UNIVERSIDAD
NACIONAL
DE COLOMBIA

Revealing brain network dynamics during the emotional state of suspense using topological data analysis

Astrid Arena Olave Herrera

Universidad Nacional de Colombia
Facultad de Ciencias
Departmento de Matemáticas
Bogotá D.C, Colombia
2021

Descubriendo las dinámicas de las redes cerebrales durante el estado emocional de suspenso usando análisis topológico de datos

Astrid Arena Olave Herrera

Tesis presentada como requisito parcial para optar al título de:
Magister en Ciencias - Matemática Aplicada

Director:
José Perea

Co-director:
Francisco Albeiro Gómez Jaramillo

Línea de investigación:
Matemáticas Aplicadas

Universidad Nacional de Colombia
Facultad de Ciencias
Departmento de Matemáticas
Bogotá D.C, Colombia
2021

Las matemáticas son el lenguaje en el que Dios escribió el universo

Acknowledgements

I would like to thank the following people, who support and encourage me through my research and my master degree:

Firstly, I would like to thank my supervisors, Dr. Francisco Gomez and Dr. Jose Perea for their guidance and support during this research, their words and experience have been invaluable throughout this study. I am extremely grateful for their encourage to pursue a Ph.D. degree. I would also like to thank Dr. Ralf Schmälzle for providing invaluable data to complete this research and Professor Gabriel Castellanos for his insights in neuroscience and make possible the bridge between his speciality and mathematics.

I am also thankful to La Universidad Nacional de Colombia, my Alma Mater, for opening the gates of knowledge. In particular, my partners in the COMBIOS group for their advice and comments on this research and my professors to encourage me in this process.

Additionally, I would like to thank my soccer team mates, for sharing with me moments to rest my mind outside my research. Also, I am greatly grateful to my friends in Sin Medida, especially Juan Camilo Lopez and Diana Toquica, for supporting me and challenging me to be a better person every day.

To conclude, my biggest thanks to you, Mom, all this work would not have been possible without you. Thanks for your support, encourage and always be there for me.

Abstract

Suspense is an affective state ubiquitous in human life, from art to quotidian events. However, little is known about the behavior of large-scale networks during suspenseful experiences. To address this question, we examined the continuous brain responses of participants watching a suspenseful movie along with a reported level of suspense from viewers. We employed sliding window analysis and Pearson correlation to measure functional connectivity states along time. Then, we used Mapper, a tool of Topological Data Analysis, to obtain a graphical representation capturing the brain's dynamical transitions across states. Our analysis revealed changes in the functional connectivity within and between Salience, Fronto-Parietal, and Default networks associated with suspense. In particular, the functional connectivity between Salience and Fronto-Parietal networks increased with the level of suspense. In contrast, the connections of both networks with the Default network decreased. Together, our findings expose the dynamical changes of functional connectivity at the network level associated with the variation of suspense and reveal topological analysis as a potentially powerful tool for studying dynamic brain networks.

Keywords: Suspense, fMRI, Dynamic functional connectivity, Topological data analysis, Mapper

Resumen

El suspenso es un estado emocional omnipresente en la vida humana, desde el arte hasta los eventos cotidianos. Sin embargo, se sabe poco sobre el comportamiento de las redes cerebrales a gran escala durante las experiencias de suspenso. Para abordar esta pregunta, examinamos continuamente las respuestas cerebrales de participantes que ven una película de suspenso junto a un reporte de los espectadores ds su nivel de suspenso. Empleamos el análisis de ventana deslizante y el índice de correlación de Pearson para medir los estados de conectividad funcional a lo largo del tiempo. Luego, usamos Mapper, una herramienta del análisis topologico de datos, para obtener una representación gráfica que captura las transiciones dinámicas del cerebro a través de los estados. Nuestro análisis reveló cambios en la conectividad funcional dentro y entre las redes saliente, fronto-parietal y por defecto asociadas con el suspenso. En particular, la conectividad funcional entre las redes saliente y fronto-parietal aumentó con el nivel de suspenso. Por el contrario, las conexiones de ambas redes con la red por defecto disminuyeron. Nuestros resultados muestran los cambios dinámicos de la conectividad funcional a nivel de red asociados con la variacion de suspenso y revelan al análisis topológico de datos como una herramienta potencialmente poderosa para estudiar la redes dinámicas del cerebro.

Palabras clave: Suspenso, fMRI, Conectividad funcional dinámica, Análisis topológico de datos, Mapper.

Contents

List of Figures

List of Tables

List of Abbreviations

AAL Automated anatomical labeling atlas. 19, 20

BNST bed nucleus of the stria terminalis. 10, 50

BOLD blood-oxygen level dependent. 7, 22

Cam-CAN Cambridge Center for Aging and Neuroscience. 17, 61

dFC dynamic functional connectivity. 2, 10, 11, 22, 23, 33, 50

FC functional connectivity. xv, 1, 3, 6, 22, 23, 30–32, 36, 37, 49, 50

fMRI functional magnetic resonance imaging. 7, 10, 11, 17, 18, 61

Isomap Isometric mapping. xv, xvi, 26, 30, 38, 39, 64, 69

LSN large-scale networks. 2–4, 6, 7, 11, 19, 31, 51

MDS Multidimensional Scaling. 27

MNI Montreal Neurological Institute. 19

MRI magnetic resonance imaging. 9, 17

PAG periaqueductal gray. 10, 50

ROIs Regions of interest. 3, 19, 33

sFC static functional connectivity. 2, 32

STM State Transition Matrix. xv, xvi, 30, 31, 41, 42, 45

SWM sliding window method. xv, 3, 22, 23, 25, 35, 36, 45, 50

TDA Topological Data Analysis. 2, 3

TR Repetition time. 19, 23, 49

WTC wavelet transform coherence. 50

1

Introduction

Researchers have been interested in the brain mechanisms underlying human emotional response since the early days of psychological science. Even though this question remained ignored beginning the second half of last century [3] in the last few decades neuroscience has again embraced emotion as a relevant research topic [4, 5].

