



UNIVERSITÉ
FRANCO
ITALIENNE

UNIVERSITÀ
ITALO
FRANCESE



Analyzing and Modeling non-equilibrium Brain Dynamics in Schizophrenia

*A thesis submitted in fulfillment of the requirements for the degree of
Physics of Complex Systems (International track)*

Supervisors:

Prof. Andrea Pagnani
Prof. Rodrigo Cofré
Prof. Alain Destexhe

Student:

Angelica Di Domenico

Academic Year 2024–2025

Abstract

Schizophrenia is a complex neuropsychiatric disorder characterized by disrupted brain dynamics, yet the underlying non-equilibrium mechanisms remain poorly understood. In this study, we employ the multivariate Ornstein–Uhlenbeck (MOU) process to model resting-state fMRI data from individuals diagnosed with schizophrenia and healthy controls, aiming to quantify group differences in irreversibility and linear response properties. By fitting the MOU process to empirical covariance matrices, we estimate subject-specific friction and diffusion matrices, which fully define the stochastic process. Our results reveal a statistically significant increase in global EPR among schizophrenia patients, indicating heightened non-equilibrium dynamics. Regional analysis identifies the left inferior temporal gyrus and left pars orbitalis as key contributors to this irreversibility. Furthermore, linear response theory reveals altered propagation of perturbations in schizophrenia, with significant differences in area-under-the-curve and latency measures, particularly involving prefrontal and parietal regions. These findings suggest that schizophrenia is associated with widespread disruption of non-equilibrium brain dynamics and highlight potential biomarkers for the disorder.

Contents

1	Introduction	1
2	Theoretical framework	3
2.1	Multivariate Ornstein-Uhlenbeck (MOU) process	3
2.2	Characterizing irreversibility	4
2.3	Linear Response Theory	5
3	Methods	7
3.1	Experimental data	7
3.2	Fit MOU process to experimental data	8
3.3	Linear response analysis	10
4	Results	11
4.1	Fit MOU process to experimental data	11
4.2	Irreversibility: matrix $\mathbf{Q}$ and entropy production rate Φ	13
4.3	Linear Response Results	14
5	Discussion	21
6	Conclusion	23
A	Experimental data	25
B	Complete derivation of the fit updates	27
	Bibliography	29

Chapter 1

Introduction

Schizophrenia is a complex neuropsychiatric disorder characterized by disruptions in thought processes, perceptions, and emotional responsiveness [1]. It is linked to a combination of genetic and neurobiological factors that impact early brain development, and typically manifests through a combination of psychotic symptoms, such as hallucinations and delusions, as well as motivational and cognitive dysfunctions [2].

Despite extensive research, the underlying neural mechanisms remain poorly understood [3]. Recent advances in neuroimaging and computational neuroscience have provided new tools to investigate the dynamical properties of brain activity in health and disease [4]. In particular, the analysis of resting-state functional magnetic resonance imaging (fMRI) data offers a window into the spontaneous fluctuations of brain activity that may reflect fundamental organizational principles of neural systems [5, 6].

A key feature of healthy brain dynamics is the delicate balance between stability and flexibility, allowing for both robust function and adaptive responses. This balance is thought to be disrupted in schizophrenia, leading to altered patterns of functional connectivity [7] and information processing [8]. Traditional approaches to analyzing fMRI data have primarily focused on static functional connectivity, which captures correlations between brain regions but does not account for the temporal dynamics and directional influences that may be crucial to understanding the disorder [9].

In this work, we employ a framework from non-equilibrium statistical physics, the multivariate Ornstein-Uhlenbeck (MOU) process, to model and analyze the directed, time-dependent interactions between brain regions in schizophrenia [10]. The MOU process provides an analytically tractable model of linear stochastic dynamics in high-dimensional systems, making it suitable for investigating whole-brain dynamics from fMRI data. Its empirical consistency with key features of the BOLD signal allows us to quantify relevant aspects of non-equilibrium brain activity [11, 12].

Specifically, we focus on measures of irreversibility and entropy production, which capture the degree to which the system's dynamics violate time-reversal symmetry, a fundamental signature of non-equilibrium processes. Recent studies have demonstrated the utility of this approach for characterizing differences in brain dynamics across states of consciousness and in various consciousness conditions [13]. Building on this work, we investigate whether similar measures can distinguish between individuals with schizophrenia and healthy controls, potentially revealing novel biomarkers of the disorder.

Moreover, we characterize how the brain responds to localized perturbations by applying linear response theory to the MOU model, which enables the estimation of

brain responses without the need for explicit stimulation [14]. This model-based approach allows us to compare effective network dynamics between individuals with schizophrenia and healthy controls, and has already been successfully employed to study altered states of consciousness [15].

In summary, our analysis focuses on three main aspects: (1) the global irreversibility of brain dynamics, quantified by the entropy production rate; (2) the regional contributions to irreversibility, assessed via the antisymmetric matrix $\mathbf{Q}$; and (3) the system's response to localized perturbations, characterized using linear response theory.

The thesis is organized as follows. Chapter 2 introduces the theoretical framework of the MOU process and the key measures used in our analysis. Chapter 3 describes the experimental data, the procedure for fitting the model to fMRI signals, and the application of linear response theory. Chapter 4 presents our results comparing brain dynamics between individuals with schizophrenia and healthy controls. Finally, Chapters 5 and 6 discuss the implications of our findings and suggest directions for future research.

By bridging concepts from non-equilibrium physics with clinical neuroscience, this work aims to provide new insights into the dynamical alterations associated with schizophrenia, potentially contributing to the development of more objective diagnostic tools and a deeper understanding of the disorder's neural basis.

Chapter 2

Theoretical framework

In this section, we provide a detailed overview of the multivariate Ornstein – Uhlenbeck (MOU) process, a fundamental stochastic model for describing linear dynamics in high-dimensional systems subject to noise. We begin by introducing the defining equations and statistical properties of the process, including its Gaussian nature and the characterization of its stationary state. We then turn to the concept of irreversibility, quantified through the antisymmetric matrix $\mathbf{Q}$ and the scalar entropy production rate Φ . We conclude with the application of linear response theory to characterize how the system reacts to localized perturbations.

2.1 Multivariate Ornstein-Uhlenbeck (MOU) process

The multivariate Ornstein-Uhlenbeck (MOU) process is the higher-dimensional extension of the Ornstein-Uhlenbeck process, defined to describe the velocity of a Brownian particle in one dimension [16, 17]. Therefore, the MOU is a diffusion process defined by N -coupled linear Langevin equations of the form:

$$\frac{dx_m(t)}{dt} = - \sum_n B_{mn} x_n(t) + \eta_m(t), \quad \langle \eta_m(t) \eta_n(t') \rangle = 2D_{mn} \delta(t - t'), \quad (2.1)$$

with $m, n = 1, \dots, N$. Adopting the vector and matrix notations, denoted using boldfaced symbols, we get [10]:

$$\frac{d\mathbf{x}(t)}{dt} = -\mathbf{B}\mathbf{x}(t) + \boldsymbol{\eta}(t), \quad \langle \boldsymbol{\eta}(t) \boldsymbol{\eta}^T(t') \rangle = 2\mathbf{D}\delta(t - t'). \quad (2.2)$$

Here, the two real $N \times N$ matrices are:

- the *diffusion matrix* $\mathbf{D}$, namely the noise covariance matrix coming from the multivariate Gaussian white noise $\boldsymbol{\eta}(t)$, which is symmetric and positive definite;
- the *friction matrix* $\mathbf{B}$, which has eigenvalues with strictly positive real parts and generally is not symmetric.

The assumptions on $\mathbf{B}$ and $\mathbf{D}$ ensure that the process relaxes exponentially to $\mathbf{x} = \mathbf{0}$ in the absence of noise and to a fluctuating stationary state with Gaussian statistics in the presence of noise. Therefore, the solution of Eq.(2.2), knowing the initial condition to be $\mathbf{x}(0)$, is given by:

$$\mathbf{x}(t) = \mathbf{G}(t)\mathbf{x}(0) + \int_0^t \mathbf{G}(t-s)\boldsymbol{\eta}(s)ds, \quad (2.3)$$

where $\mathbf{G}(t) = e^{-\mathbf{B}t}$ is the Green's function.

The process $\mathbf{x}(t)$ is Gaussian, so its statistics at any time t are entirely characterized by the first two moments: mean value $\langle \mathbf{x}(t) \rangle = \mathbf{G}(t)\mathbf{x}(0)$, and zero-lag covariance matrix $\mathbf{S}(t) = \langle \mathbf{x}(t)\mathbf{x}^T(t) \rangle$. The latter obeys the following deterministic differential equation:

$$\frac{d\mathbf{S}(t)}{dt} = 2\mathbf{D} - \mathbf{B}\mathbf{S}(t) - \mathbf{S}(t)\mathbf{B}^T. \quad (2.4)$$

The stationary state of the process is Gaussian, with zero mean and stationary covariance matrix:

$$\mathbf{S} = \lim_{t \rightarrow \infty} \mathbf{S}(t) = 2 \int_0^\infty \mathbf{G}(t)\mathbf{D}\mathbf{G}^T(t)dt, \quad (2.5)$$

and the probability density of the process in the stationary state is:

$$P(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}(\det \mathbf{S})^{1/2}} \exp\left(-\frac{1}{2}\mathbf{x}^T\mathbf{S}^{-1}\mathbf{x}\right). \quad (2.6)$$

Using Eq.(2.4), we find the *Sylvester* equation, which relates the stationary covariance matrix $\mathbf{S}$ to the matrices $\mathbf{B}$ and $\mathbf{D}$ defining the process:

$$\mathbf{B}\mathbf{S} + \mathbf{S}\mathbf{B}^T = 2\mathbf{D}. \quad (2.7)$$

We extend the definition of the covariance matrix to account for second-order moments of the process evaluated at different times. This leads to the *cross-covariance matrix*, defined as:

$$C_{mn}(t, s) = \langle x_m(t)x_n(s) \rangle, \quad \text{for } 0 \leq s \leq t. \quad (2.8)$$

This matrix captures the temporal dependencies between the components x_m and x_n of the multivariate process over time.

