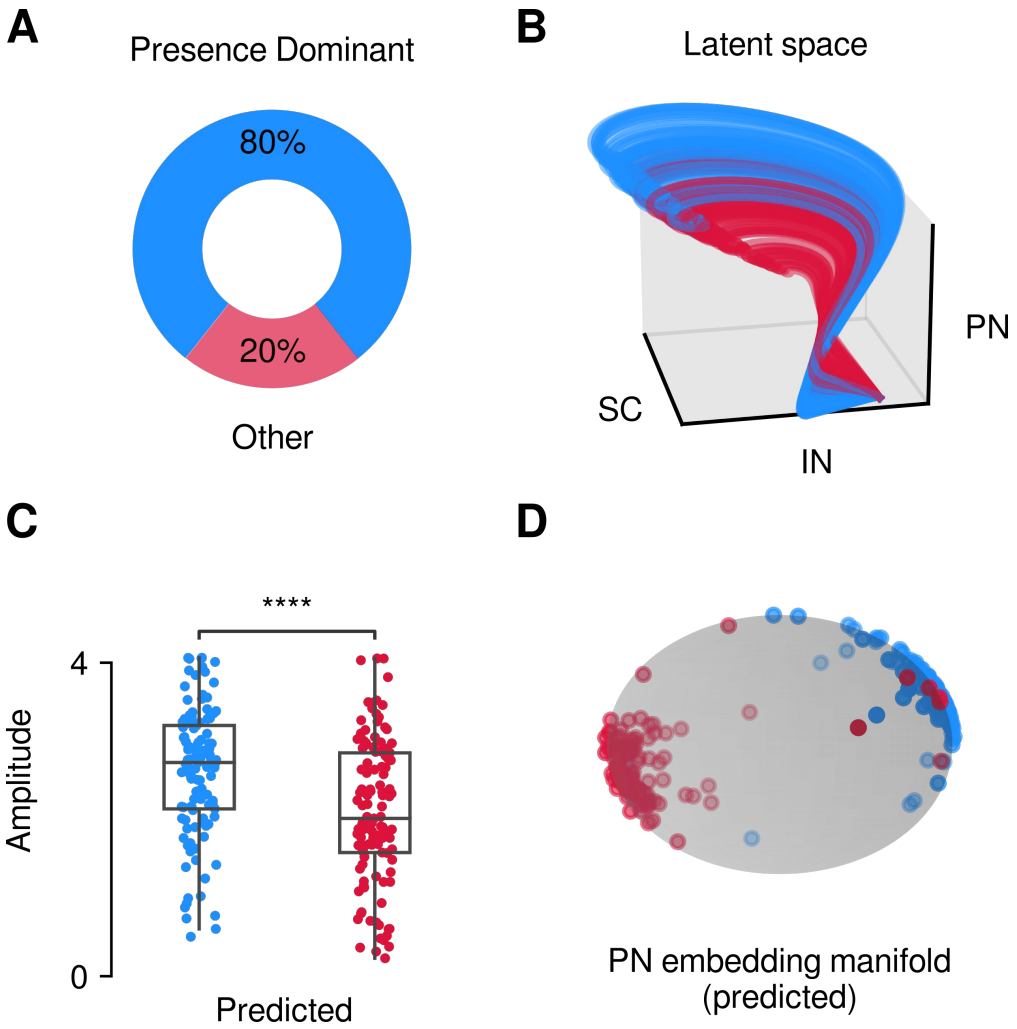

Figure 4.5.: **A**) Ratio of presence dominated sessions versus all others. **B**) Latent dynamics of the neural mass model across experimental conditions. **C**) Predicted distribution of peak PN activity. **D**) Embedding manifold of the predicted PN activities.

Indeed, comparison of the –entire dataset of– inferred PN activity revealed a near-perfect match with the peak amplitude of task ERPs, up until second order statistical moments ($p < 0.001$; Figure 4.5**C**). Moreover, embeddings of pyramidal neuron dynamics exhibited a similar pattern of isolated presence/absence clusters to that of the empirical ERPs (Figure 4.5**D**). As for the distribution of parameters, concatenation of SEs across all experimental sessions showed a significant difference between the task conditions, but only in parameter g_2, which represents the average number of synapses from SC to PN subpopulations ($p < 0.0001$, $KS = 0.28$; Figure 4.6**A**). This

synaptic connection delivers local recurrent excitation to the pyramidal subpopulation, and could thus be interpreted as the 'causal' driver of the observed elevation in ERP amplitudes during others' presence.

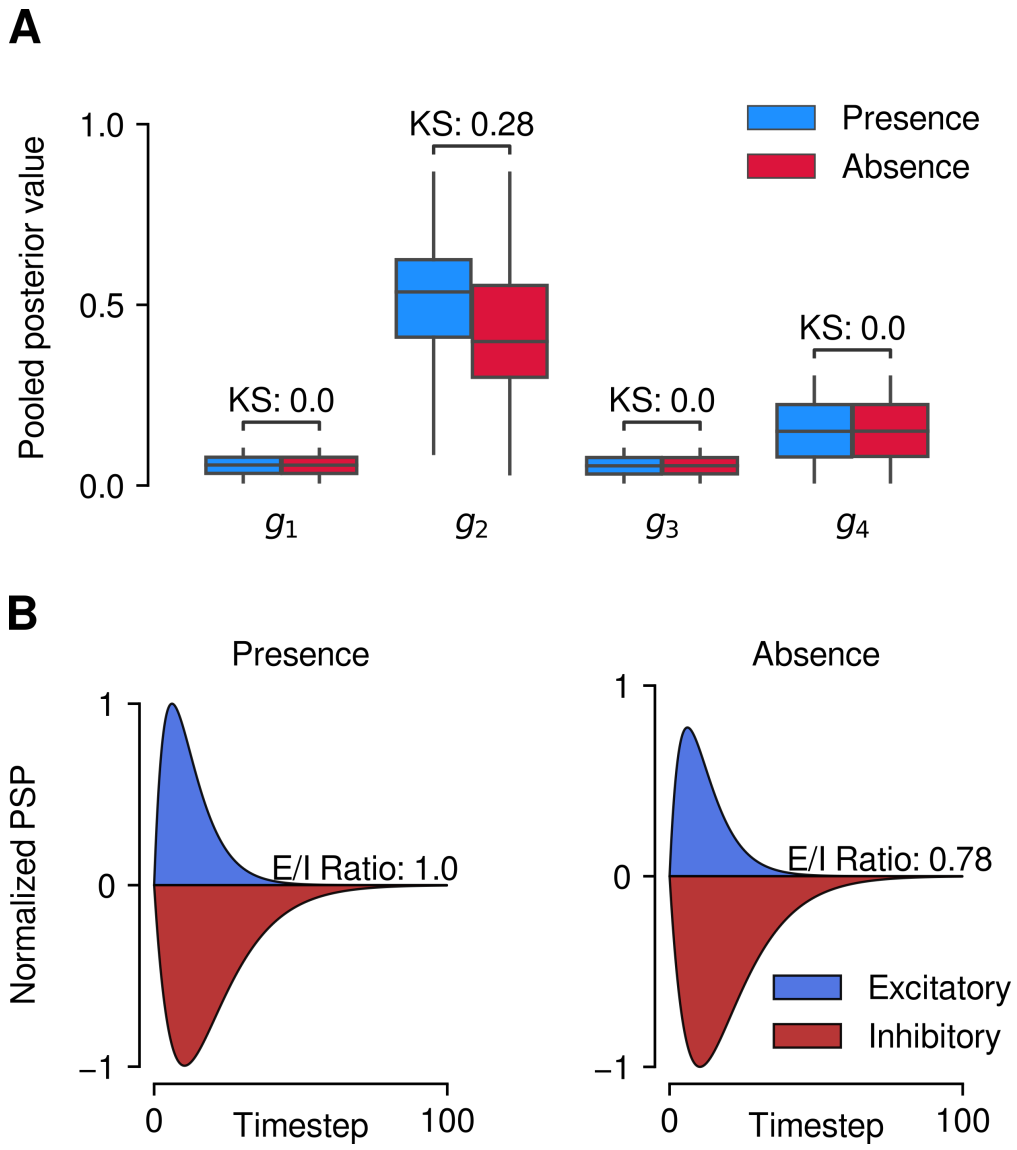

Figure 4.6.: **A**) Parameter distribution of the Jansen-Rit NMM, during presence/absence. KS represents the Kolmogorov-Smirnov distance metric, where higher values translate to a larger distance between distributions. **B**) Normalized excitatory/inhibitory post-synaptic potentials (PSPs) during absence. Note the change in E/I ratio between the conditions, as a consequence of modulations of g_2.

As timescales are held constant for all simulations, one could calculate how presence modulates the amplitude of inhibitory (i) and excitatory (e) postsynaptic potentials

(IPSPs/EPSPs; Figure 4.6**B**), by plugging in the median of the learned synaptic efficacies into their characteristic equation: $PSP_{e/i} = (g_{e/i}\frac{h_{e/i}}{\tau_{e/i}})te^{\frac{-t}{\tau_{e/i}}}$ where $g_e = g_1 + g_2 + g_3$ and $g_i = g_4$. ERPs are a measure of synaptic activity, hence increasing g_2 results in a subsequent increase in E/I ratio, which in turn could be responsible for the observed shift in evoked activity during the presence of others (Figure 4.6**B**). In this case, the potential non-identifiability of excitatory-to-inhibitory ratios can be resolved by taking the constant measures of IPSPs –across conditions– into account. I thus posit that –within the context of our task– mere presence effects are primarily driven by enhancements of excitatory synaptic efficacies (SE_e).

4.2.1. HMC validation

Unlike the microscale, the mean-field approach of the Jansen-Rit model translates to tractability of the likelihood function (despite its aforementioned degeneracy), which in turn enables one to conduct non-parametric probabilistic inferences such as HMC sampling for further validation. The computational efficiency of Hamiltonian Monte Carlo [136], however, is strongly dependent on its user-defined hyperparameters. Thus, I employed the No-U-Turn Sampler (NUTS; [86]) which is an adaptive variant of HMC. NUTS achieves this via a recursive procedure that is informed by the gradient information from the target posterior density for resolving the trajectory of the Markov Chain. This ultimately facilitates the exploration of the target distributions in complex and high-dimensional parameter spaces, albeit at a high computational cost [127, 81].

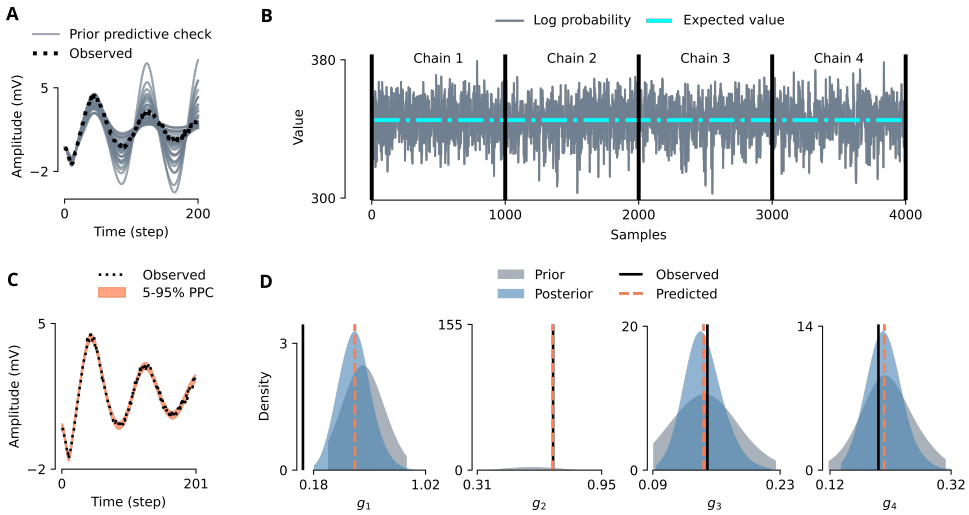

Figure 4.7.: **A**) Model simulations generated by drawing from the prior (gray) versus the observation (black). **B**) Log probability of the sampling chains. All chains were ran with random initial conditions, and eventually converged to the same stationary distribution. **C**) Simulations generated using samples drawn from posterior (orange). **D**) Posterior distribution of estimated SE. Note the large posterior shrinkage for g_2, compared to the other parameters.

To mitigate the computational burden of NUTS, I employed JAX (from NumPyro [146]), which is a high-performance computation library renowned for its capabilities in automatic differentiation and composable function transformations. JAX significantly expedited ERP simulations, which alongside a weakly informative prior distribution and robust parametrization, allowed for the convergence of the NUTS algorithm –for the dataset– in less than a minute.

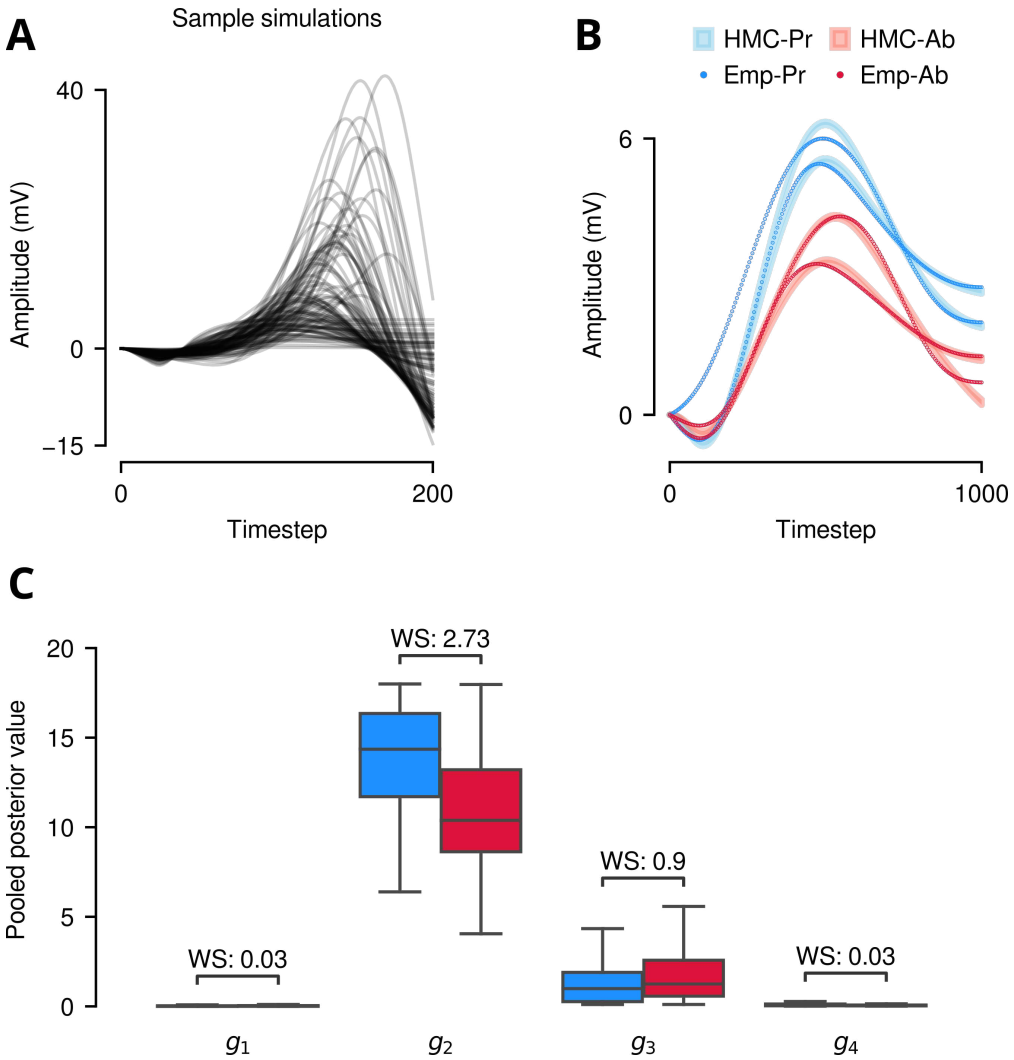

Figure 4.8.: **A**) Simulation samples generated by random sampling of the prior distribution. **B**) Time-series fit of observed ERPs from two example sessions. **C**) Inferred distribution of SE in the mesoscale model via Hamiltonian Monte-Carlo. WS represents the Wasserstein distance metric, where higher values translate to a larger distance between distributions.

