BACHELOR'S THESIS / BIOMEDICAL ENGINEERING 2024

Rebalancing the depressed brain: a whole-brain computational study on the effects of external perturbations in psilocybin and escitalopram treatments.

Marcel Socoró Garrigosa

upf. Universitat Pompeu Fabra Barcelona

Rebalancing the depressed brain: a whole-brain computational study on the effects of external perturbations in psilocybin and escitalopram treatments.

Marcel Socoró Garrigosa

Bachelor's thesis UPF 2023/2024

Thesis supervisor(s):

Prof. Gustavo Deco,

Dr. Jakub Vohryzek,

(Computational Neuroscience Group, Center for Brain and Cognition)

Universitat Pompeu Fabra Barcelona

Acknowledgments

This thesis would not have been possible without the help and support of several people.

First, I would like to express my deepest gratitude to Dr. Jakub Vohryzek, who has guided me throughout the entire process, offering invaluable insights and unwavering support. I especially appreciate your consistently positive attitude towards meeting, your openness to discussing new ideas, and your commitment to promoting good science.

I extend my sincere thanks to Prof. Gustavo Deco for providing me with the opportunity to work on such a fascinating topic. Your extensive expertise in computational neuroscience and the valuable feedback you have provided have been crucial in the development of this thesis.

Additionally, I would like to thank M.Sc. Irene Acero, Dr. Adrián Ponce, and Dr. Yonatan Sanz-Perl for assisting me with technical doubts at specific but crucial moments during the thesis.

I am deeply grateful to my parents for their continuous support throughout my degree, my sister for always being present and supportive, and the rest of my family for consistently believing in me.

Finally, a special thank you to the friends I have made at university, particularly Paul and Paula, for the countless hours spent working together this year. Your friendship and collaboration have been invaluable.

Abstract

Serotonergic psychedelics like psilocybin have been proposed as a promising avenue in the treatment of major depressive disorder (MDD). Contrary to selective serotonin reuptake inhibitors (SSRI) like escitalopram, psychedelics work through 5-HT2AR agonism, which opens a window of plasticity for psychotherapy- or neurostimulation-combined therapies. Here, we used a trial comparing psilocybin and escitalopram treatments for MDD to assess *in silico* the effect of external perturbations after pharmacological intervention. The trial included resting state fMRI scans before and one day after treatment. We built whole-brain models by fitting the data to capture the underlying causal brain mechanisms in generative effective connectivity. Then, we applied a perturbation protocol following the Dynamic Sensitivity Analysis paradigm, simulating the effects of external stimulations in parcellated regions. We evaluated the impact of perturbations by assessing regional susceptibility to change and effectivity to drive a healthy transition (both defined using static functional connectivity). We show that susceptibility is enhanced by psilocybin and reduced by escitalopram, evidencing an opening of a window of plasticity by psilocybin. However, healthy transitions are similarly achieved after both treatments, suggesting that escitalopram manages to improve perturbation effects despite differences in functional hierarchies compared to psilocybin. We elaborate that longer-term data and measures capturing the temporal richness of functional repertoires might be needed to complement these findings. Finally, we demonstrate the benefits of multi-site versus single-site perturbations and that the amygdala and the nucleus accumbens are the best perturbation targets to drive healthy transitions. Overall, the present work contributes to prior knowledge on how *in silico* perturbations can help in the treatment of MDD, serving as a foundation for larger computational studies.

Keywords

Depression, Functional magnetic resonance imaging, Whole-brain modelling, Dynamic sensitivity analysis, Neurostimulation, Psychedelics.

Preface

Throughout the degree, many subjects have captivated my interest. However, discovering the field of Computational Neuroscience was a turning point in my academic journey, as it integrated various intriguing fields I am especially passionate about. I was particularly interested in its application to neuropsychiatric disorders, and Prof. Gustavo Deco offered me the opportunity to work on depression at the Computational Neuroscience Lab within the Center for Brain and Cognition.

Major depressive disorder (MDD) is a devastating illness that acutely impairs quality of life. For instance, it is ranked as the major contributor to suicide deaths and global disability. The World Health Organization has estimated that over 300 million people suffer from depression worldwide, a number that has been rising in the last decades. First-line treatment involves pharmacological therapy, which is preferred in combination with psychotherapy. Selective serotonin reuptake inhibitors (SSRIs) such as escitalopram are the gold-standard treatment, but new antidepressants are needed to improve effectiveness and reduce side effects. Serotonergic psychedelics have recently emerged as an alternative avenue to SSRIs, having shown antidepressant action in several trials. They are thought to work by disrupting ruminative thinking and promoting insight, opening a plasticity window that could be harnessed by psychotherapy or neurostimulation strategies.

In this thesis, we use whole-brain computational models to simulate *in silico* the effect of external perturbations on psilocybin- and escitalopram-treated brains. Whole-brain models can be thought of as mathematical models that represent the brain as a network of connected nodes exhibiting local oscillations, and stimulations of a given node are simulated by perturbing the properties of its oscillations. We assess perturbation performance with two principal measures: susceptibility, which quantifies the capacity for change, and Perturbation Effectivity to Recovery (PER), which measures the ability of perturbations to drive healthy transitions.

The main findings suggest a window of plasticity is opened by psilocybin and support the use of psychotherapy or neurostimulation therapies following psilocybin and escitalopram treatments. We also explore the benefits of multi-site versus single-site stimulations, as well as which are the best areas to perturb. Overall, the thesis can be considered an initial effort into how external perturbations can be simulated using whole-brain modelling, here in the context of MDD. It is particularly relevant, provided it relies on the recent linear approximation of the whole-brain Hopf model, which allows faster computations than other methods. Ultimately, the present work aims to contribute to the advancement of *in silico* explorations of neurostimulation and combined therapies in MDD.

Index

List of figures

1. INTRODUCTION

1.1 Depression

a) Impact

Major depression disorder (MDD) is a complex and common disease that acutely limits psychosocial functioning and diminishes quality of life [1]. Characterized by recurrent depressive episodes, manifestations can encompass emotional, neurovegetative and cognitive symptoms [1]. From a diagnostic point of view, according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), anhedonia or a depressed mood are fundamental symptoms to MDD, with other lower-rank symptoms including suicidal ideation, poor concentration, sleep alteration or feelings of guilt, among others [1], [2].

Epidemiologically, MDD is estimated to affect 322 million people worldwide, a number that represents 4.4% of the global population [3]. Moreover, despite variations across countries, its overall 12-month prevalence is approximately 12% [4], evidencing its massive impact on society. Indeed, it has been reported to be the largest contributor to global disability and suicide deaths [3], and it is expected to become the leading disease burden worldwide by 2030 [5].

b) Treatment

Therapeutic options for MDD include psychological therapy, pharmacology, and neurostimulation solutions like electroconvulsive therapy (ECT), which is employed for patients that do not respond to medication [1]. For some mild cases, psychotherapy alone may suffice [1], but severe cases typically require medication, which seems to be the optimal solution when combined with psychotherapy [6].

First-line antidepressants typically involve selective serotonin reuptake inhibitors (SSRIs) [1]. These drugs work by inhibiting the serotonin transporter (SERT), which causes an increase in serotonin concentration in synaptic clefts [7]. This is consistent with the *monoamine hypothesis* of depression, whereby alterations of monoamine neurotransmitters like serotonin underlie the emergence of depressive states [7], [8]. SSRIs are typically preferred over other antidepressants due to their efficacy, safety, and tolerability [7]. Yet, they take several weeks to work, and, in some cases, they do not elicit a response [9]. A common SSRI for depression is escitalopram, the SSRI with the highest specificity for SERT [10].

Recently, serotonergic psychedelics have shown promise in treating depression [11], [12]. Evidence suggests that they work mainly but not only through 5-HT2A serotonin receptor (5-HT2AR) agonism, which increases serotonin release and opens a window of plasticity from which supportive psychotherapy could benefit [13]. However, the exact mechanism of action is not well understood [14]. A promising serotonergic psychedelic compound is psilocybin, which has demonstrated great tolerability and safety in several studies [12], [15]. Treatment efficacy has also been reported, but more evidence is needed [12].

Regarding neurostimulation, besides ECT, promising solutions include transcranial magnetic stimulation (rTMS), transcranial direct current stimulation (tDCS), or deep brain stimulation (DBS) [16]. These options are often employed in treatment-resistant depression (TRD), which affects between 20% and 30% of patients, and its combination with antidepressant medication is currently being researched [17], [18].

1.2 Neuroimaging brain connectivity

Several neuroimaging modalities remain crucial in advancing our understanding of the brain. Magnetic Resonance Imaging (MRI) is one of the most versatile in its different uses [19].

On one hand, diffusion magnetic resonance imaging (dMRI) techniques like diffusion-weighting/tensor imaging (DWI/DTI) estimate the mapping of white-matter pathways in the brain [20]. This knowledge is inferred from the anisotropic diffusion of water molecules along white-matter fibre bundles [20]. From the estimated pathways, the connectivity between brain regions can be reconstructed, giving rise to *structural connectivity* (SC) [20], [21]. However, it is important to note that diffusion tractography is restricted to mapping large-scale axonal groups, thereby unable to identify individual axons or synapses between neurons [21]. Moreover, it cannot provide information on directionality [22].

On the other hand, fMRI is a non-invasive technique measuring the blood-oxygen level-dependent (BOLD) signal [19]. This signal measures changes in oxygenation concentration, which occur in response to the transient oxygen deficit when a given region of the brain is up-regulated [23]. FMRI is therefore an indirect measure of neuronal activity [19]. It can be used to determine *functional connectivity* (FC), which refers to the statistical dependencies among neurophysiological neuronal signals and can be computed using measures like correlation or mutual information between local time series [24]. Conventionally, FC assumes temporal stationarity of functional connections between brain regions, resulting in a static FC (sFC) [25]. More recently, methods have been proposed to compute dynamic FC (dFC), considering the temporal evolution of statistical dependencies between sub-portions of time series [25], [26]. Despite the success of fMRI, its main limitation is its poor signal-to-noise ratio (SNR), which constrains many task-activation paradigms [27]. However, resting-state fMRI (rs-fMRI), where subjects are restful with their eyes closed, remains a powerful approach to overcome this limitation [27].

Besides SC and FC, there is a third type of brain connectivity named *effective connectivity* (EC). At the population level, this connectivity directly captures causal influences between brain regions, and not mere statistical dependencies in neuronal activity which might contain confounders [28]. However, this connectivity cannot be directly accessed through imaging, but rather inferred from mechanistic models of interactions able to reproduce the observed neuronal patterns of activity [28].

SC, FC and EC give measures of how different brain regions are related, either anatomically, through statistical dependence of neuronal activity, or through mechanistic causal influence, respectively. Mathematically, this information can be represented as a network of connected nodes of well-defined brain areas. The boundaries of the different

areas are obtained from a *parcellation*, usually derived from anatomical [29], functional [30] atlases, or data-driven templates [31]. Existing parcellations can range from having tens to several hundreds of regions [32]. Some popular examples include the automated anatomical labelling (AAL) with 116 cortical, subcortical, and cerebellar regions [33], the Lausanne parcellation with 66 cortical regions [34], and the Desikan-Killiany (DK) parcellation with 68 cortical regions [35]. A modified version of the latter will be relevant for this thesis.

After a parcellation is chosen, neuroimaging signals from fMRI or DTI can be parcellated [36], allowing the computation of FC or SC, respectively. FC typically involves computing correlation measures between the average time series of each region's voxels, whereas SC arises from counting the number of white matter fibres connecting voxels within distinct regions [37].

1.3 State of the art

a) Pathoconnectomics

Neuroimaging advancements in the field of structural and functional connectivity have given rise to the concept of the *connectome*, defined as a "comprehensive map of neural connections in the brain on many spatial scales" [24], [38]. The resolution provided by modern imaging techniques combined with the graph theoretical understanding of brain organisation has facilitated the discovery of key features of brain networks [24], like *small-worldness* [39] and *modularity* [40]. Graph theoretical measures have been able to properly characterise and define the topology of the healthy human brain [41], and the study of their change in neuropsychiatric diseases has emerged as a new promising research field [27], which some have coined *pathoconnectomics* [42]. The hope of pathoconnectomics is that the study of brain network properties can help in identifying biomarkers of neuropsychiatric diseases, thereby contributing to diagnosis, prognosis, or the assessment of response to interventions [24].

Important findings have been made in identifying differences between healthy and neuropsychiatric populations, both in terms of functional and structural connectivity [27]. Some of the examples include dementia [43], Alzheimer's disease [44], posttraumatic stress disorder [45], schizophrenia [46], and major depression [47], [48].

In depression, MDD states have been found to exhibit greater global synchronisation and temporal stability compared to healthy participants [49]. Moreover, brain organization has been shown to involve an overactivation of the default-mode network (DMN) as well as hypoconnectivity in the control executive network (CEN) [50], [51]. These findings support the *triple-network model* of MDD, where an enhancement of self-referential mental activity and an impoverishment of attention towards the external environment are promoted due to deficits in DMN, CEN, and the salience network (SN) [52], [53]. In other words, depressive states are characterised by a rigidity of high-level priors to any incoming information.

As mentioned earlier, a great promise in depression treatment are serotonergic psychedelics. From a dynamical system's perspective, these substances are known to enrich the brain's spatiotemporal dynamics [54]. Related evidence includes an

enlargement of the repertoire of possible FC patterns occurring across time as well as a shift in brain functional organization towards an increased large-scale integration between regions [54], [55], [56]. In depression, neuroimaging studies have found that the psychedelic experience induces decreases in functional connectivity in DMN, so they are thought to disrupt ruminative thinking and allow insight [13], [57]. These ideas are formalised in the *RElaxed Beliefs Under pSychedelics* (REBUS) model, which proposes that psychedelics would act by relaxing aberrant high-level priors and increasing the sensitivity for incoming bottom-up signalling [19], [58]. The idea of putting internal priors and bottom-up signalling on the same footing is further encapsulated in the concept of the *anarchic brain* [58].

Topological characterization of diseased brain networks using FC, SC, or related measures has proven valuable and insightful for many diseases. However, results are always descriptive and correlational, and causal links between SC and FC are missing [24]. To fully understand neuropsychiatric diseases like depression, a deeper understanding of how brain structure relates to function is needed [24].

b) Whole-brain modelling

A powerful and emerging field in computational neuroscience, helping to bridge the gap between brain structural and functional connectivity, is *whole-brain modelling* [24]. Large-scale models represent the brain as a network of different brain areas, connected through white-matter tracts (here informed by SC), according to a chosen parcellation, with the activity of a node represented by the local dynamics of that node [59], [60].

Local brain dynamics are modelled based on a trade-off between biological complexity and realism [24]. Typical neuronal dynamics models include mean-field approximations of neural mass activity like the Wilson-Cowan model [61], or phenomenological models of coupled oscillators like Kuramoto and Hopf models [62], [63], which have been shown to exhibit synchronization mechanisms [64]. Most studies consider average functional connectivity; however, since brain dynamics is marked by transitions between several states, some have started to model node dynamics with multistable solutions [65], [66].

The activity of brain regions can be simulated, enabling the computation of a simulated FC. By fitting the whole-brain model to empirical data, anatomical connections in SC can be refined, generating a measure of effective connectivity known as Generative Effective Connectivity (GEC). In this thesis, GEC optimization was achieved by using asymmetrical optimization arising from the non-reversibility INSIDEOUT method proposed in Ref. [36]. In brief, this framework considers network non-reversibility (i.e., the arrow of time) as a measure that quantifies the breaking of detailed balance, which, if obeyed, implies the disappearance of flux transitions between the system states [67]. This breaking is in turn informative of hierarchical organization, with some nodes acting more as top-down regulators and others as bottom-up signallers (i.e., there is a hierarchy between nodes in terms of their balance between intrinsic and extrinsic activity). Consequently, the INSIDEOUT method establishes that in a high-hierarchical network (i.e., highly breaking the detailed balance), its time-reversed dynamics become less plausible (i.e., the arrow of time), and therefore, hierarchy can be inferred through the level of non-reversibility between signals, a measure that can be defined as the difference between forward and reversal shifted correlation of fMRI BOLD time series [36]. GEC

optimization must therefore arise not only from fitting simulated FC to empirical FC but also from fitting the levels of simulated and empirical non-reversibility. Consequently, this simulation framework allows a mechanistic bridging between structure and function and, hence, can help in the *in silico* exploration of how damage to SC affects FC [24]. This framework is specifically known as *generative whole-brain modelling*.

Whole-brain models have been used to gain mechanistic insight into how aberrant network activity affects brain disease, for instance, by assessing the impact of damage to structural connectivity on brain function. Examples include modelling brain lesions [68], assessing integration impairment in consciousness disorders [69], and testing the disconnection hypothesis in schizophrenia [70].

In depression, a recent study by Deco *et al.* involving the use of whole-brain models has studied the pharmacological effects of psilocybin and escitalopram in brain hierarchical organization [71]. This study has found that brain hierarchy increases with escitalopram and diminishes with psilocybin, despite no major differences being present in terms of treatment efficacy. Such findings support the framework of the anarchic brain under psychedelics and open the question of which antidepressant achieves a brain organization more able to achieve a healthy control's topology, i.e., more responsive to external perturbations towards a healthy state. Elucidating favourable responsiveness to perturbations could be key to determining which antidepressant is more convenient for a combined therapy with psychotherapy or neurostimulation techniques.

c) Dynamic Sensitivity Analysis

Whole-brain computational models allow us to infer causal mechanisms underlying the pathophysiology of brain diseases. However, they do not intend to study transitions between brain states, which in a clinical setting could signify simulating the consequences of therapeutic strategies [72]. Such *in silico* explorations would not only mitigate the suboptimal efficacy of trial-and-error approaches in clinical trials but also alleviate ethical constraints associated with human experimentation [72].

Plausible transitions between various brain states can now be found *in silico* with frameworks like Dynamic Sensitivity Analysis [72], whose paradigm was introduced by Deco and colleagues in a computational study on transitions between wakefulness and sleep [73]. In the context of disease, Dynamic Sensitivity Analysis can be applied by simulating brain stimulation and assessing under which stimulations spatio-temporal dynamics are driven to that of a healthy brain.

Different properties of stimulations can be tuned depending on the underlying model for local dynamics. Examples include the type of stimulation (i.e., increasing excitation, modulating frequency, or adding noise, among others), the stimulation intensity, the location (i.e., local or global), and the duration (i.e., long-term, pulse-based or periodic) [72]. Moreover, stimulations can be applied extrinsically or intrinsically by either adding or modifying a term in the equations of the governing model, respectively. In this thesis, we will differentiate between single-site and multi-site perturbations, depending on whether only one or multiple brain regions are simultaneously perturbed.

Recent studies have applied this framework to assess stimulation strategies in neuropsychiatric disorders including Alzheimer's disease [74] and MDD [53]. In these studies, local dynamics were modelled using a phenomenological model, namely Stuart-Landau oscillators (Hopf model), and perturbations were simulated by perturbing how nodal dynamics respond to added noise [53]. In depression, *in silico* perturbations were applied to psilocybin-treated patients, finding that the areas that drive healthy transitions in responders compared to non-responders match those with a higher density of serotonergic $5HT_{2A}$ and $5HT_{1A}$ receptors [53]. This aligns with psilocin (the main metabolite of psilocybin) affinity for these receptors, with $5HT_{2A}$ agonism thought to be relevant to antidepressant function by activating plasticity signalling-cascades [53], [75]. With these findings regarded as "exploratory" [53], more studies are needed to exploit the potential of Dynamic Sensitivity Analysis in depression. For instance, how perturbations affect functional organization following antidepressant treatment could be explored to assess the *a priori* effectivity of psychotherapy- or neurostimulation-combined therapies.

1.4 Objectives

Ultimately, the goal of this thesis is to elucidate how does the depressed brain respond to external perturbations after psilocybin and escitalopram treatments. This will eventually inform on the success of psychotherapy- or neurostimulation-combined therapy following psilocybin and escitalopram treatments, which is of utmost relevance to design effective treatment strategies for depression.

To attain the main thesis objective, several secondary objectives were pursued. These included: (1) preprocessing the utilized dataset containing psilocybin and escitalopram pre- and post-treatment fMRIs, (2) using generative whole-brain models to obtain group and patient GEC matrices, (3) validating GEC computation using fitting measures and verifying known hierarchy measures—trophic coherence and directedness—, (4) *in silico* perturbing whole-brain models using single-site and multi-site approaches, (5) creating a healthy target from 100 healthy empirical FCs, and (6) assessing the success of perturbations by measuring their susceptibility and their effectivity to drive healthy transitions. Note that objectives 1-3 involved replicating results obtained by Deco *et al.* [71], whereas objectives 4-6 constituted the principal contribution to existing work.

1.5 Hypothesis

The hypothesis of this thesis is that psilocybin will increase the susceptibility of brain functional organization to external perturbations, whereas escitalopram will significantly diminish it. This proposition is grounded in the REBUS model of psychedelics and prior studies on how psilocybin and escitalopram drive distinct hierarchical reconfigurations [71]. According to this prior knowledge, a relaxation of brain hierarchy by psilocybin should enable higher susceptibility to perturbations, whereas a reinforcement of the hierarchy by escitalopram should reduce overall susceptibility. Following this hypothesis, we expect psilocybin to differentially enhance perturbations' capacity to drive healthy transitions, based on the assumption that an increased capacity for change should facilitate finding healthy transitions.