In the framework of the multivariate Ornstein–Uhlenbeck process, this quantity satisfies the relation:

$$\mathbf{C}(t, s) = \mathbf{G}(t - s) \mathbf{S}(s), \quad (2.9)$$

with the initial condition $\mathbf{C}(s, s) = \mathbf{S}(s)$ when $t = s$ [10].

In the stationary state, where the system exhibits time-translation invariance, the cross-covariance depends only on the time lag $\tau = t - s \geq 0$. Adopting the notation of [13] and consistent with [10], we define the *lagged covariance matrix* as:

$$\mathbf{S}(\tau) = \mathbf{G}(\tau) \mathbf{S}, \quad \mathbf{S}(0) = \mathbf{S}. \quad (2.10)$$

This expression highlights how the temporal structure of the process is governed by the system propagator $\mathbf{G}(\tau)$ and the stationary covariance $\mathbf{S}$.

2.2 Characterizing irreversibility

The condition for process (2.2) to be reversible, namely the *detail balance* condition is [18, 19]:

$$\mathbf{B}\mathbf{D} = \mathbf{D}\mathbf{B}^T. \quad (2.11)$$

In such a case, the stationary state of the process is an equilibrium state, and solving the *Sylvester* equation (2.7) to find $\mathbf{S}$ is relatively simple [10].

When the symmetry condition (2.11) is not satisfied, the process is irreversible, but the system can still reach a non-equilibrium stationary state. In this case, solving the *Sylvester* equation (2.7) for $\mathbf{S}$ is considerably more complicated and requires

appropriate parametrization:

$$\mathbf{L} = \mathbf{B}\mathbf{S} = \mathbf{D} + \mathbf{Q}, \quad \mathbf{L}^T = \mathbf{S}\mathbf{B}^T = \mathbf{D} - \mathbf{Q}. \quad (2.12)$$

The matrix $\mathbf{L}$ is the *Onsager* matrix [18]. Its antisymmetric part $\mathbf{Q}$ is of particular interest for our analysis, since it provides a measure of the amount of irreversibility of the process [10]. If the process is reversible, then $\mathbf{Q} = 0$, $\mathbf{L} = \mathbf{D}$, and Eq.(2.11) is satisfied.

According to one of the standard formulations of the second law of thermodynamics, an irreversible process is associated with an increase in entropy. In this work, we characterize irreversibility through the *entropy production rate* (EPR), which is defined as the amount of entropy produced per unit of time in an open system. This quantity is now widely recognized as a key indicator of non-equilibrium stationary states [20, 21]. The EPR Φ is always non-negative: it vanishes in the case of reversible dynamics and is strictly positive otherwise. As such, Φ provides a scalar measure of the irreversibility of the entire process.

In the present context, the general expression of the entropy production rate Φ per unit time is [10, 22]:

$$\Phi = \langle \mathbf{x}^T \mathbf{S}^{-1} \mathbf{Q} \mathbf{D}^{-1} \mathbf{Q} \mathbf{S}^{-1} \mathbf{x} \rangle. \quad (2.13)$$

Since the stationary state of the process is Gaussian with covariance matrix $\mathbf{S}$, we have $\langle \mathbf{x}^T \mathbf{A} \mathbf{x} \rangle = \text{tr}(\mathbf{S}\mathbf{A})$, with $\mathbf{A}$ a generic matrix. In our specific case, it gives:

$$\Phi = -\text{tr}(\mathbf{S}^{-1} \mathbf{Q} \mathbf{D}^{-1} \mathbf{Q}) \quad (2.14)$$

Using Eq.(2.12) and Eq.(2.14), we can recast Φ into the following expressions, without explicitly involving matrix $\mathbf{S}$:

$$\Phi = \text{tr}(\mathbf{B}^T \mathbf{D}^{-1} \mathbf{Q}) = -\text{tr}(\mathbf{D}^{-1} \mathbf{B} \mathbf{Q}). \quad (2.15)$$

2.3 Linear Response Theory

We apply linear response theory to the multivariate Ornstein–Uhlenbeck process, following the approaches in [14, 15]. Specifically, we investigate how the process responds to an external perturbation by analyzing the resulting change in its mean trajectory. We consider a perturbation of the form:

$$\mathbf{f}(t) = \delta(t) \mathbf{p}, \quad \text{with} \quad \mathbf{p} = \hat{\mathbf{e}}_i \in \mathbb{R}^N, \quad (2.16)$$

where $\hat{\mathbf{e}}_i$ denotes the i -th canonical basis vector. This corresponds to applying a delta-like instantaneous perturbation at time $t = 0$ to the i -th component of the process $x_i(t)$.

Under this perturbation, the dynamics of the system become:

$$\frac{d\mathbf{x}(t)}{dt} = -\mathbf{B} \mathbf{x}(t) + \boldsymbol{\eta}(t) + \mathbf{f}(t). \quad (2.17)$$

We are interested in the first-order change in the mean of the process $\mathbf{x}(t)$ induced by the perturbation. This quantity, known as the linear response, is defined as:

$$\mathbf{R}(t) := \Delta \langle \mathbf{x}(t) \rangle \approx \int_0^t \mathbf{G}(t-s) \mathbf{f}(s) ds = \mathbf{G}(t) \mathbf{p}, \quad (2.18)$$

where $\mathbf{G}(t)$ is the green function. In this formulation, $\mathbf{R}(t)$ is a vector in $\mathbb{R}^N$ representing the temporal evolution of the mean response to a unit perturbation in the i -th component of the process.

Two scalar observables characterizing $\mathbf{R}(t)$ are of particular interest in our analysis:

- The *area under the curve* (AUC) of each component of the response $\mathbf{R}(t)$.
- The *latency*, defined as the time at which each component of $\mathbf{R}(t)$ reaches its maximum.

By considering perturbations to each of the $i = 1, \dots, N$ components of the system, we obtain a full set of response vectors $\mathbf{R}^{(i)}(t)$. From these, we construct two matrices:

- $\mathbf{A} \in \mathcal{M}_{N \times N}(\mathbb{R})$, where A_{ji} is the area under the curve of the j -th component's response to a perturbation in the i -th component,
- $\mathbf{L} \in \mathcal{M}_{N \times N}(\mathbb{R})$, where L_{ji} is the latency of the j -th component's response to that same perturbation.

Chapter 3

Methods

In this section, we describe the procedure used to fit the Multivariate Ornstein–Uhlenbeck (MOU) process to fMRI data acquired from subjects in a resting state. We assume that the resting-state BOLD signal can be modeled as a stationary stochastic process. The fitting approach is based on the method introduced in [11], which has previously been applied to study changes in brain dynamics associated with varying levels of consciousness, such as the transition from wakefulness to deep sleep [13].

3.1 Experimental data

The model is fitted to fMRI data acquired from 96 subjects in a resting-state condition: 48 individuals diagnosed with schizophrenia (condition = 1) and 48 control subjects (condition = 0) [23, 24].

The analyzed BOLD fMRI signals of each subject originate from $N = 68$ anatomically defined cortical brain Regions of Interest (ROIs), based on the widely used Desikan–Killiany atlas [25]. An example of a recording from a single subject is shown in Figure 3.1. Each signal consists of a series of fMRI acquisitions, with each time point corresponding to one image acquisition. The temporal resolution is defined by the repetition time (TR), with $1 \text{ TR} = 2 \text{ s}$. Each subject’s session lasted for $T = 152$ TRs. Further details on the experimental procedure and data are provided in Appendix A.

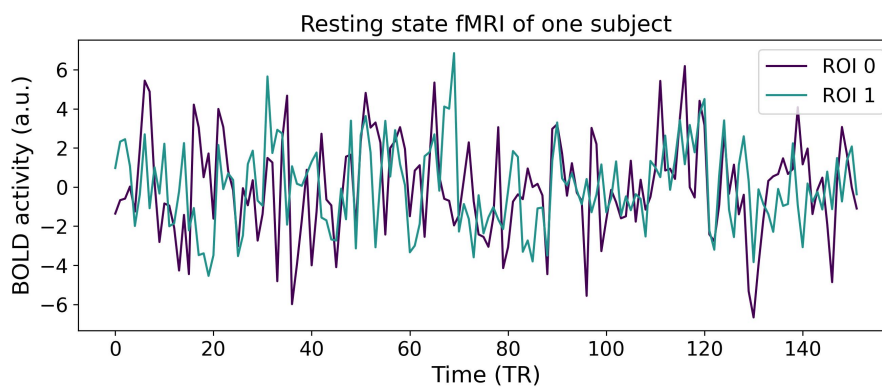


FIGURE 3.1: **Resting-state fMRI BOLD time series.** The plot shows filtered BOLD signals from subject 29 in condition 1. The signals were recorded over 152 repetition times (TRs) of 2 seconds each, corresponding to a total recording time of 304 seconds. The first two regions of interest (ROIs) from the Desikan–Killiany parcellation (out of 68 total) are shown.