One of the most powerful emotions that permeate many aspects of our life, from quotidian events to many leisure activities like reading or watching a film is *suspense*. Suspense is an affective state associated with conflict, dissonance, instability, or uncertainty regarding an emotionally significant event that on some level is not susceptible to influence or control which motivates future-oriented expectation or prediction and desire for a resolution [1].

Suspense is relevant in areas like media [6, 7], arts, [8, 9], sociology [10] and psychology [11, 12], but despite this and the increasing attention of neuroscience on emotions, to the best of our knowledge the number of studies in this field related to suspense is still small, [1, 2, 13–19], therefore the neuronal and psychological mechanisms intrinsic to suspenseful experiences are not yet fully understood, nonetheless, the path of understanding is set: "characterizing circuit interactions is believed to be key to unravelling how emotion is organized in the brain" [20].

We can understand the brain as a complex system or network, in which mental states emerge from the interaction between multiple physical and functional levels [21]. One of the most fruitful approaches to brain architecture is in terms of functional networks. Functional networks are defined as graphs whose nodes are neural elements of the brain (i.e., voxels, regions) and the edges are defined by mean of functional connectivity (FC); a measure of the statistical dependence between the neural activity of nodes [22]. However

> Most network-based neuroimaging studies utilize a *static brain network* representation which constructs a network using data from an entire scan session. In essence, these networks summarize the strength of functional connectivity between pairs of brain regions throughout a scan session. However, many changes in the brain occur at shorter time scales on the order of milliseconds

> (for neuronal activity) or seconds (for cerebral blood flow). Static network analyses are agnostic to these changes occurring at shorter time scales; however, recent interest in how networks change has led to the development of methods to examine dynamics in functional connectivity more generally [23].

Therefore, the use of *dynamic brain networks* arise to enhance the understanding time-evolving neurophysiological processes, so-called dynamic functional connectivity (dFC) in contrast to the static functional connectivity (sFC) that does not consider temporal factors. Indeed, recent research on the affective brain have demonstrated that a variety of emotional experiences are associated with dynamic interactions of large-scale networks (LSN) [24–34]. However, to our knowledge, there are not studies in suspense experiences yet.

In the last few years, neuroscientists have started to use Topological Data Analysis (TDA) for studying static brain networks [35–37] and dynamic ones [38, 39]. TDA is a recent and fast-growing field that emerged from various works in applied (algebraic) topology and computational geometry. It is mainly motivated by the idea that topology and geometry provide a powerful approach to infer robust qualitative, and sometimes quantitative information about the structure of data [40].

In particular, the topological-based method called `Mapper` [41] was used to reveal the dynamical organization of the brain associated with cognitive tasks performed over time [39]. In this work, `Mapper` proved to be a novel method that represents the brain's overall dynamical as a combinatorial object capable of (1) providing insights about how the brain dynamically adapts in a multitask paradigm (2) tracking transitions at a much faster time scale than before and (3) finding neural markers for individual differences in task performance. By contrast, little is known about the brain's temporal organization during suspense experiences.

1.1 Contributions

This study explores, for the first time, the phenomena of suspense from a network-based perspective. We hypothesize that variations in the intensity of suspense are associated with changes in the functional connectivity within and between large-scale networks.

The results of this study will contribute to understand the brain's mechanisms underlying the affective state of suspense. In addition, the results would support the approach of conceiving emotional processing at the level of dynamic interactions between large-scale

networks.

Furthermore, this work contributes to establishing the growing field of Topological Data Analysis as a source of novel and potentially powerful methods to study the brain's connectome.

1.2 Structure of thesis

The thesis is organized as follows:

Chapter 2 We define psychologically and neurologically the emotional state of suspense. Besides, we present the neural models and the dynamic interactions of LSN associated with emotions, and finally, we give the topological and statistical definition of `Mapper`.

Chapter 3 We describe our data including the selection of the Regions of interest (ROIs), we explain the sliding window method (SWM) to form the metric space and each step of `Mapper`'s algorithm. Additionally, we expound the analysis of the output graph.

Chapter 4 We present the construction of the `Mapper` graph and the results of its analysis.

Chapter 5 We characterize the graph result of `Mapper` and we identify the functional connectivity patterns associated with the suspense responsible for the graph's features. In addition, we indicate the limitations of the study and and ideas for future researches.

Chapter 6 We give the conclusions of the thesis.

2

Background

To study suspenseful experiences from a dynamic network viewpoint, we first introduce in Section 2.1 a psychological model of tension and suspense by Lehne and Koelsch. In the following Section 2.2, we present two current neurological models that associates the processes of emotions with the interactions of large-scale networks, specially, the models arise as plausible neurological explanations for the model of Lehne and Koelsch. Next, in Section 2.3 we link the LSN to suspense and related emotions to it. Finally, in Section 2.4 we introduce `Mapper` as an advantageous tool to study dynamic brain networks and show its definition.

2.1 A model of tension and suspense

Experiences of tension and suspense permeate many aspects of life, from quotidian events to many leisure activities such as reading or watching a movie. The ubiquitousness of tension and suspense suggests that they build on very basic cognitive and affective mechanisms. In fact, Lehne and Koelsch [1] argue that even if suspenseful experiences are found in different contexts, they are built on the same underlying psychological processes and consequently they have proposed a domain-independent model of tension and suspense. The model is shown in Fig. 2.1 below.