2. MATERIALS AND METHODS

2.1. Materials

a) Trial of psilocybin and escitalopram

We utilized a dataset corresponding to a phase 2 double-blind randomized controlled trial released by Cahart-Harris *et al.* in 2021 [76]. This trial comprised patients with long-standing moderate-to-severe depression and aimed to compare the pharmaceutical effects of escitalopram and psilocybin.

To be eligible for the trial, patients needed a confirmed diagnosis of MDD from a general practitioner, scoring 16 or higher on the Hamilton Depression Rating scale [77]. In total, 59 patients with MDD were recruited and randomly assigned to either the psilocybin or the escitalopram arm, with a proportion of 30 and 29, respectively. From this group, the final sample for neuroimaging lowered to 22 patients for psilocybin and 21 patients in the escitalopram arm. Two resting-state fMRI sessions with a 3-T Tim Trio scanner (Siemens) were programmed for each patient: before and after treatment, which lasted 6 weeks and 1 day. Patients on the psilocybin arm received 25 mg of psilocybin on the first dosing day (DD1) and three weeks later, on the second dosing day (DD2). The escitalopram arm received a negligible dose of 1 mg instead. After DD1, patients took daily capsules containing 10 mg of escitalopram or an inert placebo, depending on their assigned arm. Finally, the dosage was incremented to two daily capsules after the first three weeks. Further details on acquisition or exclusion criteria can be found in Ref. [76].

We accessed fMRI data from 42 patients (22 taking psilocybin and 20 taking escitalopram). Specifically, we directly worked with regional fMRI signals according to the DK80 parcellation [78], which is an extension of the Mindboggle-modified Desikan-Killiany parcellation [35] containing 62 cortical regions and 18 subcortical regions. The list of all brain regions in DK80 can be found in Table 1.

b) Human Connectome Project data

Given that the psilocybin and escitalopram trial did not include diffusion MRI scans, we accessed a standard structural connectivity (SC) template. This connectivity had been estimated in Ref. [78] from two dMRI projects of the Human Connectome Project (HCP). The aim of HCP, which is sponsored by sixteen components of the National Institutes of Health, is to elucidate anatomical and functional connectivity in health and disease by mapping the human connectome. Direct access to the DK80 SC was provided, upon its extraction using a pipeline described in Refs. [78] and [79].

Finally, we accessed HCP fMRI data from healthy subjects to construct a healthy target. We focused on 100 unrelated subjects from the project Young Adult, following Ref. [80]. Participants underwent four rs-fMRI sessions of 14 min 30 s with a repetition time of 0.72 s on a 3-T Skyra scanner (Siemens). We focused on DK80 regional fMRI signals corresponding to data from the first scanning session. Acquisition and pre-processing details can be found at the HCP website (http://www.humanconnectome.org/).

2.2. Methods

All computations presented in this thesis were performed using MATLAB (R2023b) [81].

a) Data preprocessing

Data preprocessing consisted of detrending regional fMRI BOLD signals and then filtering them with a second order Butterworth filter in the range of 0.008-0.08 Hz. Signals were filtered to avoid noise and aliasing, e.g., caused by respiratory and cardiac frequencies in the higher frequency range of the BOLD signal [82].

b) Empirical computations

Once regional fMRI signals were processed, empirical covariances $\boldsymbol{C_v}$ and empirical functional connectivity (FC) matrices $\boldsymbol{FC}$ were constructed by computing for each ij-th pair of signals $(x_i(t), x_j(t))$

$$C_{v_{ij}}^{\ empirical} = \sum_t \left(x_i(t) - \overline{x_i(t)}\right)\left(x_j(t) - \overline{x_j(t)}\right), \tag{1}$$

$$FC_{ij}^{empirical} = \frac{C_{v_{ij}}^{\ empirical}}{\sqrt{C_{v_{ii}}^{\ empirical}\, C_{v_{jj}}^{\ empirical}}}, \tag{2}$$

where $\overline{\ }$ indicates a temporal average. Similarly, forward shifted empirical covariances $\boldsymbol{C_v}(\tau)$ and shifted FC matrices $\boldsymbol{FS}(\tau)$, were obtained with:

$$C_{v_{forward_{ij}}}^{\ empirical}(\tau) = \sum_t \left(x_i(t+\tau) - \overline{x_i(t+\tau)}\right)\left(x_j(t) - \overline{x_j(t)}\right), \tag{3}$$

$$FS_{forward_{ij}}^{empirical}(\tau) = \frac{C_{v_{forward_{ij}}}^{\ empirical}(\tau)}{\sqrt{C_{v_{ii}}^{\ empirical}\, C_{v_{jj}}^{\ empirical}}}, \tag{4}$$

where τ is the shifting time, which was set to 2 times the repetition time. Note that shifting makes $\boldsymbol{C_v}_{forward}^{\ empirical}(\tau)$ and $\boldsymbol{FS}_{forward}^{empirical}(\tau)$ asymmetric matrices. Note also that normalization in $\boldsymbol{FS}_{forward}^{empirical}(\tau)$ by $C_{v_{ii}}^{\ empirical}$ and $C_{v_{jj}}^{\ empirical}$ is done without any shifting, so strictly speaking, we should refer to $\boldsymbol{FS}(\tau)$ as a normalized forward shifted covariance [71]. However, for the sake of clarity and simplicity, $\boldsymbol{FS}(\tau)$ will be loosely referred to as forward shifted functional connectivity.

c) Whole-brain model

Empirical findings were used to inform the whole-brain model. Based on prior work of Deco *et al.* [71], the dynamics of individual nodes were modelled using a Stuart-Landau oscillator model, which corresponds to the supercritical Hopf bifurcation model. Specifically, its recent linear approximation was employed, to facilitate faster computation times [83]. For an isolated node of interest j, the normal form of the supercritical Hopf bifurcation is defined in Cartesian coordinates by

$$\frac{dx_j}{dt} = (a_j - x_j^2 - y_j^2)x_j - \omega_j y_j + \beta\eta_j, \tag{5}$$

$$\frac{dy_j}{dt} = (a_j - x_j^2 - y_j^2)y_j + \omega_j x_j + \beta\eta_j, \tag{6}$$

where $\frac{\omega_j}{2\pi}$ is the intrinsic frequency of the j-th node and $\beta\eta_j$ is additive white noise with a standard deviation $\beta = 0.01$, homogeneously set for all nodes. Following Ref. [71], intrinsic frequencies were estimated from regional narrowband BOLD signals by taking their patient-averaged peak frequency. The bifurcation parameter a places the system in a noisy regime if $a < 0$ and a stable limit cycle if $a > 0$, being $a = 0$ the bifurcation point [83]. For the thesis, a slightly negative bifurcation parameter very close to the bifurcation point was used ($a = -0.02$), since it has been shown to represent an optimal working point for fitting whole-brain neuroimaging data [64].

To model the dynamics of brain regions, regional fMRI signals were modelled by x_j, and y_j remained an auxiliary variable. These variables can indeed be treated as the real and imaginary parts of a state variable $z_j = x_j + iy_j$ in the original complex coordinate formulation [83]. Moreover, since brain areas are not isolated, a term capturing external influence of other brain regions was added [71]

$$\frac{dx_j}{dt} = (a_j - x_j^2 - y_j^2)x_j - \omega_j y_j + \sum_{k=1}^{N} C_{jk}(x_k - x_j) + \beta\eta_j, \tag{7}$$

$$\frac{dy_j}{dt} = (a_j - x_j^2 - y_j^2)y_j + \omega_j x_j + \sum_{k=1}^{N} C_{jk}(y_k - y_j) + \beta\eta_j, \tag{8}$$

where C_{jk} indicates the connectivity weight between area k and area j.

Eqs. (7) and (8) can be rewritten in vector form:

$$\frac{d\boldsymbol{x}}{dt} = (\boldsymbol{a} - \boldsymbol{x}^2 - \boldsymbol{y}^2 - g\boldsymbol{s})\odot\boldsymbol{x} - \boldsymbol{\omega}\odot\boldsymbol{y} + \boldsymbol{C}\boldsymbol{x} + \beta\boldsymbol{\eta} \tag{9}$$

$$\frac{d\boldsymbol{y}}{dt} = (\boldsymbol{a} - \boldsymbol{x}^2 - \boldsymbol{y}^2 - g\boldsymbol{s})\odot\boldsymbol{y} + \boldsymbol{\omega}\odot\boldsymbol{x} + \boldsymbol{C}\boldsymbol{y} + \beta\boldsymbol{\eta} \tag{10}$$

where $\boldsymbol{x} = [x_1, \dots, x_N]^T$, $\boldsymbol{y} = [y_1, \dots, y_2]^T$, $\boldsymbol{a} = [a_1, \dots, a_N]^T$, $\boldsymbol{\omega} = [\omega_1, \dots, \omega_N]^T$, $\boldsymbol{\eta} = [\eta_1, \dots, \eta_N]^T$, $\boldsymbol{s} = [s_1, \dots, s_N]^T$ is the vector containing the strength of each node, i.e., $s_i = \sum_j C_{ij}$, $[\ \ ]^T$ represents the transpose and $\odot$ is the Hadamard element-wise product, i.e., $\mathbf{u}\odot\mathbf{v} = [\mathrm{u}_1\mathrm{v}_1, \dots, \mathrm{u}_N\mathrm{v}_N]$.

The proximity to criticality given by the bifurcation parameter allows the linearization of the dynamics around the fixed point $(\boldsymbol{x}, \boldsymbol{y}) = (\mathbf{0}, \mathbf{0})$, which for $a < 0$ is stable.

Linearization allows dropping high order terms in Eqs. (9) and (10), and the dynamics of linear fluctuations can be modelled using a Langevin stochastic differential equation

$$\frac{d}{dt}\delta\boldsymbol{u} = \boldsymbol{J}\delta\boldsymbol{u} + \boldsymbol{\eta}, \tag{11}$$

where fluctuations of $\boldsymbol{x}$ and $\boldsymbol{y}$ are described by a 2N-dimensional vector $\delta\boldsymbol{u} = [\delta\boldsymbol{x}, \delta\boldsymbol{y}]^{\boldsymbol{T}} = [\delta x_1, \dots, \delta x_N, \delta y_1, \dots, \delta y_N]^T$, $\boldsymbol{\eta}$ is a white noise 2N-dimensional vector, and $\boldsymbol{J}$ is the 2N × 2N Jacobian of the system evaluated at the fixed point $(\boldsymbol{x}, \boldsymbol{y}) = (\mathbf{0}, \mathbf{0})$, given by

$$\boldsymbol{J} = \begin{bmatrix} \boldsymbol{J_{xx}} & \boldsymbol{J_{xy}} \\ \boldsymbol{J_{yx}} & \boldsymbol{J_{yy}} \end{bmatrix}, \tag{12}$$

where $\boldsymbol{J_{xx}}, \boldsymbol{J_{xy}}, \boldsymbol{J_{yx}}, \boldsymbol{J_{yy}}$ are N × N matrices given by $\boldsymbol{J_{xx}} = \boldsymbol{J_{yy}} = \boldsymbol{C} + diag(\boldsymbol{a} - \boldsymbol{s})$ and $\boldsymbol{J_{yx}} = -\boldsymbol{J_{xy}} = diag(\boldsymbol{\omega})$, being $diag(\boldsymbol{v})$ a diagonal matrix with a vector $\boldsymbol{v}$ on the diagonal [71], [83].

This equation helps to find the temporal evolution $\delta\boldsymbol{u}$, from which network statistics can be extracted. For instance, the covariance of fluctuations $\boldsymbol{C_v^{simulated}} = \langle \delta\boldsymbol{u}\delta\boldsymbol{u}^{\boldsymbol{T}} \rangle$ is here relevant, since it will allow to define simulated FC. From Eq. (11), the differential equation governing $\boldsymbol{C_v}$ can be deduced

$$\frac{d\boldsymbol{C_v}}{dt} = \boldsymbol{J}\boldsymbol{C_v} + \boldsymbol{C_v}\boldsymbol{J}^{\boldsymbol{T}} + \boldsymbol{Q_n}, \tag{13}$$

where $\boldsymbol{Q_n}$ is the noise covariance matrix $Q_n = \langle \boldsymbol{\eta}\boldsymbol{\eta}^{\boldsymbol{T}} \rangle$. If we $\frac{d\boldsymbol{C_v}}{dt}$ is set to zero, the stationary covariance can be obtained by solving:

$$\boldsymbol{J}\boldsymbol{C_v} + \boldsymbol{C_v}\boldsymbol{J}^{\boldsymbol{T}} + \boldsymbol{Q_n} = \mathbf{0}, \tag{14}$$

This is a Lyapunov equation that has a unique solution provided that $\boldsymbol{J}$ is asymptotically stable. It can be solved using the eigen-decomposition of the Jacobian matrix $\boldsymbol{J}$ [83], [84]. Here, it was solved using MATLAB solver *sylvester()*.

Note that the stationary covariance gives a measure of how fluctuations from different regions covary at stationarity, i.e., when the stable origin $\boldsymbol{z} = \mathbf{0}$ is reached. The stationary covariance can be used to calculate the forward lagged covariance, defined as $\boldsymbol{C_{v_{forward}}}(\tau) = \langle \delta\boldsymbol{u}(t+\tau)\delta\boldsymbol{u}(t)^{\boldsymbol{T}} \rangle$, using:

$$\boldsymbol{C_{v_{forward}}}(\tau) = \boldsymbol{e}^{\boldsymbol{\tau J}}\boldsymbol{C_{v_{forward}}}(0) \tag{15}$$

Where $\boldsymbol{C_{v_{forward}}}(0) = \boldsymbol{C_v}$ is the static covariance (i.e., with zero lag). For deeper mathematical insight, see Ref. [83]. Since we are only interested in the real part of network statistics, simulated covariance $\boldsymbol{C_v}^{simulated}(\tau)$ and simulated shifted covariance $\boldsymbol{C_{v_{forward}}}^{simulated}(\tau)$ matrices will refer to the matrices arising from taking the first N rows and

N columns of $\boldsymbol{C_{v_{forward}}}(\tau)$ and $\boldsymbol{C_v}$, respectively. Analogously to empirical computations, simulated static FC $\boldsymbol{FC}^{simulated}$ and shifted FC $\boldsymbol{FS}^{simulated}_{forward}(\tau)$ can be obtained with:

$$FC_{ij}^{simulated} = \frac{C_{v_{ij}}^{simulated}}{\sqrt{C_{v_{ii}}^{simulated} C_{v_{jj}}^{simulated}}} \quad (16)$$

$$FS_{forward_{ij}}^{simulated}(\tau) = \frac{C_{v_{forward_{ij}}}^{simulated}(\tau)}{\sqrt{C_{v_{ii}}^{simulated} C_{v_{jj}}^{simulated}}} \quad (17)$$

d) Model optimization

The Hopf whole-brain model takes a matrix $\boldsymbol{C}$ and simulates on top of it the dynamics the different brain regions, allowing the computation of both static and shifted simulated FC matrices through network statistic measures of real-part fluctuations.

Here, patient-specific SC matrices were not available, so a generic SC template was used to construct the initial $\boldsymbol{C}$. More specifically, $\boldsymbol{C}$ was set as a scaled version of SC, ensuring a maximal value of 0.2. This was done to ensure consistence with prior literature set of parameters [64].

Through an iterative process, models were then optimized. At each iteration, $\boldsymbol{C}$ weights were refined so as to achieve maximal resemblance between $FC_{ij}^{empirical}$ and $FC_{ij}^{simulated}$ and between $FS_{forward,ij}^{empirical}(\tau)$ $FS_{forward,ij}^{simulated}$, using as learning rates $\alpha = 0.0004$ and $\varsigma = 0.0001$:

$$C_{ij} = C_{ij} + \alpha\left(FC_{ij}^{empirical} - FC_{ij}^{simulated}\right) + \varsigma\left(FS_{forward,ij}^{empirical}(\tau) - FS_{forward,ij}^{simulated}(\tau)\right) \quad (18)$$

Every 100 iterations, a total error was calculated, given by:

$$error_{total} = FC_{ij}^{empirical} - FC_{ij}^{simulated} + FS_{forward,ij}^{empirical}(\tau) - FS_{forward,ij}^{simulated}(\tau) \quad (19)$$

The optimization loop was exited whenever the error varied less than 0.1% with regards to the previous error computed. The resulting optimized $\boldsymbol{C}$ matrix is called Generative Effective Connectivity (GEC) and contains the effective conductivity values for each existing pair of anatomical connections and not just the dMRI based density of fibres [71].

Fitting $FS_{forward,ij}^{simulated}(\tau)$ to $FS_{forward,ij}^{empirical}(\tau)$ is intended to fit the levels of non-reversibility, a measure that has been used to describe the hierarchy of brain networks [36] and whose formal definition is introduced in MA1. This fitting is hence essential to preserve hierarchical organization in GEC matrices. To accelerate the fitting process, $\boldsymbol{C}$ optimization was divided in two stages. First, $\boldsymbol{C}$ was fitted to group average empirical FC

matrices, obtaining group GECs for each of the defined groups: psilocybin before treatment (PB), psilocybin after treatment (PA), escitalopram before treatment (EB) and escitalopram after treatment (EA). Second, group GECs were used as initial $\boldsymbol{C}$ matrices to fit patient-specific models, using individual empirical FCs as targets for simulated FCs. Fig. 1 shows a simplified graphical representation of the pipeline for GEC generation through whole-brain modelling.

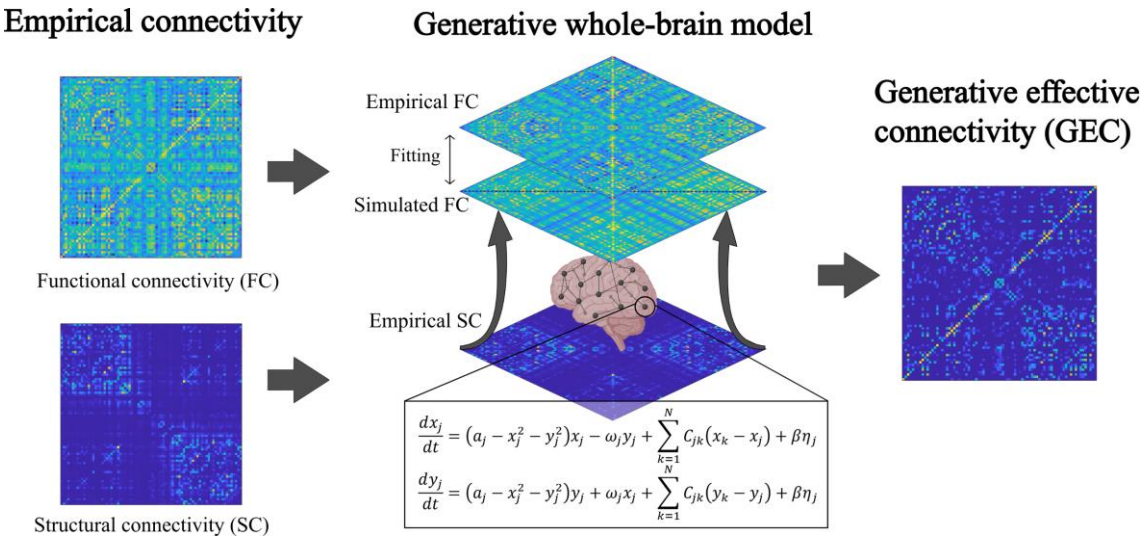

Figure 1. Pipeline for GEC generation. Once empirical FC and SC are obtained, a generative whole-brain model using Stuart-Landau oscillators is used to obtain a GEC matrix. Based on the empirical SC, the model first generates a simulated FC, which is then fitted to empirical FC in an iterative process. The connectivity matrix is thereby modified until convergence, obtaining GEC.

e) Model validation

Whole-brain models were validated in a two-fold approach. First, the convergence of the fitting process was assessed by tracking the total error. Moreover, the correlation between $\boldsymbol{FC}^{empirical}$ and $\boldsymbol{FC}^{simulated}$ as well as between $\boldsymbol{FS}^{empirical}(\tau)$ and $\boldsymbol{FS}^{simulated}(\tau)$ was computed.

For further validation, the topology of GEC matrices was then analysed using some network-theoretical measures. Since Deco and colleagues [71] analysed the effect of psilocybin and escitalopram on functional hierarchical organization using the same data, the same hierarchy measures were computed, namely *hierarchical levels* (or trophic levels) and *trophic coherence* (or directedness) adapted for directed networks [85]. Given that this still represents a verification step with the current literature and does not constitute the main contribution of the thesis, the formal definition of both measures can be found in MA2.

f) Perturbation protocol

Once GECs were calculated for all the depressed patients, new whole-brain models were built through the lenses of external perturbation. In this case, patient-specific GECs were used to inform models' connectivity, but this time, models carried some parameter disturbance to simulate the effect of an external perturbation. For each patient, single-site perturbations were simulated by increasing white noise standard deviation β in a particular Stuart-Landau oscillator (see Eqs. (7) and (8)), corresponding to a particular node or brain area. These perturbations can be understood as driving local oscillations to a noisier regime. For each patient, all nodes were individually perturbed, with a repertoire of noise levels $\boldsymbol{\beta}$ ranging from 0.02 to 0.5 in steps of 0.01 ($\boldsymbol{\beta_{single}} = [0.02, 0.03, \dots, 0.5]$). In the context of perturbation, explored noise levels $\boldsymbol{\beta}$ were also referred to as *perturbation intensities* or *stimulation intensities*.