For each subject, we compute empirical covariance matrices with a zero lag $\hat{\mathbf{S}}(0)$, and a lag of 1 TR $\hat{\mathbf{S}}(1)$. The elements of these two matrices are defined as follows:

$$\hat{S}_{ij}(\tau) = \frac{1}{T-2} \sum_{t=1}^{T-1} [x_i(t) - \bar{x}_i] [x_j(t+\tau) - \bar{x}_j] \quad \text{with } i, j = 1, \dots, N \quad (3.1)$$

where $\bar{x}_i = \frac{1}{T} \sum_{t=1}^T x_i(t)$ denotes the empirical mean of the signal $x_i(t)$, and $\tau = 0, 1$.

By normalizing $\hat{\mathbf{S}}(0)$, we obtain the functional connectivity (FC) matrix, which captures the Pearson correlation coefficients between each pair of ROIs (see Figure 3.2b):

$$FC_{ij} = \frac{\hat{S}_{ij}(0)}{\sqrt{\hat{S}_{ii}(0) \hat{S}_{jj}(0)}} \quad \text{with } i, j = 1, \dots, N. \quad (3.2)$$

Another relevant measure derived from neuroimaging data is the structural connectivity (SC) matrix, which quantifies the strength of anatomical connections between pairs of ROIs. In this study, we use structural connectivity data (see Figure 3.2a) obtained from unrelated healthy adults in the Human Connectome Project (HCP) [26, 27].

3.2 Fit MOU process to experimental data

The choice of the MOU model is motivated by its balance between simplicity and suitability for modeling brain activity. Its linear dynamics allow for tractable analytical calculations, while its structure is well-suited for whole-brain fMRI data, where signals are recorded from a parcellation of the cortex into distinct Regions of Interest (ROIs). Notably, the MOU dynamics generate exponentially decaying autocovariances, which closely resemble the temporal profiles observed in empirical BOLD signals. The model can be interpreted as a network in which spontaneous fluctuations are locally generated at each node and then propagate through the anatomical or functional connections across the brain. Despite these advantages, the MOU model presents two main limitations: it assumes stationarity of the fMRI signals over time and does not incorporate an explicit hemodynamic component to account for the transformation from neural activity to BOLD responses [12].

To apply this framework to our data, we fit the MOU process to the fMRI time series of each subject (an example is shown in Figure 3.1). The fitting procedure follows the approach introduced in [11]¹, which adjusts the covariance structure of the MOU process to match the empirical covariances defined in Equation (3.1). The objective is to estimate, for each subject, the friction matrix $\mathbf{B}$ and the diffusion matrix $\mathbf{D}$.

Since each covariance matrix has dimension $N \times N$, we reduce the number of free parameters to make the fitting procedure tractable by imposing constraints on $\mathbf{B}$ and $\mathbf{D}$. Specifically, we assume that the diffusion matrix $\mathbf{D}$ is diagonal, which corresponds to assuming that the noise is not only temporally white but also spatially uncorrelated: $\langle \eta_m(t) \eta_n(t') \rangle = 2D_{mm} \delta_{mn} \delta(t - t')$.

To constrain the structure of the friction matrix $\mathbf{B}$, we apply a subject-specific topological mask. Such a mask is defined as the logical OR between the subject's functional connectivity matrix and the structural connectivity matrix (both of which

¹The notation used in this report corresponds to that of [11], with the following substitutions: $\mathbf{B} \leftrightarrow -\mathbf{J}$ and $\mathbf{D} \leftrightarrow \frac{1}{2}\Sigma$.

are binarized using a matrix-specific threshold), where the diagonal is forced to be zero. The threshold values were chosen to strike a balance between the performance metrics of the fitting procedure, specifically the model error and the goodness of fit, described in the next Section 4.1. An example of a topological mask is shown in Figure 3.2c.

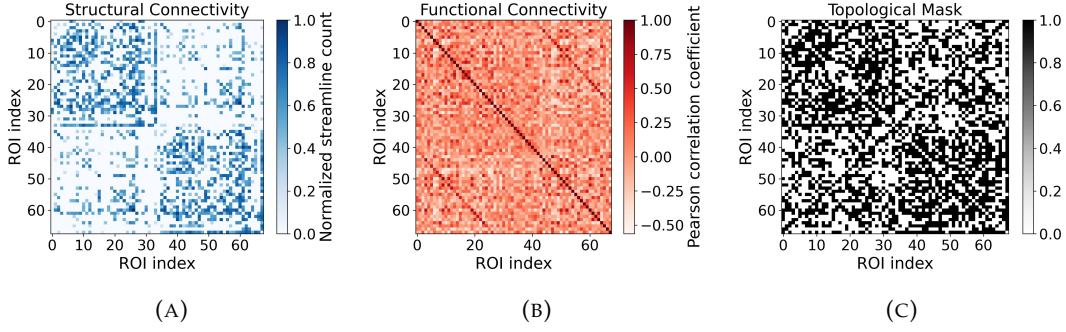


FIGURE 3.2: **Structural and Functional Connectivity, and Combined Topological Mask.** (A) Normalized structural connectivity (SC) matrix computed from ENIGMA diffusion MRI data using the ENIGMA Toolbox [28], based on the Desikan–Killiany atlas with 68 cortical and subcortical regions of interest (ROIs). Each element represents the normalized streamline count between pairs of ROIs, reflecting the estimated white-matter connectivity strength. (B) Empirical functional connectivity matrix for subject 29 in condition 1, derived from resting-state BOLD signals. The matrix entries correspond to Pearson correlation coefficients between pairs of ROIs. (C) Topological mask for subject 29 in condition 1, computed as the logical OR between the structural connectivity matrix in subfigure (A) (thresholded at 0.9) and the functional connectivity matrix in subfigure (B) (thresholded at 0.1)

We employ a gradient descent algorithm to iteratively update the matrices $\mathbf{B}$ and $\mathbf{D}$ to minimize the following loss function:

$$L = \|\Delta \mathbf{S}(0)\|^2 + \|\Delta \mathbf{S}(1)\|^2 \quad (3.3)$$

where $\Delta \mathbf{S}(\tau) = \hat{\mathbf{S}}(\tau) - \mathbf{S}(\tau)$, with $\hat{\mathbf{S}}(\tau)$ defined in Equation (3.1), and $\mathbf{S}(\tau)$ computed using the theoretical model (Equation (2.10)).

The updates for $\mathbf{B}$ and $\mathbf{D}$ are derived by exploiting Equations (2.7) and (2.10). In particular, the Sylvester equation (2.7) is solved using the Bartels–Stewart algorithm [29].

More specifically, the update for the diffusion matrix $\mathbf{D}$ is obtained by perturbing $\mathbf{B}$, $\mathbf{S}(0)$ in Equation (2.7) and keeping the first order expansion, given by:

$$\Delta \mathbf{D} \approx \epsilon_D \left(\mathbf{B} \Delta \mathbf{S}(0) + \Delta \mathbf{S}(0) \mathbf{B}^T \right), \quad (3.4)$$

where ϵ_D absorbs constants and acts as the learning rate.

While the friction matrix $\mathbf{B}$ is updated by perturbing $\mathbf{S}(0)$ and $\mathbf{S}(1)$ and keeping the first order expansion, given by:

$$\Delta \mathbf{B} \approx \epsilon_B \mathbf{S}(0)^{-1} \left(\Delta \mathbf{S}(0) - \Delta \mathbf{S}(1) e^{\mathbf{B}^T} \right), \quad (3.5)$$

where ϵ_B is the learning rate for $\mathbf{B}$.

In summary, the update $\Delta \mathbf{B}$ exploits the sensitivity of $\mathbf{S}(1)$ to changes in $\mathbf{B}$, while $\Delta \mathbf{D}$ is derived from the constraint given by the Sylvester Equation (2.7). The learning rates ϵ_B and ϵ_D are chosen to ensure convergence and numerical stability during the optimization process.

3.3 Linear response analysis

In experimental neuroscience, external perturbations have been used to probe brain dynamics in different levels of consciousness, for instance, through the combination of Transcranial Magnetic Stimulation (TMS) and EEG recordings [30].

Inspired by this approach, it is of interest to investigate how the brain would respond to perturbations using model-based methods, applied to resting states recordings, which avoid the need for actual stimulation. This is also possible with the model used in our study, the Multivariate Ornstein–Uhlenbeck (MOU) process, by applying Linear Response Theory. Such an approach has previously been employed to compare brain dynamics between healthy individuals and patients in altered states of consciousness, such as coma [15].

In this work, we adopt a similar strategy to investigate localized differences in brain dynamics between control subjects and individuals diagnosed with schizophrenia. After the fitting procedure described in Section 3.2, we obtain for each subject a matrix $\mathbf{B}$ and a matrix $\mathbf{D}$, which together define a subject-specific Multivariate Ornstein–Uhlenbeck (MOU) process that models the observed fMRI BOLD signals. To characterize how these signals respond to localized perturbations, we use linear response theory, as introduced in Section 2.3. Specifically, we analyze the system’s response to a delta-like perturbation applied to a single brain region i . This theoretical perturbation can be interpreted as a proxy for real external stimulation of the brain, which would manifest as a Dirac delta in the recorded BOLD signal.

Figure 3.3 shows an example of the response function $\mathbf{R}^{(i)}(t)$, as defined in Equation (2.18), following a perturbation of region $i = 47$ in a specific subject. For each subject and condition, we compute the response functions for all regions $i = 1, \dots, 68$. From these, we derive two matrices that characterize the response for each subject: the area under the curve $\mathbf{A}$ and the latency $\mathbf{L}$, as introduced in Section 2.3.