Similar to all other instances, synthetic validation was done in order to assess the robustness of the inference process. Synaptic efficacies were drawn from a gamma

distribution $\Gamma(\alpha, \beta)$:

Parameter	g_1	g_2	g_3	g_4
α	18.16	29.9	29.14	30.770
β	33.33	50.00	20.00	142.86

Table 4.2.: Prior distribution of synaptic efficacies in HMC.

Four HMC chains were run per-ERP, with a transitory stage of $n_{warmup} = 2000$, and a sampling stage of $n_{samples} = 1000$ (max-tree-depth = 12, target-acceptance-probability = 0.6). Subsequent to monitoring the convergence diagnostics –in order to check for reliability of estimates– (Gelman-Rubin $\hat{R}$; [70]), the median of the sampled posteriors for each evoked potential was used to compute the manifold embeddings and latent dynamics. As evident, HMC not only generates posterior distribution of synaptic efficacies which closely matches that of SBI (Figure 4.8**C**), but also provide excellent fits to task ERPs (Figure 4.8**B**). The only visible difference between the distributions sampled via HMC and SBI is with regards to the parameter g_3, which is due to the fact that HMC sequentially marches through the steps of time series, without explicit focus on experimentally-relevant summary statistics (e.g. maximum amplitude), resulting in parameter distributions that best represent the 'entirety' of the time series.

4.3. Macroscale inference

While integration can provide useful insights into the efficacy of brain network connectivity [28, 193], it cannot be used as a direct metric of synaptic efficacy. Within the confines of this work, probabilistic inference of evoked potentials revealed local recurrent excitation (g_2) as the most probably driver of social facilitation in dorsolateral prefrontal and anterior cingulate cortices. It is therefore reasonable to posit that conspecific presence leads to elevation of this recurrent excitation within attentional networks.

To test this hypothesis, I simulated the coupled Jansen-Rit whole-brain model ($n = 20000$) with parameters sampled from a truncated uniform prior distribution. Prior range $g_2 \in [101.25, 110.7]$ was set such that all simulations would generate alpha oscillations. All simulations lasted 3000 ms, with the initial 2000 ms removed, and the last 1000 ms of which were retained from which functional connectivity and integration values (per-network) were computed in order to serve as summary statistics.

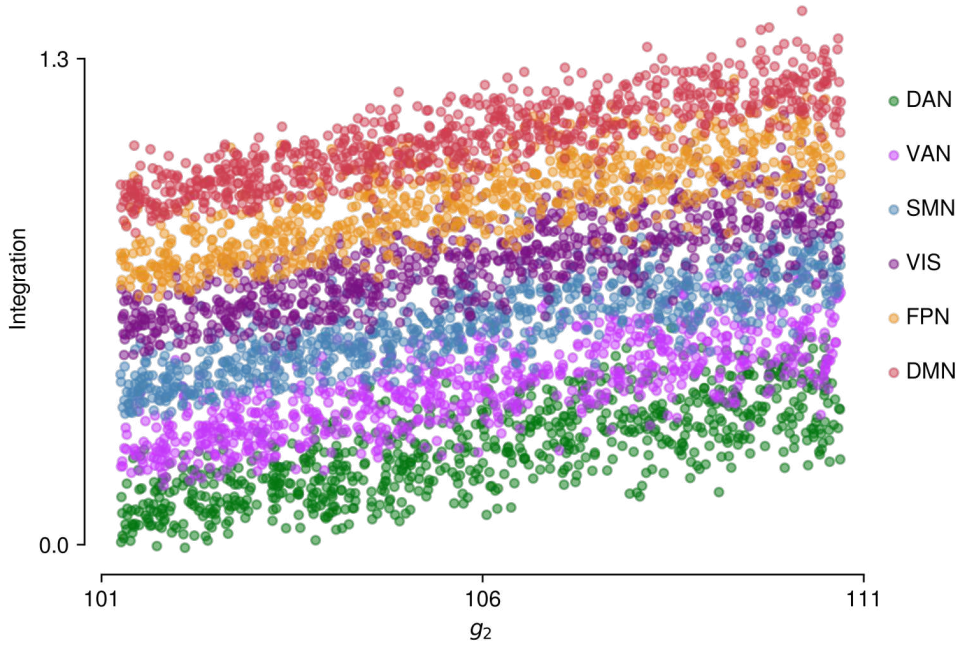

Figure 4.9.: Network integration measures (summary statistics) across the functional brain networks from the Schaefer atlas [161].

Drawn parameter sets and integration values were subsequently called upon in order to train the (NSF) neural density estimator. Similar to previous scales, synthetic validation was conducted before inferring SEs from the task data (Figure 4.10). This validation demonstrated a close similarity between observations and predicted SE values, and their respective simulated functional connectivities (Figure 4.10). Finally, empirical integration values were given to the trained density estimator in order to

obtain joint posterior distributions of SE for each subject, condition, and functional network.

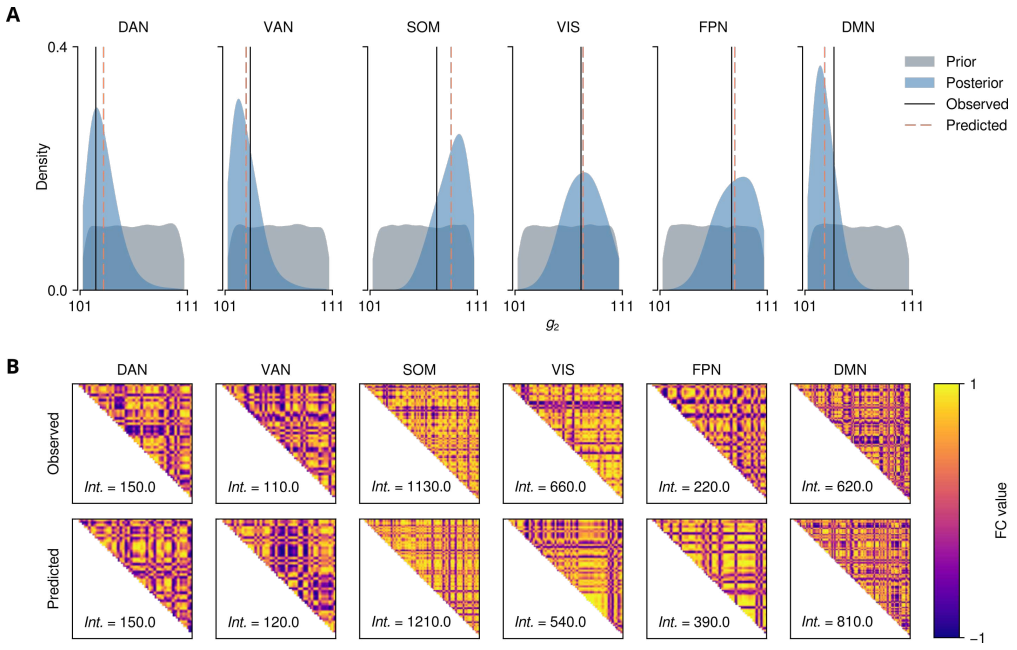

Figure 4.10.: **A**) Posterior predictive check of predicted versus observed macroscale SE across all functional networks.**B**) Posterior predictive check of observed (black) versus predicted (orange) observed macroscale SE across all functional networks.

The results of this inference were indeed surprising. Pooling the estimated posteriors across all subjects revealed a significant elevation of synaptic efficacy during conspecific presence within dorsal attention network, ventral attention network, and frontoparietal network ($p < 0.001$, $WS = 0.52$, $WS = 0.25$, $WS = 0.42$, respectively; Figure 4.11(top panel)). As for the male subject group, the presence of others significantly decreased SE within DAN and VAN, but only a negligible decrease in FPN ($p < 0.001$, $WS = 0.14$, $WS = 0.22$, $WS = 0.08$, respectively; Figure 4.11(bottom panel). Interestingly, the median –and pooled– distribution of SE across all networks bore an astonishing likeness to the task integration mapping (Figure 4.12).

Since neural measures should optimally be perused alongside behavioral measures [73], I calculated between-condition ratios (i.e. percent change) of task performance (average paddle speed) and average synaptic efficacy, in order to compute the correlation between brain and behavioral metrics. The results of this correlation proved to be vindicating. First, synaptic efficacy and task performance were significantly correlated within DAN and VAN in the female subject group ($\rho_{Pearson} = 0.79$, $\rho_{Spearman} = 0.45$ for DAN, and $\rho_{Pearson} = 0.73$, $\rho_{Spearman} = 0.51$ for VAN; Figure 4.13).

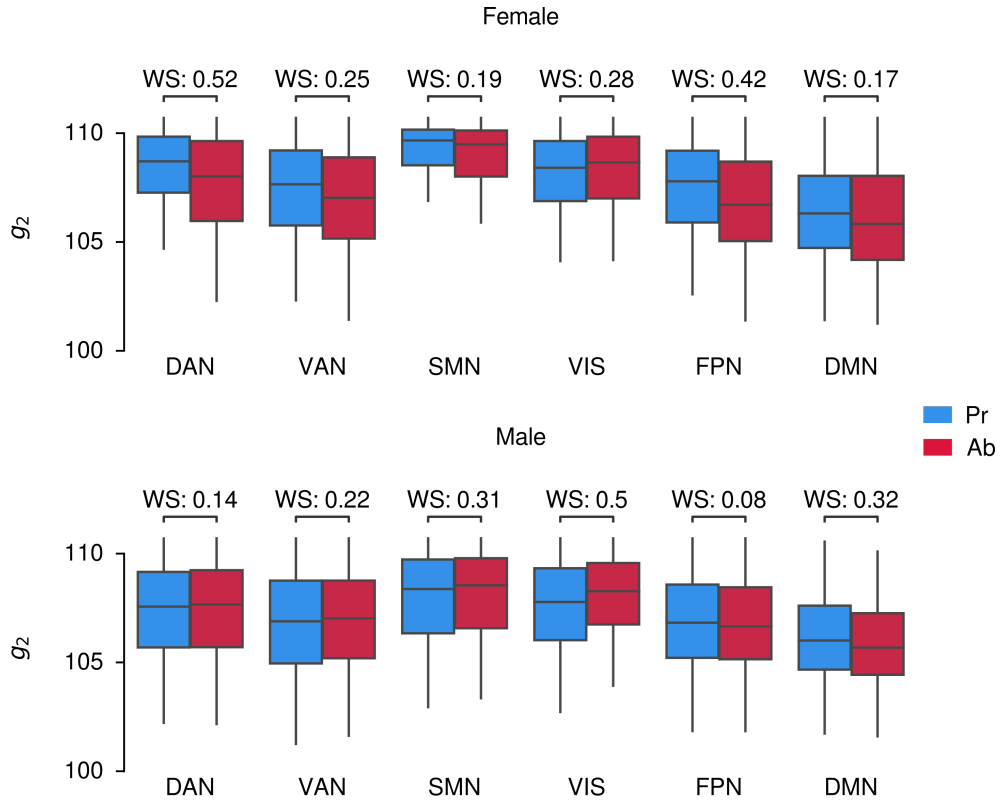

Figure 4.11.: Posterior distribution of excitatory SE (g_2) within the functional networks in the female (top), and male (bottom) subject groups.

FPN exhibited a high –albeit linear– correlation with behavioral performance, which could not be replicated on ranked data using the Spearman test ($\rho_{Pearson} = 0.68$, $\rho_{Spearman} = 0.12$; Figure 4.13). Analogous to FPN, correlations of SE and performance metrics in SMN and DMN, did not survive after ranking the datasets ($\rho_{Spearman} = (0.31, 0.09)$ for SMN, and DMN, respectively; Figure 4.13). The non-facilitated subject group (males) showed no significant correlation of synaptic efficacy with behavioral performance in any of the attentional networks (Figure 4.13, lower panels). Interestingly, performance seemed to be negatively correlated with the synaptic efficacy of the visual network in both groups, albeit more strongly within the female subject group ($\rho_{Spearman} = (-0.54, -0.34)$ for the female and male subject groups, respectively; Figure 4.13).

Attentional modulation of visual stimuli –and not their mere perception–, could thus contribute to the observed effects of mere presence (with regards to our task). However, the limited number of subjects per group constrains the replicability of the above findings, and needs to be further explored in future investigations. In a nutshell, the above results provide support for the potential role of synaptic efficacy in presence-induced attentional modulations.

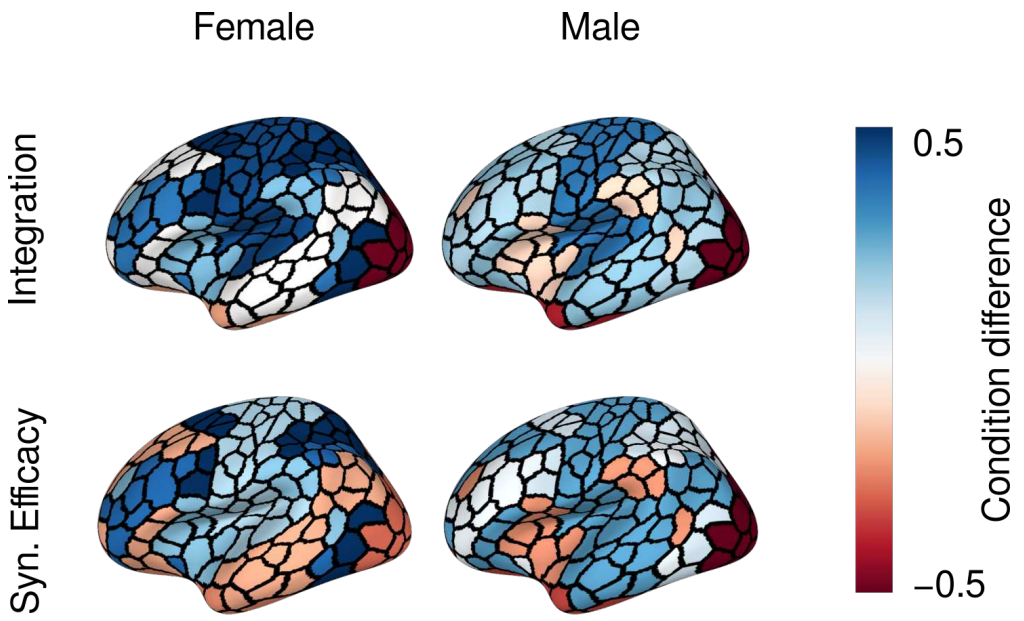

Figure 4.12.: Condition difference heatmaps of empirical observations (integration) versus predicted synaptic efficacies across the subject groups.