Perturbed whole-brain models generated perturbed simulated FCs, and various measurements were used to assess perturbation performance. First, *susceptibility S* measured the dissimilarity between perturbed and unperturbed FCs. In other words, it assessed the brain's *capacity to change* after being perturbed with a given intensity in a given brain area. Understood in these terms, susceptibility pretended to inform on the *plasticity* of brain functional organization.

Two metrics were defined, one using the mean squared error (MSE) or *error-based* and another using Pearson's correlation (CORR) or *correlation-based*:

$$S_{MSE} = mse\left(\boldsymbol{FC}^{simulated}_{perturbed}, \boldsymbol{FC}^{simulated}_{unperturbed}\right), \tag{20}$$

$$S_{CORR} = 1 - corr\left(\boldsymbol{FC}^{simulated}_{perturbed}, \boldsymbol{FC}^{simulated}_{unperturbed}\right), \tag{21}$$

where mean squared error and correlation functions $mse(\boldsymbol{A}, \boldsymbol{B})$ and $corr(\boldsymbol{A}, \boldsymbol{B})$ between two matrices **A** and **B** were respectively defined as

$$mse(\boldsymbol{A}, \boldsymbol{B}) = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1|j\neq i}^{N} \left(A_{ij} - B_{ij}\right)^2 \tag{22}$$

$$corr(\boldsymbol{A}, \boldsymbol{B}) = \frac{\sum_{i=1}^{N} \sum_{j=1|j\neq i}^{N} \left(A_{ij} - \overline{A_{ij}}\right)\left(B_{ij} - \overline{B_{ij}}\right)}{\sqrt{\sum_{i=1}^{N} \sum_{j=1|j\neq i}^{N} \left(A_{ij} - \overline{A_{ij}}\right)^2 \sum_{i=1}^{N} \sum_{j=1|j\neq i}^{N} \left(B_{ij} - \overline{B_{ij}}\right)^2}} \tag{23}$$

Note that FC self-connections (i.e., $i = j$) are here dismissed, since they are trivially constant and equal to one in any correlation-based FC matrix [86].

The motivation in using Pearson's correlation is that it has been extensively used as a measure of FC similarity, proving successful in applications like identifying individuals out of a large group [87], [88]. Here, it is rather used to construct a measure of decorrelation (by taking $1 - corr$ in S_{CORR}), provided the definition of susceptibility. Note that the differences between S_{MSE} and S_{CORR} rely in that S_{MSE} quantifies the mean difference in FC weights and is therefore sensitive to the scale of the data, whereas S_{CORR} captures how weights change in relation to their average, insensitive to the scale of the

data. This makes both metrics complementary in that they detect different ways of diverging from unperturbed FC.

Susceptibility, which captures functional malleability, has the limitation of not reporting whether the change exerted by perturbations is *favourable* (or *beneficial*) or not —here understood as promoting a healthy transition—. The assessment of *in silico* stimulations was therefore complemented with another measure, termed Perturbation Effectivity to Recovery (PER). This concept has been described in prior computational studies [74], and aims to capture the effectivity of perturbations in driving the brain towards a healthy state. Here, a *healthy state* was approximated by the empirical FC arising from the average of 100 healthy empirical FC matrices. As mentioned in *Materials*, fMRI DK80-parcellated signals from 100 unrelated subjects part of the HCP Young Adult were accessed, and empirical FCs were computed as in Eq. (2). The resulting average healthy FC will be broadly referred to as *healthy target* or *healthy FC*.

Similarly to susceptibility, error- and correlation-based PER measures were defined:

$$PER_{MSE} = 1 - mse\left(\boldsymbol{FC}^{simulated}_{perturbed}, \boldsymbol{FC}^{empirical}_{healthy}\right) \tag{24}$$

$$PER_{CORR} = corr\left(\boldsymbol{FC}^{simulated}_{perturbed}, \boldsymbol{FC}^{empirical}_{healthy}\right) \tag{25}$$

Note that in this case the interest remains in getting similar to healthy FC, so PER_{MSE} instead of PER_{CORR} was corrected (by taking $1 - mse$ in PER_{MSE}) to account for a similarity and not an error measure.

Furthermore, to measure the strict effect that perturbations have towards driving healthy transitions, analogous baseline similarity measures were computed between unperturbed FCs and the healthy target, termed Baseline Similarity to Recovery (BSR):

$$BSR_{MSE} = 1 - mse\left(\boldsymbol{FC}^{simulated}_{unperturbed}, \boldsymbol{FC}^{empirical}_{healthy}\right) \tag{26}$$

$$BSR_{CORR} = corr\left(\boldsymbol{FC}^{simulated}_{unperturbed}, \boldsymbol{FC}^{empirical}_{healthy}\right) \tag{27}$$

Perturbations were considered *successful* or *beneficial* whenever PER scored higher than BSR, i.e., whenever FC was driven closer to the healthy target. Consequently, PER−BSR differences were used to assess the benefit of perturbations, and a particular focus was placed on how this benefit changes with pharmacology. Moreover, across the set of regions perturbed and stimulation intensities used, the *best perturbation* was defined for each patient as the perturbation giving the highest PER (or PER−BSR) value. The framework of the perturbation protocol and how susceptibility, PER and BSR relate to each other is illustrated in Fig. 2.

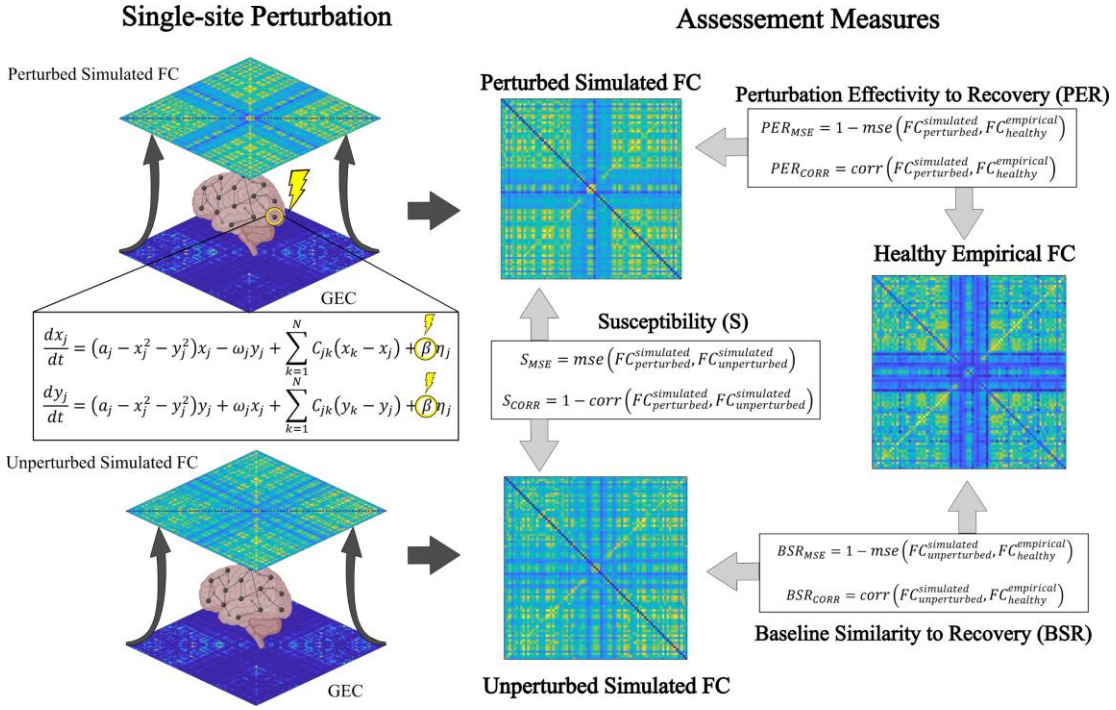

Figure 2. Perturbation protocol. Once patient-specific GECs have been computed using generative whole-brain modelling (see Fig. 1), these matrices are used to inform a new whole-brain model, which is systematically perturbed. This figure illustrates single-site perturbations, where the noise level β (originally set to 0.01) is modified in a specific brain region, set to a given stimulation intensity. The repertoire of brain regions and perturbation intensities utilized ground the perturbational map for each subject (see Fig. A4 for an example using the median susceptibility across patients). For a given perturbation, whole-brain modelling allows to project a perturbed simulated FC. Susceptibility is defined as a measure of dissimilarity between unperturbed and perturbed simulated FCs. Perturbation effectivity to recovery (PER) is defined as a similarity metric between perturbed simulated and a healthy empirical FC. Finally, baseline similarity to recovery (BSR) measures the similarity between the unperturbed simulated FC and the healthy target.

Last, the benefits of multi-site stimulation versus single-site stimulation were explored. Since the number of possible multi-site perturbations scales exponentially with the number of areas perturbed, a greedy perturbation strategy was implemented, choosing at each greedy level the area that maximizes PER when perturbed. The greedy level was set to 20 areas, and the algorithm was run first using PER_{MSE} maximization and then PER_{CORR} maximization. For each greedy level, all eligible areas —initially all of them— were perturbed with a stimulation intensity repertoire $\boldsymbol{\beta}$ ranging from 0.01 to 0.1 in steps of 0.01 ($\boldsymbol{\beta_{multi}} = [0.01, 0.02, \ldots, 0.1]$). The pair of area and noise level showing the highest PER value was kept, with that area being removed from the eligible set. Noise levels above 0.1 were here not considered to reduce computation time, after realising high PER values are not frequent above this threshold (see Fig. A9 and Fig. A10).

g) Connectivity measures

Results from the perturbation protocol were correlated to connectivity measures at the level of nodes. This was particularly studied for susceptibility, to elucidate whereas high perturbation effects happen when stimulating highly connected nodes. The two connectivity measures studied were effective weight and structural strength.

For a node n in a GEC network, its in-weight ω_n^{in} and out-weight ω_n^{out} were first defined by:

$$\omega_n^{in} = \sum_m GEC_{nm} \qquad (28) \qquad \omega_n^{out} = \sum_m GEC_{mn} \qquad (29)$$

The effective weight of that node was then computed by summing inward and outward contributions:

$$u_n = \omega_n^{in} + \omega_n^{out} \qquad (30)$$

For SC, where symmetry reflects the non-directionality of information, nodal strength was defined as:

$$s_n = \sum_m SC_{nm} \qquad (31)$$

Note that due to the lack of patient-specific SC, structural nodal strength could only be computed for the general SC template, whereas for effective weight, patient-specific data was available for all psilocybin and escitalopram before-treatment and after-treatment groups (PB, PA, EB, EA). This enabled studying the effect of psilocybin and escitalopram on effective weight.

h) Statistical testing

For a given measure of interest, statistical tests were performed to assess statistical significance between groups. Essentially, such tests involved assessing the effects of treatment (i.e., PB with PA and EB with EA), as well as comparing differences between treatment arms (i.e., PA with EA and PB with PA). Due to the relatively small sample size (22 patients for psilocybin and 20 for escitalopram), non-parametric Wilcoxon tests were conducted, not assuming Gaussianity in distributions. Paired Wilcoxon tests were used to study treatment effects, while unpaired Wilcoxon tests were employed to assess differences between arms. For both cases, Wilcoxon tests were implemented using 5000 permutations to approximate the null hypothesis distribution from the data.

Statistical significance was declared for p-values below 0.05 ($p < 0.05$). In figures, different levels of significance were denoted using the following asterisk code: *** for $p < 0.001$, ** for $p < 0.01$, and * for $p < 0.05$.

3. RESULTS

3.1. Generative whole-brain modelling

Patient-specific generative effective connectivity (GEC) matrices were computed for psilocybin and escitalopram arms before and after treatment (PB, PA, EB, EA), and an important verification step was to assess the quality of the fittings. For this, correlation between simulated and empirical findings was computed, for both functional connectivity (FC) and shifted functional connectivity (FS) (Fig. 3).

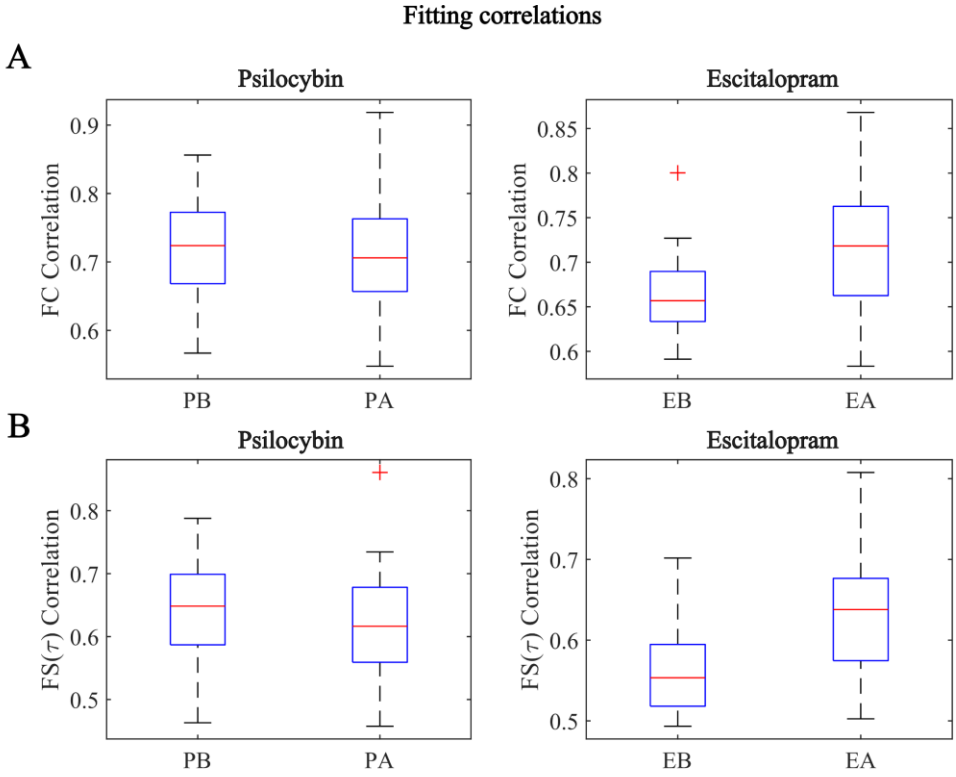

Figure 3. Correlations between simulated and empirical data. A) Correlations between simulated and empirical FC matrices are plotted, showing medians above 0.65 for both psilocybin (left) and escitalopram (right) groups before and after treatment (PB, PA, EB, EA). **B)** Analogous fitting correlations are shown using FS, exhibiting lower but still moderate correlation values for all groups.

To further confirm a proper fitting optimization, the temporal evolution of the total fitting error was tracked. Fig. A1 shows the evolution of the total error for group and individual fitting —for individual fitting the average across subjects is shown.

Upon GECs obtention, a further validation step was to compute hierarchical measures for each group and compare the results with those of Ref. [71]. Fig. 4 shows trophic coherence distributions for psilocybin and escitalopram arms before and after treatment. Results indicate a significant decrease in trophic coherence for psilocybin ($p<0.01$, paired Wilcoxon test with 5000 permutations) and significant increase in escitalopram's arm ($p<0.001$, paired Wilcoxon test with 5000 permutations). Hierarchical levels of brain regions were also examined, showing a consistent diminishment with psilocybin and a less clear trend in escitalopram (see Fig. A2). Overall, the results are consistent with prior findings [71].

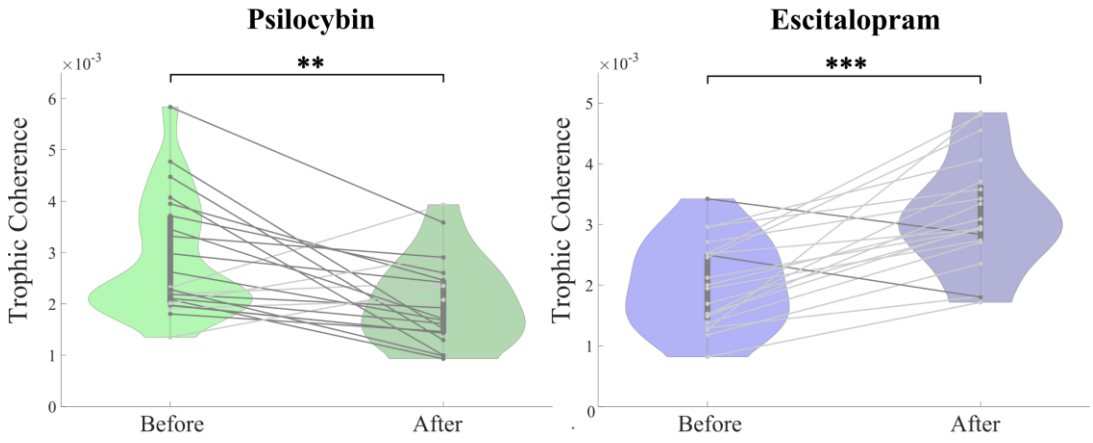

Figure 4. Opposite effects of psilocybin and escitalopram on trophic coherence. Paired violin plots show trophic coherence distributions for psilocybin (left) and escitalopram (right) groups before and after treatment. Light and dark trajectories indicate increases and a decrease in trophic coherence, respectively. Significant differences between pre-treatment and post-treatment groups were demonstrated ($p<0.01$ for psilocybin, $p<0.001$ for escitalopram, paired Wilcoxon tests with 5000 permutations), with psilocybin decreasing and escitalopram increasing trophic coherence.

3.2. Perturbation protocol

Before running the perturbation protocol, the healthy FC target was computed (see Fig. A3). Subsequently, the perturbation protocol was carried out, computing susceptibility and PER measures for each perturbation.

a) Susceptibility

Susceptibility (S) was computed for all perturbations, defining a susceptibility map for each patient (see Fig. A4 for group median maps). To extract statistics across patients, a first approach consisted of computing the average susceptibility across all perturbations —both across brain regions and stimulation intensities— and looking at group differences (Fig. 5). Results revealed that despite significant different group baselines, psilocybin and escitalopram modify in opposite directions the average susceptibility, with psilocybin increasing it and escitalopram decreasing it. These changes were statistically significant for both drugs using S_{MSE} ($p<0.01$ for psilocybin, $p<0.001$ for escitalopram, paired Wilcoxon tests with 5000 permutations) and only for psilocybin using S_{CORR} ($p<0.05$, paired Wilcoxon test with 5000 permutations).

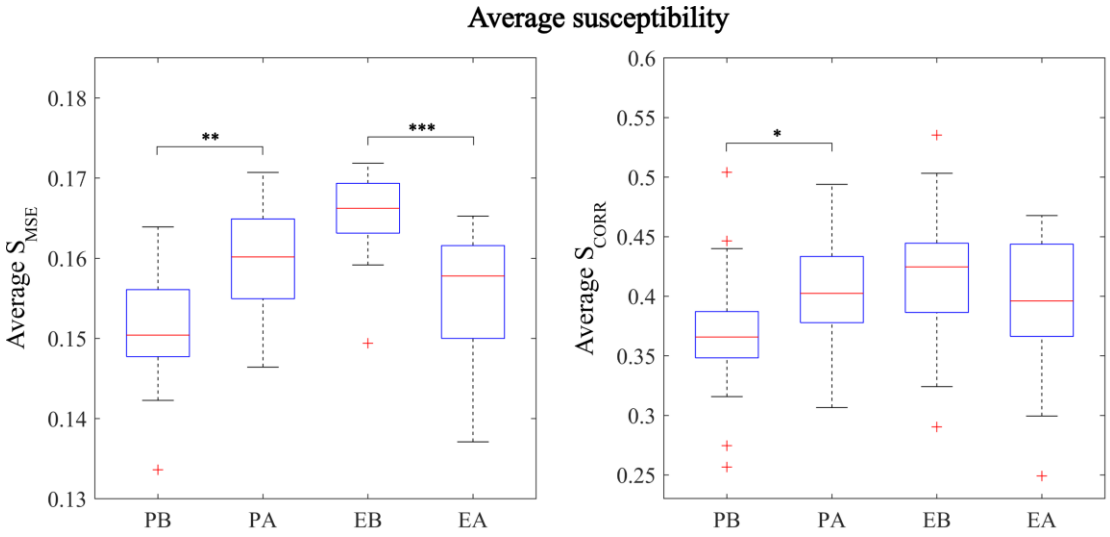

Figure 5. Opposite effects of psilocybin and escitalopram on average susceptibility. The figure shows the group distribution of average susceptibility to perturbations (S) among patients (i.e., the average is computed across all perturbed areas and all perturbation intensities), both using S_{MSE} (left) and S_{CORR} (right). Using S_{MSE}, psilocybin significantly increases susceptibility ($p<0.01$, paired Wilcoxon test with 5000 permutations) whereas escitalopram significantly decreases it ($p<0.001$, paired Wilcoxon test with 5000 permutations). Using S_{CORR} instead (right), a similar trend is present, with psilocybin significantly increasing susceptibility ($p<0.05$, paired Wilcoxon test with 5000 permutations) and escitalopram showing a decrease that is not significant.

After elucidating the overall effects of psilocybin and escitalopram on susceptibility, the noise level was fixed and the number of areas increasing susceptibility with treatment were counted. The goal was to determine whether increases of susceptibility associated to psilocybin occur broadly across areas or are due to a more local phenomenon. To decide which intensities to use, the dependency of susceptibility on noise levels was tracked (see Fig. A5), and a low ($\beta = 0.04$) and a high ($\beta = 0.4$) noise level were defined. While the high noise level was arbitrarily defined, the low one was selected as the noise level that overall resulted in highest Perturbation Effectivity to Recovery (PER) values (see Fig. A6), to account for relevant perturbations from a therapeutic perspective. Fig. 6A shows the number of areas increasing susceptibility for the low noise level, and Fig. A7 for the high noise level.