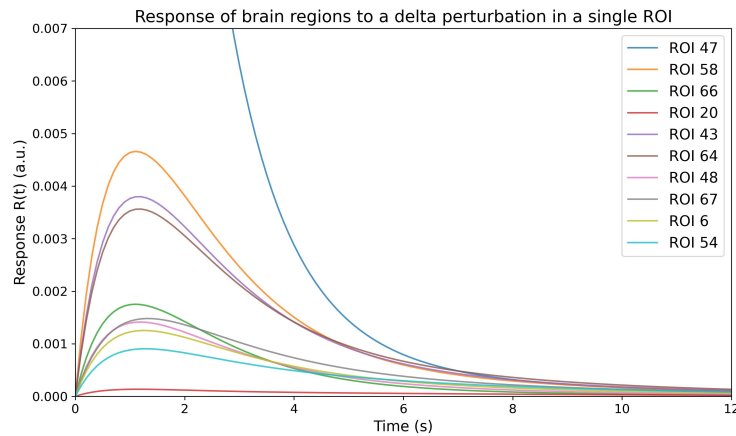


FIGURE 3.3: **Response of brain regions to a delta perturbation in a single ROI.** Response functions of the brain regions of subject 29 in condition 1 following a perturbation in region 47. For clarity, only the 10 regions with the highest area under the curve (AUC) and lowest latency are shown.

Chapter 4

Results

In this section, we present the results obtained by fitting the Multivariate Ornstein - Uhlenbeck (MOU) model to the resting-state fMRI data of both control subjects and patients diagnosed with schizophrenia. The analysis is structured into three main parts. First, we evaluate the quality of the model fitting procedure using two performance measures: model error and goodness of fit.

Second, we assess the degree of time-irreversibility in brain dynamics by computing the irreversibility matrix $\mathbf{Q}$ and the associated entropy production rate (EPR) Φ . This includes both a global comparison between the two populations and a regional analysis to identify local contributions to the observed differences.

Finally, we apply linear response theory to characterize how brain regions react to localized perturbations. We analyze both the area under the curve and the latency of the modeled responses to uncover specific perturbation-response connections that differ significantly between the two groups. This approach allows us to detect subtle changes in directed interactions that are not captured by global measures alone.

4.1 Fit MOU process to experimental data

We fit the MOU to the data of each subject, and we use two matrices of performance to evaluate the process: Model Error and Goodness of Fit.

The model error is defined as:

$$\text{Model error} = \frac{1}{2} [\|\mathbf{S}(0) - \hat{\mathbf{S}}(0)\|_F + \|\mathbf{S}(1) - \hat{\mathbf{S}}(1)\|_F]. \quad (4.1)$$

where $\|A\|_F = \sqrt{\sum_{i,j} A_{ij}^2}$ is the Frobenius norm.

While the goodness of fit is computed by first vectorising the upper-triangular part (excluding the diagonal) of each covariance matrix,

$$\mathbf{s}_{\text{model}} = \text{vec}[\mathbf{S}(\ell)], \quad \mathbf{s}_{\text{data}} = \text{vec}[\hat{\mathbf{S}}(\ell)], \quad \text{with } \ell \in \{0, 1\}, \quad (4.2)$$

to then subsequently compute the Pearson Coefficient:

$$r_\ell = \frac{\text{COV}(\mathbf{s}_{\text{model}}, \mathbf{s}_{\text{data}})}{\sigma_{\text{model}} \sigma_{\text{data}}}, \quad \ell \in \{0, 1\}, \quad (4.3)$$

And average the two results:

$$\text{Goodness of fit} = \frac{1}{2} (r_0 + r_1). \quad (4.4)$$

The model error quantifies the element-wise discrepancy between the model-generated and empirical covariance matrices and is sensitive to both pattern and scale differences. A lower value indicates a better match. Conversely, the Pearson correlation coefficient is scale-invariant and evaluates the similarity in the spatial pattern of connectivity between the two matrices. Therefore, a higher value of goodness of fit, ideally close to 1, indicates a better reproduction of the empirical covariance structure.

Figure 4.1 summarizes the results of the fitting procedure in the 48 subjects in both conditions. As shown, each subject achieves a goodness of fit above 0.6, while the model error remains mostly below 0.8, indicating a generally good agreement between the model and empirical data.

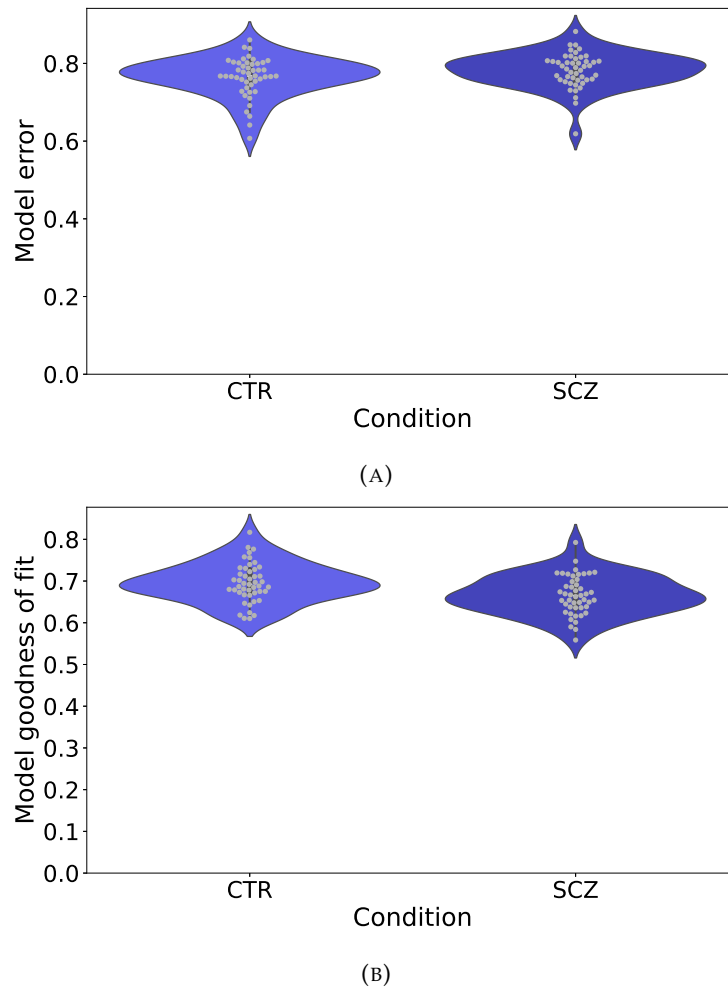


FIGURE 4.1: Model Error and Goodness of Fit Across Subjects and Conditions. **(A)** Model error of the overall fitting procedure. Each dot represents a subject in a specific condition, while the violin plots illustrate the distribution of model errors across subjects for the two conditions (SCZ = schizophrenia, CTR = control). **(B)** Model goodness of fit, measured as the average Pearson correlation between the model and empirical covariance matrices. Each dot corresponds to a subject in a specific condition, and the violin plots summarize the distribution across subjects in the two groups.

4.2 Irreversibility: matrix $\mathbf{Q}$ and entropy production rate Φ

Using the subject-specific parameters obtained from the fitting procedure described in Section 3.2, namely the matrices $\mathbf{B}$ and $\mathbf{D}$, we compute the matrix $\mathbf{Q}$, which we name *irreversibility matrix* for each subject. From $\mathbf{Q}$, we then derive the corresponding entropy production rate (EPR) Φ using Equation (2.15).

To assess differences in the global irreversibility of the process between the two populations, we compare the distributions of EPR values across the two groups. Since the data are independent and no assumption was made on their Gaussianity, we employ the non-parametric Mann–Whitney U test to statistically evaluate the group differences.

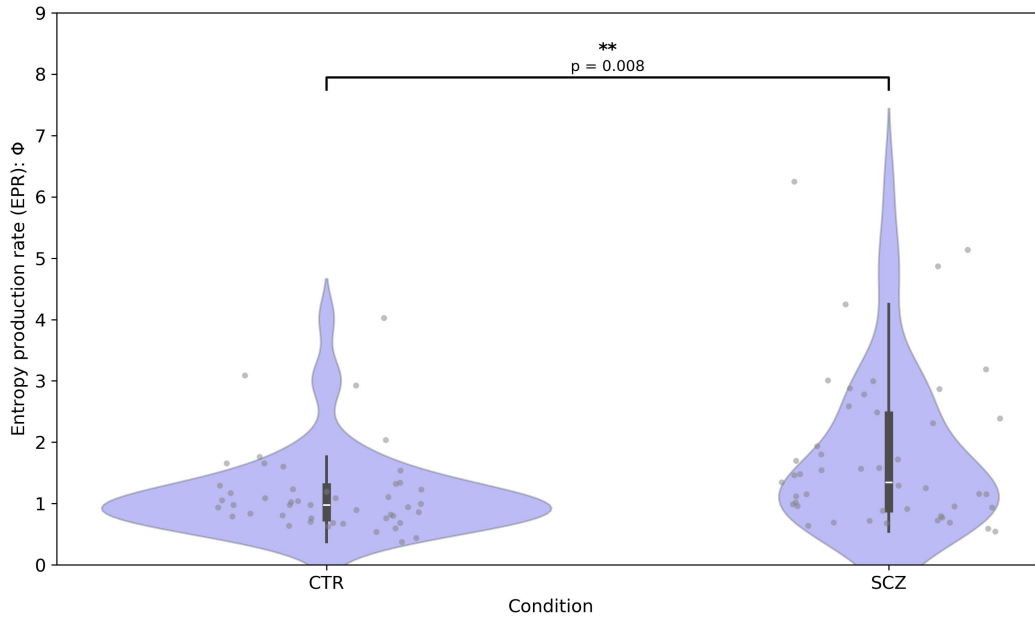


FIGURE 4.2: MOU Entropy Production Rate of each individual in each condition. Results of the Mann–Whitney U test comparing the distributions of entropy production rate (EPR) between the two groups: subjects diagnosed with schizophrenia (SCZ) and control subjects (CTR). The test reveals a statistically significant difference with a p-value of $p = 0.008$, below the significance threshold of $p < 0.05$ (corresponding to a 95% confidence level). Each point corresponds to one subject; the horizontal line in the middle of each distribution indicates the population median. Asterisks denote the level of statistical significance, with one to three stars indicating increasing levels of significance.