> According to this model, experiences of tension originate from the perception of an initiating event that is associated with *conflict, instability, dissonance,or uncertainty* which triggers future-directed processes of *prediction, expectation and anticipation* (modulated by previous knowledge, situational factors, or personality).These predictive processes create a space of possible outcome events (note that these anticipated outcome events can be conscious or unconscious, and more or less specific). A divergence between the affective values of anticipated events (i.e., their desirability) then results in an experience of tension [1].

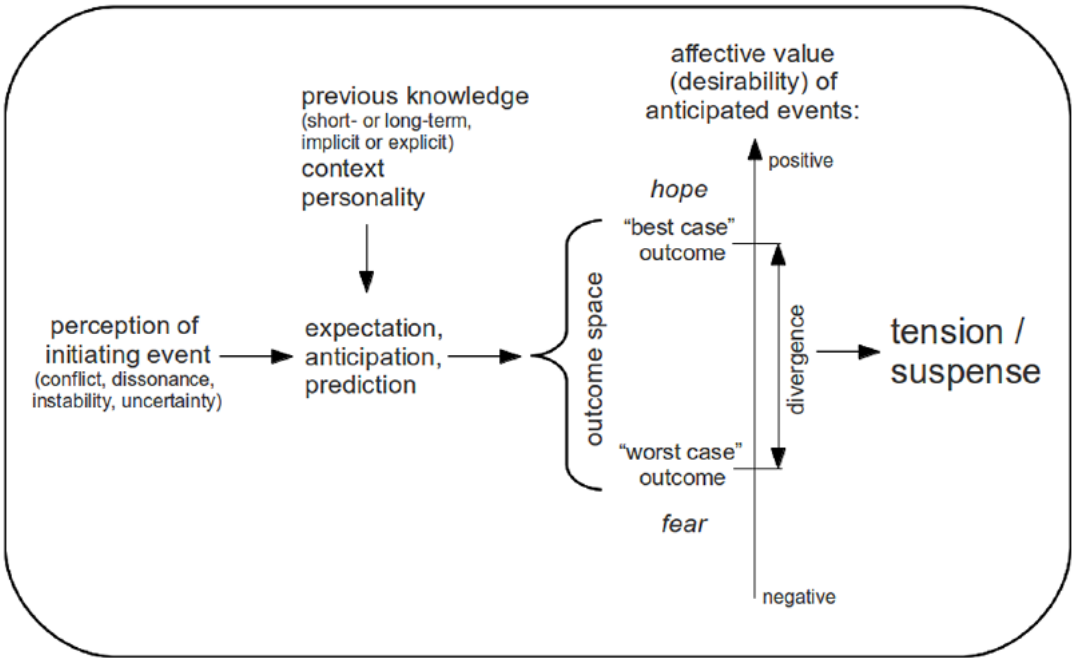


Figure 2.1: Model of tension and suspense proposed by Lehne and Koelsch. An initiating event associated with conflict, dissonance, instability, or uncertainty triggers processes of expectation, anticipation, and prediction that depend on previous knowledge, context, and personality factors. The future-directed processes generate a space of anticipated outcome events that vary with regard to their affective values/desirability: a positive outcome is associated with hope and negative outcomes with fear. The divergence between the opposite outcomes leads to the subjective experience of tension. (Reprinted from [1])

The key components of this model are:

Conflict, dissonance and instability Experiences with this characteristic, create a yearning for more stable or consonant states.

Uncertainty Uncertainty is an essential component underlying tension experience (although its exact role is a matter of debate, see the paradox of suspense [10]). Usually it takes the form of an implicit or explicit question (e.g., the classic "Whodunit?" in a detective story), triggering an experience of tension that resolves when an answer to the question is provided.

Expectation, prediction and anticipation Events that unfold in time are constantly evaluated against a background of predictions that are continuously updated during their temporal evolution. The anticipated events with positive valence elicit emotions

of hope, whereas negative events create fear (or anxiety if the anticipated events are more diffuse and not clearly specified) and both states can co-occur causing both positive or negative outcomes are possible.

Emotional significance of anticipated events The anticipated events need to have some emotional significance (e.g., a medical diagnosis, job interview, etc.) to generate tension or suspense. Whereas this is relatively obvious for real life experiences, it is not the case for the tension experiences created by many forms of media entertainment such as music or movies. Although, for instance, in narrative plots, this can be explained by processes of identification and empathy with the characters of the plot, making them on some level, personally significant.

Lack of control During a tension event there is an inability to influence the course of the situation. This sense of lack of control may generate a feeling of helplessness that can add to the experience of tension. This is particular clear in suspense movies when viewers are aware of an inevitable danger for the protagonist but are unable to warn him, then adding more tension to the viewer.

Nonetheless, the model proposed by Lehne and Koelsch is a psychological model of tension and suspense, as we are interested in the neural mechanisms underlying suspenseful experiences, in the following section we present some neurological models of emotions.

2.2 A network-based approach of emotions

The current hypothesis on the emotional brain asserts that instances of any emotion category are not expected to be specifically related to increased activation in any single brain region or set of regions, indeed, brain regions traditionally associated with emotions (e.g. amygdala, insula, hypothalamus among others) show consistent activation in other mental phenomena (e.g. memory, prospection, empathy among others) [42]. Instead, emotions are the result of the dynamic interaction of the LSN, influencing and shaping one another in real time according to the principles of constraint satisfaction [5].

Large-scale networks

Large-scale networks are defined as a collection of interconnected brain areas that interact to perform circumscribed functions [43]. They have been successfully delineated in resting-state FC. This approach examines synchronized patterns of spontaneous oscillations in

blood-oxygen level dependent (BOLD) signal measured at rest with functional magnetic resonance imaging (fMRI) [44].