For both noise levels, more areas increased their susceptibility to perturbation with psilocybin treatment than with escitalopram. This was especially true using S_{CORR}, where a higher difference between treatments was found. The prevalence among subjects of areas increasing susceptibility was rendered (Fig. 6B), showing consistent (>0.8) increases in temporal, parietal and frontal areas with psilocybin treatment.

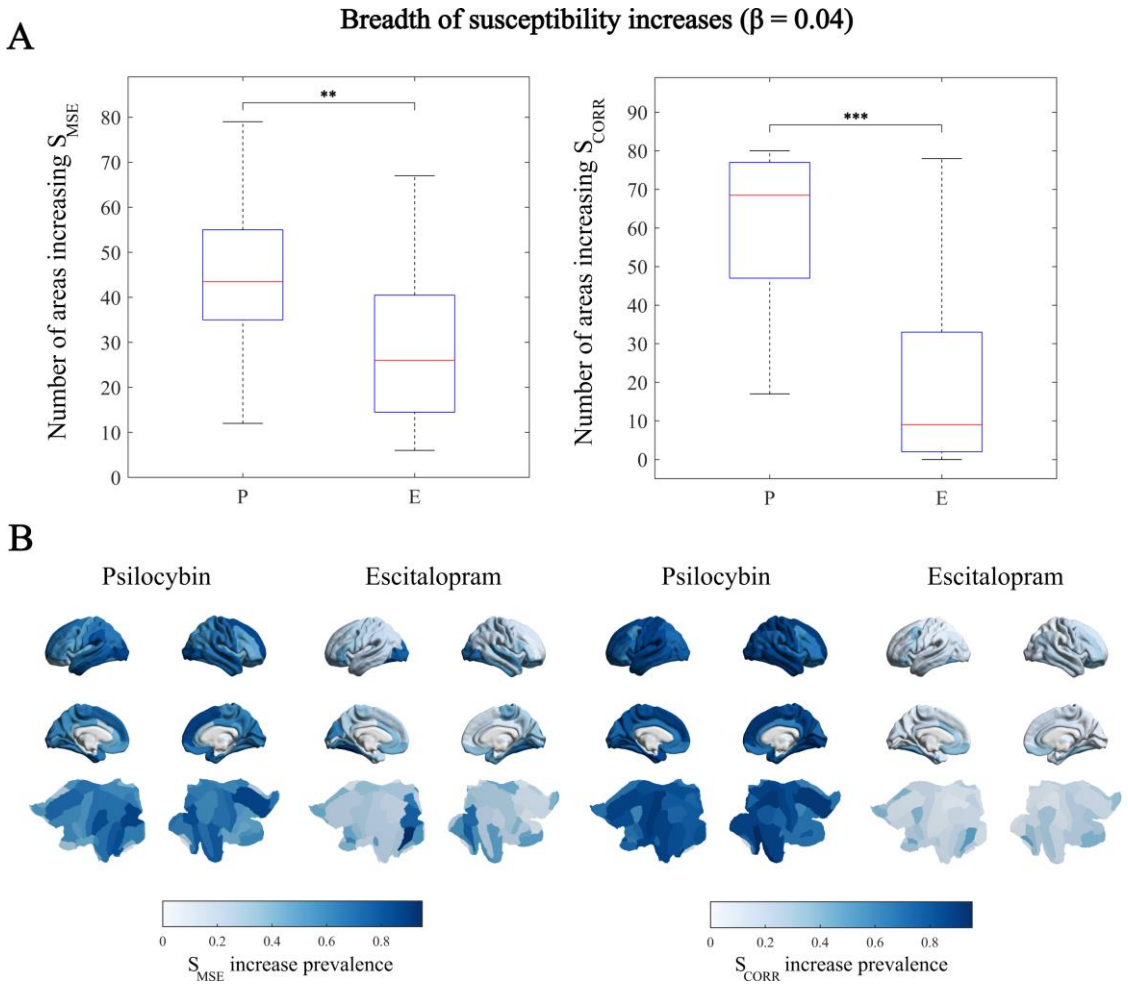

Figure 6. Breadth of susceptibility increases with a low perturbation intensity. The figure shows that with a low perturbation intensity (β = 0.04), susceptibility differentially increases in more areas with psilocybin treatment, compared to escitalopram, where decreases are predominant. **A)** Number of regions increasing susceptibility using S_{MSE} (left) and S_{CORR} (right). Psilocybin-treated brains show significantly higher values than escitalopram for both S_{MSE} ($p<0.01$, Wilcoxon test with 5000 permutations) and S_{CORR} (right, $p<0.001$, Wilcoxon test with 5000 permutations). **B)** The renderings show the regional prevalence of susceptibility increases across patients, for both S_{MSE} (left) and S_{CORR} (right).

The susceptibility of each node was further correlated with associated connectivity measures to determine whether the node's ability to change functional organization when perturbed correlates with its degree of connectivity within the network. To eliminate variations associated to different noise levels, the stimulation intensity of perturbations was fixed to β = 0.04. First, regional susceptibility values were selected as the median values across all patients, and they correlated with structural nodal strengths using the template structural connectivity (SC) matrix (Fig. 7A). Results indicate a moderate-to-high correlation (R=0.417 using S_{MSE} and R=0.662 using S_{CORR}).

Afterwards, regional susceptibility changes produced by psilocybin and escitalopram treatments were correlated with the corresponding changes in effective weight (Fig. 7B). Since here effective weight changes were available for each patient, the correlated values were constructed taking the medians representing each treatment arm. Results show that there is a moderate correlation for both S_{MSE} and S_{CORR}, and interestingly, using S_{CORR},

increases in effective weight do not necessarily imply increases in susceptibility, and vice versa.

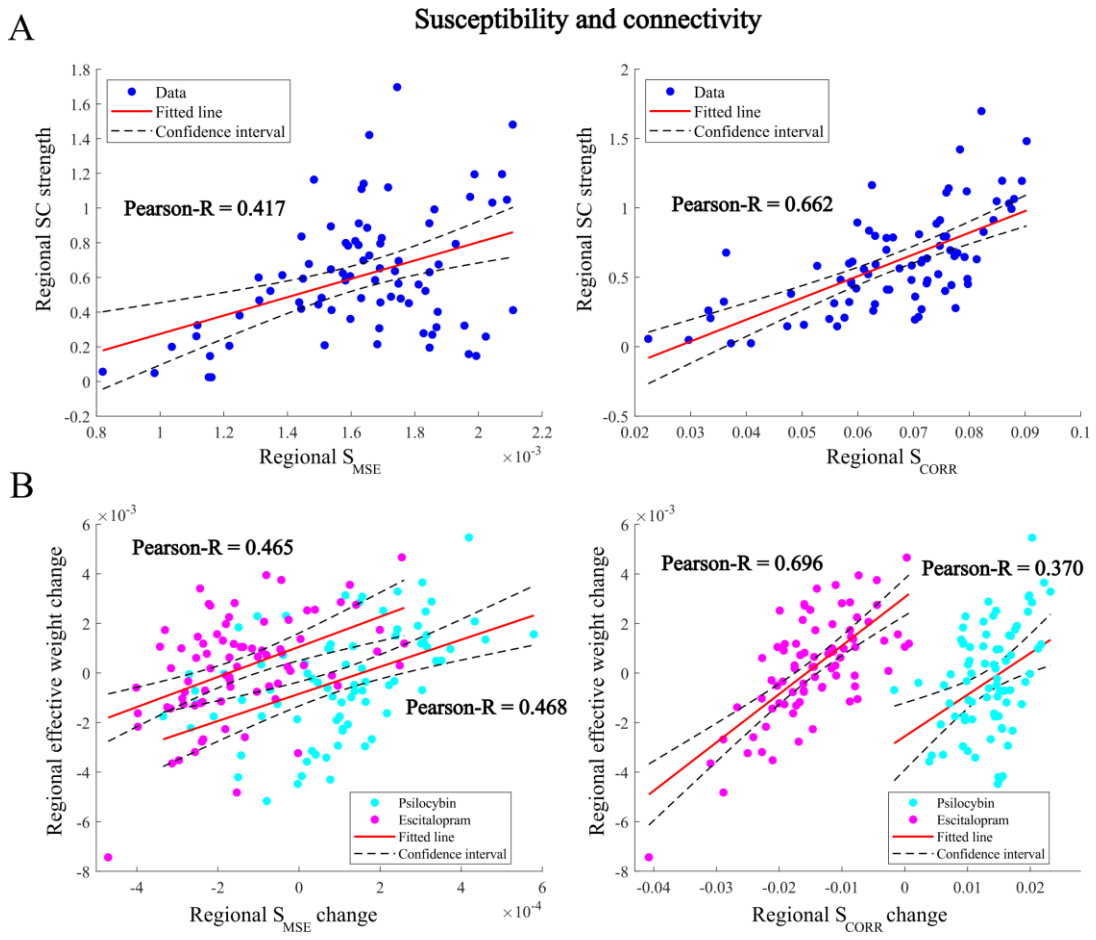

Figure 7. Susceptibility correlates with structural strength and effective weight. A) Susceptibility measures S_{MSE} (left) and S_{CORR} (right) regionally correlate with structural nodal strength (R = 0.417; R = 0.662, respectively). Susceptibility values correspond to the median values across all patients, for perturbations with a stimulation intensity β = 0.04. **B)** Regional susceptibility changes and regional effective connectivity changes due to both psilocybin (turquoise) and escitalopram (magenta) treatments correlate using S_{MSE} (left, R = 0.468 for psilocybin; R = 0.465 for escitalopram) and S_{CORR} (right, R = 0.370 for psilocybin; R = 0.696 for escitalopram) measures. Here, correlated regional values were extracted by taking the median values across patients of the same arm, and again, susceptibility values correspond to perturbations with a stimulation intensity β = 0.04.

b) Perturbation Effectivity to Recovery

To strictly measure the impact of perturbations towards reaching a healthy state, Baseline Similarity to Recovery (BSR) for each patient was first computed (see Fig. A8). Afterwards, Perturbation Effectivity to Recovery (PER) was assessed for all perturbations, and perturbational maps —across brain areas and noise levels— were resumed using the median map across groups (see Fig. A9). Importantly, these maps comprise only noise levels below 0.1, since higher noise levels demonstrated very low PER values (see Fig. A10). Afterwards, the distributions of best perturbations were examined in opposition to BSR values (Fig. 8A).

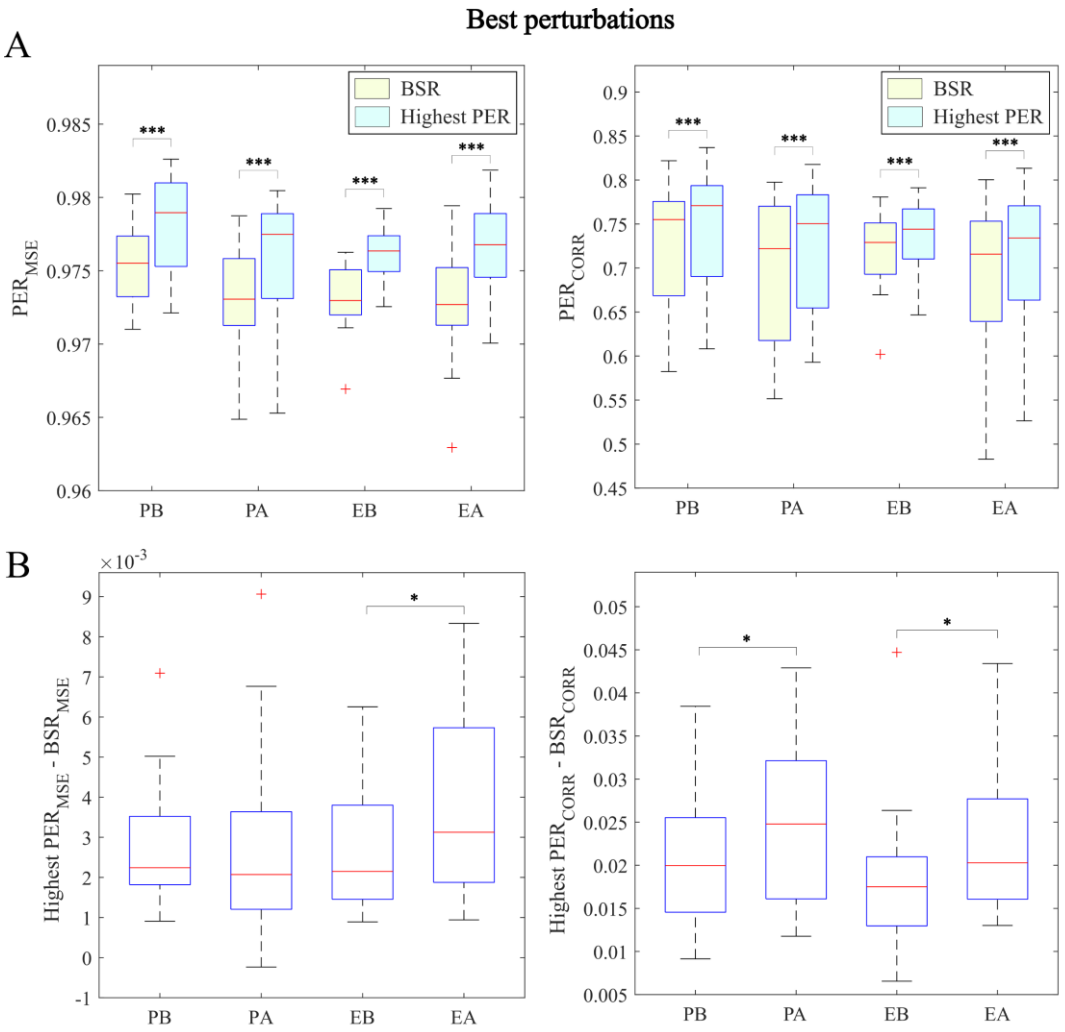

Figure 8. **Best perturbations drive significant healthy transitions. A)** Best perturbations manage to significantly drive the brain towards a state healthier than non-perturbed brains. This is consistent for all groups, both using BSR_{MSE}/PER_{MSE} (left) and BSR_{CORR}/PER_{CORR} (right) metrics ($p<0.001$, paired Wilcoxon tests with 5000 permutations). **B)** The beneficial effect of perturbations significantly increases after treatment. With PER_{CORR} (right), this is consistent for both psilocybin and escitalopram ($p<0.05$, paired Wilcoxon tests with 5000 permutations), whereas using PER_{MSE}, the trend is significant only for escitalopram ($p<0.05$, paired Wilcoxon test with 5000 permutations).

It was found that best perturbations provide significantly higher PER values than BSR, both using error-based and correlation-based measures. The distributions of the difference PER–BSR were also examined (Fig. 8B), and results indicated higher values following pharmacological treatment. These differences were significant for both error and correlation-based measures in escitalopram ($p<0.05$, paired Wilcoxon test with 5000 permutations) and only using correlation for psilocybin ($p<0.05$, paired Wilcoxon test with 5000 permutations).

A further question of interest was to examine the breadth of perturbation success in each group (Fig. 9A), where perturbation success was defined for an area if PER values were higher than their corresponding BSR. For each patient, the noise was fixed to that of the best perturbation in all areas. Results showed that all groups similarly benefit from perturbations, with escitalopram-treated patients showing a significant increase in their broadness profiles ($p<0.05$, paired Wilcoxon tests with 5000 permutations). The

difference in treatment effects was further studied by taking the number of areas whose perturbation was beneficial after treatment (PER>BSR) and whose PER–BSR difference increased with treatment (Fig. 9B). Differences between psilocybin and escitalopram were here not significant, although a subtle trend seems to indicate that escitalopram is more broadly enhancing perturbation benefits.

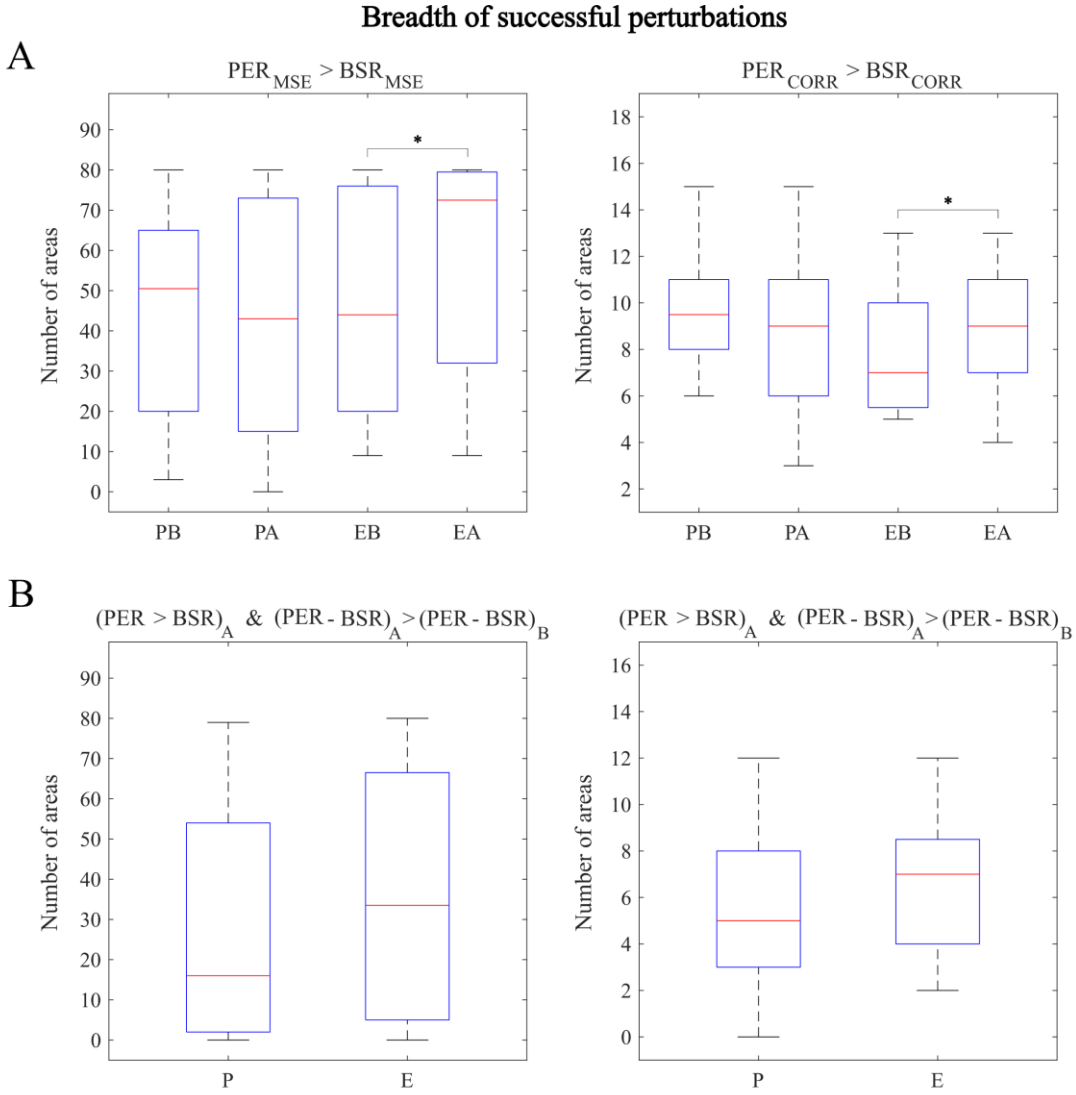

Figure 9. Breadth of healthy transitions with patient best perturbation intensity. A) The number of areas whose PER is higher than its BSR are shown for all groups, using error (left) and correlation-based (right) metrics. The noise level of perturbations was here homogeneously set across areas but patient-specific, and it was selected as that allowing the patient's highest PER. Significance was only found in escitalopram treatment for error and correlation-based metrics ($p<0.05$, paired Wilcoxon tests with 5000 permutations). **B)** The number of areas whose perturbation beneficial effects are enhanced with treatment ($(PER–BSR)_A > (PER–BSR)_B$) and whose PER is higher than its BSR (PER>BSR) after treatment is shown, both using error (left) and correlation-based (right) metrics. Although a trend seems to favour more areas in escitalopram, no significance is demonstrated.

Afterwards, the greedy multi-site strategy was implemented, finding for each patient a combination of 20 areas perturbed with a given stimulation intensity (see Fig. A11 for greedy trajectories). Multi-site perturbations were compared to best single-site perturbations (Fig. 10), showing significant higher PER–BSR values ($p<0.001$, paired

Wilcoxon tests with 5000 permutations). Last, within multi-site perturbations, no significant differences were found between before-after distributions, although a trend seems to favour higher values after treatment.