As shown in Figure 4.2, we find a statistically significant difference in the EPR Φ values between the subjects in each condition: subjects diagnosed with schizophrenia exhibit significantly higher entropy production rates compared to the control group. These findings suggest that the entropy production rate may serve as a global neural marker associated with schizophrenia.

To investigate the regional contributions to the global difference in irreversibility, reflected in the significant difference in the entropy production rate of the two conditions, we analyze the differences in the $\mathbf{Q}$ matrices between the two groups.

Specifically, to assess the contribution of each Region of Interest (ROI) to global irreversibility, we define the *nodal irreversibility* of node i as:

$$V_i^Q = \sum_j |Q_{ij}|, \quad i = 1, \dots, 68. \quad (4.5)$$

This yields, for each subject, a vector $\mathbf{V}^Q$ whose components provide a proxy for the local irreversibility associated with each ROI.

To identify ROIs with significant differences in nodal irreversibility between the two populations, we perform a Mann–Whitney U test independently on each component of $\mathbf{V}^Q$. Given the multiple hypothesis tests involved, we apply the Bonferroni correction to control the risk of false positives and maintain the reliability of the results.

This analysis reveals statistically significant differences in two specific ROIs, highlighted in Figure 4.3.

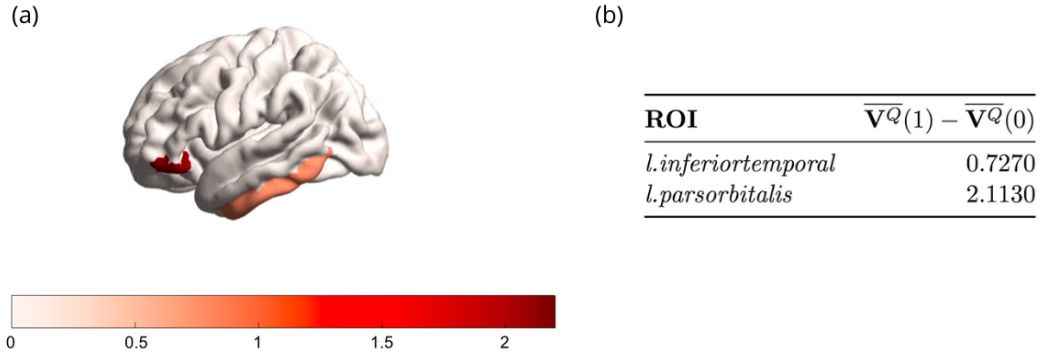


FIGURE 4.3: **Statistically significant differences in nodal irreversibility $\mathbf{V}^Q$.** (a) Cortical surface visualization with regions defined by the Desikan–Killiany atlas. The color map represents the group-level difference in nodal irreversibility: $\overline{\mathbf{V}^Q}(1) - \overline{\mathbf{V}^Q}(0)$, where (1) indicates the schizophrenia group and (0) the control group. Only regions exhibiting statistically significant differences after Bonferroni correction are shown. (b) Table listing the names of the ROIs with statistically significant group differences in nodal irreversibility, along with the corresponding differences in group means.

These two brain regions provide the most significant local contributions to the observed global difference in the entropy production rate. However, they are not solely responsible for this global effect: repeating the comparison of the EPR associated with the two conditions, after excluding these ROIs, we still find a statistically significant difference. This suggests that the increase in irreversibility observed in schizophrenia is distributed across multiple brain regions rather than being localized to a few specific areas.

4.3 Linear Response Results

We analyze the response matrices $\mathbf{A}$ and $\mathbf{L}$ derived from the Linear Response Theory framework, as described in Section 3.3, comparing control subjects and individuals with schizophrenia across all region pairs. Our focus is on statistically significant differences in both the magnitude, quantified by the area under the curve (AUC), and the timing of responses, quantified by the latency.

Therefore, we compare the elements of the response-based matrices $\mathbf{A}$ and $\mathbf{L}$ across the two populations. Specifically, we test for statistically significant differences in $\mathbf{A}_{ij}$ and $\mathbf{L}_{ij}$, which respectively quantify the magnitude and latency of the modeled response in region j following a perturbation in region i . This analysis aims to reveal alterations in the directional propagation of activity between ROIs in patients diagnosed with schizophrenia.

For each pair (i, j) , we perform a Mann–Whitney U test independently on the values of $\mathbf{A}_{ij}$ or $\mathbf{L}_{ij}$ across subjects in the two groups. As in the analysis of $\mathbf{V}^Q$ in Section 4.2, we apply the Bonferroni correction to account for multiple comparisons.

The analysis of the magnitude matrix $\mathbf{A}$ reveals statistically significant differences between the two populations in specific perturbation–response pairs, as illustrated in the directed graph in Figure 4.4. The connections shown in the graph correspond to those listed in Table 4.1, which reports the names of the region pairs along with the corresponding differences in response magnitude between control subjects and patients diagnosed with schizophrenia. Positive values indicate stronger responses in control subjects compared to the schizophrenia group. In summary, the model predicts that for specific perturbation–response pairs, individuals with schizophrenia exhibit a reduced response magnitude following a delta-like perturbation to the BOLD signals.

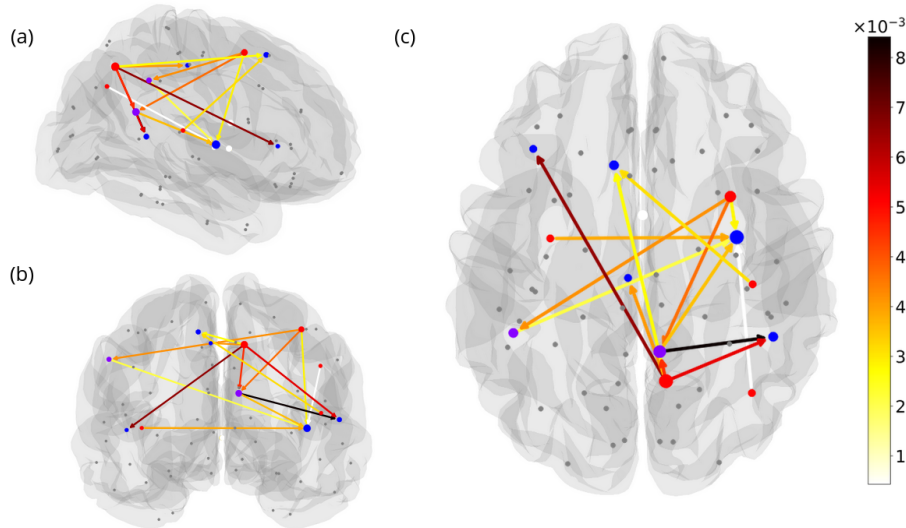


FIGURE 4.4: Significant Differences in Area Under the Curve (AUC) between control subjects and schizophrenic patients. The plot highlights statistically significant differences in the area under the curve (AUC) of the response functions between the two populations. Only statistically significant pairs are shown; all other entries were set to zero. The color map represents the value of the mean difference of the AUC in the two conditions. Specifically, for each pair (i, j) we compute the average of the AUC in the subjects in each condition (condition 0 = control, condition 1 = schizophrenia), and then we compute the respective difference: $\bar{\mathbf{A}}_{ij}(0) - \bar{\mathbf{A}}_{ij}(1)$. Node colors indicate directionality: red for perturbed nodes, blue for response nodes, and purple for nodes acting in both roles, depending on the connection. Subfigures (a), (b), and (c) show sagittal ($x = 0$), coronal ($y = 0$), and axial ($z = 0$) projections, respectively.

TABLE 4.1: **Region pairs showing the largest AUC response differences.** Each entry lists a perturbed–response ROI pair exhibiting a significant difference in the area under the curve (AUC) of the response function between groups. Values are reported as $\times 10^{-3}$ in arbitrary units (a.u.).

Perturbed ROI	Response ROI	Mean AUC Difference
l.supramarginal	r.insula	2.5
l.insula	r.insula	3.9
r.caudalmiddlefrontal	l.supramarginal	4.2
r.caudalmiddlefrontal	r.isthmuscingulate	4.6
r.caudalmiddlefrontal	r.insula	3.1
r.inferiorparietal	r.insula	1.1
r.isthmuscingulate	r.bankssts	8.3
r.isthmuscingulate	r.insula	3.7
r.precuneus	l.parstriangularis	6.9
r.precuneus	l.posteriorcingulate	4.1
r.precuneus	l.superiorfrontal	3.0
r.precuneus	r.bankssts	6.0
r.precuneus	r.isthmuscingulate	5.2
r.transversetemporal	l.superiorfrontal	3.4

Moreover, analysis of the in-degree and out-degree of the directed graph provides further insights into the organization of perturbation–response interactions. Notably, the right precuneus and the right caudal middle frontal gyrus emerge as key driver regions, exerting influence over multiple targets when perturbed, as shown in Table 4.3 and represented in Figure 4.5. On the other hand, the right insula appears prominently as a receiver, exhibiting consistent and widespread responses to a variety of perturbations, as reported in Table 4.2 and represented in Figure 4.5.