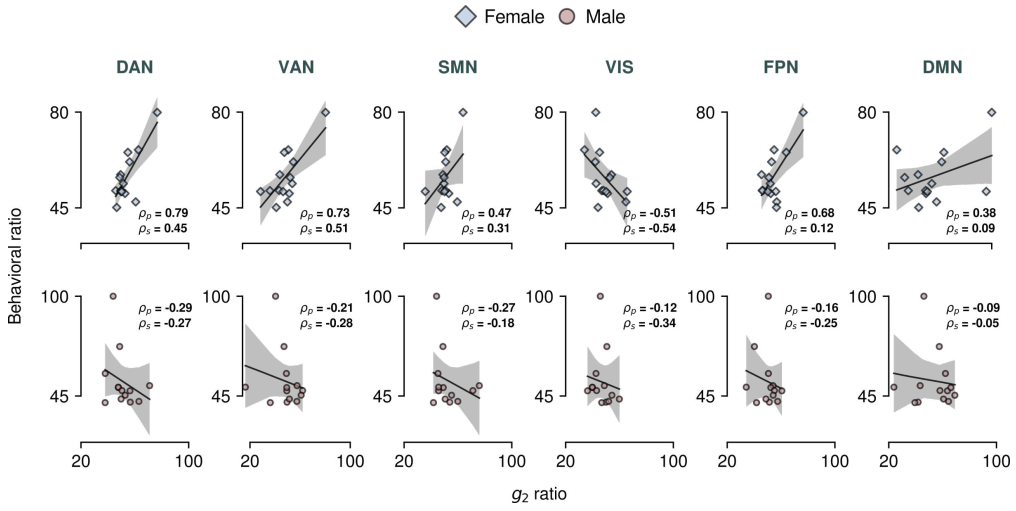

Figure 4.13.: Correlation of between-condition ratios of predicted SE versus task performance (average paddle speed) in the two subject groups.

5. Discussion

5.1. Summary

Social cognition is central to human cognition, and as such, a large body of work has been dedicated to studying the contribution of social stimuli to behavioral and neural dynamics [111, 104, 38]. Most of said work however, centers around explicit communications between social agents. Awareness of others' presence in the immediate environment –without explicit interactions– constitutes one of the most fundamental components of social cognition. Presence-induced improvements in task performance have recently been posited to originate from attentional modulation of context-oriented neurons [44, 131], but our understanding of their neurobiological underpinnings is still extremely insufficient. This could partially stem from the challenge of constructing a robust framework that could consolidate evidence obtained from mere presence research, into a unified perspective.

Fortuitously, advents in Bayesian learning, such as simulation-based inference (SBI; [33, 74, 81, 82, 84, 7]) have significantly aided in bridging this gap. Simulation-based inference effectively eliminates the need for employment of gradient-based variants of probabilistic inference, which are –severely– computationally inefficient at the scale of single neuron or whole-brain recordings. The amortized strategy conducted in SBI enables one to re-employ an identical neural network for both synthetic validation and inference on empirical observations, further bolstering the robustness of this framework [74, 84]. Through SBI, we have aimed to form an agnostic and exceedingly verifiable perspective of SFI effects that not only operates across brain scales, but is also informed by meaningful empirical summary statistics (such as neural firing rates or integration).

The first step in our work revolved around analysis of the behavioral and neural data from two social facilitation tasks, comprising an association learning task in Rhesus monkeys and a lateral interception task in humans. We were able to identify highly informative low-dimensional data features which were strongly modulated in the presence of conspecifics, but also highly correlated with behavioral performance in both tasks; In the association learning task, presence of others significantly elevated single-neuron firing rates and evoked potential amplitudes, both of which were highly correlated with task performance (i.e. accuracy or learning speed). As for lateral interception, presence of others significantly improved behavioral responses (i.e. average paddle speed), which was highly correlated with measures of integration in attentional networks, but only in the facilitated (i.e. female) subject group.

After careful configuration of suitable generative models, we leveraged deep neural

density estimators in order to investigate the link between modulations of synaptic efficacy and attention during the presence/absence of conspecifics across three spatiotemporal brain scales. With regards to single neurons (microscale), conspecific presence strongly increased the ratio of inhibitory to excitatory weights (i.e. excitatory synaptic efficacy; SE_e) in dorsolateral prefrontal and anterior cingulate cortices; regions which have been strongly linked to attentional modulation. The above findings were further corroborated by the results obtained in evoked potentials, where we saw a significant elevation of local recurrent excitation (i.e. SE_e) to the pyramidal neuron subpopulation during the presence condition. This presence-induced facilitation of learning lends support to the attentional framework of SFI effects, and is in line with previous findings on the interaction between attentional modulation and connectivity in said regions; Not only attention shifting is reported to heavily rely on the degree of effective connectivity between the two regions [110], but alterations to said connectivity has been implicated in clinical measures of impaired cognitive control and inattention [23, 76]. Moreover, the specialization of synapses in anterior cingulate hint at its large involvement in increasing the signal-to-noise ratio of dorsolateral regions during difficult cognitive tasks, with reports of disruption of these mechanisms in neuropathologies such as major depressive disorder and schizophrenia [125, 198].

But how does the presence of conspecifics affect synaptic efficacy within functional brain networks? Akin to the findings of the preceding scales, we observed an increase in synaptic efficacy during conspecific presence, which strongly matched the group-dependent patterns in empirical measures –of integration–. In a nutshell, we observed a strong correlation of task performance with synaptic efficacy across attentional networks such as DAN, VAN, and FPN, almost mirroring the correlation trends observed in integration. Surprisingly, the frontoparietal network strayed from this trend when ranked correlations were computed, potentially stemming from the insensitivity of the density estimation process to subtle changes of integration within FPN. This insensitivity might be due to the use of average template connectivity and functional atlas in the simulations, as opposed to an individualized approach. Indeed, the functional topography of FPN has been reported to vary between individuals, which highlights the necessity of precision mapping in future investigations [120, 31].

Regardless, the above findings suggest that influence of task attentional demands strongly influences the level of communication between attentional networks, particularly between DAN and VAN [188]. Such concerted activation might encompass the proposed function of attention as a transporter of response decisions –on currently focused targets–, towards higher-level regions/networks which deal with response execution [57], and is typically linked with changes in FC [177, 150]. This is further corroborated by the involvement of DAN in orienting attention and the contribution of VAN during its attentional reorientation [188]. Said findings suggest that modulations of coherence within attentional networks mediate the observed SFI effects in the lateral interception task. The elevation of synaptic efficacy of attentional networks might stem from a more efficacious and coordinated mode of processing, eventually resulting in improved motor performance. In a nutshell, we posit that presence-induced SFI effects might predominantly originate from enhancement of SE_e in attentional

brain networks, the effects of which ripple across brain scales.

5.2. Arousal versus attentional views of SFI effects

What if the observed elevations in behavioral performance in the association learning and lateral interception tasks simply stem from arousal? A possible response is that cognitive signals obtained from observed movements –of others– reportedly modulate neural populations involved in processing said movements selectively, rather than having a more diffuse arousal-like effect on the visual system [176]. Regardless, evidence does point to the existence of significant interplay between attentional modulation and arousal; For one, research suggests that arousal can significantly modulate visual attention, in that high-arousal stimuli –regardless of valence– tend to narrow the focus of attention [3, 137]. The origins of visual stimuli is also said to influence attention, with task-oriented arousing stimuli enhancing it, and task-irrelevant ones impairing it [204]. From an neurobiological perspective, this influence of arousal to attentional modulation in visual tasks might implicate the involvement of thalamus and visual cortex [189], with said influence being reportedly mediated via noradrenergic systems, particularly alpha2 adrenoreceptors [95]. Decomposing the unique contribution of arousal versus attentional modulation to the observed improvements in behavioral performance thus necessitates a significantly more granular investigation.

5.3. Homeostatic framework of SE within the social brain

Combining the existence of context-oriented (i.e. social) neurons with our findings of altered excitatory/inhibitory (E/I) balance during presence, hint at potential perturbations to the activity of said neurons in socially-oriented neuropathologies, such as schizophrenia and autism spectrum disorder (ASD). Fluctuations of E/I balance in biological neural networks have indeed been implicated in these pathologies [200]. For instance, modulations of E/I balance in ASD –as a result of connectivity abnormalities– reportedly perturb social interactions and attentional processes [203]. ASD also alters functional connectivity between social brain cortices and cerebellum [138]. Similarly, connectivity appears to be significantly perturbed in schizophrenia across numerous brain networks, with reported reductions in measures of resting-state segregation between the executive control, default mode, and dorsal attention networks [155, 195]. Within the social brain networks, such impairments are supposedly compensated for through an increase in activation of association cortices and frontal control networks [135]. Note that while the gradient of these perturbations appears to differ between ASD and schizophrenia, with the former reportedly exhibiting diminished activity within the social-cognitive network, compared to the latter [53], the exact relationship between connectivity alterations and said pathologies is still poorly understood [59, 60].

A related question is that variations in neural activity could simply stem from neuromodulation, rather than enhancement of synaptic efficacy. It is therefore paramount to underline that neuromodulation is famously conceptualized as context-sensitive changes in neural excitability and presynaptic efficacy [162, 65]. Biologically, this process is said to be mediated by neuromodulators, such as neuropeptides and biogenic amines, which act through G protein-coupled receptors to modify neuronal dynamics, excitability, and synaptic efficiency. These modulators can dramatically affect both neuronal excitability and synaptic function [8]; Consider the stomatogastric ganglion of crustaceans for instance, where a 'hybrid' chemical-electrical synapse between lateral pyloric and pyramidal neurons can switch from depolarizing to hyperpolarizing in response to dopamine [170, 55]. Dopamine also enhances calcium influx into presynaptic terminals of Drosophila taste receptor neurons under starvation. This gating of sensory input effectively modulates the excitability of these neurons [92, 156]. Similarly, in the mammalian retina, neuromodulators like nitric oxide and dopamine regulate electrical synapses between amacrine cells and cone bipolar cells, effectively decoupling these connections in high light conditions and modifying information flow through retinal circuitry [196].

In essence, one can think of experimental conditions inducing neuromodulators to curate a subset of synapses for activity, which alter neuronal excitability and presynaptic efficacy, in order to structure information flow on the scale of seconds to minutes (revolving around second messengers and GCPRs) [8].

5.4. Limitations and future directions

To reiterate, research surrounding the biological correlates of others' presence in humans and other higher primates is extremely lacking. My aim throughout this endeavor was been to address this sparsity via the proposed probabilistic inference framework across three brain scales; an endeavor which, to the best of my knowledge, has rarely been undertaken in the field of computational neuroscience, and never with regards to SFI effects.

No matter the novelty and scope of the work presented here, there are certain limitations that need to be contended with, so as to provide a faithful perspective on social facilitation. Albeit highly informative in terms of providing long-term insights into SFI effects, the longitudinal setup of the association learning task was constrained in terms of representation of sex and number of subjects. Further research in non-human primates should be also conducted in order to investigate the spatial composition of context-dependent social/asocial neurons across the entire brain. While we attempted to relieve some of the abovementioned constraints in the lateral-interception task, there is still ample ground for improvement of future mere presence tasks. For one, the study would have immensely benefited from a larger participant pool, despite our attempt in setting the sum total of participants to a reasonable amount (for visuomotor studies). Social facilitation research inherently necessitates a sizable participant pool, as many intra/inter-individual effects must be controlled for, in order to conduct

a thorough examination of *mere* presence effects.

Second, sheer inclusion of the two sexes does not equate to a careful control paradigm; To elaborate, we believe that the absence of behavioral modulations in the presence of others in the male subject group could stem from the perception of rank in the immediate social milieu (i.e. task settings [6]). On the other hand, this absence might not stem from any confounding factor, as previous research has outlined significant sex differences in social cognition and brain function. Women generally show greater interest in social information, empathy, and positive evaluations of faces compared to men [152, 151]. These differences are observed in various aspects of social cognition, including face processing, emotion recognition, response to infant faces, and processing of social interactions [152]. Nevertheless, one should take care as to not generalize the lack of improvements during the presence condition in the male (or any) subject group as an inherent difference between the subject groups in every facilitation task.

Third, it is essential for future SFI investigations to address functional lateralization of brain networks. With regards to cognitive process which are relevant to this work for instance, the left hemisphere –or parcels therefrom– has been linked to motor control, and the right hemisphere to visuospatial attention [85, 103]. In addition, the left hemisphere is said to be more efficient in network organization, while the right hemisphere reportedly exhibits higher measures of betweenness centrality and small-worldness [22]. Hemispheric asymmetries are also observed in various social neuropathologies such as schizophrenia [184]. All of the above findings could bear widely disparate contributions to our proposed attentional perspective of SFI effects, which outline the need for further investigations into SFI effects that specifically address this lateralization.

Lastly, it is paramount to specify that although SFI effects range from barely detectable to extremely significant, recent meta-studies have provided substantial evidence for the validity of SFI observations across species and tasks [18], enabling us to treat social facilitation as a viable proxy of mere presence effects. Nonetheless, it is prudent to simultaneously strive for developing alternative frameworks with the intention of attaining a more profound understanding of mere presence of others.

Conclusion

This study delved into the neurobiological bases of social facilitation, the enhancement of performance due to the presence of others. Through a comparative analysis involving both monkeys and humans, we presented evidence suggesting that mere presence effect are mediated by heightened synaptic efficacy within attentional brain networks. In monkeys engaged in an associative learning task, the presence of an observer resulted in a selective increase in excitatory synaptic efficacy (SE_e) within the dorsolateral prefrontal cortex (dlPFC) and anterior cingulate cortex (ACC), brain regions pivotal for attentional processes. Human participants, performing a lateral interception task under observation, exhibited facilitated performance, characterized by accelerated average paddle speeds, in a specific subject group. This enhancement was correlated with elevated SE_e within the dorsal attention network (DAN) and ventral attention network (VAN), implying a more synchronized and efficient mode of information processing. These findings offer a fresh perspective on social facilitation, emphasizing the role of attentional networks and synaptic efficacy. Moreover, the observed alterations in the excitatory-inhibitory (E/I) balance raise the possibility of connections to social neurodevelopmental disorders like autism and schizophrenia. Future research should aim to investigate the generalizability of these findings across a wider array of tasks, subject populations, and social contexts. In conclusion, this study illuminates the neural underpinnings of social facilitation and lays the groundwork for a deeper foray into the intricate interplay between mere presence, attentional modulation, and brain network dynamics.

Bibliography

[1] Oriol Abril-Pla, Virgile Andreani, Colin Carroll, et al. “PyMC: a modern, and comprehensive probabilistic programming framework in Python”. In: *PeerJ Computer Science* 9 (2023), e1516 (cit. on p. 47).

[2] Juan A. Acebrón, Luis L. Bonilla, Conrado J. Pérez Vicente, et al. “The Kuramoto model: A simple paradigm for synchronization phenomena”. In: *Reviews of Modern Physics* 77 (2005), pp. 137–185. URL: https://api.semanticscholar.org/CorpusID:8877676 (cit. on p. 67).

[3] Erkin Asutay and Daniel Västfjäll. “Exposure to arousal-inducing sounds facilitates visual search”. In: *Scientific Reports* 7.1 (2017) (cit. on p. 89).

[4] Maria Augustinova and Ludovic Ferrand. “The influence of mere social presence on Stroop interference: New evidence from the semantically-based Stroop task”. In: *Journal of Experimental Social Psychology* 48.5 (2012), pp. 1213–1216 (cit. on p. 21).

[5] João CB Azzi, Angela Sirigu, and Jean-René Duhamel. “Modulation of value representation by social context in the primate orbitofrontal cortex”. In: *Proceedings of the National Academy of Sciences* 109.6 (2012), pp. 2126–2131 (cit. on p. 40).