In the present investigation, we consider three of the six LSN defined in [45]. The regions considered for each network are presented in section 3.2:

Salience network This network is involved in the identification of important or salient information and may include internally generated (i.e. remembered) information. At the same time, it has been suggested that this network mediates the connection between Default and Fronto-Parietal networks. In [44], it is part of a bigger network, the *Midcingulo-Insular* network which also comprehends the Ventral-Attention and Dorsal-Attention networks.

Fronto-Parietal Task Control Network Also known as *central executive*, *executive control* or *Lateral Frontoparietal* network. The functions of this system include executive functions, such as goal-oriented cognition, working memory, inhibition and task switching.

Default Network Also known *task-negative* or *Medial Frontoparietal* network. Even though there is not consensus in its primary functions, it is likely to be involved in the formation, temporal binding, and dynamic reconfiguration of associative representations based on current goal states, the detection of the associative relevance of internal and external stimuli, providing value coding, semantic associations, monitoring the environment, imagination, future-thinking and contextual associative processing.

The network-based approach of emotions is asserted by two different models.

A constructionist model

The psychological constructionist approach of emotion, proposed by Lindquist, Barret et al. [5, 42, 46], assumes that emotions are psychological events that emerge out of more basic psychological operations (psychological primitives) that are not specific to emotion. They suggest a set of four *basic psychological operations* that are a first approximation to psychological primitives:

Core affect Is the representation of sensations from inside the body that can be experienced as a bodily symptom or as feelings of pleasure/displeasure with some degree of

arousal. It is associated with limbic and paralimbic brain regions within the Salience network.

Exteroceptive sensation Is the representation of sensations from outside the body (e.g., visual, auditory, tactile, and olfactory sensations). It is associated with modal and hetero-modal sensory cortices.

Conceptualization This is an automatic and effortless process which links perceptions of sensory input from the world with input from the body to create a meaningful psychological moment. Is is believed to be associated with the midline cortical, lateral, prefrontal, and temporal regions within the Default network.

Emotion words In their view, emotion categories are abstract categories that are socially constructed and humans use words as the glue that holds the category together, thus, emotion words anchor emotion categories which work in hand with conceptualization. It is related to language-relevant brain regions.

Executive attention This operation helps to determine which conceptual representations are utilized to make meaning out of that state, and which are suppressed. It is supported by the Fronto-Parietal network

In this way, to construct emotional experiences: sensations and feelings from the body (core affect) and from the world (exteroceptive sensation) are made meaningful by past experiences, including knowledge about the emotion categories encoded in language (conceptualization and emotion words). This process is regulated by executive control.

Particularly, the model of tension of suspense of Lehne and Koelsch is compatible with the constructionist model: the conflict, dissonance and instability of situations (exteroceptive sensation) combined with the the sensation of uncertainty and expectation (core affect) that are made meaningful by its emotional significance and the lack of control generate the affective state of suspense.

A dynamic emergent processes model

On the other hand, Pessoa [20, 47, 48] proposes the following model:

> The brain basis of emotion involves large-scale cortical/subcortical networks that are distributed and sensitive to bodily signals. The high degree of signal distribution and integration provides a nexus for the intermixing of information

> related to perception, cognition, emotion, motivation, and action. Importantly, the functional architecture consists of multiple overlapping networks that are highly dynamic and context sensitive. Thus, how a given brain region affiliates with a specific network shifts as a function of task demands and brain state [20].

As Pessoa explains, there are important differences between the two models: In the constructionist model, basic operations map onto distributed networks, notably those characterized during resting-state functional magnetic resonance imaging. In the present model, there are no domain-general basic operations. The mind-brain is not built from a set of finite primitives, but from dynamic emergent processes. Here, the proposed functionally integrated systems are flexible and dynamic, thus highly context dependent, for example, the cortex-amygdala system, it does not have a core function like "affect generation", instead, its particular functional state determines how it will contribute to multiple mental operations, which involve not only arousal, vigilance, and novelty, but also attention, value determination, and decision making more broadly.

The Pessoa's model emphasize the inherent dynamic nature of connections within and between networks in function of emotional context. Furthermore, the brain regions involved in emotional processes may in principle, remain the same trough diverse emotions, and are the temporal properties of the connections and syncronization between regions that would determine the uniqueness of each emotion [4]. Thus, we could expect to find evidence of interactions related to several emotions as we present in next section.

2.3 Large scale networks in suspense

2.3.1 Dynamic brain networks of related emotions

Various researches that have found evidence of large-scale networks interactions related to several emotions like anger, sadness, fear, disgust and happiness [24–34]. Especially, we find guiding results of mental states related to suspense like anxiety, anticipation and stress.

For example, Hermans et. al. [26] discovered that acute stress prompts a reallocation of resources to the Salience network, promoting fear and vigilance, at the cost of the Fronto-Parietal network. After stress subsides, resource allocation to these two networks reverses, which normalizes emotional reactivity and enhances higher-order cognitive processes important for long-term survival, demonstrating the dynamical changes of network organization.

Furthermore, Najafi, Kinnision and Pessoa in [33] analyzed the dFC within and between the Default, Fronto-Parietal and Salience and subcortical regions during anxious anticipation. To stimulate periods of anxious trough a visual stimuli: two dots slowly meandered on a screen randomly, when they collide and unpleasant mild electric shock was delivered. In this way the scanning time was divided into periods of approach (dots moving closer) and periods of retreat (dots moving apart). During periods of retreat the Salience network exhibited the strongest chances during approach and/or retreat periods, the positive functional connections increased during approach, and decreased during retreat. In the same way, positive dFC between the Salience and Fronto-Parietal networks increased during approach and decreased during retreat. However, the opposite was found between the Salience and the Default networks, it decreases during approach and increases during retreat. At the same time, the negative functional connections between the Salience and the Default networks displayed an opposite pattern, thus, an increase during approach periods (and a decrease during retreat).