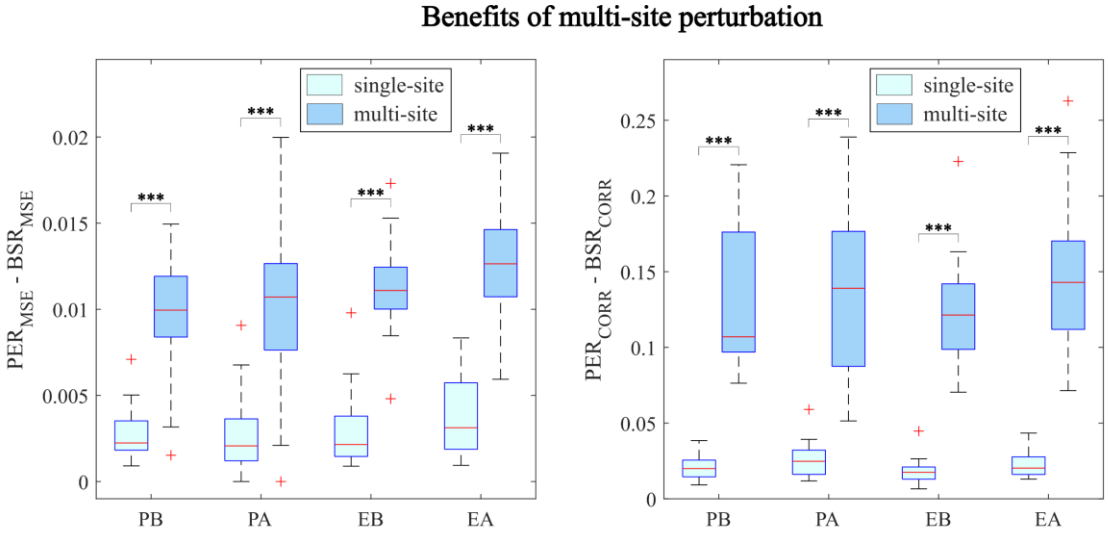

Figure 10. Benefits of multi-site versus single-site perturbations. Stimulation of multiple areas significantly improves perturbation benefits towards approaching a healthy FC compared to single-site strategies. For all groups and using both PER_{MSE} and PER_{CORR}, multi-site perturbations were able to reach higher PER–BSR values compared to the best single-site ($p<0.001$, paired Wilcoxon tests with 5000 permutations).

A final question was to elucidate which areas integrate best stimulation solutions, for both the single- and greedy multi-site approaches. For the sake of simplicity, only best stimulations using PER_{CORR} were considered (Fig. 11). For single-area perturbations, the prevalence of selected areas across patients was measured for all groups. The areas with higher prevalence were the right amygdala and the nucleus accumbens (left and right). Other areas included the left putamen, the left globus pallidus internus, and the right medial orbitofrontal cortex, and their presence was exclusive of before-treatment strategies. The same analysis was carried out considering multi-stimulations, this time measuring the presence of areas in 20-area multi-stimulation strategies (Fig. 11B). Overall, the repertoire of areas more frequently perturbed is consistent across groups, and coincides with best single-stimulation areas, adding areas like the anterior cingulate or the globus pallidus externa. A detailed report of recruited areas prevalence can be found in Fig. A12.

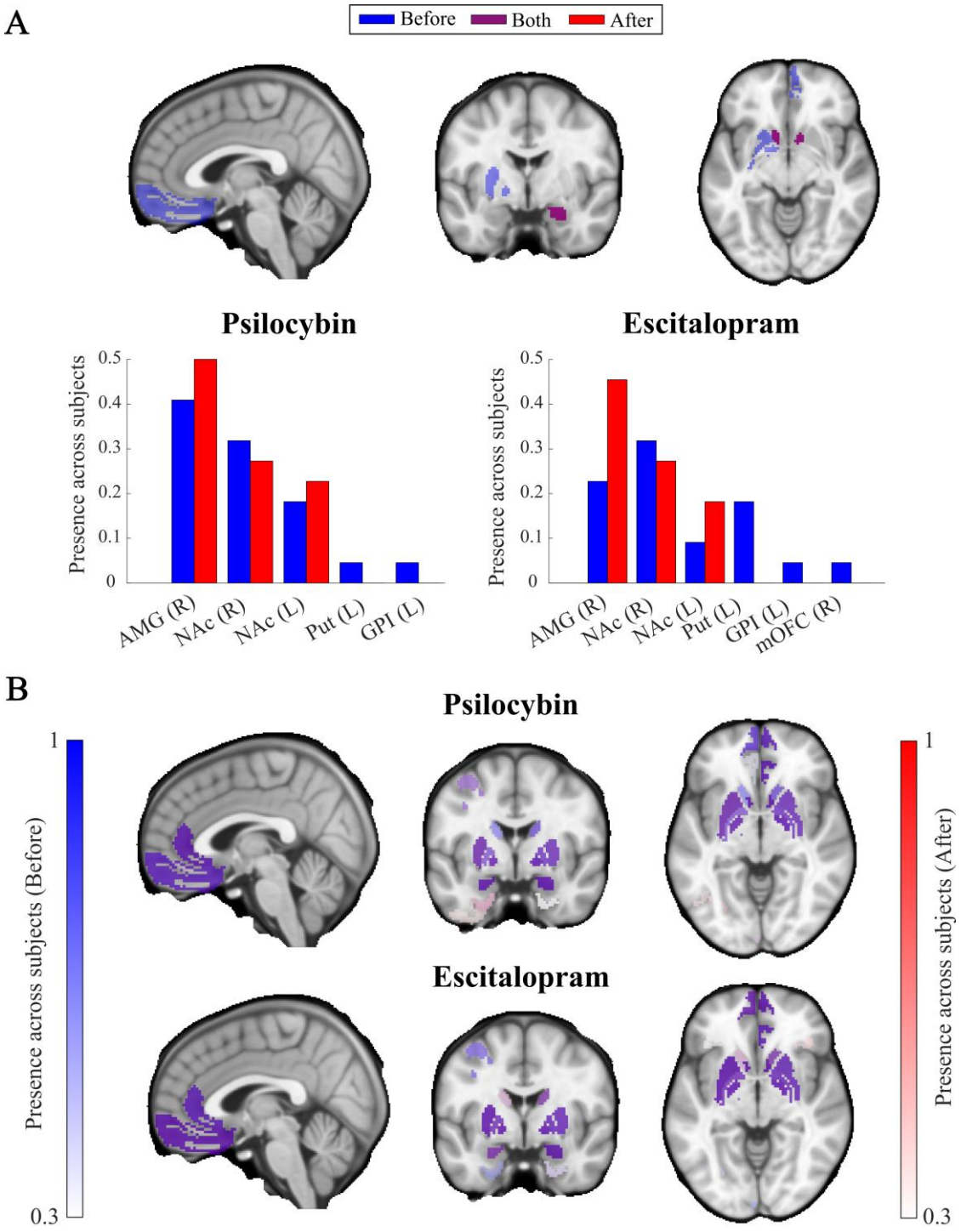

Figure 11. Best stimulation areas. A) At the top, parasagital (left), coronal (middle) and axial (right) slices of the brain show the areas corresponding to best single-site perturbations. The right amygdala (AMG) and the nucleus accumbens (NAc) are part of pre- and post-treatment stimulation strategies (purple), and the left putamen (Put), the left globus pallida internus (GPI) and the right medial orbitofrontal cortex (mOFC) are only present in stimulation strategies before treatment (blue). At the bottom, the presence of those areas among subject perturbational strategies is shown for psilocybin (left) and escitalopram (right) groups before (blue) and after (red) treatment. **B)** The presence of areas in greedy multi-site stimulations is directly shown on parasagital (left), coronal (middle) and axial (right) brain slices, for psilocybin (up) and escitalopram (bottom) groups before (blue) and after (red) treatment.

4. DISCUSSION

The aim of this thesis was to assess the effects of external perturbations on psilocybin and escitalopram treatments. First, we used correlation between DK80-parcellated fMRI signals to construct empirical functional connectivity (FC) matrices for each patient. Second, we built whole-brain models to enable the projection of simulated FC from connectivity matrices. In this case, since patient-specific structural connectivity (SC) was not available, all connectivity matrices were set to a generic SC template. Then, through an optimization process split into group and patient fitting, connectivity weights were refined for each patient, seeking maximal resemblance between its simulated and empirical FC. In this way, we obtained patient-specific generative effective connectivity (GEC) matrices, capturing causal brain mechanisms. These matrices were further used to assess topological differences between groups and verify consistency with prior studies. We then ran a perturbation protocol *in silico* by regionally increasing the noise level in whole-brain models, this time directly informed by GECs (see *Methods*). Since perturbations were reflected in perturbed simulated FC matrices, we defined susceptibility (S) as a measure of dissimilarity with their unperturbed counterparts, and Perturbation Effectivity to Recovery (PER) as a measure of similarity to an empirical healthy target. To strictly measure the effect of perturbations in promoting recovery, we compared PER with Baseline Similarity to Recovery (BSR), defined as the similarity between unperturbed simulated FC and the healthy target. Error- (MSE) and correlation-based (CORR) metrics were used for S, PER and BSR, and we used as a healthy target the average of 100 healthy FC matrices from the Young Adult project of the Human Connectome Project. The perturbation protocol explored all single-site perturbations with various stimulation intensities, and later, a greedy multi-site protocol was used to estimate the benefits of multi-site stimulation.

Results from the fitting process of whole-brain models (Fig. 3) revealed that the fitting was successful both at the group and patient level. Regarding topological findings, two opposite trends were identified (Fig. 4), with psilocybin exerting a diminishment of trophic coherence and escitalopram inducing a significant increase. This is consistent with prior findings [71] and aligns with the REBUS principle, where psychedelics flatten the hierarchy of brain organization. These findings validate generative whole-brain models, providing a solid basis for implementing the perturbation protocol.

4.1. Susceptibility

Findings from the perturbation protocol indicate that across all noise levels and regions explored, susceptibility to perturbations is on average increased following psilocybin treatment, whereas for escitalopram, the opposite trend is observed (Fig. 5). In other words, psilocybin makes brain functional organization more malleable, whereas escitalopram strengthens its rigidity. The same trend was confirmed after fixing the stimulation intensity (Fig. 6 and Fig. A7), with notably more areas increasing susceptibility with psilocybin compared to escitalopram. This links with the flattening of brain hierarchy produced by psilocybin (Fig. 4) [71], where a general enhancement of bottom-up signalling across areas is consistent with a broad susceptibility increase. We interpret increases in susceptibility as an opening of a window of plasticity, which has already been proposed to be a major role of psychedelics' 5-HT2AR signalling [13]. For escitalopram, reduced functional susceptibility to perturbations can analogously be linked

to known increases in trophic coherence (Fig. 4) [71], and we suggest it may relate to the emotional blunting associated with SSRIs [89].

Last, we explored the relationship between regional connectivity and susceptibility (Fig. 7). High correlation values were found between structural nodal strength and susceptibility (Fig. 7A). This was expected, since a more connected network should be able to produce higher changes in the rest of areas if perturbed. Furthermore, the change in susceptibility produced by treatment was found to be correlated with the change in effective weights (Fig. 7B). However, positive changes in effective weight do not necessarily imply positive changes in susceptibility, and vice versa. This means that the relationship between regional susceptibility and connectivity is not straightforward, and that may rely on the intuition that the impact of a perturbation is not only related to the immediate connectivity of the perturbed node but also to the connectivity of the rest of the network.

4.2. Perturbation Effectivity to Recovery

To truly assess the beneficial impact of perturbations, we assessed PER results by comparing them to corresponding BSR values, prior to perturbation. BSR results indicate that the psilocybin group before treatment is significantly closer to the healthy target compared to the rest of groups (Fig. A8). This suggests that psilocybin is driving FC away from healthy FC, which is counterintuitive given the antidepressant effects of psilocybin [12]. However, what is here being accounted for as a healthy state is a healthy static FC matrix, which may not fully represent what a nondepressed healthy state is. Indeed, by using static FC, dynamical information about how functional correlation patterns evolve over time is lost, and it may be relevant in this context. Future studies should therefore contemplate using dynamic measures like metastability to construct susceptibility and PER measures, as implemented in Ref. [74] in the context of Alzheimer's disease. Besides this possibility, the psychedelic-induced state is in fact different from a normal healthy state, so we should indeed expect some functional differences between them [90]. We hypothesize that one-day-after-treatment fMRI scans may underlie the observed differences and that longer-term fMRI following treatment should show FC matrices closer to the healthy target. Nonetheless, the mentioned difference is not significant for PER_{CORR}.

Once the perturbation protocol was run, the best perturbations across noise levels and areas were selected, and their PER values were found to be significantly higher than their respective baselines (Fig. 8A). This means that, if properly chosen, perturbations exert an improvement towards approaching the healthy target, i.e., they can be useful. Especially interesting is the case of groups after treatment, where PER values significantly higher than BSR indicate that perturbations meaningfully complement pharmacological treatment. Aside from inherent comparison limitations, these results align with the benefits of combined therapy in MDD, with psychotherapy —which could be understood as a type of external perturbation— and pharmacological treatment having shown better clinical outcomes than pharmacology alone [6]. Moreover, when analysing the distributions of PER–BSR differences, significantly higher values were found for treated patients compared to untreated conditions (Fig. 8B). Together, these findings not only support that stimulation is complementary to pharmacology, but also that stimulation

effects are synergistically enhanced by it. Yet, using PER_{MSE}–BSR_{MSE}, statistical significance was only demonstrated in escitalopram treatment.

As with susceptibility, a subsequent analysis implied fixing the noise level, here to that of the best perturbation in each patient (Fig. 9). First, we examined how many areas benefit from perturbation (i.e., positive PER–BSR differences) (Fig. 9A). With prior knowledge of broader susceptibility increases for psilocybin compared to escitalopram (Fig. 6), the expectations were to find more areas benefiting from perturbations after psilocybin treatment. However, counterintuitive results were found, with no differences produced by psilocybin treatment and significant increases associated with escitalopram. Potential explanations may rely on the fact that a broadly enhanced susceptibility by psilocybin makes many regions more sensitive to perturbations, which, if not properly adjusted —as in here with a homogeneous stimulation intensity— drive perturbed FC further instead of closer to the target. Also, we suggest the number of beneficial perturbations may change if studying longer-term fMRI data and using more complex perturbational approaches, by better exploiting the perturbational landscape in later post-treatment stages. We believe that, despite evident limitations in the comparison, therapeutic effects like those conveyed by psychotherapy would be better approximated, complementing the present results. Additionally, we computed the number of areas improving their perturbation benefits with treatment (Fig. 9B), always ensuring a positive benefit after treatment. Although a trend favours escitalopram, we found no statistical differences between escitalopram and psilocybin treatments.

Finally, by running a greedy multi-site perturbation protocol, we showed that perturbation strategies involving the stimulation of multiple areas outperform single-site best perturbations, resulting in significantly higher PER values (Fig. 10). This demonstrates the benefits of multi-site versus single-site stimulations, which were expected, provided MDD has been shown to affect multiple brain networks [52]. The enhancing effects of pharmacology were preserved with multi-site stimulations, although statistical significance was not demonstrated. This seems to suggest that the advantage of combined therapy is diminished the more complex perturbations can be.

4.3. Best areas to perturb

Last, the most suitable areas for stimulation were examined (Fig. 11) for all groups. In best single-site perturbations, the areas more consistently perturbed were the right amygdala and the nucleus accumbens, with no major differences between groups (Fig. 11A). This suggests that similar perturbational strategies can be applied for depressed patients before and after treatment, with scarce differences existing between treatments. The nucleus accumbens is known to be involved in the reward circuitry [91] and to play a crucial role in MDD patients with anhedonia [92]. Results therefore align with the literature and support the recent interest in targeting this area through DBS stimulation [16], [93]. On the other hand, major depression has also been linked to alterations in the amygdala [94], which is associated with emotion regulation, among others [95]. Stimulating this area could therefore notably improve depressive symptoms, a strategy that has been recently implemented using real-time fMRI neurofeedback [96]. Interestingly, results here exclusively point to the right amygdala. This supports existing studies on amygdala lateralization, which, for instance, regard the right amygdala as being more linked to automated emotional processing versus the left amygdala as being more

implicated in intentional mood control [97], [98]. Last, the heterogeneity among patients in identifying the optimal brain stimulation target supports the personalization of brain stimulation solutions, a concept already underscored for MDD [99].

Finally, we analysed for each patient the 20 areas selected by greedy multi-site perturbations. To get a sense of the importance of each area, we plotted their prevalence across patient multi-site strategies (see Fig. 11B). Interestingly, areas recruited more consistently are very similar across groups and resemble those of single-site perturbations. They include the amygdala and the nucleus accumbens, but also the globus pallidus internus and externa, the putamen, and the medial orbitofrontal cortex (mOFC), which were considered the best areas to stimulate in single-site perturbations in some patients (Fig. 11A). Note that the mentioned regions are subcortical, which limits their applicability in non-invasive brain stimulation (NIBS) techniques. The medial orbitofrontal cortex (mOFC) stands as the main cortical exception, so we suggest it as the best target for NIBS. This is coherent with prior knowledge, with mOFC being related to reward and mOFC reduced functional connectivity being related to depression [100] .

4.4. Limitations and future work

There are some limitations inherent to this work. First, fMRI data from only 42 patients was available (22 in the psilocybin arm and 20 taking escitalopram), a number that should be larger for increased validation of the results. This is especially important given the known heterogeneity attributed to MDD [1], and it could allow studying relationships like the one between PER before treatment and psychological outcomes, to assess whether PER can predict responsiveness to pharmacological treatment. Second, the linear approximation of the Hopf whole-brain model was used instead of the non-linear alternative. This introduces some degree of imprecision, which was justified by the reduction in computational cost, but future work should also consider the non-linear model. In line with this idea, other perturbation protocols could be defined, for instance, perturbing the bifurcation parameter in Eqs. (7) and (8) instead of the noise level or exploring more extensively the landscape of multi-site stimulations. Last, the way of assessing susceptibility and perturbation effectivity to recovery (PER) could be expanded, including dynamic functional properties and not just static FC measures. Measures like metastability or Probable Metastable Substrates (PMS) could be utilized, as introduced in Refs. [74] and [53], respectively. Finally, as earlier mentioned, this computational study could be complemented by examining fMRI signals some months after treatment, and not immediately upon treatment completion. The present work should be viewed as an initial inroad into the *in silico* exploration of neurostimulation solutions for MDD using Dynamic Sensitivity Analysis.

5. CONCLUSION

In this thesis, we elucidated how the brain responds to external perturbations after psilocybin and escitalopram treatments. By implementing a perturbation protocol, we showed that susceptibility overall increases with psilocybin treatment, suggesting the opening of a window of plasticity. Using escitalopram, susceptibility decreases, indicating that the brain is functionally more rigid to perturbations, which is in line with an increased brain hierarchy and possibly with the emotional blunting associated with SSRIs. We also showed that this opposite trend holds when fixing the stimulation intensity and that susceptibility regionally correlates with structural and effective connectivity measures. We also demonstrated the therapeutic potential of perturbations, which, when properly tailored, can effectively facilitate healthy transitions, even for pharmacologically treated patients, highlighting the benefits of combined therapy. Moreover, the therapeutic effects of perturbations were significantly enhanced by psilocybin and escitalopram treatments, revealing a synergy that should be studied in future studies. However, when constraining all areas to the same patient-best stimulation intensity, only escitalopram demonstrated increasing the number of areas driving healthy transitions. We propose that a therapeutic advantage of psilocybin's window of plasticity might be reflected in the long term and that one-day-after-treatment fMRI scans may be limiting the present study. Using a greedy multi-site perturbation protocol, we then demonstrated the benefits of multi-site stimulation versus single-site stimulation, with similar improvements by psilocybin and escitalopram treatments. Finally, we showed that the best areas to rebalance healthy FC are the right amygdala and the nucleus accumbens, and that the best cortical area is the medial orbitofrontal cortex, in line with prior studies on neurostimulation therapy for MDD treatment. Since minor differences were observed in single- and multi-stimulation best perturbation strategies between groups, we suggest that the success of stimulation might rely more on personalization than on the treatment arm. Future work should study larger samples, eventually including longer-term data and using dynamic functional measures like metastability. The present work should be viewed as an initial inroad into the *in silico* exploration of neurostimulation solutions for MDD using Dynamic Sensitivity Analysis.

BIBLIOGRAPHY

[1] G. S. Malhi and J. J. Mann, 'Depression', *The Lancet*, vol. 392, no. 10161, pp. 2299–2312, Nov. 2018, doi: 10.1016/S0140-6736(18)31948-2.

[2] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association, 2013. doi: 10.1176/appi.books.9780890425596.

[3] World Health Organization, 'Depression and other common mental disorders: global health estimates.' World Health Organization: Geneva., 2017.

[4] R. C. Kessler and E. J. Bromet, 'The Epidemiology of Depression Across Cultures', *Annu Rev Public Health*, vol. 34, no. 1, pp. 119–138, Mar. 2013, doi: 10.1146/annurev-publhealth-031912-114409.

[5] World Health Organization, 'The Global Burden of Disease: 2004 Update'. 2008.

[6] P. Cuijpers, M. Sijbrandij, S. L. Koole, G. Andersson, A. T. Beekman, and C. F. Reynolds, 'Adding psychotherapy to antidepressant medication in depression and anxiety disorders: a meta-analysis', *World Psychiatry*, vol. 13, no. 1, pp. 56–67, Feb. 2014, doi: 10.1002/wps.20089.

[7] A. Chu and W. Roopma, 'Selective Serotonin Reuptake Inhibitors', in *StatPearls*, Treasure Island: StatPearls Publishing, 2023.

[8] M. Hamon and P. Blier, 'Monoamine neurocircuitry in depression and strategies for new treatments', *Prog Neuropsychopharmacol Biol Psychiatry*, vol. 45, pp. 54–63, Aug. 2013, doi: 10.1016/j.pnpbp.2013.04.009.