[6] Michela Balconi and Silvia Pagani. “Social hierarchies and emotions: Cortical prefrontal activity, facial feedback (EMG), and cognitive performance in a dynamic interaction”. In: *Social Neuroscience* 10 (2015), pp. 166 –178. URL: https://api.semanticscholar.org/CorpusID:33481293 (cit. on p. 91).

[7] Nina Baldy, Martin Breyton, Marmaduke M Woodman, et al. “Inference on the Macroscopic Dynamics of Spiking Neurons”. In: *Neural Computation* (2024), pp. 1–43. ISSN: 0899-7667. DOI: 10.1162/neco_a_01701. URL: https://doi.org/10.1162/neco_a_01701 (cit. on p. 87).

[8] Cornelia I Bargmann. “Beyond the connectome: how neuromodulators shape neural circuits”. In: *Bioessays* 34.6 (2012), pp. 458–465 (cit. on p. 90).

[9] Robert S Baron, Danny Moore, and Glenn S Sanders. “Distraction as a source of drive in social facilitation research.” In: *Journal of personality and social psychology* 36.8 (1978), p. 816 (cit. on p. 21).

[10] Roberto Barrio and Andrey L. Shilnikov. “Parameter-sweeping techniques for temporal dynamics of neuronal systems: case study of Hindmarsh-Rose model”. In: *Journal of Mathematical Neuroscience* 1 (2011), pp. 6 –6. URL: https://api.semanticscholar.org/CorpusID:508826 (cit. on p. 27).

[11] MD Bauman and CM Schumann. "Advances in nonhuman primate models of autism: Integrating neuroscience and behavior". In: *Experimental neurology* 299 (2018), pp. 252–265 (cit. on p. 38).

[12] Mark A Beaumont, Wenyang Zhang, and David J Balding. "Approximate Bayesian computation in population genetics". In: *Genetics* 162.4 (2002), pp. 2025–2035 (cit. on p. 33).

[13] Chen Beer and Omri Barak. "Revealing and reshaping attractor dynamics in large networks of cortical neurons". In: *bioRxiv* (2023) (cit. on p. 36).

[14] Mikhail Belkin and Partha Niyogi. "Laplacian eigenmaps for dimensionality reduction and data representation". In: *Neural computation* 15.6 (2003), pp. 1373–1396 (cit. on pp. 112, 113).

[15] Clément Belletier, Alice Normand, and Pascal Huguet. "Social-facilitation-and-impairment effects: From motivation to cognition and the social brain". In: *Current Directions in Psychological Science* 28.3 (2019), pp. 260–265 (cit. on pp. 18, 20, 21, 66).

[16] Michael Betancourt. "A conceptual introduction to Hamiltonian Monte Carlo". In: *arXiv preprint arXiv:1701.02434* (2017) (cit. on pp. 32, 33).

[17] Christopher M Bishop and Nasser M Nasrabadi. *Pattern recognition and machine learning*. Vol. 4. 4. Springer, 2006 (cit. on p. 44).

[18] Charles F Bond and Linda J Titus. "Social facilitation: a meta-analysis of 241 studies." In: *Psychological bulletin* 94.2 (1983), p. 265 (cit. on pp. 18, 91).

[19] Michael Breakspear. "Dynamic models of large-scale brain activity". In: *Nature Neuroscience* 20 (2017), pp. 340–352. URL: https://api.semanticscholar.org/CorpusID:20538785 (cit. on p. 62).

[20] Farran Briggs, George R Mangun, and W Martin Usrey. "Attention enhances synaptic efficacy and the signal-to-noise ratio in neural circuits". In: *Nature* 499.7459 (2013), pp. 476–480 (cit. on p. 31).

[21] Nicolas Brunel. "Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons". In: *Journal of computational neuroscience* 8 (2000), pp. 183–208 (cit. on pp. 63, 64).

[22] Karen Caeyenberghs and Alexander Leemans. "Hemispheric lateralization of topological organization in structural brain networks". In: *Human Brain Mapping* 35.9 (2014), pp. 4944–4957 (cit. on p. 91).

[23] Weidong Cai, Kristi R. Griffiths, Mayuresh S. Korgaonkar, et al. "Inhibition-related modulation of salience and fronto-parietal networks predicts cognitive control ability and inattention symptoms in children with ADHD". In: *Molecular psychiatry* 26 (2019), pp. 4016 –4025. URL: https://api.semanticscholar.org/CorpusID:204963880 (cit. on p. 88).

[24] Steve WC Chang. "An emerging field of primate social neurophysiology: Current developments". In: *eneuro* 4.5 (2017) (cit. on p. 23).

[25] Rishidev Chaudhuri, Berk Gerçek, Biraj Pandey, et al. "The population dynamics of a canonical cognitive circuit". In: *bioRxiv* (2019) (cit. on p. 36).

[26] Patrick Chen and Weizhe Hong. "Neural circuit mechanisms of social behavior". In: *Neuron* 98.1 (2018), pp. 16–30 (cit. on p. 32).

[27] Xinyin Chen and Doran C. French. "Children's social competence in cultural context." In: *Annual review of psychology* 59 (2008), pp. 591–616. URL: https://api.semanticscholar.org/CorpusID:11836481 (cit. on p. 20).

[28] Jessica R Cohen and Mark D'Esposito. "The segregation and integration of distinct brain networks and their relationship to cognition". In: *Journal of Neuroscience* 36.48 (2016), pp. 12083–12094 (cit. on p. 83).

[29] Jessica R. Cohen and Mark D'Esposito. "The Segregation and Integration of Distinct Brain Networks and Their Relationship to Cognition". In: *The Journal of Neuroscience* 36 (2016), pp. 12083 –12094. URL: https://api.semanticscholar.org/CorpusID:11949878 (cit. on p. 57).

[30] Blake J Cook, Andre DH Peterson, Wessel Woldman, et al. "Neural field models: historical perspectives and recent advances". In: *arXiv preprint arXiv:2103.10554* (2021) (cit. on p. 70).

[31] Maurizio Corbetta and Gordon L. Shulman. "Human cortical mechanisms of visual attention during orienting and search". In: *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 353.1373 (1998). Ed. by G. W. Humphreys, J. Duncan, and A. Treisman, pp. 1353–1362 (cit. on p. 88).

[32] Natalia I. Cordova, Alexa Tompary, and Nicholas B. Turk-Browne. "Attentional modulation of background connectivity between ventral visual cortex and the medial temporal lobe". In: *Neurobiology of Learning and Memory* 134 (2016), pp. 115–122. ISSN: 1074-7427. DOI: https://doi.org/10.1016/j.nlm.2016.06.011. URL: https://www.sciencedirect.com/science/article/pii/S1074742716300867 (cit. on p. 31).

[33] Kyle Cranmer, Johann Brehmer, and Gilles Louppe. "The frontier of simulation-based inference". In: *Proceedings of the National Academy of Sciences* 117.48 (2020), pp. 30055–30062 (cit. on pp. 33, 34, 87).

[34] Lee Cronk. "Culture's Influence on Behavior: Steps Toward a Theory". In: *Evolutionary Behavioral Sciences* 11 (2017), 36–52. URL: https://api.semanticscholar.org/CorpusID:151435721 (cit. on p. 20).

[35] Katalin Csilléry, Michael GB Blum, Oscar E Gaggiotti, et al. "Approximate Bayesian computation (ABC) in practice". In: *Trends in ecology & evolution* 25.7 (2010), pp. 410–418 (cit. on p. 33).

[36] Ding Cui and Koen Nelissen. "Two-monkey fMRI setup for investigating multi-faceted aspects of social cognition and behavior involving a real-live conspecific". In: *NeuroImage* 255 (2022), p. 119187 (cit. on p. 22).

[37] Catherine Culot, Gaia Corlazzoli, Carole Fantini-Hauwel, et al. “The relation between task-relatedness of anxiety and metacognitive performance”. In: *Consciousness and cognition* 94 (2021), p. 103191 (cit. on pp. 20, 49).

[38] JP Curley, CL Jensen, R Mashoodh, et al. “Social influences on neurobiology and behavior: epigenetic effects during development”. In: *Psychoneuroendocrinology* 36.3 (2011), pp. 352–371 (cit. on p. 87).

[39] Antea D’Andrea, Federico Chella, Tom R Marshall, et al. “Alpha and alpha-beta phase synchronization mediate the recruitment of the visuospatial attention network through the Superior Longitudinal Fasciculus”. In: *NeuroImage* 188 (2019), pp. 722–732 (cit. on p. 54).

[40] Teruko Danjo. “Allocentric representations of space in the hippocampus”. In: *Neuroscience research* 153 (2020), pp. 1–7 (cit. on pp. 22, 23).

[41] Olivier David, Stefan J Kiebel, Lee M Harrison, et al. “Dynamic causal modeling of evoked responses in EEG and MEG”. In: *NeuroImage* 30.4 (2006), pp. 1255–1272 (cit. on p. 67).

[42] Olivier David, James M Kilner, and Karl J Friston. “Mechanisms of evoked and induced responses in MEG/EEG”. In: *Neuroimage* 31.4 (2006), pp. 1580–1591 (cit. on pp. 29, 45, 67).

[43] Arnaud Delorme. “EEG is better left alone”. In: *Scientific reports* 13.1 (2023), p. 2372 (cit. on p. 52).

[44] Marie Demolliens, Faiçal Isbaine, Sylvain Takerkart, et al. “Social and asocial prefrontal cortex neurons: a new look at social facilitation and the social brain”. In: *Social cognitive and affective neuroscience* 12.8 (2017), pp. 1241–1248 (cit. on pp. 18, 23, 37, 39, 41, 42, 45, 58, 87).

[45] Christophe Destrieux, Bruce Fischl, Anders Dale, et al. “Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature”. In: *Neuroimage* 53.1 (2010), pp. 1–15 (cit. on p. 53).

[46] Sally S. Dickerson, Peggy J Mycek, and Frank P. Zaldivar. “Negative social evaluation, but not mere social presence, elicits cortisol responses to a laboratory stressor task.” In: *Health psychology : official journal of the Division of Health Psychology, American Psychological Association* 27 1 (2008), pp. 116–21. URL: https://api.semanticscholar.org/CorpusID:12439300 (cit. on p. 18).

[47] Shaul Druckmann and Dmitri B. Chklovskii. “Neuronal Circuits Underlying Persistent Representations Despite Time Varying Activity”. In: *Current Biology* 22.22 (2012), pp. 2095–2103 (cit. on p. 36).

[48] Florence Dumas, Pascal Huguet, and Edwige Ayme. “Social context effects in the Stroop task: When knowledge of one’s relative standing makes a difference”. In: *Current psychology letters. Behaviour, brain & cognition* 16, Vol. 2, 2005 (2005) (cit. on p. 21).

[49] Barnaby D. Dunn, T. Dalgleish, and Andrew D. Lawrence. "The somatic marker hypothesis: A critical evaluation". In: *Neuroscience & Biobehavioral Reviews* 30 (2006), pp. 239–271. URL: https://api.semanticscholar.org/CorpusID:207087890 (cit. on p. 20).

[50] Conor Durkan, Artur Bekasov, Iain Murray, et al. "Neural spline flows". In: *Advances in Neural Information Processing Systems* 32 (2019), pp. 7511–7522 (cit. on p. 34).

[51] Daniel Durstewitz, Quentin J. M. Huys, and Georgia Koppe. "Psychiatric Illnesses as Disorders of Network Dynamics." In: *Biological psychiatry. Cognitive neuroscience and neuroimaging* (2018). URL: https://api.semanticscholar.org/CorpusID:52288970 (cit. on p. 27).

[52] Egidio D'Angelo and Viktor Jirsa. "The quest for multiscale brain modeling". In: *Trends in neurosciences* 45.10 (2022), pp. 777–790 (cit. on pp. 22, 66).

[53] Shaun M. Eack, Jessica A. Wojtalik, Matcheri S. Keshavan, et al. "Social-cognitive brain function and connectivity during visual perspective-taking in autism and schizophrenia". In: *Schizophrenia Research* 183 (2017), pp. 102–109. URL: https://api.semanticscholar.org/CorpusID:3724501 (cit. on p. 89).

[54] Gerald M Edelman and Joseph A Gally. "Degeneracy and complexity in biological systems". In: *Proceedings of the national academy of sciences* 98.24 (2001), pp. 13763–13768 (cit. on pp. 31, 67).

[55] JUDITH S Eisen and E Marder. "A mechanism for production of phase shifts in a pattern generator". In: *Journal of neurophysiology* 51.6 (1984), pp. 1375–1393 (cit. on p. 90).

[56] Amirhossein Esmaeili, Marie Demolliens, Marjolein Viersen, et al. "Probing other's Presence: Probabilistic Inference Across Brain Scales Reveals Enhanced Excitatory Synaptic Efficacy". In: *bioRxiv* (2024) (cit. on p. 3).

[57] Uri Feintuch and Asher Cohen. "Visual attention and coactivation of response decisions for features from different dimensions". In: *Psychological Science* 13.4 (2002), pp. 361–369 (cit. on p. 88).

[58] Xiaoli Feng, Xujun Wu, Ryan J Morrill, et al. "Social correlates of the dominance rank and long-term cortisol levels in adolescent and adult male rhesus macaques (Macaca mulatta)". In: *Scientific reports* 6.1 (2016), p. 25431 (cit. on p. 38).

[59] Alex Fornito and Edward T. Bullmore. "Reconciling abnormalities of brain network structure and function in schizophrenia". In: *Current Opinion in Neurobiology* 30 (2015), pp. 44–50. URL: https://api.semanticscholar.org/CorpusID:7683965 (cit. on p. 89).

[60] Jennifer H. Foss-Feig, Brendan D. Adkinson, Jie Lisa Ji, et al. "Searching for Cross-Diagnostic Convergence: Neural Mechanisms Governing Excitation and Inhibition Balance in Schizophrenia and Autism Spectrum Disorders". In: *Biological Psychiatry* 81 (2017), pp. 848–861. URL: https://api.semanticscholar.org/CorpusID:4551528 (cit. on p. 89).

[61] Jan Fousek, Giovanni Rabuffo, Kashyap Gudibanda, et al. "Symmetry breaking organizes the brain's resting state manifold". In: *bioRxiv* (2022). URL: https://api.semanticscholar.org/CorpusID:245735486 (cit. on p. 36).

[62] Karl Friston, Jérémie Mattout, Nelson Trujillo-Barreto, et al. "Variational free energy and the Laplace approximation". In: *Neuroimage* 34.1 (2007), pp. 220–234 (cit. on p. 29).

[63] Karl J Friston. "Functional and effective connectivity in neuroimaging: a synthesis". In: *Human brain mapping* 2.1-2 (1994), pp. 56–78 (cit. on pp. 31, 57).

[64] Karl J Friston. "Functional and effective connectivity: a review". In: *Brain connectivity* 1.1 (2011), pp. 13–36 (cit. on pp. 31, 56).

[65] Karl J. Friston, Christian Buechel, Gereon Rudolf Fink, et al. "Psychophysiological and Modulatory Interactions in Neuroimaging". In: *NeuroImage* 6 (1997), pp. 218–229. URL: https://api.semanticscholar.org/CorpusID:571344 (cit. on p. 90).

[66] Karl J Friston, Lee Harrison, and Will Penny. "Dynamic causal modelling". In: *Neuroimage* 19.4 (2003), pp. 1273–1302 (cit. on pp. 28–30).