It is important to note that this is not a global effect. The observed differences are localized to specific directed connections, rather than reflecting a widespread reduction in response magnitude across the brain. To verify this, we perform an additional analysis: for each subject, we compute the average of the entire matrix $\mathbf{A}$, obtaining a scalar value $A = \langle \mathbf{A} \rangle_{ij}$. Comparing these values across the two populations, we do not observe a statistically significant difference.

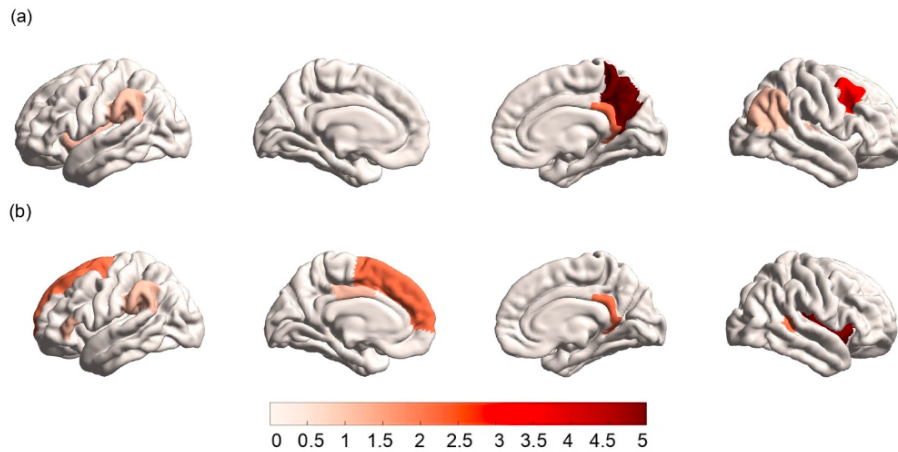


FIGURE 4.5: **Brain regions with significant difference in AUC, ranked based on nodal degree.** The figure shows the cortical surface with regions defined by the Desikan–Killiany atlas. **(a)** ROIs ranked based on the out-degree represented in the color map. The regions are the ones listed in Table 4.3. **(b)** ROIs ranked based on the in-degree represented in the color map. The regions are the ones listed in Table 4.2.

TABLE 4.2: **Brain regions ranked by in-degree.** Each ROI is ranked by its in-degree, defined as the number of perturbations from other regions that elicit a significant response in the ROI.

ROI	In-degree
r.insula	5
r.isthmuscingulate	2
r.bankssts	2
l.superiorfrontal	2
l.supramarginal	1
l.parstriangularis	1
l.posteriorcingulate	1

TABLE 4.3: **Brain regions ranked by out-degree.** Each ROI is ranked by its out-degree, defined as the number of other regions that exhibit significant response differences when the listed ROI is perturbed.

ROI	Out-degree
r.precuneus	5
r.caudalmiddlefrontal	3
r.isthmuscingulate	2
l.insula	1
l.supramarginal	1
r.inferiorparietal	1
r.transversetemporal	1

In addition to the analysis of response AUC, we investigate the latency of the system's reaction, which encodes a different aspect of the dynamics; specifically, the rapidity or delay with which the system responds to external perturbation. While the area under the curve reflects the overall magnitude of the response, the latency focuses on its temporal characteristics.

Figure 4.6 illustrates the connections for which the latency differs significantly between the two groups. Each edge represents a perturbation-response pair, and the edge color indicates the group-level difference in average latency values: $\bar{L}_{ij}(0) - \bar{L}_{ij}(1)$. Only connections with an average latency difference above 1.7s are shown for visualization clarity.

The corresponding region pairs are listed in Table 4.4, region pairs are listed in Table 5, which highlights the most delayed responses. These differences reveal that, for specific connections, the response in control subjects is not only stronger (as previously shown) but also slower compared to patients with schizophrenia. This suggests that the dynamical propagation of information may be altered in schizophrenia, potentially reflecting a shift toward faster but less robust responses.

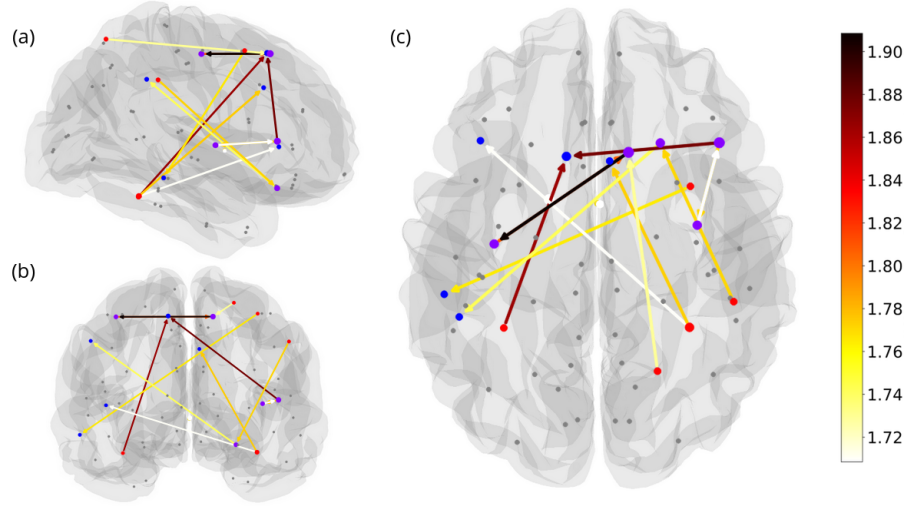


FIGURE 4.6: **Significant Differences in Latency between control subjects and schizophrenic patients.** The plot highlights statistically significant differences in the latency of the response functions between the two populations, computed as the group-level difference for each connection: $\bar{L}_{ij}(0) - \bar{L}_{ij}(1)$. Only statistically significant pairs are shown; all other entries were set to zero. For visualization purposes, we display only the connections with a mean latency difference greater than 1.7. Node colors indicate directionality: red for perturbed nodes, blue for response nodes, and purple for nodes acting in both roles, depending on the connection. Subfigures (a), (b), and (c) show sagittal ($x = 0$), coronal ($y = 0$), and axial ($z = 0$) projections, respectively.

TABLE 4.4: **Region pairs with the largest differences in response latency.** Each entry lists a perturbed–response ROI pair showing a significant difference in the latency of the response function between groups, higher than 1.7 s.

Perturbed ROI	Response ROI	Mean Latency Difference (s)
l.fusiform	l.superiorfrontal	1.87
l.precentral	r.parstriangularis	1.70
l.precentral	r.superiorfrontal	1.81
r.caudalmiddlefrontal	l.middletemporal	1.77
r.fusiform	l.parstriangularis	1.71
r.fusiform	r.caudalanteriorcingulate	1.77
r.lateralorbitofrontal	l.supramarginal	1.74
r.parstriangularis	l.superiorfrontal	1.88
r.parstriangularis	r.insula	1.77
r.superiorfrontal	l.precentral	1.91
r.superiorparietal	r.superiorfrontal	1.73
r.supramarginal	r.lateralorbitofrontal	1.77
r.supramarginal	r.superiorfrontal	1.69
r.insula	r.parstriangularis	1.71

We also analyze the out-degree and in-degree of the directed graph based on latency differences. As shown in Table 4.5 and represented in Figure 4.7, the right pars triangularis region acts as a major source of faster responses when perturbed. Conversely, Table 4.6 shows that the right pars triangularis, the left precentral, and right superior frontal gyrus are the most frequent targets of short-latency effects, indicating their consistent role as fast responders to external perturbations.

Finally, we test whether the observed differences in latency are global or localized. By averaging the latency values across the entire matrix for each subject and comparing the group-level means, we do not observe a statistically significant difference. This confirms that, as with the area under the curve, latency alterations are specific to particular perturbation–response connections, rather than being a general property of the system.

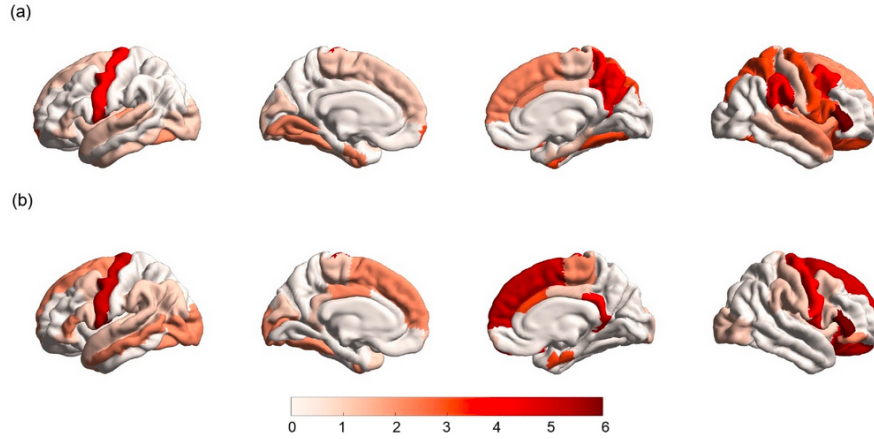


FIGURE 4.7: **Brain regions with significant difference in latency, ranked based on nodal degree.** The figure shows the cortical surface with regions defined by the Desikan–Killiany atlas. **(a)** ROIs ranked based on the out-degree represented in the color map. The regions are the ones listed in Table 4.5. **(b)** ROIs ranked based on the in-degree represented in the color map. The regions are the ones listed in Table 4.6.

TABLE 4.5: **Perturbed regions ranked by out-degree.** ROIs are ranked by out-degree, defined as the number of other regions showing significant latency differences in response to their perturbation. Only regions with an out-degree greater than 4 are shown.