On the other hand, the dFC between the Salience network and subcortical regions like amygdala, periaqueductal gray (PAG) , and the bed nucleus of the stria terminalis (BNST) increased during approach and decreased during retreat, also the BNST became more central for people with greater anxiety [27]. This reveals how network organization unfolds with time during periods of anxious anticipation.

2.3.2 Brain structures related to suspense

In the last few years, researchers have identified some brain structures involved in the experience of suspense through different experiments. It is worthy to highlight the variety of sources to provoke suspense in each investigation and how the found brain activity is consistent with the network-based approach.

- A functional magnetic resonance imaging (fMRI) study [15], investigating the relation of neural activity with tension experiences during **music listening** revealed that tension is associated to neuronal activity in the pars orbitalis of the inferior frontal gyrus and the amygdala and pointed to a possible functional role of the orbitofrontal cortex and amygdala in linking processes of expectancy and prediction to affective experience. These regions are linked to the *Salience* and the *Default* networks.
- Another fMRI investigation [16] , this time, exploring the neural correlates of suspense during the **reading of a literary text**, found out suspense is related to activity in the medial frontal cortex, posterior temporal and temporo-parietal regions, as well

as the dorsolateral prefrontal cortex along the inferior frontal sulcus including the inferior frontal gyrus and premotor cortex. These regions are associated with the *Fronto-Parietal* and the *Default* network indicating that the emotional experience of suspense depends on brain areas associated with mentalizing, social cognition, predictive inference and possibly cognitive control.

- Recently, using functional magnetic resonance imaging scans of people **watching a suspenseful movie** [2] is obtained a highest correlation between suspense ratings and the activation of the midcingulate gyrus, the angular gyrus and the right lateral prefrontal cortex regions and moderately high correlation with the anterior cingulate, the bilateral anterior insula, bilateral frontal cortices, the precuneus, medial prefrontal cortex, and the bilateral angular gyri. These regions are prominent parts of the *Salience* and the *Fronto-Parietal* networks, hence, being linked to the integration of sensory, emotional, and cognitive information, and particularly to processes related to anxious apprehension and anticipatory anxiety.

2.4 The Mapper approach to dynamic brain networks

From this perspective, it is clear that the brain is an inherently dynamic system, LSN dynamically evolve and reconfigure as a function of context. Not only the research on the emotional brain moved towards a dynamic-network based approach, the research on cognition, behavior, and consciousness took the same path to uncover the brain network reconfiguration over time. [23, 39, 49].

Many methodologies have been developed to investigate dFC [50]. In particular, as mentioned in the introduction, Saggar et al. [39] analyzed brain dynamical organization associated with ongoing cognition using `Mapper`, a topological based method for data analysis introduced in 2007 [41].

With the intention of testing the efficacy of this methodology in estimating a representation of the brain's dynamical organization and capturing transitions in the whole brain activity, they employed an fMRI dataset with known ground truth about the timing of transitions between mental states as dictated by tasks (working memory, arithmetic operations, and a visuospatial search task) along with blocks of rest.

As a result, the output given by `Mapper` "can be conceptualized as a low-dimensional depiction of how the brain dynamically evolved across different functional configurations during the scan" [39]. This approach tracks both within- and between-task transitions at a much faster time scale (4-9 s) than before. In addition, the `Mapper`'s output of all

participants indicates there is more consistency in the whole-brain functional configurations during the math and memory task compared to the less demanding resting state, also, participants with specialized whole-brain configurations for different tasks were those with the highest overall task performance. Furthermore, it is possible to identify the underlying patterns of brain activity putatively responsible for the observed topological features in the output graph.

In contrast to other techniques, this method does not arbitrarily collapse data in space or time and has the ability to solve fundamental issues like:

- Uncovering the temporal and spatial scales that best capture clinically and behaviorally relevant brain dynamics.
- Understanding whether the dynamical landscape of possible configurations is best conceptualized as continuous or discrete.
- Recognizing what constitutes healthy and aberrant dynamics.

Altogether, `Mapper` is "a novel method to distill the complex brain dynamics associated with ongoing cognition into a set of interactive and behaviorally relevant representations by taking full advantage of the original temporal and spatial scales of the data" [39].

This result derives from the fact that `Mapper` "can be used to reduce high dimensional data sets into simplicial complexes with far fewer points which can capture topological and geometric information at a specified resolution". In general, `Mapper` has certain advantages over other methods of data analysis such that:

- It can be applied to any data embedded into a metric space non-necessarily Euclidean. It is especially applicable when metrics are not derived from traditional ones.
- The method has the ability to capture details even in a large data set.
- This approach is invariant under small deformations, thus, is more robust to noise than other methods.
- The output can have various scales of resolution, which is useful in distinguishing between real features and artifacts, since features will persist through different levels of resolution.
- The output as a combinatorial object is helpful to distinguish the relations between the data, e.g., similarity or periodicity, and is easier to manipulate computationally.

- Its properties of coordinate and deformation invariance make it suitable for examining data across participants and projects.

Therefore, `Mapper` has also been applied to various problems such as the characterization of states in biomolecular folding pathways [51, 52], mutational profiles in breast-cancer patients [53], discovering cord and brain injuries [54] and the distribution of malaria in Colombia [55].

We describe here the `Mapper` method starting from its mathematical definition and following its application to data.