[67] KJ Friston and C Büchel. "Attentional modulation of effective connectivity from V2 to V5/MT in humans". In: *Proceedings of the National Academy of Sciences* 97.13 (2000), pp. 7591–7596 (cit. on pp. 28, 31).

[68] Teresa Garcia-Marques and Alexandre Fernandes. "Meta-Analysis of Social Presence Effects on Stroop Task Performance." In: *Psychological reports* (2024), p. 332941241227150. URL: https://api.semanticscholar.org/CorpusID:267334431 (cit. on pp. 19–21).

[69] A Gelman, C Robert, N Chopin, et al. "Bayesian Data Analysis.[Google Scholar]". In: (1995) (cit. on p. 32).

[70] Andrew Gelman and Donald B Rubin. "Inference from iterative simulation using multiple sequences". In: *Statistical science* 7.4 (1992), pp. 457–472 (cit. on p. 82).

[71] Mónika Gergelyfi, Benvenuto Jacob, Etienne Olivier, et al. "Dissociation between mental fatigue and motivational state during prolonged mental activity". In: *Frontiers in Behavioral Neuroscience* 9 (2015). URL: https://api.semanticscholar.org/CorpusID:2767784 (cit. on p. 20).

[72] Manesh Girn, Roni Setton, Gary R Turner, et al. "The "limbic network," comprising orbitofrontal and anterior temporal cortex, is part of an extended default network: Evidence from multi-echo fMRI". In: *Network Neuroscience* 8 (2024), pp. 860 –882 (cit. on p. 54).

[73] Alex Gomez-Marin, Joseph J Paton, Adam R Kampff, et al. "Big behavioral data: psychology, ethology and the foundations of neuroscience". In: *Nature neuroscience* 17.11 (2014), pp. 1455–1462 (cit. on p. 84).

[74] Pedro J Gonçalves, Jan-Matthis Lueckmann, Michael Deistler, et al. "Training deep neural density estimators to identify mechanistic models of neural dynamics". In: *Elife* 9 (2020), e56261 (cit. on pp. 34, 35, 87).

[75] Pulin Gong and Cees van Leeuwen. "Distributed Dynamical Computation in Neural Circuits with Propagating Coherent Activity Patterns". In: *PLoS Computational Biology* 5.12 (2009). Ed. by Karl J. Friston, e1000611 (cit. on p. 36).

[76] Ivan Grahek, Amitai Shenhav, Sebastian Musslick, et al. "Motivation and cognitive control in depression". In: *Neuroscience & Biobehavioral Reviews* 102 (2019), pp. 371–381 (cit. on p. 88).

[77] Alexandre Gramfort, Martin Luessi, Eric Larson, et al. "MEG and EEG data analysis with MNE-Python". In: *Frontiers in neuroscience* 7 (2013), p. 70133 (cit. on p. 53).

[78] Alexandre Gramfort, Martin Luessi, Eric Larson, et al. "MNE software for processing MEG and EEG data". In: *neuroimage* 86 (2014), pp. 446–460 (cit. on p. 53).

[79] Deanna J Greene, Eric Mooshagian, Jonas T Kaplan, et al. "The neural correlates of social attention: automatic orienting to social and nonsocial cues". In: *Psychological Research PRPF* 73 (2009), pp. 499–511 (cit. on p. 22).

[80] ET Hall. *The Hidden Dimension. New Garden City*. 1966 (cit. on p. 49).

[81] Meysam Hashemi, AN Vattikonda, Viktor Sip, et al. "The Bayesian Virtual Epileptic Patient: A probabilistic framework designed to infer the spatial map of epileptogenicity in a personalized large-scale brain model of epilepsy spread". In: *NeuroImage* 217 (2020), p. 116839 (cit. on pp. 80, 87).

[82] Meysam Hashemi, Anirudh N Vattikonda, Jayant Jha, et al. "Amortized Bayesian inference on generative dynamical network models of epilepsy using deep neural density estimators". In: *Neural Networks* 163 (2023), pp. 178–194 (cit. on pp. 34, 87, 114).

[83] Meysam Hashemi, Anirudh N Vattikonda, Viktor Sip, et al. "On the influence of prior information evaluated by fully Bayesian criteria in a personalized whole-brain model of epilepsy spread". In: *PLoS computational biology* 17.7 (2021), e1009129 (cit. on p. 28).

[84] Meysam Hashemi, Abolfazl Ziaee-mehr, Marmaduke Woodman, et al. “Probabilistic Inference on Virtual Brain Models of Disorders”. In: *bioRxiv* (2024), pp. 2024–02 (cit. on p. 87).

[85] Pierre-Yves Hervé, Laure Zago, Laurent Petit, et al. “Revisiting human hemispheric specialization with neuroimaging”. In: *Trends in Cognitive Sciences* 17.2 (2013), pp. 69–80 (cit. on p. 91).

[86] Matthew D Hoffman, Andrew Gelman, et al. “The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo.” In: *J. Mach. Learn. Res.* 15.1 (2014), pp. 1593–1623 (cit. on p. 80).

[87] Christopher J Honey, Olaf Sporns, Leila Cammoun, et al. “Predicting human resting-state functional connectivity from structural connectivity”. In: *Proceedings of the National Academy of Sciences* 106.6 (2009), pp. 2035–2040 (cit. on p. 70).

[88] Pascal Huguet, Florence Dumas, and Jean-M Monteil. “Competing for a desired reward in the Stroop task: when attentional control is unconscious but effective versus conscious but ineffective.” In: *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale* 58.3 (2004), p. 153 (cit. on p. 21).

[89] Pascal Huguet, Marie P Galvaing, Jean M Monteil, et al. “Social presence effects in the Stroop task: further evidence for an attentional view of social facilitation.” In: *Journal of personality and social psychology* 77.5 (1999), p. 1011 (cit. on pp. 18, 20, 21).

[90] Raoul Huys, Dionysios Perdikis, and Viktor K Jirsa. “Functional architectures and structured flows on manifolds: a dynamical framework for motor behavior.” In: *Psychological review* 121.3 (2014), p. 302 (cit. on p. 36).

[91] Aapo Hyvärinen and Erkki Oja. “Independent component analysis: algorithms and applications”. In: *Neural networks* 13.4-5 (2000), pp. 411–430 (cit. on p. 113).

[92] Hidehiko K Inagaki, Shlomo Ben-Tabou De-Leon, Allan M Wong, et al. “Visualizing neuromodulation in vivo: TANGO-mapping of dopamine signaling reveals appetite control of sugar sensing”. In: *Cell* 148.3 (2012), pp. 583–595 (cit. on p. 90).

[93] Faical Isbaine, Marie Demolliens, Abdelouahed Belmalih, et al. “Learning by observation in the macaque monkey under high experimental constraints”. In: *Behavioural brain research* 289 (2015), pp. 141–148 (cit. on p. 37).

[94] Eugene M. Izhikevich. “Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting”. In: 2006. URL: https://api.semanticscholar.org/CorpusID:18522847 (cit. on pp. 25, 26, 61).

[95] Coull J. “Modulation of attention by noradrenergic alpha2-agents varies according to arousal level.” In: *Drug News and Perspectives* (2001) (cit. on p. 89).

[96] Pranav Jagtap and Vaibhav A Diwadkar. "Effective connectivity of ascending and descending frontalthalamic pathways during sustained attention: Complex brain network interactions in adolescence". In: *Human Brain Mapping* 37.7 (2016), pp. 2557–2570 (cit. on p. 31).

[97] Ben H Jansen and Vincent G Rit. "Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns". In: *Biological cybernetics* 73.4 (1995), pp. 357–366 (cit. on p. 67).

[98] Mainak Jas, Denis A Engemann, Yousra Bekhti, et al. "Autoreject: Automated artifact rejection for MEG and EEG data". In: *NeuroImage* 159 (2017), pp. 417–429 (cit. on p. 53).

[99] Jianrong Jia, Ying Fan, and Huan Luo. "Alpha-Band Phase Modulates Bottom-up Feature Processing." In: *Cerebral cortex* (2021). URL: https://api.semanticscholar.org/CorpusID:237243967 (cit. on p. 54).

[100] Viktor Jirsa. "Structured flows on manifolds as guiding concepts in brain science". In: *Selbstorganisation–ein Paradigma für die Humanwissenschaften: Zu Ehren von Günter Schiepek und seiner Forschung zu Komplexität und Dynamik in der Psychologie* (2020), pp. 89–102 (cit. on p. 35).

[101] Viktor Jirsa, Huifang Wang, Paul Triebkorn, et al. "Personalised virtual brain models in epilepsy". In: *The Lancet Neurology* 22.5 (2023), pp. 443–454 (cit. on p. 70).

[102] Yohan J John, Kayle S Sawyer, Karthik Srinivasan, et al. "It's about time: Linking dynamical systems with human neuroimaging to understand the brain". In: *Network Neuroscience* 6.4 (2022), pp. 960–979 (cit. on pp. 26, 27).

[103] Goulven Josse and Nathalie Tzourio-Mazoyer. "Hemispheric specialization for language". In: *Brain Research Reviews* 44.1 (2004), pp. 1–12 (cit. on p. 91).

[104] Tobias Kalenscher. "Social Neuroscience: How the Brain Attends to the Joys and Pains of Others". In: *Current Biology* 30.19 (2020), R1076–R1078 (cit. on p. 87).

[105] Sheida Kazemi and Yousef Jamali. "Phase synchronization and measure of criticality in a network of neural mass models". In: *Scientific reports* 12.1 (2022), p. 1319 (cit. on p. 70).

[106] JA Scott Kelso. "On the coordination dynamics of (animate) moving bodies". In: *Journal of Physics: Complexity* 3.3 (2022), p. 031001 (cit. on p. 36).

[107] Mikail Khona and Ila R. Fiete. "Attractor and integrator networks in the brain". In: *Nature Reviews Neuroscience* 23.12 (2022), pp. 744–766 (cit. on p. 36).

[108] Lyle Kingsbury, Shan Huang, Tara Raam, et al. "Cortical representations of conspecific sex shape social behavior". In: *Neuron* 107.5 (2020), pp. 941–953 (cit. on pp. 31, 32).

[109] Karl Christoph Klauer, Julia Herfordt, and Andreas Voss. "Social presence effects on the Stroop task: Boundary conditions and an alternative account". In: *Journal of Experimental Social Psychology* 44.2 (2008), pp. 469–476 (cit. on p. 21).

[110] Hirohito M. Kondo, Naoyuki Osaka, and Mariko Osaka. "Cooperation of the anterior cingulate cortex and dorsolateral prefrontal cortex for attention shifting". In: *NeuroImage* 23 (2004), pp. 670–679. URL: https://api.semanticscholar.org/CorpusID:16979638 (cit. on p. 88).

[111] Emel Küçük and David Pitcher. "Visual neuroscience: A brain area tuned for processing social interactions". In: *Current Biology* 34.2 (2024), R53–R55 (cit. on p. 87).

[112] L Lapicque. "Recherches quantitatives sur l'excitation electrique des nerfs". In: *J. Physiol. Paris* 9 (1907), pp. 620–635 (cit. on p. 61).

[113] Eric et al. Larson. *MNE-Python*. 2024. DOI: 10.5281/zenodo.10999175. URL: https://doi.org/10.5281/zenodo.10999175 (cit. on p. 53).

[114] Scott Linderman, Matthew Johnson, Andrew Miller, et al. "Bayesian learning and inference in recurrent switching linear dynamical systems". In: *Artificial intelligence and statistics*. PMLR. 2017, pp. 914–922 (cit. on p. 45).

[115] Muriel Lobier, J Matias Palva, and Satu Palva. "High-alpha band synchronization across frontal, parietal and visual cortex mediates behavioral and neuronal effects of visuospatial attention". In: *Neuroimage* 165 (2018), pp. 222–237 (cit. on p. 54).

[116] Yi-An Ma, Yuansi Chen, Chi Jin, et al. "Sampling can be faster than optimization". In: *Proceedings of the National Academy of Sciences* 116.42 (2019), pp. 20881–20885 (cit. on pp. 27, 32).

[117] Laurens Van der Maaten and Geoffrey Hinton. "Visualizing data using t-SNE." In: *Journal of machine learning research* 9.11 (2008) (cit. on p. 112).

[118] Colin M MacLeod. "Half a century of research on the Stroop effect: an integrative review." In: *Psychological bulletin* 109.2 (1991), p. 163 (cit. on p. 21).

[119] Elisa Magosso, Giulia Ricci, and Mauro Ursino. "Modulation of brain alpha rhythm and heart rate variability by attention-related mechanisms". In: *AIMS Neuroscience* 6 (2019), pp. 1 –24. URL: https://api.semanticscholar.org/CorpusID:86586564 (cit. on p. 54).

[120] Scott Marek and Nico U. F. Dosenbach. "The frontoparietal network: function, electrophysiology, and importance of individual precision mapping". In: *Dialogues in Clinical Neuroscience* 20.2 (2018), pp. 133–140 (cit. on p. 88).

[121] Hazel Markus. "The effect of mere presence on social facilitation: An unobtrusive test". In: *Journal of Experimental Social Psychology* 14.4 (1978), pp. 389–397 (cit. on p. 40).

[122] Brenda McCowan, Brianne A. Beisner, John P. Capitanio, et al. "Network Stability Is a Balancing Act of Personality, Power, and Conflict Dynamics in Rhesus Macaque Societies". In: *PLoS ONE* 6 (2011). URL: https://api.semanticscholar.org/CorpusID:17070998 (cit. on p. 38).

[123] Leland McInnes, John Healy, Nathaniel Saul, et al. "UMAP: Uniform Manifold Approximation and Projection". In: *The Journal of Open Source Software* 3.29 (2018), p. 861 (cit. on p. 112).

[124] AR Mclntosh and Francisco Gonzalez-Lima. "Structural equation modeling and its application to network analysis in functional brain imaging". In: *Human brain mapping* 2.1-2 (1994), pp. 2–22 (cit. on p. 29).

[125] Maria Medalla and Helen Barbas. "Synapses with Inhibitory Neurons Differentiate Anterior Cingulate from Dorsolateral Prefrontal Pathways Associated with Cognitive Control". In: *Neuron* 61 (2009), pp. 609–620. URL: https://api.semanticscholar.org/CorpusID:8358383 (cit. on p. 88).

[126] Edda van Meurs, Jona Greve, and Bernd Strauss. "Moving in the presence of others - a systematic review and meta-analysis". In: *International Review of Sport and Exercise Psychology* (2021). URL: https://api.semanticscholar.org/CorpusID:242401316 (cit. on p. 18).

[127] Betancourt Michael, Byrne Simon, Livingstone Sam, et al. "The geometric foundations of Hamiltonian Monte Carlo". In: *Bernoulli (Andover.)* 23.4A (2017), pp. 2257–2298 (cit. on pp. 32, 33, 80).

[128] Earl K. Miller and T. J. Buschman. "Cortical circuits for the control of attention". In: *Current Opinion in Neurobiology* 23 (2012), pp. 216–222. URL: https://api.semanticscholar.org/CorpusID:14723093 (cit. on p. 56).

[129] Maha Mnif, Soufien Chikh, and Mohamed Jarraya. "Effect of Social Context on Cognitive and Motor Behavior: A Systematic Review". In: *Journal of Motor Behavior* 54.5 (2022), pp. 631–647 (cit. on p. 51).