Perturbed ROI	Out-degree
r.parstriangularis	6
r.caudalmiddlefrontal	4
l.precentral	4
r.precuneus	4
r.supramarginal	4

TABLE 4.6: **Response regions ranked by in-degree.** ROIs are ranked by in-degree, defined as the number of perturbations from other regions that elicit significant latency changes in the listed ROI. Only regions with an in-degree greater than 4 are shown.

Response ROI	In-degree
r.parstriangularis	6
l.precentral	5
r.superiorfrontal	5
r.lateralorbitofrontal	4
r.precentral	4
r.isthmuscingulate	4

Chapter 5

Discussion

We employ concepts of non-equilibrium physics to investigate alterations in brain dynamics in schizophrenia by modeling resting-state fMRI data with the multivariate Ornstein–Uhlenbeck (MOU) process. Our analysis reveals that individuals diagnosed with schizophrenia exhibit greater global irreversibility in brain dynamics, as indicated by a significantly higher entropy production rate (EPR). This aligns with the growing literature linking neuropsychiatric disorders to non-equilibrium processes [12, 31]. The higher EPR in schizophrenia may be due to altered feedback mechanisms or aberrant synaptic efficiency, leading to less predictable neural activity [32].

We investigated the brain-regional contributions to the global difference in irreversibility by examining the matrix $\mathbf{Q}$. Our findings show that the left inferior temporal gyrus and the left pars orbitalis are key contributors to the difference in the entropy production rate of the two populations. Such regions are known to play a role in semantic processing and executive function, cognitive domains that are often impaired in schizophrenia [33–35]. However, the persistence of a significant global difference in EPR, even after excluding these regions, implies a distributed nature of the irreversibility alteration, consistent with the disorder’s whole-brain dysconnectivity [36].

The linear response analysis further supports this interpretation by revealing directional, pairwise differences in how perturbations propagate through the brain network. Notably, the precuneus and insula emerged as hubs of dysregulated connectivity, echoing their roles in self-referential processing and salience detection. Furthermore, the observed latency differences in prefrontal–parietal interactions may reflect a temporal dysregulation of information integration, potentially underlying the cognitive fragmentation characteristic of the disorder. Importantly, both AUC and latency effects were found to be connection-specific rather than global, meaning that localized perturbations significantly affect the responses of particular region pairs without altering the overall network dynamics.

The MOU model employed in this work provides analytical tractability and captures the stationary structure of empirical BOLD signals. However, it has several limitations. Most notably, it assumes a diagonal diffusion matrix $\mathbf{D}$, thereby neglecting spatial correlations in the noise. This simplification may underestimate interactions among ROIs and misrepresent the true structure of neural variability. Additionally, the assumption of stationarity in BOLD signals remains unverifiable due to experimental constraints.

Moreover, since the MOU is linear, it doesn’t take into account non-Gaussian or nonlinear dynamics, which could play a role in investigating the dynamical brain differences in schizophrenia. Finally, the model does not incorporate an explicit hemodynamic component to account for the transformation from neural activity to BOLD responses [12].

Future studies could take into account the hemodynamic nature of the BOLD activity and use more biophysically based models to connect these whole-brain dynamics observed differences with the biological mechanisms underlying them. Moreover, structural connectivity constraints could be integrated more rigorously, and also time-resolved EPR explored.

Chapter 6

Conclusion

This study shows the utility of tools from non-equilibrium statistical physics in revealing whole-brain dynamical features of schizophrenia. By modeling fMRI data with the MOU process, we identified increased irreversibility and altered linear response properties in patients, supporting the hypothesis that schizophrenia involves a higher degree of non-equilibrium dynamics in neural processes.

The entropy production rate and the irreversibility matrix provide, respectively, global and local metrics that may aid future diagnostic methods. In contrast, the perturbative approach offers insight into how localized perturbations propagate across the brain network, which could inform therapeutic strategies.

These findings underscore the value of physics-inspired tools in clinical neuroscience and pave the way for further investigations into dynamical biomarkers of psychiatric disorders.

Appendix A

Experimental data

The experimental data used in this study were obtained from the publicly available dataset associated with the UCLA Consortium for Neuropsychiatric Phenomics (CNP) LA5c Study [37, 38]. The dataset includes resting-state fMRI (rs-fMRI) scans from both cognitively healthy control participants and individuals diagnosed with schizophrenia (SCZ). Each subject's raw BOLD rs-fMRI data consisted of 152 time frames acquired over 304 seconds. From these volumes, noise-corrected time series of 152 samples were extracted for 68 cortical regions covering both hemispheres, based on the Desikan–Killiany atlas. The final sample comprised 48 patients diagnosed with schizophrenia (12 female, mean age: 36.5 ± 8.8) and 48 healthy controls (16 female, mean age: 36.6 ± 8.9). Diagnostic criteria, behavioral assessments, imaging acquisition parameters, and preprocessing steps are described in detail in the associated publications [23, 37]. The processed data and scripts used for analysis are available at the corresponding GitHub repository [24].

Appendix B

Complete derivation of the fit updates

We compute the updates for $\mathbf{B}$ and $\mathbf{D}$ in order to minimize the loss function in Equation (3.2).

Using Equation (2.7) we find the update for $\Delta\mathbf{D}$.

We start from Sylvester Equation (2.7):

$$\mathbf{B}\mathbf{S}(0) + \mathbf{S}(0)\mathbf{B}^T = 2\mathbf{D}. \quad (\text{B.1})$$

Perturb $\mathbf{B}$, $\mathbf{S}(0)$, and $\mathbf{D}$ by $\Delta\mathbf{B}$, $\Delta\mathbf{S}(0)$, and $\Delta\mathbf{D}$, respectively. The first-order expansion gives:

$$\Delta\mathbf{B}\mathbf{S}(0) + \mathbf{B}\Delta\mathbf{S}(0) + \Delta\mathbf{S}(0)\mathbf{B}^T + \mathbf{S}(0)\Delta\mathbf{B}^T = 2\Delta\mathbf{D}. \quad (\text{B.2})$$

Rearranging for $\Delta\mathbf{D}$:

$$\Delta\mathbf{D} = \frac{1}{2} \left(\Delta\mathbf{B}\mathbf{S}(0) + \mathbf{B}\Delta\mathbf{S}(0) + \Delta\mathbf{S}(0)\mathbf{B}^T + \mathbf{S}(0)\Delta\mathbf{B}^T \right). \quad (\text{B.3})$$

Assuming symmetry and neglecting cross-terms involving $\Delta\mathbf{B}$, we approximate:

$$\Delta\mathbf{D} = \epsilon_D \left(\mathbf{B}\Delta\mathbf{S}(0) + \Delta\mathbf{S}(0)\mathbf{B}^T \right), \quad (\text{B.4})$$

where ϵ_D absorbs constants and acts as the learning rate.

Using Equation (2.10), we find the update for $\Delta\mathbf{B}$. Starting from the time-lagged covariance in Equation (2.10):

$$\mathbf{S}(1) = \mathbf{S}(0)e^{-\mathbf{B}^T}. \quad (\text{B.5})$$

Perturb $\mathbf{B}$ and $\mathbf{S}(0)$, leading to:

$$\Delta\mathbf{S}(1) = \Delta\mathbf{S}(0)e^{-\mathbf{B}^T} - \mathbf{S}(0)e^{-\mathbf{B}^T}\Delta\mathbf{B}^T. \quad (\text{B.6})$$

Rearrange for $\Delta\mathbf{B}$:

$$\mathbf{S}(0)e^{-\mathbf{B}^T}\Delta\mathbf{B}^T = \Delta\mathbf{S}(0)e^{-\mathbf{B}^T} - \Delta\mathbf{S}(1). \quad (\text{B.7})$$

Multiply both sides by $\left(\mathbf{S}(0)e^{-\mathbf{B}^T}\right)^{-1} = e^{\mathbf{B}}\mathbf{S}(0)^{-1}$:

$$\Delta\mathbf{B}^T = e^{\mathbf{B}}\mathbf{S}(0)^{-1} \left(\Delta\mathbf{S}(0)e^{-\mathbf{B}^T} - \Delta\mathbf{S}(1) \right). \quad (\text{B.8})$$

Take the transpose (noting $\mathbf{S}(0)$, $\Delta\mathbf{S}(0)$, and $\Delta\mathbf{S}(1)$ are symmetric):

$$\Delta\mathbf{B} = \epsilon_B \mathbf{S}(0)^{-1} \left(\Delta\mathbf{S}(0) - \Delta\mathbf{S}(1)e^{\mathbf{B}^T} \right), \quad (\text{B.9})$$

where ϵ_B is the learning rate for $\mathbf{B}$.