2.4.1 The Topological Mapper

Definition 2.1. An *abstract simplicial complex* $\mathcal{A}$ is a finite collection of sets such that if $\alpha \in \mathcal{A}$ and $\beta \subset \alpha$ implies $\beta \in \mathcal{A}$.

We call the sets in $\mathcal{A}$ *simplices* and they are finite. The *dimension* of a simplex α is dim $\alpha = |\alpha| - 1$ where $|\alpha|$ is the cardinal of the set α. The simplex α it is called a *k-simplex* if its dimension is k. We define the dimension of the complex as the maximum dimension of any of its simplices. The *vertex set* is the union of all simplices, $V(\mathcal{A}) = \bigcup_{\alpha\in\mathcal{A}} \alpha$. [56].

Definition 2.2. A *cover* of a set X is a collection of open sets $\mathcal{U} = \{U_\alpha\}_{\alpha\in A}$ such that $X \subset \bigcup_{i\in I} U_i$

Let X be a topological space and $\mathcal{U} = \{U_\alpha\}_{\alpha\in A}$ a finite cover of X.

Definition 2.3. The *nerve of the cover* $\mathcal{U}$ is the simplicial complex $N(\mathcal{U})$ whose vertex set is the index set A, and where any subset $\{\alpha_0, ..., \alpha_k\} \subset A$ spans a k-simplex in $N(\mathcal{U})$ if and only if $U_{\alpha_0} \cap ... \cap U_{\alpha_k} \neq \emptyset$

Suppose there is a continuous map $f : X \to Z$. We called f a filter function and Z a parameter or filter space. Let Z be equipped with a covering $\mathcal{U} = \{U_\alpha\}_{\alpha\in A}$, again for some finite indexing set A. Since f is continuous, the sets $f^{-1}(U_\alpha)$ also form an open covering of X.

For each α, $f^{-1}(U_\alpha)$ is decomposed in its path connected components, so we write $f^{-1}(U_\alpha) = \bigcup_{i=1}^{c_\alpha} V_\alpha(i)$, where c_α is the number of connected components in $f^{-1}(U_\alpha)$. As $f^{-1}(U_\alpha)$ is a cover of X, the union of all connected components is too. We note as $f_X^*(\mathcal{U})$ the covering of X obtained this way.

Definition 2.4. The *Mapper* construction $M_X(\mathcal{U}, f)$ of X using f and $\mathcal{U}$ is the simplicial complex defined as

$$M_X(\mathcal{U}, f) := N(f_X^*(\mathcal{U}))$$

According to the nerve theorem for `Mapper` [57] for suitable open covers, the space X is homotopy equivalent to Mapper construction of X. Therefore, we could expect to recover meaningful geometrical and topological information of X from the output of Mapper.

2.4.2 The Statistical Mapper

It is analogous to the topological `Mapper` but is developed precisely for point cloud data. "The main idea in passing from the topological version to the statistical version is that *clustering* should be regarded as the statistical version of the geometric notion of partitioning a space into its connected components" [41]. Consider a finite point cloud data. We can split the construction of `Mapper` in the following steps:

1. Choose a meaningful distance for the point cloud data. In this way, we form a finite metric space X to represent the cloud data.
2. Define a filter function $f : X \to Z$ to some metric space Z.
3. Select a finite covering $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ of Z.
4. Decompose each $f^{-1}(U_\alpha)$ into clusters using a clustering algorithm $\mathcal{C}$, we note as $V_\alpha(i)$ the i-th cluster of $f^{-1}(U_\alpha)$ and c_α a index set of the clusters. Note that $f_X^*(\mathcal{U}) = \{V_\alpha(i)\}_{\alpha \in A, i \in c_\alpha}$ is a finite cover of X.
5. Construct the nerve $N(f_X^*(\mathcal{U}))$.

As a remark, in the literature, the first step is not part of the `Mapper` construction, nevertheless, we consider it is a fundamental step when working with data.

As a result, we have a simplicial complex, theoretically speaking of a certain dimension k depending on the intersection of the open sets, however, in practice only the 0- and 1-simplices obtained in the nerve are kept. In this way, we form the 1-nerve complex of Mapper which it is structured like a graph. This graph is subjected to further investigation depending on the context of the analysis.

It is worthy to note the `Mapper` output is highly dependent on the filter function and the covering used. We show this fact in the following example

Example 2.5. Let X be a sample of the unitary circle center at the origin of $\mathbb{R}^2$ with a small amount of noise. Let the parameter space be the interval $[0, 1]$. To show the dependency of the `Mapper` construction on the filter and the covering we define two filter functions $f_i(x) = \|x - p_i\|_2$ with p_1 the left most point in the data and $p_2 = (0, 0)$. To find a covering, we divide the range of the filters into a set of smaller intervals which overlap. This gives us two parameters to control the resolution, the number of intervals (N) and the percentage overlap between successive intervals (p). In this case, we form $\mathcal{U}_2$ with $N = 2$ and $p = 20\%$ and $\mathcal{U}_3$ with $N = 3$ and $p = 30\%$. We show the results in Figure 2.2

Note that $M_X(\mathcal{U}_2, f_1)$ and $M_X(\mathcal{U}_3, f_2)$ fail to resemble the homotopy type of the unitary circle, meanwhile, $M_X(\mathcal{U}_3, f_1)$ is homotopy equivalent to the unitary circle, giving the best approximation for the sample data.