[130] Laura Molina-García and Arantza Barrios. "Sex differences in learning—shared principles across taxa". In: *Current Opinion in Physiology* 6 (2018). Sex Differences, pp. 65–74. ISSN: 2468-8673. DOI: https://doi.org/10.1016/j.cophys.2018.05.004. URL: https://www.sciencedirect.com/science/article/pii/S2468867318300762 (cit. on p. 38).

[131] Elisabetta Monfardini, Jérôme Redouté, Fadila Hadj-Bouziane, et al. "Others' sheer presence boosts brain activity in the attention (but not the motivation) network". In: *Cerebral Cortex* 26.6 (2016), pp. 2427–2439 (cit. on pp. 48, 87).

[132] Elisabetta Monfardini, Amélie J Reynaud, Jérôme Prado, et al. "Social modulation of cognition: Lessons from rhesus macaques relevant to education". In: *Neuroscience & Biobehavioral Reviews* 82 (2017), pp. 45–57 (cit. on pp. 23, 66).

[133] Rosalyn J. Moran, Dimitris A. Pinotsis, and Karl J. Friston. "Neural masses and fields in dynamic causal modeling". In: *Frontiers in Computational Neuroscience* 7 (2013). ISSN: 1662-5188. DOI: 10.3389/fncom.2013.00057. URL: https://www.frontiersin.org/journals/computational-neuroscience/articles/10.3389/fncom.2013.00057 (cit. on p. 66).

[134] George Mountoufaris, Aditya Nair, Bin Yang, et al. "A line attractor encoding a persistent internal state requires neuropeptide signaling". In: *Cell* (2024) (cit. on p. 45).

[135] Prerona Mukherjee, Heather C. Whalley, James W. McKirdy, et al. "Altered amygdala connectivity within the social brain in schizophrenia." In: *Schizophrenia bulletin* 40 1 (2014), pp. 152–60. URL: https://api.semanticscholar.org/CorpusID:8027589 (cit. on p. 89).

[136] Radford M. Neal. "MCMC Using Hamiltonian Dynamics". In: *Handbook of Markov Chain Monte Carlo* 54 (2010), pp. 113–162 (cit. on pp. 32, 80).

[137] Jianguang Ni, Huihui Jiang, Yixiang Jin, et al. "Dissociable Modulation of Overt Visual Attention in Valence and Arousal Revealed by Topology of Scan Path". In: *PLoS ONE* 6.4 (2011). Ed. by Jan Lauwereyns, e18262 (cit. on p. 89).

[138] Giusy Olivito, Silvia Clausi, Fiorenzo Laghi, et al. "Resting-State Functional Connectivity Changes Between Dentate Nucleus and Cortical Social Brain Regions in Autism Spectrum Disorders". In: *The Cerebellum* 16 (2016), pp. 283 – 292. URL: https://api.semanticscholar.org/CorpusID:2092016 (cit. on p. 89).

[139] Marius Pachitariu, Shashwat Sridhar, and Carsen Stringer. "Solving the spike sorting problem with Kilosort". In: *BioRxiv* (2023), pp. 2023–01 (cit. on p. 111).

[140] George Papamakarios, Eric T Nalisnick, Danilo Jimenez Rezende, et al. "Normalizing Flows for Probabilistic Modeling and Inference." In: *J. Mach. Learn. Res.* 22.57 (2021), pp. 1–64 (cit. on p. 34).

[141] European Parliament and Council of the European Union. *Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the Protection of Animals Used for Scientific Purposes*. Sept. 22, 2010. URL: http://data.europa.eu/eli/dir/2010/63/oj (cit. on p. 38).

[142] Anand Patil, David Huard, and Christopher J Fonnesbeck. "PyMC: Bayesian stochastic modelling in Python". In: *Journal of statistical software* 35.4 (2010), p. 1 (cit. on p. 47).

[143] Jonathan Peirce, Jeremy R Gray, Sol Simpson, et al. "PsychoPy2: Experiments in behavior made easy". In: *Behavior research methods* 51 (2019), pp. 195–203 (cit. on p. 49).

[144] William D Penny. "Comparing dynamic causal models using AIC, BIC and free energy". In: *Neuroimage* 59.1 (2012), pp. 319–330 (cit. on p. 30).

[145] Andre D. H. Peterson, Blake J. Cook, Wessel Woldman, et al. "Neural Field Models: A mathematical overview and unifying framework". In: *Mathematical Neuroscience and Applications* (2021). URL: https://api.semanticscholar.org/CorpusID:245335431 (cit. on p. 70).

[146] Du Phan, Neeraj Pradhan, and Martin Jankowiak. "Composable effects for flexible and accelerated probabilistic programming in NumPyro". In: *arXiv preprint arXiv:1912.11554* (2019) (cit. on pp. 33, 81).

[147] Ajay S Pillai and Viktor K Jirsa. "Symmetry breaking in space-time hierarchies shapes brain dynamics and behavior". In: *Neuron* 94.5 (2017), pp. 1010–1026 (cit. on p. 35).

[148] Judith Platania and Gary P Moran. "Social facilitation as a function of the mere presence of others". In: *The Journal of social psychology* 141.2 (2001), pp. 190–197 (cit. on p. 48).

[149] Victoria N Poole, Meghan E Robinson, Omar Singleton, et al. "Intrinsic functional connectivity predicts individual differences in distractibility". In: *Neuropsychologia* 86 (2016), pp. 176–182 (cit. on p. 58).

[150] Amy L Proskovec, Elizabeth Heinrichs-Graham, Alex I Wiesman, et al. "Oscillatory dynamics in the dorsal and ventral attention networks during the reorienting of attention". In: *Human brain mapping* 39.5 (2018), pp. 2177–2190 (cit. on p. 88).

[151] Alice Mado Proverbio. "Sex differences in social cognition: The case of face processing". In: *Journal of Neuroscience Research* 95.1-2 (2016), pp. 222–234 (cit. on p. 91).

[152] Alice Mado Proverbio. "Sex differences in the social brain and in social cognition". In: *Journal of Neuroscience Research* 101.5 (2021), pp. 730–738 (cit. on p. 91).

[153] Ryan Remedios, Ann Kennedy, Moriel Zelikowsky, et al. "Social behaviour shapes hypothalamic neural ensemble representations of conspecific sex". In: *Nature* 550.7676 (2017), pp. 388–392 (cit. on p. 32).

[154] Kay A. Robbins, Jon Touryan, Tim R. Mullen, et al. "How Sensitive are EEG Results to Preprocessing Methods: A Benchmarking Study". In: *bioRxiv* (2020). URL: https://api.semanticscholar.org/CorpusID:212933787 (cit. on p. 52).

[155] Edmund T Rolls, Marco Loh, Gustavo Deco, et al. "Computational models of schizophrenia and dopamine modulation in the prefrontal cortex". In: *Nature Reviews Neuroscience* 9.9 (2008), pp. 696–709 (cit. on p. 89).

[156] Cory M Root, Kang I Ko, Amir Jafari, et al. "Presynaptic facilitation by neuropeptide signaling mediates odor-driven food search". In: *Cell* 145.1 (2011), pp. 133–144 (cit. on p. 90).

[157] James Benedict Rowe, Karl J. Friston, Richard S. J. Frackowiak, et al. "Attention to Action: Specific Modulation of Corticocortical Interactions in Humans". In: *NeuroImage* 17 (2001), pp. 988–998. URL: https://api.semanticscholar.org/CorpusID:54320547 (cit. on p. 31).

[158] Matthew D Sacchet, Roan A LaPlante, Qian Wan, et al. "Attention drives synchronization of alpha and beta rhythms between right inferior frontal and primary sensory neocortex". In: *Journal of neuroscience* 35.5 (2015), pp. 2074–2082 (cit. on p. 54).

[159] Glenn S. Sanders. "Self-presentation and drive in social facilitation". In: *Journal of Experimental Social Psychology* 20 (1984), pp. 312–322. URL: https://api.semanticscholar.org/CorpusID:144365136 (cit. on p. 20).

[160] Paula Sanz-Leon, Stuart A Knock, Andreas Spiegler, et al. "Mathematical framework for large-scale brain network modeling in The Virtual Brain". In: *Neuroimage* 111 (2015), pp. 385–430 (cit. on p. 70).

[161] Alexander Schaefer, Ru Kong, Evan M Gordon, et al. "Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI". In: *Cerebral cortex* 28.9 (2018), pp. 3095–3114 (cit. on pp. 54, 69, 70, 83).

[162] Gabriele Scheler. "Neuromodulation influences synchronization and intrinsic read-out". In: *F1000Research* 7 (2018) (cit. on p. 90).

[163] Michael Schirner, Simon Rothmeier, Viktor K Jirsa, et al. "An automated pipeline for constructing personalized virtual brains from multimodal neuroimaging data". In: *NeuroImage* 117 (2015), pp. 343–357 (cit. on p. 70).

[164] Steffen Schneider, Jin Hwa Lee, and Mackenzie Weygandt Mathis. "Learnable latent embeddings for joint behavioural and neural analysis". In: *Nature* 617.7960 (2023), pp. 360–368 (cit. on p. 46).

[165] Rens van de Schoot, Sarah Depaoli, Ruth King, et al. "Bayesian statistics and modelling". In: *Nature Reviews Methods Primers* 1.1 (2021), p. 1 (cit. on pp. 27, 28).

[166] Erik De Schutter. "Computational Modeling Methods for Neuroscientists". In: 2009. URL: https://api.semanticscholar.org/CorpusID:60100586 (cit. on pp. 63, 64).

[167] Yayoi Sekiguchi and Toshimichi Hata. "Effects of the mere presence of conspecifics on the motor performance of rats: Higher speed and lower accuracy". In: *Behavioural processes* 159 (2019), pp. 1–8. URL: https://api.semanticscholar.org/CorpusID:56140119 (cit. on p. 18).

[168] Aslihan Selimbeyoglu, Christina K. Kim, Masatoshi Inoue, et al. "Modulation of prefrontal cortex excitation/inhibition balance rescues social behavior in CNTNAP2-deficient mice". In: *Science Translational Medicine* 9 (2017). URL: https://api.semanticscholar.org/CorpusID:13771981 (cit. on p. 31).

[169] Dinkar Sharma, Rob Booth, Rupert Brown, et al. “Exploring the temporal dynamics of social facilitation in the Stroop task”. In: *Psychonomic Bulletin & Review* 17 (2010), pp. 52–58 (cit. on p. 21).

[170] Karen A Sigvardt and Brian Mulloney. “Sensory alteration of motor patterns in the stomatogastric nervous system of the spiny lobster Panulirus interruptus”. In: *Journal of Experimental Biology* 97.1 (1982), pp. 137–152 (cit. on p. 90).

[171] Fernando H. Lopes da Silva. “Functional localization of brain sources using EEG and/or MEG data: volume conductor and source models.” In: *Magnetic resonance imaging* 22 10 (2004), pp. 1533–8 (cit. on p. 53).

[172] Vikaas S Sohal and John LR Rubenstein. “Excitation-inhibition balance as a framework for investigating mechanisms in neuropsychiatric disorders”. In: *Molecular psychiatry* 24.9 (2019), pp. 1248–1257 (cit. on p. 31).

[173] Hansem Sohn and Devika Narain. “Neural implementations of Bayesian inference”. In: *Current Opinion in Neurobiology* 70 (2021), pp. 121–129 (cit. on pp. 28, 32).

[174] Estate M Sokhadze, Manuel F Casanova, Emily L Casanova, et al. “Event-related potentials (ERP) in cognitive neuroscience research and applications”. In: *NeuroRegulation* 4.1 (2017), pp. 14–14 (cit. on p. 45).

[175] James F. Sorce and Gregory Fouts. “Level of Motivation in Social Facilitation of a Simple Task”. In: *Perceptual and Motor Skills* 37 (1973), pp. 567 –572. URL: https://api.semanticscholar.org/CorpusID:41981010 (cit. on p. 20).

[176] Thomas A. Sorensen and Daniel Barratt. “Is threat the only modulator of attentional selectivity? Redefining the Easterbrook hypothesis”. In: *Frontiers in Psychology* 5 (2014) (cit. on p. 89).

[177] Sara Spadone, Stefania Della Penna, Carlo Sestieri, et al. “Dynamic reorganization of human resting-state networks during visuospatial attention”. In: *Proceedings of the National Academy of Sciences* 112.26 (2015), pp. 8112–8117 (cit. on p. 88).

[178] Nicolas Spatola, Clément Belletier, Pierre Chausse, et al. “Improved cognitive control in presence of anthropomorphized robots”. In: *International Journal of Social Robotics* 11 (2019), pp. 463–476 (cit. on p. 21).

[179] Olaf Sporns. “Network attributes for segregation and integration in the human brain”. In: *Current opinion in neurobiology* 23.2 (2013), pp. 162–171 (cit. on p. 58).

[180] Valery V. Starovoitov and Yu. I. Golub. “Data normalization in machine learning”. In: *Informatics* (2021). URL: https://api.semanticscholar.org/CorpusID:244221434 (cit. on p. 74).

[181] Johannes Stelzer, Eric Lacosse, Jonas Bause, et al. “Brainglance: visualizing group level MRI data at one glance”. In: *Frontiers in Neuroscience* 13 (2019), p. 972 (cit. on p. 54).

[182] Klaas Enno Stephan, Will D Penny, Rosalyn J Moran, et al. "Ten simple rules for dynamic causal modeling". In: *Neuroimage* 49.4 (2010), pp. 3099–3109 (cit. on pp. 28–30).

[183] Michael J. Strube. "What Did Triplett Really Find? A Contemporary Analysis of the First Experiment in Social Psychology". In: *The American Journal of Psychology* 118.2 (2005), pp. 271–286. ISSN: 0002-9556. DOI: 10.2307/30039059. eprint: https://scholarlypublishingcollective.org/uip/ajp/article-pdf/118/2/271/1876041/30039059.pdf. URL: https://doi.org/10.2307/30039059 (cit. on p. 18).

[184] Yöney T. "Hemispheric Specialization and Psychopathology". In: (2001) (cit. on p. 91).

[185] Roberta Tatti, Melissa S Haley, Olivia K Swanson, et al. "Neurophysiology and regulation of the balance between excitation and inhibition in neocortical circuits". In: *Biological psychiatry* 81.10 (2017), pp. 821–831 (cit. on p. 63).

[186] Corinne Teeter, Ramakrishnan Iyer, Vilas Menon, et al. "Generalized leaky integrate-and-fire models classify multiple neuron types". In: *Nature communications* 9.1 (2018), p. 709 (cit. on p. 61).

[187] Alvaro Tejero-Cantero, Jan Boelts, Michael Deistler, et al. "sbi: A toolkit for simulation-based inference". In: *Journal of Open Source Software* 5.52 (2020), p. 2505. DOI: 10.21105/joss.02505. URL: https://doi.org/10.21105/joss.02505 (cit. on pp. 33, 34).

[188] Annalisa Tosoni, Paolo Capotosto, Antonello Baldassarre, et al. "Neuroimaging evidence supporting a dual-network architecture for the control of visuospatial attention in the human brain: a mini review". In: *Frontiers in Human Neuroscience* 17 (2023) (cit. on p. 88).

[189] J I. Tracy, F Mohamed, S Faro, et al. "The effect of autonomic arousal on attentional focus". In: *NeuroReport* 11.18 (2000), pp. 4037–4042 (cit. on p. 89).

[190] Aki Vehtari, Andrew Gelman, and Jonah Gabry. "Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC". In: *Statistics and computing* 27 (2017), pp. 1413–1432 (cit. on p. 30).

[191] Amit Vinograd, Aditya Nair, Joseph Kim, et al. "Causal evidence of a line attractor encoding an affective state". In: *Nature* (2024), pp. 1–3 (cit. on p. 45).