Combining the results, the final parameter updates are the ones reported in Equations (3.5) (3.4):

$$\Delta\mathbf{B} = \epsilon_B \mathbf{S}(0)^{-1} \left[\Delta\mathbf{S}(0) - \Delta\mathbf{S}(1)e^{\mathbf{B}^T} \right], \quad (\text{B.10})$$

$$\Delta\mathbf{D} = \epsilon_D \left(\mathbf{B}\Delta\mathbf{S}(0) + \Delta\mathbf{S}(0)\mathbf{B}^T \right). \quad (\text{B.11})$$

Bibliography

- [1] LS Pilowsky. "Understanding schizophrenia". In: *BMJ: British Medical Journal* 305.6849 (1992), p. 327.
- [2] Murray R. et al. Kahn R. Sommer I. "Schizophrenia". In: *Nat Rev Dis Primers* 1, 15067 (2015). DOI: <https://doi.org/10.1038/nrdp.2015.67>.
- [3] Akbarian S et al. Tandon R Nasrallah H. "The schizophrenia syndrome, circa 2024: What we know and how that informs its nature." In: *Schizophr Res.* (2024 Feb). DOI: [10.1016/j.schres.2023.11.015](https://doi.org/10.1016/j.schres.2023.11.015).
- [4] Andreas Meyer-Lindenberg. "From maps to mechanisms through neuroimaging of schizophrenia". In: *Nature* 468.7321 (2010), pp. 194–202.
- [5] Arun KM et al. Smitha KA Akhil Raja K. "Resting state fMRI: A review on methods in resting state connectivity analysis and resting state networks." In: *Neuroradiol J.* (2017). DOI: [10.1177/1971400917697342](https://doi.org/10.1177/1971400917697342).
- [6] Martijn P Van Den Heuvel and Alex Fornito. "Brain networks in schizophrenia". In: *Neuropsychology review* 24 (2014), pp. 32–48.
- [7] Julia M Sheffield and Deanna M Barch. "Cognition and resting-state functional connectivity in schizophrenia". In: *Neuroscience & Biobehavioral Reviews* 61 (2016), pp. 108–120.
- [8] Vaughan Carr and Jocelyn Wale. "Schizophrenia: An information processing model". In: *Australian and New Zealand Journal of Psychiatry* 20.2 (1986), pp. 136–155.
- [9] Edoardo G. Spinelli et al. Massimo Filippi. "Resting State Dynamic Functional Connectivity in Neurodegenerative Conditions: A Review of Magnetic Resonance Imaging Findings". In: *Front. Neurosci.* (20 June 2019). DOI: <https://doi.org/10.3389/fnins.2019.00657>.
- [10] Claude Godrèche and Jean-Marc Luck. "Characterising the nonequilibrium stationary states of Ornstein-Uhlenbeck processes". In: *Journal of Physics A: Mathematical and Theoretical* 52.3 (Jan. 2019). arXiv:1807.00694 [cond-mat], p. 035002. ISSN: 1751-8113, 1751-8121. DOI: [10.1088/1751-8121/aaf190](https://doi.org/10.1088/1751-8121/aaf190). URL: <http://arxiv.org/abs/1807.00694> (visited on 05/05/2025).
- [11] Matthieu Gilson et al. "Estimation of Directed Effective Connectivity from fMRI Functional Connectivity Hints at Asymmetries of Cortical Connectome". In: *PLoS Computational Biology* 12.3 (2016), e1004762. DOI: [10.1371/journal.pcbi.1004762](https://doi.org/10.1371/journal.pcbi.1004762). URL: <https://doi.org/10.1371/journal.pcbi.1004762>.
- [12] Pallarés V et al. Gilson M Zamora-López G. "Model-based whole-brain effective connectivity to study distributed cognition in health and disease." In: *Netw Neurosci.* (2020 Apr 1). DOI: [doi:10.1162/netn_a_00117](https://doi.org/10.1162/netn_a_00117).
- [13] Matthieu Gilson et al. "Entropy production of multivariate Ornstein-Uhlenbeck processes correlates with consciousness levels in the human brain". eng. In: *Physical Review E* 107.2-1 (Feb. 2023), p. 024121. ISSN: 2470-0053. DOI: [10.1103/PhysRevE.107.024121](https://doi.org/10.1103/PhysRevE.107.024121).
- [14] Gorka Zamora-López and Matthieu Gilson. "An integrative dynamical perspective for graph theory and the analysis of complex networks". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 34.4 (Apr. 2024). ISSN: 1089-7682. DOI: [10.1063/5.0202241](https://doi.org/10.1063/5.0202241). URL: <http://dx.doi.org/10.1063/5.0202241>.

- [15] Gilson M et al. Panda R López-González A. “Whole-brain analyses indicate the impairment of posterior integration and thalamo-frontotemporal broadcasting in disorders of consciousness.” In: *Hum Brain Mapp* (2023 Aug 1).
- [16] G. E. Uhlenbeck and L. S. Ornstein. “On the Theory of the Brownian Motion”. In: *Phys. Rev.* 36 (5 Sept. 1930), pp. 823–841. DOI: [10.1103/PhysRev.36.823](https://link.aps.org/doi/10.1103/PhysRev.36.823). URL: <https://link.aps.org/doi/10.1103/PhysRev.36.823>.
- [17] Ming Chen Wang and G. E. Uhlenbeck. “On the Theory of the Brownian Motion II”. In: *Rev. Mod. Phys.* 17 (2-3 Apr. 1945), pp. 323–342. DOI: [10.1103/RevModPhys.17.323](https://link.aps.org/doi/10.1103/RevModPhys.17.323). URL: <https://link.aps.org/doi/10.1103/RevModPhys.17.323>.
- [18] Gardiner C.W. *Handbook of Stochastic Methods for Physics, Chemistry and Natural Sciences*. Springer, 2004.
- [19] Melvin Lax. “Fluctuations from the Nonequilibrium Steady State”. In: *Rev. Mod. Phys.* 32 (1 Jan. 1960), pp. 25–64. DOI: [10.1103/RevModPhys.32.25](https://link.aps.org/doi/10.1103/RevModPhys.32.25). URL: <https://link.aps.org/doi/10.1103/RevModPhys.32.25>.
- [20] Prigogine I. *Introduction to Thermodynamics of Irreversible Processes*. New York: Wiley, 1961.
- [21] S. R. De Groot and P. Mazur. *Non-Equilibrium Thermodynamics*. Amsterdam: North Holland, 1962.
- [22] Gong G. L. Qian M. P. Qian M. “The Reversibility and The Entropy Production of Markov Processes”. In: *Contemporary Mathematics* 118 225 (1991).
- [23] Parkes L et al. Bryant AG Aquino K. “Extracting interpretable signatures of whole-brain dynamics through systematic comparison”. In: *PLoS Comput Biol.* (2024).
- [24] Dynamics and Neural Systems. *fMRI Features Disorders*. https://github.com/DynamicsAndNeuralSystems/fMRI_FeaturesDisorders. 2024.
- [25] Fischl B et al. Desikan RS Ségonne F. “An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest.” In: *NeuroImage* 31.3 (2006). DOI: [10.1016/j.neuroimage.2006.01.021](https://doi.org/10.1016/j.neuroimage.2006.01.021).
- [26] Stephen M. Smith et al. David C. Van Essen. “The WU-Minn Human Connectome Project: An overview”. In: *NeuroImage* 80 (2013), pp. 62–79. DOI: [10.1016/j.neuroimage.2013.05.041](https://doi.org/10.1016/j.neuroimage.2013.05.041).
- [27] Wilson JA et al. Glasser MF Sotiropoulos SN. “The minimal preprocessing pipelines for the Human Connectome Project.” In: *NeuroImage* 80 (2013). DOI: [10.1016/j.neuroimage.2013.04.127](https://doi.org/10.1016/j.neuroimage.2013.04.127).
- [28] Park By. et al. Larivière S. Paquola C. “The ENIGMA Toolbox: multiscale neural contextualization of multisite neuroimaging datasets.” In: *Nature Methods* 18.7 (2021). DOI: [10.1038/s41592-021-01186-4](https://doi.org/10.1038/s41592-021-01186-4).
- [29] Timme Marc and Casadiego Jose. “Revealing networks from dynamics: an introduction”. In: *Journal of Physics A: Mathematical and Theoretical* 47.34 (Aug. 2014), p. 343001. DOI: [10.1088/1751-8113/47/34/343001](https://doi.org/10.1088/1751-8113/47/34/343001). URL: <https://dx.doi.org/10.1088/1751-8113/47/34/343001>.
- [30] Rosanova M et al. Casali AG Gosseries O. “A theoretically based index of consciousness independent of sensory processing and behavior.” In: *Sci Transl Med.* (2013 Aug 14). DOI: [10.1126/scitranslmed.3006294](https://doi.org/10.1126/scitranslmed.3006294).
- [31] Sebastian M Geli et al. “Non-equilibrium whole-brain dynamics arise from pairwise interactions”. In: *Cell Reports Physical Science* 6.3 (2025).
- [32] Genevieve J Yang et al. “Altered global brain signal in schizophrenia”. In: *Proceedings of the National Academy of Sciences* 111.20 (2014), pp. 7438–7443.
- [33] Monte S Buchsbaum. “The frontal lobes, basal ganglia, and temporal lobes as sites for schizophrenia”. In: *Schizophrenia Bulletin* 16.3 (1990), pp. 379–389.
- [34] Motoaki Nakamura et al. “Orbitofrontal volume deficit in schizophrenia and thought disorder”. In: *Brain* 131.1 (2008), pp. 180–195.

- [35] Yuka Yasuda et al. "Brain morphological and functional features in cognitive subgroups of schizophrenia". In: *Psychiatry and clinical neurosciences* 74.3 (2020), pp. 191–203.
- [36] Friston KJ. Stephan KE Baldeweg T. "Synaptic plasticity and dysconnection in schizophrenia." In: *Biol Psychiatry*. (2006 May 15). DOI: [10.1016/j.biopsych.2005.10.005](https://doi.org/10.1016/j.biopsych.2005.10.005).
- [37] Triplett W. et al. Poldrack R. Congdon E. "A phenome-wide examination of neural and cognitive function." In: *Sci Data* 3 (2016). DOI: <https://doi.org/10.1038/sdata.2016.110>.
- [38] Cannon T et al. Bilder R Poldrack R. *UCLA Consortium for Neuropsychiatric Phenomics LA5c Study*. <https://openneuro.org/datasets/ds000030/versions/00016>. 2020.