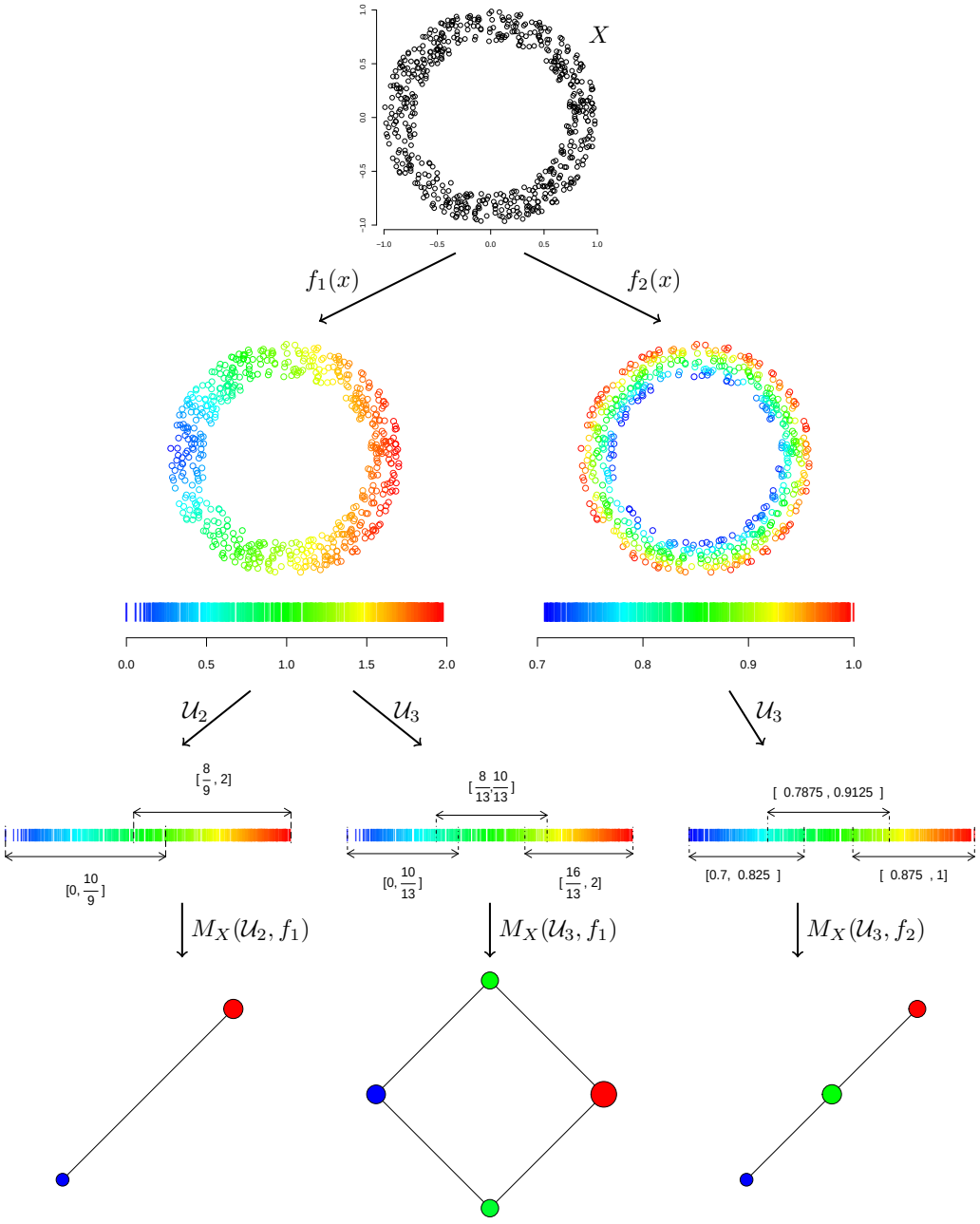


Figure 2.2: Mapper ring data example. The data is sampled from a noisy circle. We use two different filters and two open covers dividing the range of the filters. For each interval we compute the clustering of the points lying in each interval and connect the clusters whenever they have non empty conversation. At the bottom are the simplicial complexes formed by the combination.

3

Materials and methods

3.1 Data acquisition

We analyzed an open dataset acquired by the Cambridge Center for Aging and Neuroscience (Cam-CAN) [58, 59].

Participants

A total number of 700 participants were selected to take part of this study including 100 individuals in each decile age 18-87 and an equal number of men and women. Eligibility measures included cognitive health, meeting hearing and English language requirements, and being eligible for MRI scanning. [59] The final sample for the fMRI time series analysis including 492 participants.

Stimuli

Films are widely used to elicit emotion in a variety of research studies [60], like no other artistic resource, films have the ability to engage attention and evoke powerful and memorable emotions, not only individually but shared across an audience. Perhaps the most notorious example of this is the genre of suspense. In spite of the different ways humans react to even identical stimuli, "it seems fair to say that film-watchers undergo the closest thing to a global or common response when viewing stress-causing moments within films" [18]. Therefore, movies are well suited to advance in this project since the changes of network connectivity can be temporally related to the intensity rates of suspense and they are more likely to be easily identified through all subjects.

The stimulus is an edited version of Alfred Hitchcock's "Bang! You're Dead", a black and white television drama portraying a little boy that plays around with a charged revolver thinking it is a toy gun. The film has been previously used in fMRI studies on

synchronization of the brain responses across viewers ([2], [61]), indeed, the suspenseful film evoked a global response at a neural level.

In the study, the full 25-minute episode was condensed to 8 minutes while maintaining the narrative. Participants were instructed to watch, listen, and pay attention to the movie (they were not aware of its title) [58].

Continuous Response Measurement Sample

A total of 22 human raters were recruited to provide continuous response measures [62]. While viewing the same edited clip from the fMRI sample, participants were directed to "continuously evaluate the degree of suspense" they were experiencing using an online rating tool [63] . We use the averaged degree of reported suspense calculated in [2]. We present the degree of suspense through time in Figure 3.1.