[192] Amit Vinograd, Aditya Nair, Scott W Linderman, et al. "Intrinsic dynamics and neural implementation of a hypothalamic line attractor encoding an internal behavioral state". In: *bioRxiv* (2024), pp. 2024–05 (cit. on p. 36).

[193] Rong Wang, Mianxin Liu, Xinhong Cheng, et al. "Segregation, integration, and balance of large-scale resting brain networks configure different cognitive abilities". In: *Proceedings of the National Academy of Sciences* 118.23 (2021), e2022288118 (cit. on p. 83).

[194] Roman M Wittig, Catherine Crockford, Anja Weltring, et al. "Social support reduces stress hormone levels in wild chimpanzees across stressful events and everyday affiliations". In: *Nature communications* 7.1 (2016), p. 13361 (cit. on p. 40).

[195] Neil D. Woodward, Baxter P. Rogers, and Stephan Heckers. "Functional resting-state networks are differentially affected in schizophrenia". In: *Schizophrenia Research* 130 (2011), pp. 86–93. URL: https://api.semanticscholar.org/CorpusID:12336084 (cit. on p. 89).

[196] Xiao-Bo Xia and Stephen L Mills. "Gap junctional regulatory mechanisms in the AII amacrine cell of the rabbit retina". In: *Visual neuroscience* 21.5 (2004), pp. 791–805 (cit. on p. 90).

[197] Wang Xiao-Jing. "Theoretical and computational neuroscience: Attractor network models". In: (2008) (cit. on p. 36).

[198] Ting Ye, Jing Peng, Binbin Nie, et al. "Altered functional connectivity of the dorsolateral prefrontal cortex in first-episode patients with major depressive disorder". In: *European journal of radiology* 81.12 (2012), pp. 4035–4040 (cit. on p. 88).

[199] BT Thomas Yeo, Fenna M Krienen, Jorge Sepulcre, et al. "The organization of the human cerebral cortex estimated by intrinsic functional connectivity". In: *Journal of neurophysiology* (2011) (cit. on p. 54).

[200] Ofer Yizhar, Lief E. Fenno, Matthias Prigge, et al. "Neocortical excitation/inhibition balance in information processing and social dysfunction". In: *Nature* 477 (2011), pp. 171–178. URL: https://api.semanticscholar.org/CorpusID:4359464 (cit. on pp. 31, 89).

[201] Kyoko Yoshida, Nobuhito Saito, Atsushi Iriki, et al. "Social error monitoring in macaque frontal cortex". In: *Nature neuroscience* 15.9 (2012), pp. 1307–1312 (cit. on p. 40).

[202] Robert B Zajonc. "Social Facilitation: A solution is suggested for an old unresolved social psychological problem." In: *Science* 149.3681 (1965), pp. 269–274 (cit. on p. 18).

[203] Basilis Zikopoulos and Helen Barbas. "Altered neural connectivity in excitatory and inhibitory cortical circuits in autism". In: *Frontiers in Human Neuroscience* 7 (2013). URL: https://api.semanticscholar.org/CorpusID:6974702 (cit. on p. 89).

[204] Andras N. Zsidó. "The effect of emotional arousal on visual attentional performance: a systematic review". In: *Psychological Research* 88.1 (2023), pp. 1–24 (cit. on p. 89).

Appendix

A. Validation of single-neuron clustering

Regardless of the choice of theoretical framework, the context-dependent trends in the empirical microscale data need to be robust, as they lay the rationale and foundation of this entire endeavor. Thus, I set out to re-validate the previously-observed trends in single-neurons.

This validation began with spike-sorting a subset of the experimental sessions, in order to ensure the propriety and presence of context-oriented neurons. The previous analysis also outline the potential existence of 'neutral' neurons, whose activity was not modulated in any experimental condition. For this, I employed the Kilosort [139] spike-sorting algorithm. A cursory visual glance at the results of this attempt hints at the existence of both context-based and context-free neurons in dlPFC and ACC (Figure A1**A**).

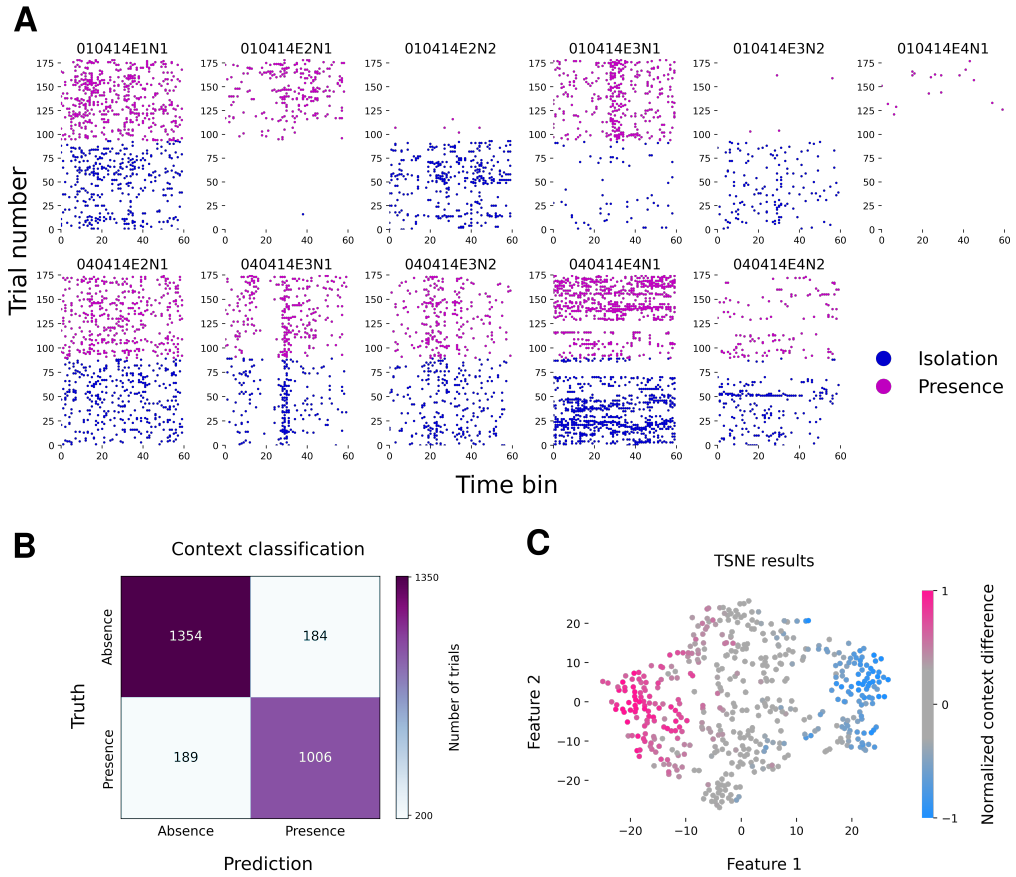

Figure A1.: **A**) Spike-sorting of spike times via Kilosort corroborates the existence of context-specific and 'neutral' neurons. **B**) Classification of the trial condition from firing rate pattern. The social context of majority of trials can be reliably predicted from firing rate characteristics. **C**) Dimensionality reduction via t-SNE confirms the choice of firing rate as a robust predictor of task condition.

In order to formally verify the abovementioned clusters, I conducted logistic regression over the firing rate obtained from pooled single-neuron trials (lbfgs), which was able to identify the 'class' of each neural trial purely via it's firing rate (Figure A1**B**). Since the results of future analyses heavily depended on acquiring representative pools of presence and absence for the entire dataset, it was necessary to show evidence that this context-based activity is maintained across sessions.

For this, I fed the session-averaged normalized neural activity (i.e. firing rate) of an aggregate comprising of all neuron categories, into t-SNE (t-distributed Stochastic Neighbor Embedding), a manifold extraction technique [117]. t-SNE is a dimensionality reduction technique that represents high-dimensional data in a lower-dimensional space, preserving local structure. It constructs probability distributions over pairs of high-dimensional points, assigning higher probabilities to similar points and lower probabilities to dissimilar ones. A similar distribution is defined for points in the low-dimensional map, and the algorithm minimizes the Kullback-Leibler divergence between the two distributions through iterative updates of point coordinates [117]. t-SNE's strength lies in its ability to capture essential data features agnostically, providing a global view of the data structure, including clusters, outliers, and other important features.

The embeddings obtained from t-SNE were then colored using a color map created from the difference of activity across experimental conditions. As seen in Figure A1**C**, not only can we observe a clear separation between social and asocial neurons, but we also see the appearance of neutral (i.e. context-free) neurons. These results together provide sufficient evidence for the robustness of neural classification in the experimental data.

B. Robustness of ERP amplitude as a feature

In addition to CEBRA, a large array of manifold extraction tests were ran on evoked potentials, as the pivotal data feature in this scale consisted of their condition-based amplitude differential. Four methods were employed in this validation, comprising t-SNE, UMAP (Uniform Manifold Approximation and Projection), spectral embedding (Laplacian eigenmaps), and ICA (independent component analysis) (Figure A2).

UMAP attempts to preserve the global structure of high-dimensional data [123]. Unlike methods such as t-SNE, it employs a geometric approach based on manifold learning, by representing each point with multiple embeddings and uses a probabilistic metric to balance local and global information. The algorithm iteratively refines these embeddings, resulting in an optimal embedding that captures the global topology of the data [123]. UMAP is therefore computationally efficient and effective in capturing complex structures, making it suitable for large datasets and providing an intuitive and interpretable representation [123].

Spectral embedding on the other hand, leverages the spectral properties of the graph Laplacian matrix to identify meaningful clusters or structure in high-dimensional data [14]. It is particularly useful for network or community detection problems.

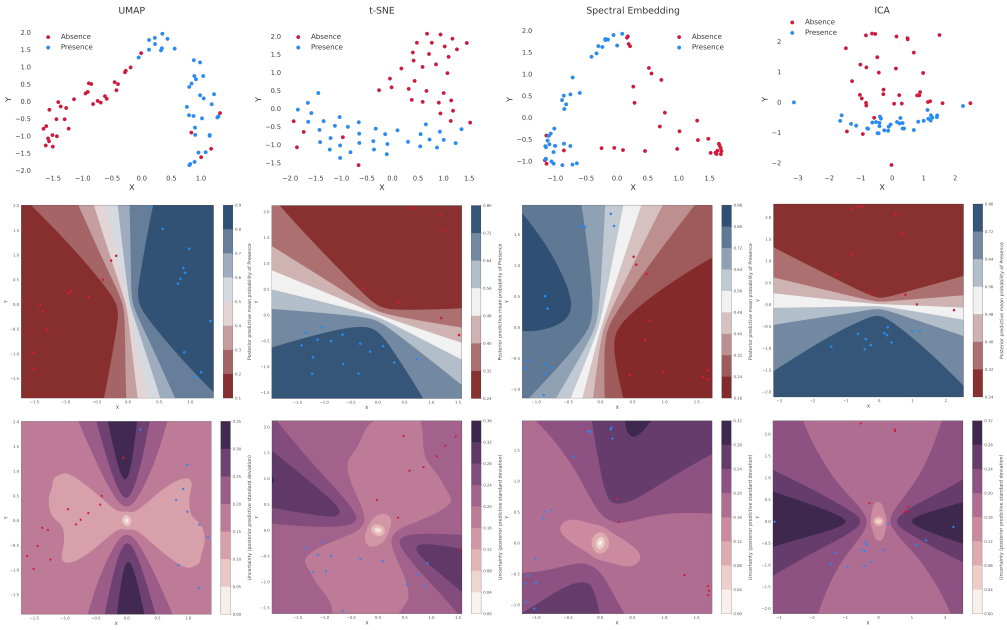

Figure A2.: **Top row:** Results of embedding algorithms. Note the clear separation of the majority of ERPs based on task condition. **Middle row:** Quantification of the probability of embeddings belonging to the presence condition. **Bottom row:** Uncertainty map of an embedding belonging to (n)either condition. This map paints a similar picture as the probability map in the middle row, but based on another metric.

The Laplacian matrix encodes the connectivity structure of the graph, and by projecting data onto its eigenvectors, we can reduce dimensionality while preserving global relationships. This approach effectively identifies clusters separated by significant gaps in the eigenvectors, corresponding to boundaries between communities [14].

Last but not least, independent component analysis (ICA) attempts to separate mixed signals into their original independent components [91]. It assumes the mixed data is a linear combination of additive, non-Gaussian, and independent components. ICA finds the maximum likelihood estimate of the weight vector to recover the original components. It is effective in separating mixed signals and handling linear and nonlinear relationships [91]. However, it relies on specific assumptions and is sensitive to noise and outliers, hence it is recommended to preprocess the data prior to running the algorithm.

As evident from Figure A2, all algorithms are able to clearly separate the condition of evoked potentials based on the amplitude (albeit with some variance in the degree of separation). Similar to CEBRA, probability and uncertainty maps were generated from a subset each algorithm's embeddings, all of which corroborate our findings with CEBRA.

C. Degeneracy of a full Jansen-Rit NMM

The Jansen-Rit NMM –as previously mentioned– exhibits heavy degeneracy when completely unconstrained (i.e. whilst varying all ten parameters to generate simulations). To reiterate, this phenomenon occurs when multiple distinct sets of model parameters can generate identical or nearly identical output dynamics. Degeneracy leads to an explosion of possible parameter combinations that yield similar results, making it challenging to uniquely identify the underlying neural mechanisms from observed brain activity.

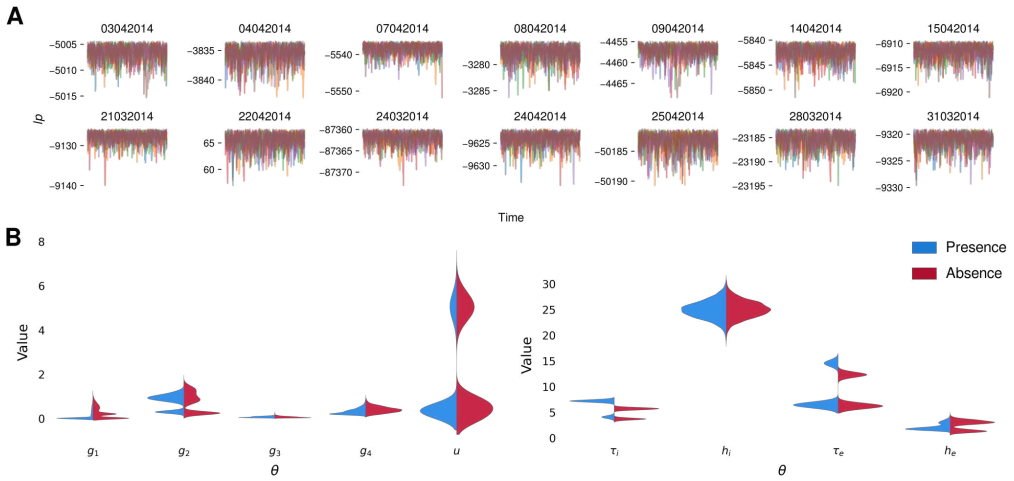

Figure A3.: Convergence and degeneracy of full-parameter HMC. **A**) Log probability (LP) of HMC chains for a sample of experimental sessions. The similarity of sampling chains' LP values –in each session– can be used as a qualitative metric for the uniformity of convergence. **B**) Posterior distribution of all 10 model parameters for a single run (i.e. session) of HMC. Note the high degree of degeneracy present for the majority of the parameters.