[9] A. Cipriani *et al.*, 'Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis', *The Lancet*, vol. 391, no. 10128, pp. 1357–1366, Apr. 2018, doi: 10.1016/S0140-6736(17)32802-7.

[10] A. N. Edinoff *et al.*, 'Selective Serotonin Reuptake Inhibitors and Adverse Effects: A Narrative Review', *Neurol Int*, vol. 13, no. 3, pp. 387–401, Aug. 2021, doi: 10.3390/neurolint13030038.

[11] F. Palhano-Fontes *et al.*, 'Rapid antidepressant effects of the psychedelic ayahuasca in treatment-resistant depression: a randomized placebo-controlled trial', *Psychol Med*, vol. 49, no. 4, pp. 655–663, Mar. 2019, doi: 10.1017/S0033291718001356.

[12] R. L. Carhart-Harris *et al.*, 'Psilocybin with psychological support for treatment-resistant depression: an open-label feasibility study', *Lancet Psychiatry*, vol. 3, no. 7, pp. 619–627, Jul. 2016, doi: 10.1016/S2215-0366(16)30065-7.

[13] R. Carhart-Harris and D. Nutt, 'Serotonin and brain function: a tale of two receptors', *Journal of Psychopharmacology*, vol. 31, no. 9, pp. 1091–1120, Sep. 2017, doi: 10.1177/0269881117725915.

[14] M. I. Husain *et al.*, 'Serotonergic psychedelics for depression: What do we know about neurobiological mechanisms of action?', *Front Psychiatry*, vol. 13, Feb. 2023, doi: 10.3389/fpsyt.2022.1076459.

[15] R. L. Carhart-Harris *et al.*, 'Psilocybin with psychological support for treatment-resistant depression: six-month follow-up', *Psychopharmacology (Berl)*, vol. 235, no. 2, pp. 399–408, Feb. 2018, doi: 10.1007/s00213-017-4771-x.

[16] R. Karrouri, Z. Hammani, R. Benjelloun, and Y. Otheman, 'Major depressive disorder: Validated treatments and future challenges', *World J Clin Cases*, vol. 9, no. 31, pp. 9350–9367, Nov. 2021, doi: 10.12998/wjcc.v9.i31.9350.

[17] H. Akhtar, F. Bukhari, M. Nazir, M. N. Anwar, and A. Shahzad, 'Therapeutic Efficacy of Neurostimulation for Depression: Techniques, Current Modalities,

and Future Challenges', *Neurosci Bull*, vol. 32, no. 1, pp. 115–126, Feb. 2016, doi: 10.1007/s12264-015-0009-2.
[18] C. Cusin and D. D. Dougherty, 'Somatic therapies for treatment-resistant depression: ECT, TMS, VNS, DBS', *Biol Mood Anxiety Disord*, vol. 2, no. 1, p. 14, Dec. 2012, doi: 10.1186/2045-5380-2-14.
[19] J. Vohryzek, 'Spacetime connectomics for describing and modelling complex dynamics of human brain states', University of Oxford, 2021.
[20] H. Johansen-Berg and M. F. S. Rushworth, 'Using Diffusion Imaging to Study Human Connectional Anatomy', *Annu Rev Neurosci*, vol. 32, no. 1, pp. 75–94, Jun. 2009, doi: 10.1146/annurev.neuro.051508.135735.
[21] P. Hagmann, P. E. Grant, and D. A. Fair, 'MR connectomics: a conceptual framework for studying the developing brain', *Front Syst Neurosci*, vol. 6, 2012, doi: 10.3389/fnsys.2012.00043.
[22] D. K. Jones and M. Cercignani, 'Twenty-five pitfalls in the analysis of diffusion MRI data', *NMR Biomed*, vol. 23, no. 7, pp. 803–820, Aug. 2010, doi: 10.1002/nbm.1543.
[23] G. H. Glover, 'Overview of Functional Magnetic Resonance Imaging', *Neurosurg Clin N Am*, vol. 22, no. 2, pp. 133–139, Apr. 2011, doi: 10.1016/j.nec.2010.11.001.
[24] G. Deco and M. L. Kringelbach, 'Great Expectations: Using Whole-Brain Computational Connectomics for Understanding Neuropsychiatric Disorders', *Neuron*, vol. 84, no. 5, pp. 892–905, Dec. 2014, doi: 10.1016/J.NEURON.2014.08.034.
[25] M. Filippi, E. G. Spinelli, C. Cividini, and F. Agosta, 'Resting State Dynamic Functional Connectivity in Neurodegenerative Conditions: A Review of Magnetic Resonance Imaging Findings', *Front Neurosci*, vol. 13, Jun. 2019, doi: 10.3389/fnins.2019.00657.
[26] S. S. Menon and K. Krishnamurthy, 'A Comparison of Static and Dynamic Functional Connectivities for Identifying Subjects and Biological Sex Using Intrinsic Individual Brain Connectivity', *Sci Rep*, vol. 9, no. 1, p. 5729, Apr. 2019, doi: 10.1038/s41598-019-42090-4.
[27] M. Greicius, 'Resting-state functional connectivity in neuropsychiatric disorders', *Curr Opin Neurol*, vol. 21, no. 4, pp. 424–430, Aug. 2008, doi: 10.1097/WCO.0b013e328306f2c5.
[28] K. J. Friston, 'Functional and Effective Connectivity: A Review', *Brain Connect*, vol. 1, no. 1, pp. 13–36, Jan. 2011, doi: 10.1089/brain.2011.0008.
[29] A. Schaefer *et al.*, 'Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI', *Cerebral Cortex*, vol. 28, no. 9, pp. 3095–3114, Sep. 2018, doi: 10.1093/cercor/bhx179.
[30] N. Tzourio-Mazoyer *et al.*, 'Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain', *Neuroimage*, vol. 15, no. 1, pp. 273–289, Jan. 2002, doi: 10.1006/nimg.2001.0978.
[31] C. F. Beckmann and S. M. Smith, 'Probabilistic Independent Component Analysis for Functional Magnetic Resonance Imaging', *IEEE Trans Med Imaging*, vol. 23, no. 2, pp. 137–152, Feb. 2004, doi: 10.1109/TMI.2003.822821.
[32] R. C. Craddock *et al.*, 'Imaging human connectomes at the macroscale', *Nat Methods*, vol. 10, no. 6, pp. 524–539, Jun. 2013, doi: 10.1038/nmeth.2482.
[33] N. Tzourio-Mazoyer *et al.*, 'Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-

Subject Brain', *Neuroimage*, vol. 15, no. 1, pp. 273–289, Jan. 2002, doi: 10.1006/nimg.2001.0978.
[34] Patric Hagmann, 'From diffusion MRI to brain connectomics.', École Polytechnique Fédérale de Lausanne, 2005.
[35] R. S. Desikan *et al.*, 'An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest', *Neuroimage*, vol. 31, no. 3, pp. 968–980, Jul. 2006, doi: 10.1016/j.neuroimage.2006.01.021.
[36] M. L. Kringelbach, Y. S. Perl, E. Tagliazucchi, and G. Deco, 'Toward naturalistic neuroscience: Mechanisms underlying the flattening of brain hierarchy in movie-watching compared to rest and task', *Sci Adv*, vol. 9, no. 2, Jan. 2023, doi: 10.1126/sciadv.ade6049.
[37] P. Moghimi, A. T. Dang, Q. Do, T. I. Netoff, K. O. Lim, and G. Atluri, 'Evaluation of functional MRI-based human brain parcellation: a review', *J Neurophysiol*, vol. 128, no. 1, pp. 197–217, Jul. 2022, doi: 10.1152/jn.00411.2021.
[38] O. Sporns, G. Tononi, and R. Kötter, 'The Human Connectome: A Structural Description of the Human Brain', *PLoS Comput Biol*, vol. 1, no. 4, p. e42, 2005, doi: 10.1371/journal.pcbi.0010042.
[39] D. J. Watts and S. H. Strogatz, 'Collective dynamics of "small-world" networks', *Nature*, vol. 393, no. 6684, pp. 440–442, Jun. 1998, doi: 10.1038/30918.
[40] M. E. J. Newman, 'Modularity and community structure in networks', *Proceedings of the National Academy of Sciences*, vol. 103, no. 23, pp. 8577–8582, Jun. 2006, doi: 10.1073/pnas.0601602103.
[41] E. Bullmore and O. Sporns, 'Complex brain networks: graph theoretical analysis of structural and functional systems', *Nat Rev Neurosci*, vol. 10, no. 3, pp. 186–198, Mar. 2009, doi: 10.1038/nrn2575.
[42] M. Rubinov and E. Bullmore, 'Fledgling pathoconnectomics of psychiatric disorders', *Trends Cogn Sci*, vol. 17, no. 12, pp. 641–647, Dec. 2013, doi: 10.1016/j.tics.2013.10.007.
[43] S. A. R. B. Rombouts *et al.*, 'Model-free group analysis shows altered BOLD FMRI networks in dementia', *Hum Brain Mapp*, vol. 30, no. 1, pp. 256–266, Jan. 2009, doi: 10.1002/hbm.20505.
[44] M. A. A. Binnewijzend *et al.*, 'Resting-state fMRI changes in Alzheimer's disease and mild cognitive impairment', *Neurobiol Aging*, vol. 33, no. 9, pp. 2018–2028, Sep. 2012, doi: 10.1016/j.neurobiolaging.2011.07.003.
[45] A. KARL, M. SCHAEFER, L. MALTA, D. DORFEL, N. ROHLEDER, and A. WERNER, 'A meta-analysis of structural brain abnormalities in PTSD', *Neurosci Biobehav Rev*, vol. 30, no. 7, pp. 1004–1031, 2006, doi: 10.1016/j.neubiorev.2006.03.004.
[46] M. P. van den Heuvel and A. Fornito, 'Brain Networks in Schizophrenia', *Neuropsychol Rev*, vol. 24, no. 1, pp. 32–48, Mar. 2014, doi: 10.1007/s11065-014-9248-7.
[47] M. D. Greicius *et al.*, 'Resting-State Functional Connectivity in Major Depression: Abnormally Increased Contributions from Subgenual Cingulate Cortex and Thalamus', *Biol Psychiatry*, vol. 62, no. 5, pp. 429–437, Sep. 2007, doi: 10.1016/j.biopsych.2006.09.020.
[48] L. Wang, D. F. Hermens, I. B. Hickie, and J. Lagopoulos, 'A systematic review of resting-state functional-MRI studies in major depression', *J Affect Disord*, vol. 142, no. 1–3, pp. 6–12, Dec. 2012, doi: 10.1016/j.jad.2012.04.013.

[49] M. Demirtaş *et al.*, 'Dynamic functional connectivity reveals altered variability in functional connectivity among patients with major depressive disorder', *Hum Brain Mapp*, vol. 37, no. 8, pp. 2918–2930, Aug. 2016, doi: 10.1002/hbm.23215.
[50] G. Northoff, 'Spatiotemporal psychopathology I: No rest for the brain's resting state activity in depression? Spatiotemporal psychopathology of depressive symptoms', *J Affect Disord*, vol. 190, pp. 854–866, Jan. 2016, doi: 10.1016/j.jad.2015.05.007.
[51] R. H. Kaiser, J. R. Andrews-Hanna, T. D. Wager, and D. A. Pizzagalli, 'Large-Scale Network Dysfunction in Major Depressive Disorder', *JAMA Psychiatry*, vol. 72, no. 6, p. 603, Jun. 2015, doi: 10.1001/jamapsychiatry.2015.0071.
[52] V. Menon, 'Large-scale brain networks and psychopathology: a unifying triple network model', *Trends Cogn Sci*, vol. 15, no. 10, pp. 483–506, Oct. 2011, doi: 10.1016/j.tics.2011.08.003.
[53] J. Vohryzek *et al.*, 'Brain dynamics predictive of response to psilocybin for treatment-resistant depression', *Brain Commun*, vol. 6, no. 2, Mar. 2024, doi: 10.1093/braincomms/fcae049.
[54] J. Vohryzek *et al.*, 'The flattening of spacetime hierarchy of the *N,N* -dimethyltryptamine brain state is characterized by harmonic decomposition of spacetime (HADES) framework', *Natl Sci Rev*, vol. 11, no. 5, Apr. 2024, doi: 10.1093/nsr/nwae124.
[55] G. Petri *et al.*, 'Homological scaffolds of brain functional networks', *J R Soc Interface*, vol. 11, no. 101, p. 20140873, Dec. 2014, doi: 10.1098/rsif.2014.0873.
[56] E. Tagliazucchi, R. Carhart-Harris, R. Leech, D. Nutt, and D. R. Chialvo, 'Enhanced repertoire of brain dynamical states during the psychedelic experience', *Hum Brain Mapp*, vol. 35, no. 11, pp. 5442–5456, Nov. 2014, doi: 10.1002/hbm.22562.
[57] D. Nutt and R. Carhart-Harris, 'The Current Status of Psychedelics in Psychiatry', *JAMA Psychiatry*, vol. 78, no. 2, p. 121, Feb. 2021, doi: 10.1001/jamapsychiatry.2020.2171.
[58] R. L. Carhart-Harris and K. J. Friston, 'REBUS and the Anarchic Brain: Toward a Unified Model of the Brain Action of Psychedelics', *Pharmacol Rev*, vol. 71, no. 3, pp. 316–344, Jul. 2019, doi: 10.1124/pr.118.017160.
[59] A. Pathak, D. Roy, and A. Banerjee, 'Whole-Brain Network Models: From Physics to Bedside', *Front Comput Neurosci*, vol. 16, May 2022, doi: 10.3389/fncom.2022.866517.
[60] G. Deco, V. Jirsa, A. R. McIntosh, O. Sporns, and R. Kötter, 'Key role of coupling, delay, and noise in resting brain fluctuations', *Proceedings of the National Academy of Sciences*, vol. 106, no. 25, pp. 10302–10307, Jun. 2009, doi: 10.1073/pnas.0901831106.
[61] H. R. Wilson and J. D. Cowan, 'Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons', *Biophys J*, vol. 12, no. 1, pp. 1–24, Jan. 1972, doi: 10.1016/S0006-3495(72)86068-5.
[62] J. Cabral *et al.*, 'Cognitive performance in healthy older adults relates to spontaneous switching between states of functional connectivity during rest', *Sci Rep*, vol. 7, no. 1, p. 5135, Jul. 2017, doi: 10.1038/s41598-017-05425-7.
[63] C. J. Honey *et al.*, 'Predicting human resting-state functional connectivity from structural connectivity', *Proceedings of the National Academy of Sciences*, vol. 106, no. 6, pp. 2035–2040, Feb. 2009, doi: 10.1073/pnas.0811168106.

[64] G. Deco, M. L. Kringelbach, V. K. Jirsa, and P. Ritter, 'The dynamics of resting fluctuations in the brain: metastability and its dynamical cortical core', *Sci Rep*, vol. 7, no. 1, p. 3095, Jun. 2017, doi: 10.1038/s41598-017-03073-5.
[65] G. Deco, A. Ponce-Alvarez, D. Mantini, G. L. Romani, P. Hagmann, and M. Corbetta, 'Resting-State Functional Connectivity Emerges from Structurally and Dynamically Shaped Slow Linear Fluctuations', *Journal of Neuroscience*, vol. 33, no. 27, pp. 11239–11252, Jul. 2013, doi: 10.1523/JNEUROSCI.1091-13.2013.
[66] E. C. A. Hansen, D. Battaglia, A. Spiegler, G. Deco, and V. K. Jirsa, 'Functional connectivity dynamics: Modeling the switching behavior of the resting state', *Neuroimage*, vol. 105, pp. 525–535, Jan. 2015, doi: 10.1016/j.neuroimage.2014.11.001.
[67] C. W. Lynn, E. J. Cornblath, L. Papadopoulos, M. A. Bertolero, and D. S. Bassett, 'Broken detailed balance and entropy production in the human brain', *Proceedings of the National Academy of Sciences*, vol. 118, no. 47, Nov. 2021, doi: 10.1073/pnas.2109889118.
[68] J. Alstott, M. Breakspear, P. Hagmann, L. Cammoun, and O. Sporns, 'Modeling the Impact of Lesions in the Human Brain', *PLoS Comput Biol*, vol. 5, no. 6, p. e1000408, Jun. 2009, doi: 10.1371/journal.pcbi.1000408.
[69] R. Panda *et al.*, 'Whole-brain analyses indicate the impairment of posterior integration and thalamo-frontotemporal broadcasting in disorders of consciousness', *Hum Brain Mapp*, vol. 44, no. 11, pp. 4352–4371, Aug. 2023, doi: 10.1002/hbm.26386.
[70] J. Cabral, M. Kringelbach, and G. Deco, 'Functional Graph Alterations in Schizophrenia: A Result from a Global Anatomic Decoupling?', *Pharmacopsychiatry*, vol. 45, no. S 01, pp. S57–S64, May 2012, doi: 10.1055/s-0032-1309001.
[71] G. Deco, Y. Sanz Perl, S. Johnson, N. Bourke, R. L. Cahart-Harris, and M. L. Kringelbach, 'Pharmacological effects on brain dynamics in depression: Different hierarchical reconfigurations by psilocybin and escitalopram [Manuscript submitted for publication]', 2024.
[72] J. Vohryzek *et al.*, 'Dynamic sensitivity analysis: Defining personalised strategies to drive brain state transitions via whole brain modelling', *Comput Struct Biotechnol J*, vol. 21, pp. 335–345, 2023, doi: 10.1016/j.csbj.2022.11.060.
[73] G. Deco *et al.*, 'Awakening: Predicting external stimulation to force transitions between different brain states', *Proceedings of the National Academy of Sciences*, vol. 116, no. 36, pp. 18088–18097, Sep. 2019, doi: 10.1073/pnas.1905534116.
[74] J. Vohryzek *et al.*, 'Design of effective personalised perturbation strategies for enhancing cognitive intervention in Alzheimer's disease', *bioRxiv*, p. 2023.04.20.537688, Jan. 2023, doi: 10.1101/2023.04.20.537688.
[75] L. A. Desouza *et al.*, 'The Hallucinogenic Serotonin2A Receptor Agonist, 2,5-Dimethoxy-4-Iodoamphetamine, Promotes cAMP Response Element Binding Protein-Dependent Gene Expression of Specific Plasticity-Associated Genes in the Rodent Neocortex', *Front Mol Neurosci*, vol. 14, Dec. 2021, doi: 10.3389/fnmol.2021.790213.
[76] R. Carhart-Harris *et al.*, 'Trial of Psilocybin versus Escitalopram for Depression', *New England Journal of Medicine*, vol. 384, no. 15, pp. 1402–1411, Apr. 2021, doi: 10.1056/NEJMoa2032994.
[77] M. Hamilton, 'A RATING SCALE FOR DEPRESSION', *J Neurol Neurosurg Psychiatry*, vol. 23, no. 1, pp. 56–62, Feb. 1960, doi: 10.1136/jnnp.23.1.56.