In order to elaborate on this degeneracy, we ran HMC on all 10 parameters. As seen in Figure A3**B**, even the posterior distribution from a single session exhibits strong multi-modality, which means that the evoked potentials could have been generated from different combinations of parameters (i.e. degeneracy). Notice that the sampling process for all HMC chains has converged successfully, hence eliminating non-convergence as a reason for said multi-modality (Figure A3**A**). We can formally quantify this degeneracy via calculation of inter-parameter correlations. Figure A4**A, B** clearly demonstrates the high correlation of JR model parameters. High correlations between model parameters are a strong indicator of degeneracy as they imply that changes in one parameter can be compensated for by changes in another, resulting in nearly identical model predictions. This can significantly complicate the sampling process via gradient-based approaches, as they are particularly susceptible to slow mixing and divergence in the presence of strong inter-parameter correlations [82].

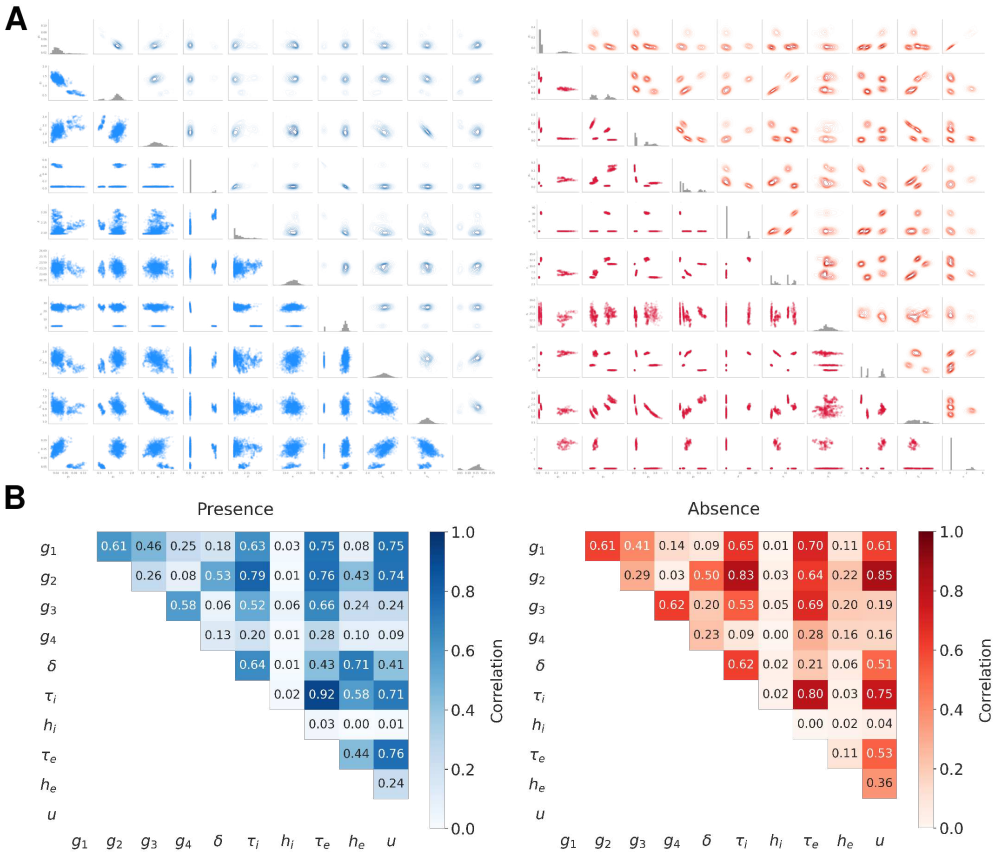

Figure A4.: **A**) Pairgrid plot of JR parameters comprised of KDE (upper triangular), scatter (lower triangular), and a 2D distribution (diagonal) plots of model parameters. Note the high degree of multi-modality present across numerous parameter pairs in both conditions. **B**) Correlation matrices of JR parameters across conditions, clearly illustrating the existence of extreme degeneracy present in the –fully unconstrained– JR model.

D. Optimization of external input in coupled JR NMMs

Simulations generated from the whole-brain JR network can be subject to a significant amount of stochasticity due to the implementation of an external gaussian input, in such a way that simulating the model with the exact parameters can lead to different measures of synchrony (e.g. integration). Since sum of functional connectivity as a data feature is an integral part of the macroscale inference, parameter sweeps were ran by varying the global connectivity and external current strength parameters. Fortunately, analysis of global $\sum FC$ indicated a suitable input strength with $\mu = 295$ (Figure A5). The standard deviation of this input was set to 0 in order to further reduce stochasticity.

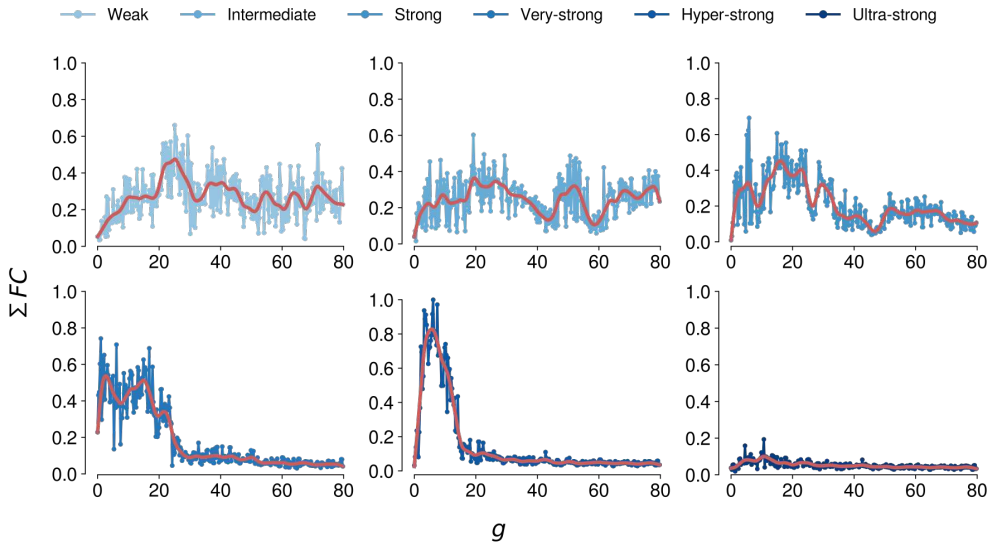

Figure A5.: Macroscale external current optimization via parameter sweep. Integration values outline regions in the parameter space that are less prone to stochasticity (i.e. regions where integration can be modeled as a strictly increasing function of connectivity).

E. Computation of source signal-to-noise ratio

What if our reported correlation of integration within attentional networks with behavioral performance were simply due to differences in signal-to-noise ratio (SNR) of the two experimental conditions? In order to rule out this possibility, source SNR of all functional networks were computed via the MNE package.

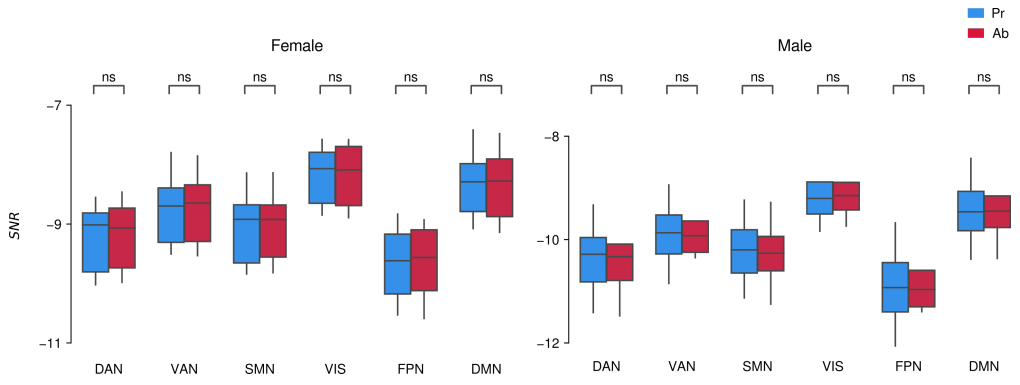

Figure A6.: Signal-to-noise ratio of source activity for the female (top) and male (bottom) subject groups. All networks measure similar values of SNR across both conditions.

SNR across conditions was not siginificantly different in any of the functional net-

works, which in turn can serve as further corroboration for the observed modulations of network coherence in attentional networks such as DAN and VAN (Figure A6).

F. Sensitivity analysis of generative models across brain scales

One critique of the results outlined here would entail that said findings might simply stem from hyper-sensitivity of model outputs to small parameters. Put another way, the estimation of SE from empirical observations might be put to jeopardy, given evidence towards generation of highly disparate outputs from small perturbation to parameter values.

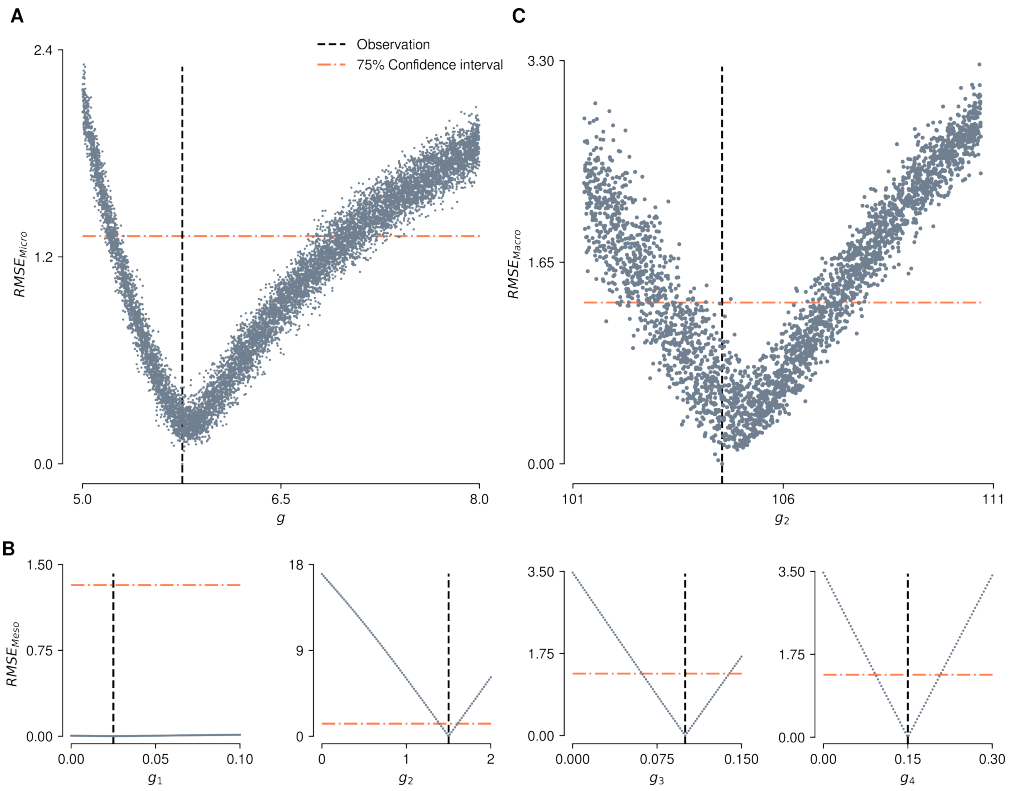

Figure A7.: **A**) Sensitivity analysis of the Brunel spiking network. **B**) Sensitivity analysis of the Jansen-Rit NMM **C**) Sensitivity analysis of the whole-brain coupled JR model.

To address this, we performed sensitivity analysis for all three of the aforementioned generative models. One conducts this analysis by fixing all but one of model parameters. Subsequently, the model is simulated based on sampled/binned parameter values from a suitable range (a range similar to that of the prior distribution in our case). Finally, root mean square error (RMSE) of simulated time series (or features

thereof) versus the ground-truth time series (obtained from the ground-truth parameter value) is computed. Thus, the variability in the model's error estimates can be derived using critical values from the chi-square distribution (df = 1), where the threshold Δ_α corresponds to the desired confidence level. These thresholds define likelihood-based regions with confidence level α. As evident from (Figure A7), RMSE consistently crosses the 75% confidence interval across all brain scales, which provides further support for the robustness of the inference process.

G. Hypothesis testing of empirically-informed inference

Irrespective of all synthetic measures which aim to ensure the robustness of the inference process, the distribution of empirical observations could also significantly influence the learned parameters.

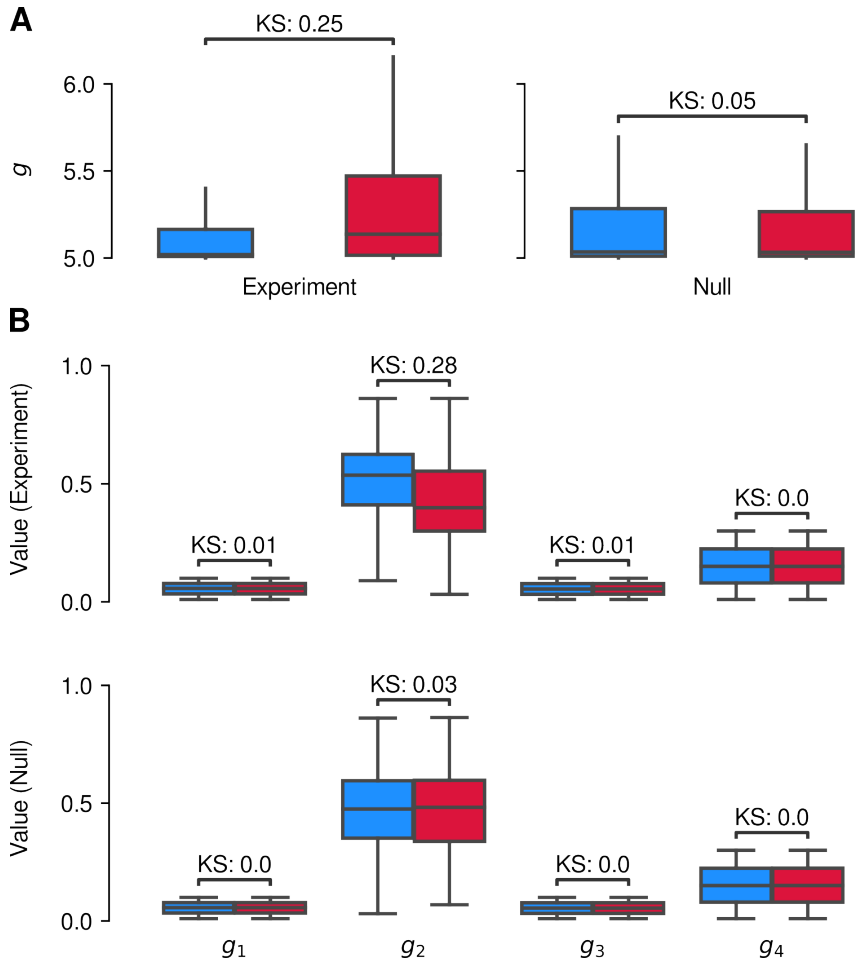

Figure A8.: **A**) The effect of data permutation in inference of microscale SE. **B**) The effect of data permutation in inference of mesoscale synaptic efficacies.

Thus, we conducted the inference process both on the normal experimental data, and on a control dataset. The control (i.e. null) dataset was generated via shuffling (i.e. randomizing the order of) the experimental trials/sessions. Figure A8 clearly demonstrates that data permutation disrupted the original pattern of data distribution in both microscale and mesoscale observations.