[78] G. Deco, D. Vidaurre, and M. L. Kringelbach, ‘Revisiting the global workspace orchestrating the hierarchical organization of the human brain’, *Nat Hum Behav*, vol. 5, no. 4, pp. 497–511, Jan. 2021, doi: 10.1038/s41562-020-01003-6.
[79] A. Horn, W. Neumann, K. Degen, G. Schneider, and A. A. Kühn, ‘Toward an electrophysiological “sweet spot” for deep brain stimulation in the subthalamic nucleus’, *Hum Brain Mapp*, vol. 38, no. 7, pp. 3377–3390, Jul. 2017, doi: 10.1002/hbm.23594.
[80] J. Vohryzek, G. Deco, B. Cessac, M. L. Kringelbach, and J. Cabral, ‘Ghost Attractors in Spontaneous Brain Activity: Recurrent Excursions Into Functionally-Relevant BOLD Phase-Locking States’, *Front Syst Neurosci*, vol. 14, Apr. 2020, doi: 10.3389/fnsys.2020.00020.
[81] The MathWorks Inc., ‘MATLAB, version 9.14 (R2023b)’. Natick, Massachusetts, 2023.
[82] E. Glerean, J. Salmi, J. M. Lahnakoski, I. P. Jääskeläinen, and M. Sams, ‘Functional Magnetic Resonance Imaging Phase Synchronization as a Measure of Dynamic Functional Connectivity’, *Brain Connect*, vol. 2, no. 2, pp. 91–101, Apr. 2012, doi: 10.1089/brain.2011.0068.
[83] A. Ponce-Alvarez and G. Deco, ‘The Hopf whole-brain model and its linear approximation’, *Sci Rep*, vol. 14, no. 1, p. 2615, Jan. 2024, doi: 10.1038/s41598-024-53105-0.
[84] G. Deco, A. Ponce-Alvarez, P. Hagmann, G. L. Romani, D. Mantini, and M. Corbetta, ‘How Local Excitation-Inhibition Ratio Impacts the Whole Brain Dynamics’, *Journal of Neuroscience*, vol. 34, no. 23, pp. 7886–7898, Jun. 2014, doi: 10.1523/JNEUROSCI.5068-13.2014.
[85] R. S. MacKay, S. Johnson, and B. Sansom, ‘How directed is a directed network?’, *R Soc Open Sci*, vol. 7, no. 9, p. 201138, Sep. 2020, doi: 10.1098/rsos.201138.
[86] R. Liégeois, A. Santos, V. Matta, D. Van De Ville, and A. H. Sayed, ‘Revisiting correlation-based functional connectivity and its relationship with structural connectivity’, *Network Neuroscience*, vol. 4, no. 4, pp. 1235–1251, Jan. 2020, doi: 10.1162/netn_a_00166.
[87] M. Venkatesh, J. Jaja, and L. Pessoa, ‘Comparing functional connectivity matrices: A geometry-aware approach applied to participant identification’, *Neuroimage*, vol. 207, p. 116398, Feb. 2020, doi: 10.1016/j.neuroimage.2019.116398.
[88] E. S. Finn *et al.*, ‘Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity’, *Nat Neurosci*, vol. 18, no. 11, pp. 1664–1671, Nov. 2015, doi: 10.1038/nn.4135.
[89] D. Marazziti *et al.*, ‘Emotional Blunting, Cognitive Impairment, Bone Fractures, and Bleeding as Possible Side Effects of Long-Term Use of SSRIs.’, *Clin Neuropsychiatry*, vol. 16, no. 2, pp. 75–85, Apr. 2019.
[90] L. Roseman, R. Leech, A. Feilding, D. J. Nutt, and R. L. Carhart-Harris, ‘The effects of psilocybin and MDMA on between-network resting state functional connectivity in healthy volunteers’, *Front Hum Neurosci*, vol. 8, May 2014, doi: 10.3389/fnhum.2014.00204.
[91] A. M. Klawonn and R. C. Malenka, ‘Nucleus Accumbens Modulation in Reward and Aversion’, *Cold Spring Harb Symp Quant Biol*, vol. 83, pp. 119–129, 2018, doi: 10.1101/sqb.2018.83.037457.
[92] Y. Hu, C. Zhao, H. Zhao, and J. Qiao, ‘Abnormal functional connectivity of the nucleus accumbens subregions mediates the association between anhedonia and

major depressive disorder', *BMC Psychiatry*, vol. 23, no. 1, p. 282, Apr. 2023, doi: 10.1186/s12888-023-04693-0.
[93] B. H. Bewernick *et al.*, 'Nucleus Accumbens Deep Brain Stimulation Decreases Ratings of Depression and Anxiety in Treatment-Resistant Depression', *Biol Psychiatry*, vol. 67, no. 2, pp. 110–116, Jan. 2010, doi: 10.1016/j.biopsych.2009.09.013.
[94] K. Perlman *et al.*, 'A systematic meta-review of predictors of antidepressant treatment outcome in major depressive disorder', *J Affect Disord*, vol. 243, pp. 503–515, Jan. 2019, doi: 10.1016/j.jad.2018.09.067.
[95] M. L. Phillips *et al.*, 'Identifying Predictors, Moderators, and Mediators of Antidepressant Response in Major Depressive Disorder: Neuroimaging Approaches', *American Journal of Psychiatry*, vol. 172, no. 2, pp. 124–138, Feb. 2015, doi: 10.1176/appi.ajp.2014.14010076.
[96] K. D. Young *et al.*, 'Randomized Clinical Trial of Real-Time fMRI Amygdala Neurofeedback for Major Depressive Disorder: Effects on Symptoms and Autobiographical Memory Recall', *American Journal of Psychiatry*, vol. 174, no. 8, pp. 748–755, Aug. 2017, doi: 10.1176/appi.ajp.2017.16060637.
[97] J. E. Murphy, J. A. Yanes, L. A. J. Kirby, M. A. Reid, and J. L. Robinson, 'Left, right, or bilateral amygdala activation? How effects of smoothing and motion correction on ultra-high field, high-resolution functional magnetic resonance imaging (fMRI) data alter inferences', *Neurosci Res*, vol. 150, pp. 51–59, Jan. 2020, doi: 10.1016/j.neures.2019.01.009.
[98] M. Dyck, J. Loughead, T. Kellermann, F. Boers, R. C. Gur, and K. Mathiak, 'Cognitive versus automatic mechanisms of mood induction differentially activate left and right amygdala', *Neuroimage*, vol. 54, no. 3, pp. 2503–2513, Feb. 2011, doi: 10.1016/j.neuroimage.2010.10.013.
[99] J. Gogulski *et al.*, 'Personalized Repetitive Transcranial Magnetic Stimulation for Depression', *Biol Psychiatry Cogn Neurosci Neuroimaging*, vol. 8, no. 4, pp. 351–360, Apr. 2023, doi: 10.1016/j.bpsc.2022.10.006.
[100] E. T. Rolls, W. Cheng, and J. Feng, 'The orbitofrontal cortex: reward, emotion and depression', *Brain Commun*, vol. 2, no. 2, Jul. 2020, doi: 10.1093/braincomms/fcaa196.

A | SUPPORTING INFORMATION

List of contents

1. Mathematical Appendix (MA)

MA1. Non-reversibility

Fitting $FS_{forward,ij}^{simulated}(\tau)$ to $FS_{forward,ij}^{empirical}(\tau)$ has been shown to be equivalent to fitting the levels of non-reversibility $NR^{simulated}$ and $NR^{empirical}$, which can be defined as

$$NR^{simulated}(\tau) = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1|_{j\neq i}}^{N} \left(FS_{forward,ij}^{simulated}(\tau) - FS_{reversal,ij}^{simulated}(\tau) \right)^2 \quad \text{(A1)}$$

$$NR^{empirical}(\tau) = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1|_{j\neq i}}^{N} \left(FS_{forward,ij}^{empirical}(\tau) - FS_{reversal,ij}^{empirical}(\tau) \right)^2 \quad \text{(A2)}$$

where

$$FS_{reversal_{ij}}^{empirical}(\tau) = \frac{C_{v_{reversal_{ij}}}^{empirical}(\tau)}{\sqrt{C_{v_{ii}}^{empirical} \, C_{v_{jj}}^{empirical}}}, \quad \text{(A3)}$$

$$FS_{reversal_{ij}}^{simulated}(\tau) = \frac{C_{v_{reversal_{ij}}}^{simulated}(\tau)}{\sqrt{C_{v_{ii}}^{simulated} \, C_{v_{jj}}^{simulated}}}, \quad \text{(A4)}$$

and

$$\boldsymbol{C}_{\boldsymbol{v}_{reversal}}^{empirical}(\tau) = \sum_{t} \left(x_i^{(r)}(t+\tau) - \overline{x_i^{(r)}(t+\tau)} \right) \left(x_j^{(r)}(t) - \overline{x_j^{(r)}(t)} \right), \quad \text{(A5)}$$

$$\boldsymbol{C}_{\boldsymbol{v}_{reversal}}^{simulated}(\tau) = \langle \delta \boldsymbol{u}^{(r)}(t+\tau) \delta \boldsymbol{u}^{(r)}(t)^{\boldsymbol{T}} \rangle, \quad \text{(A6)}$$

being $x^r(t)$ and $\delta \boldsymbol{u}^{(r)}(t)$ the reversed versions of $x(t)$ and $\delta \boldsymbol{u}(t)$, respectively, (i.e., with inverted temporal ordering), and $C_v^{empirical} = C_{v_{reversal}}^{empirical}(0)$, $C_v^{simulated} = C_{v_{reversal}}^{simulated}(0)$.

MA2. Hierarchical levels and trophic coherence

Hierarchical levels (or trophic levels) and trophic coherence (or directedness) measures of generative effective connectivity (GEC) matrices were computed following Refs. [85] and [71].

For each node n in GEC network, its in-weight ω_n^{in} and out-weight ω_n^{out} are given by:

$$\omega_n^{in} = \sum_m GEC_{nm} \qquad \text{(A7)} \qquad \omega_n^{out} = \sum_m GEC_{mn} \qquad \text{(A8)}$$

From these measures the effective weight u_n and the imbalance v_n can be defined, capturing the sum and the difference between the inward and outward flow of information in node n, respectively:

$$u_n = \omega_n^{in} + \omega_n^{out} \qquad \text{(A9)} \qquad v_n = \omega_n^{in} - \omega_n^{out} \qquad \text{(A10)}$$

Node hierarchical levels h_n comprised in the N-dimensional vector $\boldsymbol{h} = [h_1, \dots, h_N]^T$ can be subsequently computed by solving the following linear system of equations:

$$\boldsymbol{\Lambda h} = \boldsymbol{v} \qquad \text{(A11)}$$

Where $\boldsymbol{v} = [v_1, \dots, v_N]$ and $\boldsymbol{\Lambda}$ is the weighted graph-Laplacian operator, defined by

$$\boldsymbol{\Lambda} = diag(\boldsymbol{u}) - \boldsymbol{C} - \boldsymbol{C^T} \qquad \text{(A12)}$$

Upon computation of trophic levels, the network's global directionality can be assessed by computing its directedness (or trophic coherence) using:

$$F_0 = 1 - \frac{\sum_{mn} \omega_{mn} \ (h_n - h_m - 1)^2}{\sum_{mn} \omega_{mn}} \qquad \text{(A13)}$$

A network with $F_0 = 1$ is considered maximally coherent (i.e., with all trophic level differences $h_n - h_m$ equal to 1), whereas $F_0 = 0$ indicates maximal incoherence (i.e., with all trophic level differences $h_n - h_m$ equal to 0). High F_0 values indicate greater degree of hierarchical organization.

2. Supplementary tables

Left hemisphere		Right hemisphere	
Index	**Brain area**	Index	**Brain area**
1	left caudal anterior cingulate	**41**	right hippocampus
2	left caudal middle frontal	**42**	right amygdala
3	left cuneus	**43**	right stn
4	left entorhinal	**44**	right gpi
5	left fusiform	**45**	right gpe
6	left inferior parietal	**46**	right putamen
7	left inferior temporal	**47**	right caudate
8	left isthmus cingulate	**48**	right nucleus accumbens
9	left lateral occipital	**49**	right thalamus
10	left lateral orbitofrontal	**50**	right insula
11	left lingual	**51**	right transverse temporal
12	left medial orbitofrontal	**52**	right supramarginal
13	left middle temporal	**53**	right superior temporal
14	left parahippocampal	**54**	right superior parietal
15	left paracentral	**55**	right superior frontal
16	left pars opercularis	**56**	right rostral middle frontal
17	left pars orbitalis	**57**	right rostral anterior cingulate
18	left pars triangularis	**58**	right precuneus
19	left pericalcarine	**59**	right precentral
20	left postcentral	**60**	right posterior cingulate
21	left posterior cingulate	**61**	right postcentral
22	left precentral	**62**	right pericalcarine
23	left precuneus	**63**	right pars triangularis
24	left rostral anterior cingulate	**64**	right pars orbitalis
25	left rostral middle frontal	**65**	right pars opercularis
26	left superior frontal	**66**	right paracentral
27	left superior parietal	**67**	right parahippocampal
28	left superior temporal	**68**	right middle temporal
29	left supramarginal	**69**	right medial orbitofrontal
30	left transverse temporal	**70**	right lingual
31	left insula	**71**	right lateral orbitofrontal
32	left thalamus	**72**	right lateral occipital
33	left nucleus accumbens	**73**	right isthmus cingulate
34	left caudate	**74**	right inferior temporal
35	left putamen	**75**	right inferior parietal
36	left gpe	**76**	right fusiform
37	left gpi	**77**	right entorhinal
38	left stn	**78**	right cuneus
39	left amygdala	**79**	right caudal middle frontal
40	left hippocampus	**80**	right caudal anterior cingulate

Table A1. DK80 parcellation brain regions. Cortical (white) and subcortical (light gray) regions are shown.

3. Supplementary Figures

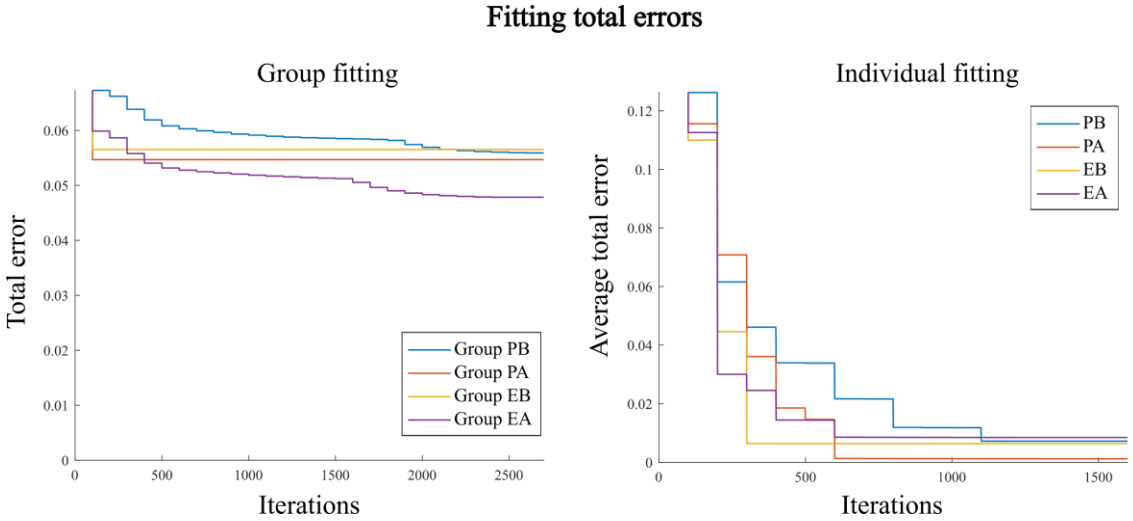

Figure A1. Fitting total errors. The temporal evolution of total errors for group (left) and individual fitting (right) is shown. For individual fitting, the average error across participants of the same group is shown. The error was arbitrarily initialized to 10000 and updated each 100 iterations until convergence (defined when the relative difference of the new error with respect to the previous error was less than 0.1%).

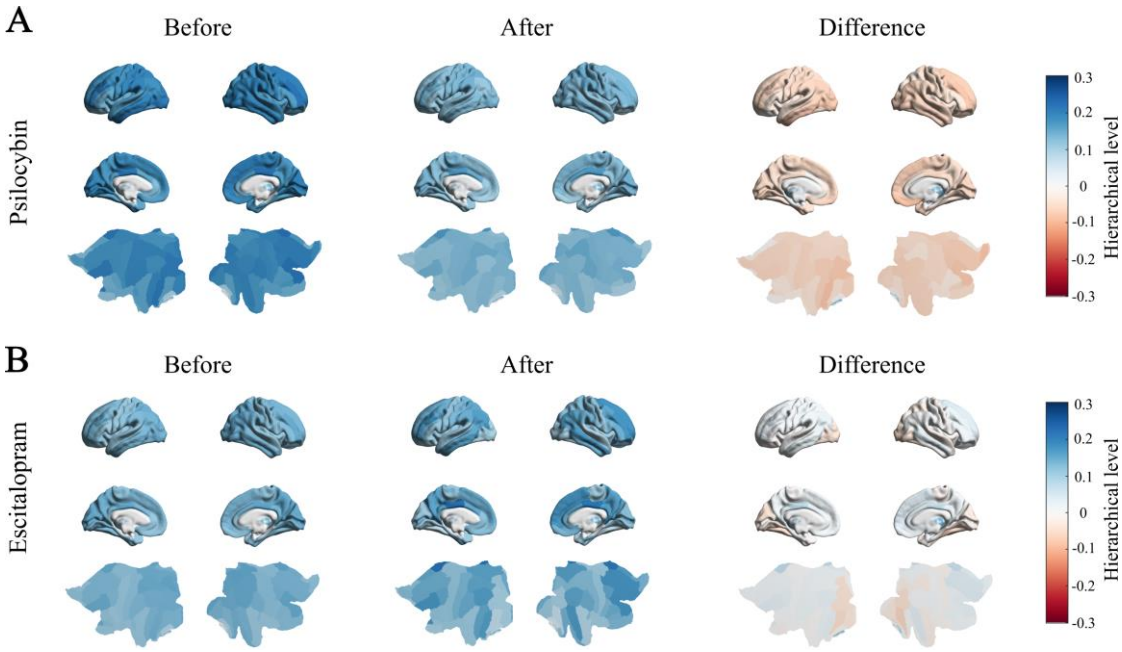

Figure A2. Hierarchical levels. The renderings show the median regional hierarchical levels across patients before (left) and after (centre) treatment, as well as the after–before difference (right). **A)** Hierarchical levels are shown for psilocybin treatment. **B)** Hierarchical levels are shown for escitalopram treatment.

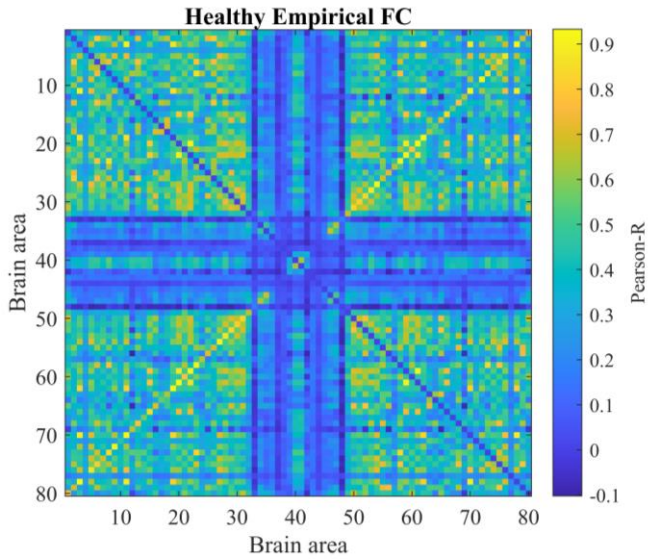

Figure A3. **Healthy target.** The figure shows the healthy empirical FC extracted from averaging empirical FC matrices from 100 unrelated subjects of the HCP Young Adult project. The diagonal is deliberately set to zero to enhance visualization.

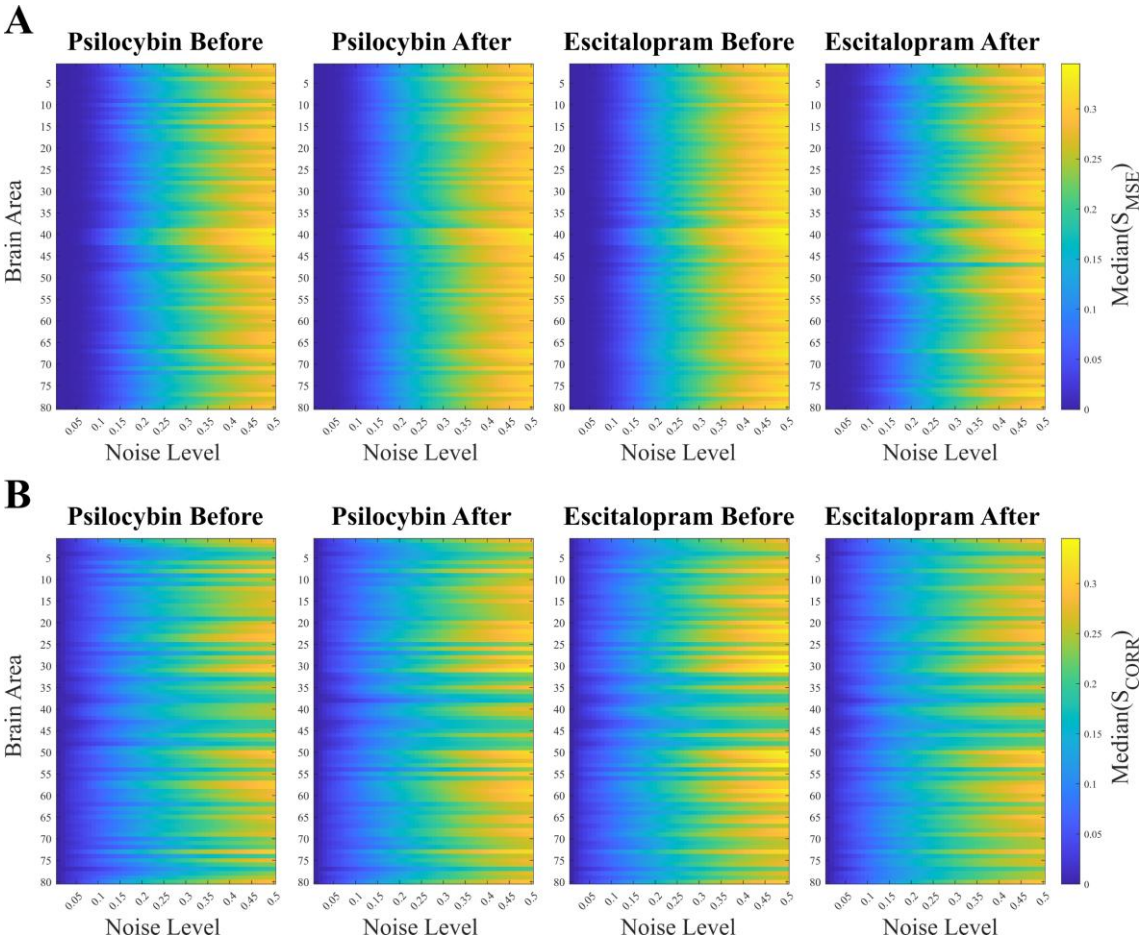

Figure A4. Susceptibility median perturbational maps. A) For all perturbations, the median susceptibility S_{MSE} across patients of a same group is shown. **B)** The same maps are shown using S_{CORR} instead of S_{MSE}.

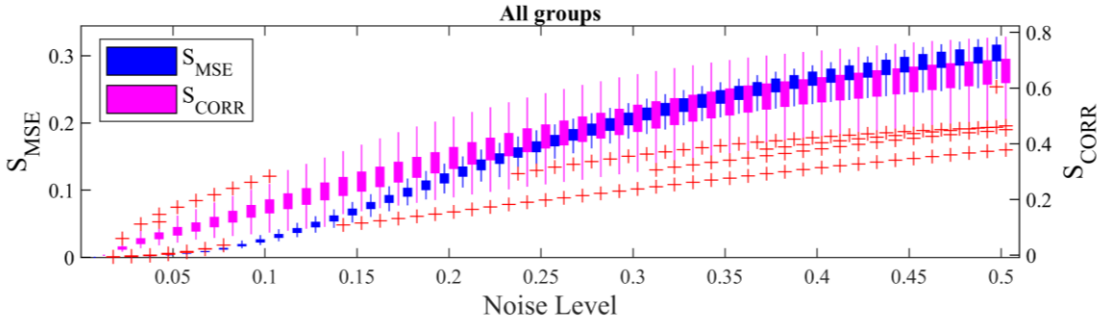

Figure A5. **Susceptibility across noise levels.** The patient distribution of the median perturbation susceptibility across areas is plotted as a function of the noise level. The distributions represent all subjects, whether treated or not and regardless of the antidepressant arm.

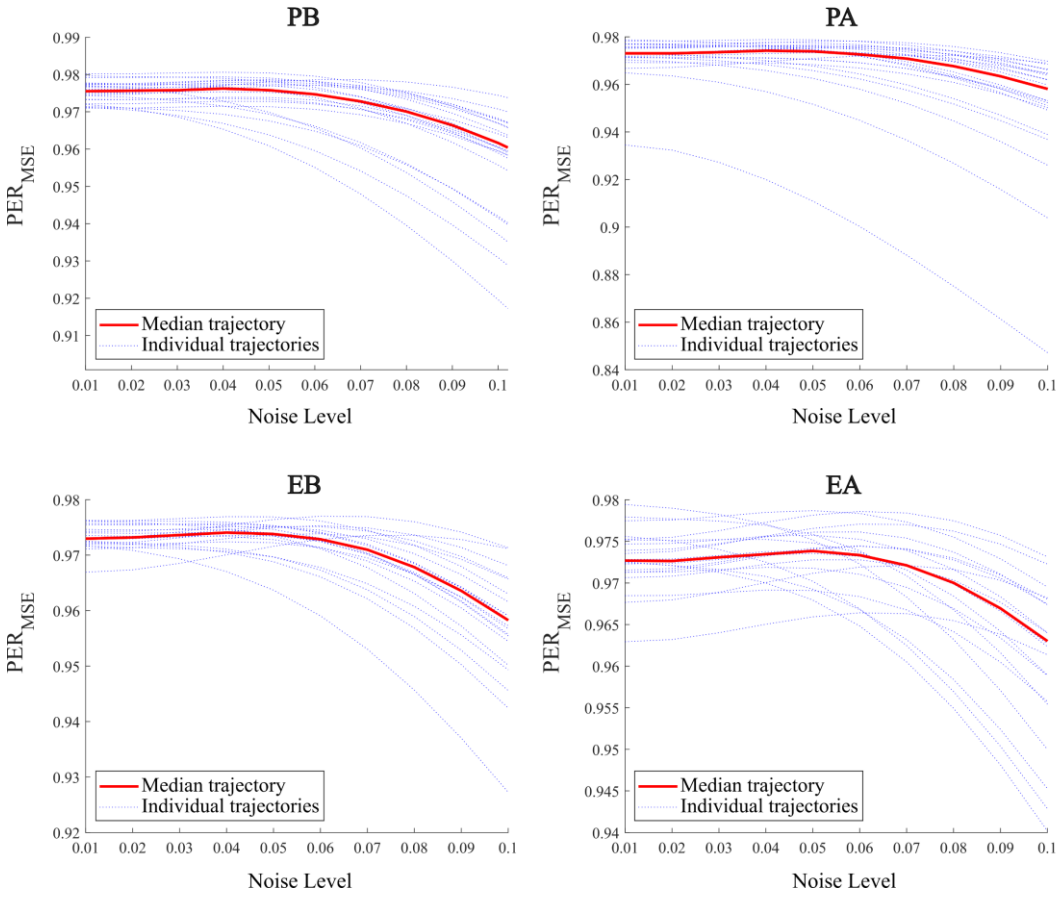

Figure A6. Noise levels associated to highest PER values. For each patient, the area-average perturbation effectivity to recovery (PER) using PER_{MSE} is shown across noise levels. For PB, PA, and EB, the median patient trajectory finds its maximum at a noise level $\beta = 0.04$. For EA, the maximum happens at $\beta = 0.05$.

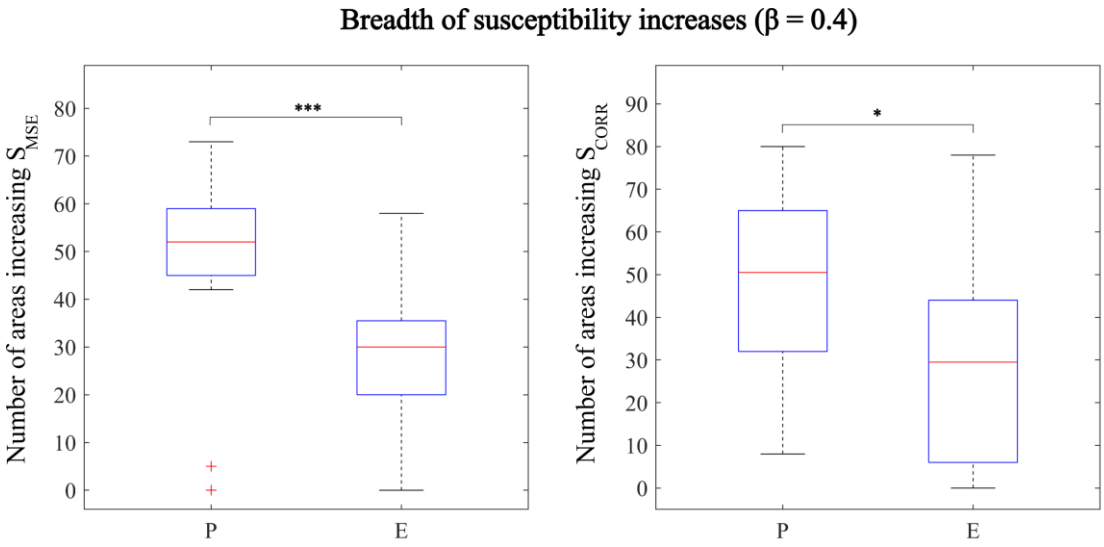

Figure A7. **Breadth of susceptibility increases with a high perturbation intensity.** Fixing the perturbation intensity to a high value (β = 0.4), susceptibility increases in more areas with psilocybin treatment compared to escitalopram. Statistical significance was found for both S_{MSE} (left, $p<0.001$, Wilcoxon test with 5000 permutations) and S_{CORR} (right, $p<0.05$, Wilcoxon test with 5000 permutations).

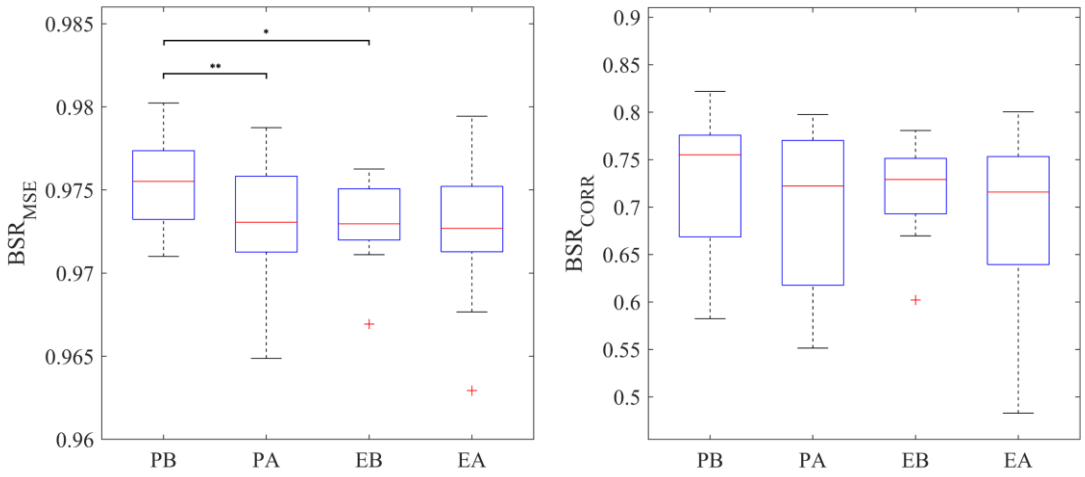

Figure A8. Baseline similarity to recovery. Baseline Similarity to Recovery (BSR) using error- (left) and correlation-based (right) metrics is shown for all groups. Using BSR_{MSE}, BSR was significantly higher for PB compared to PA ($p<0.01$, paired Wilcoxon test with 5000 permutations) and EB ($p<0.05$, Wilcoxon test with 5000 permutations) groups. A similar trend is shown using BSR_{CORR}, although no significance was found between groups.

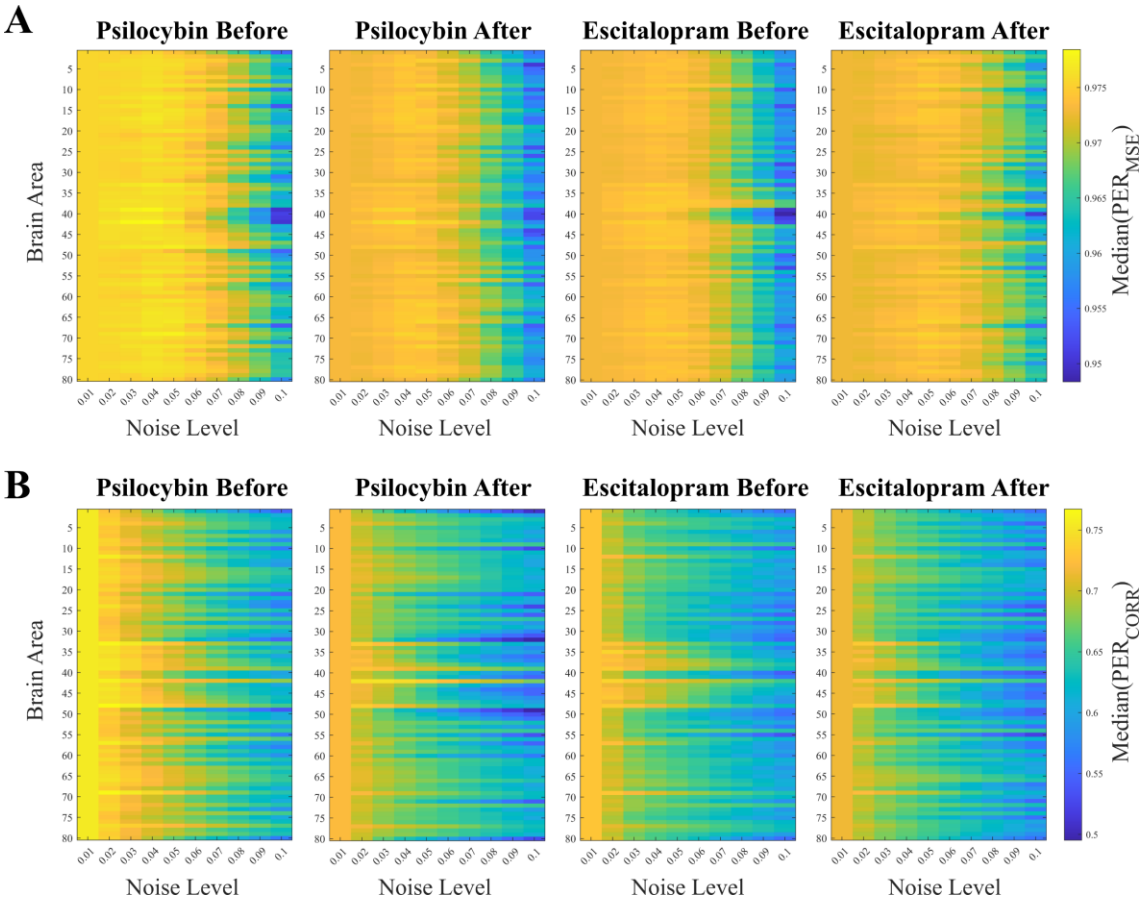

Figure A9. PER median perturbational maps. A) For all perturbations, the median PER using mean PER_{MSE} across patients of a same group is shown. **B)** The same procedure is applied using PER_{CORR}.

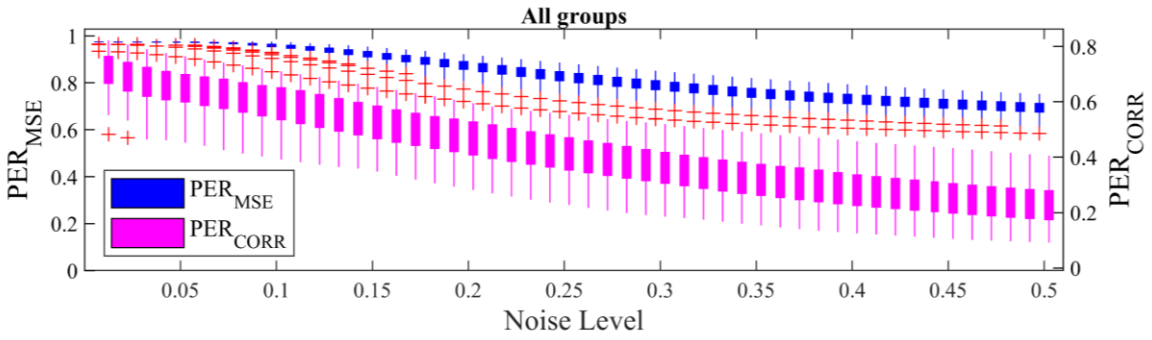

Figure A10. PER across noise levels. The patient distribution of the median perturbation effectivity to recovery (PER) across areas is plotted as a function of the noise level. The distributions represent all subjects, whether treated or not and regardless of the antidepressant arm.

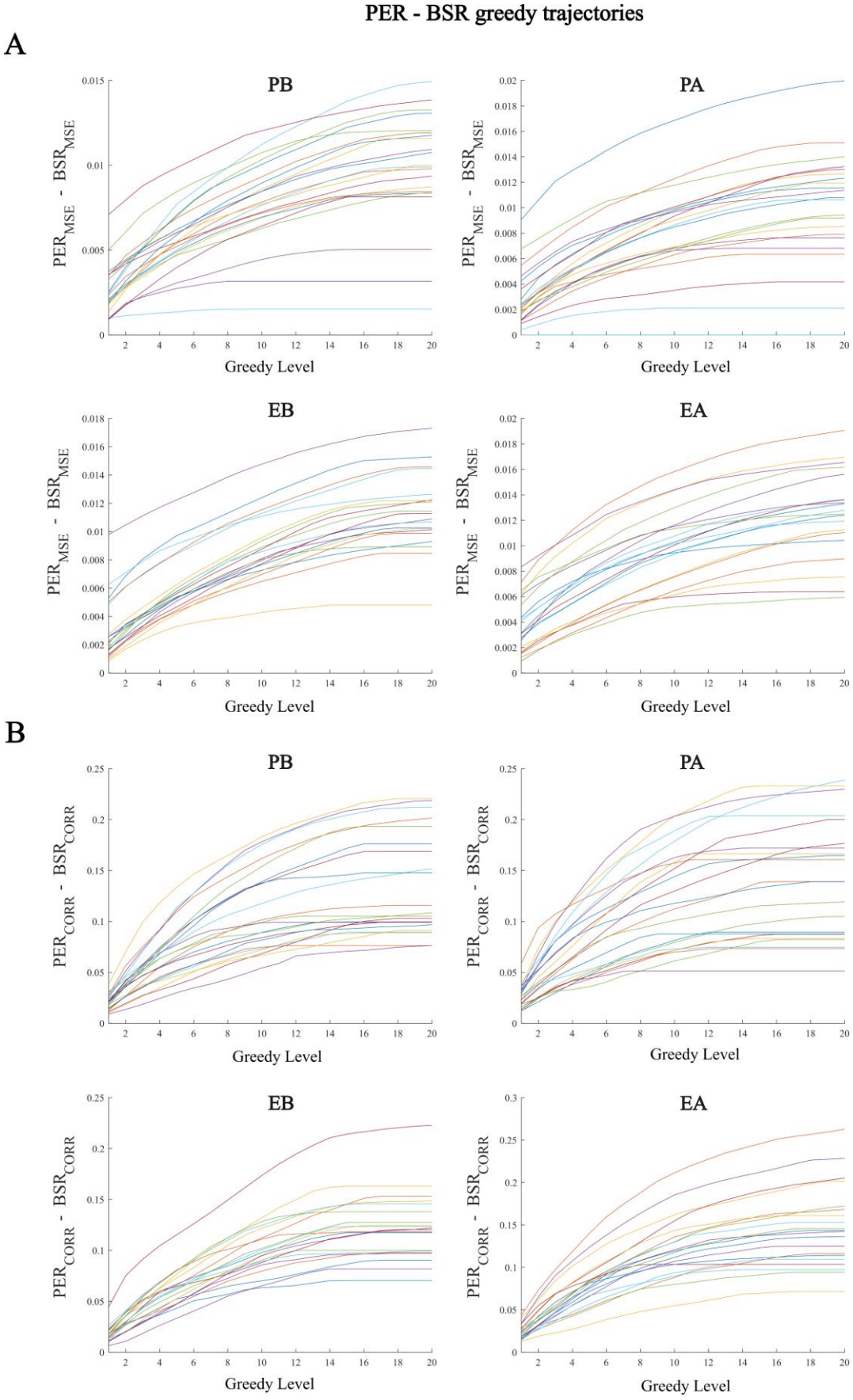

Figure A11. Greedy multi-site perturbation trajectories. A) The figure shows $PER_{MSE} - BSR_{MSE}$ trajectories for all PB (upper left), PA (upper right), EB (bottom left) and EA (bottom right) groups arising from PER_{MSE} maximization. **B)** Analogous $PER_{CORR} - BSR_{CORR}$ trajectories are shown for the greedy algorithm using PER_{CORR} maximization

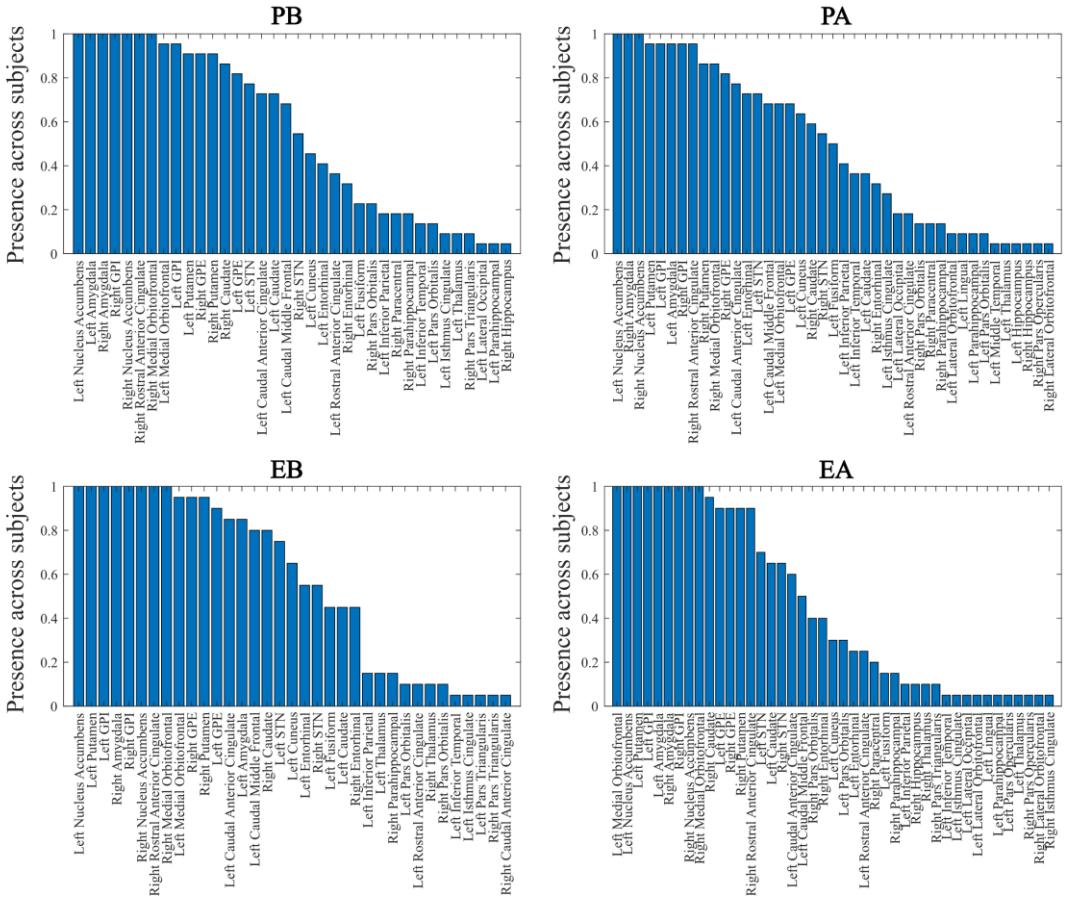

Figure A12. **Brain regions included in greedy multi-site perturbations.** The figure shows the presence of regions in the 20-area greedy multi-site perturbation strategy across patients, for both psilocybin and escitalopram arms before and after treatment (PB, PA, EB, EA).