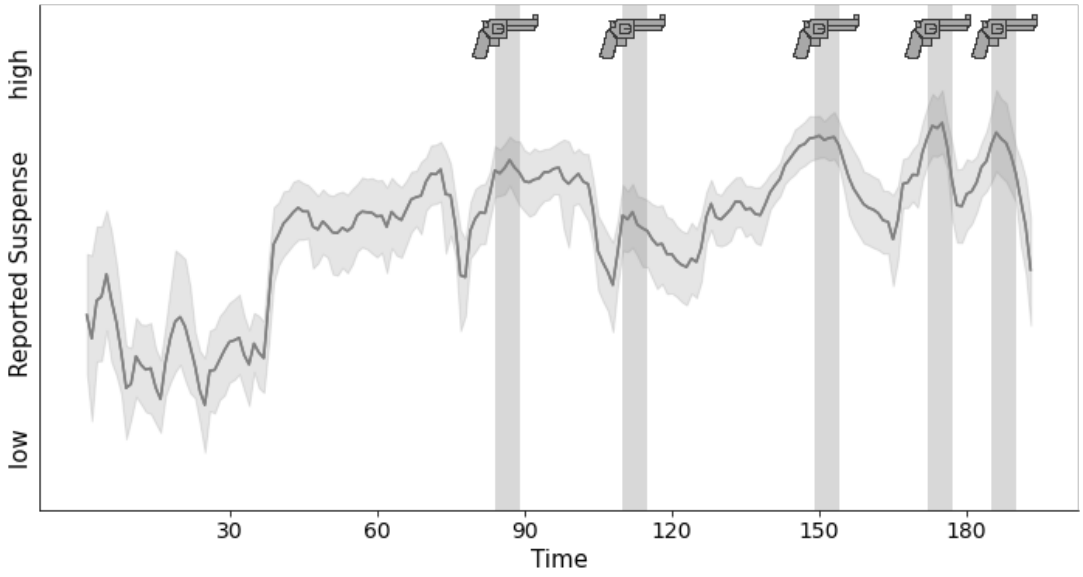


Figure 3.1: Continuous reports of suspense. The group-averaged time series of continuous ratings of suspense.The gray-shaded vertical blocks indicate movie scenes comprising close-ups of the gun when the kid has it aimed at a person. Replicated from [2].

fMRI scanning and processing

A total of 193 volumes were acquired using a 3T Siemens TIM Trio System with a 32-channel head coil. For the functional scan, T2*-weighted echo planar images (EPIs) were

acquired using a multi-echo sequence (Repetition time (TR) = 2.47 seconds; 5 echoes; flip angle 78 deg; 32 axial slices; field of view = 192 mm × 192 mm; voxel-size = 3 mm × 3 mm × 4.44 mm) with an acquisition time of 8 minutes and 13 seconds. More details are in the report on [58].

The imaging data were processed using the nipype framework [64] including motion correction, slice-time correlation, co-registration and nonlinear normalization to the MNI template. Functional data were detrended and high-pass filtered at 0.01 Hz and time series were extracted using NiLearn [65]. From the original dataset of 646 participants some participants were excluded if the data was incomplete or exhibited abnormal behavior resulting in a final sample of 492 participants with 193 volumes of 231.502 voxels. We use the processed data from [2].

3.2 Regions of interest

The whole volume is divided in regions using brain parcellations. Brain parcellations provide fundamental insights into the organizational principles of the human brain and are great strategies of data reduction, enabling information from hundreds of thousands of voxels or vertices to be compressed into manageable sets of regions reflecting distinct entities [66].

The chosen parcellation is the Shen et al. atlas [67], a groupwise graph-theory-based parcellation approach to define nodes (regions) for network analysis. The atlas defined in [68] comprehends 268 regions covering the whole brain. Additionally, the regions are grouped into nine networks: Somato-Motor, Cingular-opercular, Auditory, Default, Visual, Frontal-Parietal, Salience, Ventral-Attention and Dorsal-Attention [45]. The information of each region is retrieved from the BioImage Suite Web [69].

In accordance with the previous researches on brain structures related to suspense and dynamic interactions of LSN in anxious anticipation and stress presented in Section 2.2 we focus on Salience, Fronto-Parietal and Default networks. Therefore, our Regions of interest (ROIs) are from these networks, excluding some regions whose anatomical regions do not correspond to the ones commonly linked to these networks in the literature. Nevertheless, we maintained the temporal regions in the Fronto-Parietal and Default networks.

Ultimately, we considered 74 regions from Shen Atlas. The distribution of the regions by network and lobes are in the Appendix A (Table A1) . The corresponding anatomical regions from the Automated anatomical labeling atlas (AAL) for each network are shown in Table 3.1. Note that some regions are present in more than one network, this detail

arises from the fact the atlases are different, meanwhile AAL is based on anatomy, the Shen Atlas gives priority to connectivity, thus the regions can overlap.

Salience			
1	Angular Gyrus R	7	Frontal Orbital Cortex
2	Paracingulate Gyrus	8	Dorsolateral superior Frontal Gyrus
3	Inferior Frontal Gyrus	9	Medial superior Frontal Gyrus L
4	Middle Frontal Gyrus	10	Cingulate Gyrus ant. Division R
5	Insular Cortex	11	Supramarginal Gyrus post. division R
6	Frontal Pole	12	Frontal Operculum Cortex
Fronto-Parietal			
1	Frontal Pole	6	Inferior Frontal Gyrus pars opercularis
2	Superior Frontal Gyrus	7	Inferior Frontal Gyrus pars triangularis
3	Middle Frontal Gyrus	8	Middle Temporal Gyrus R
4	Supramarginal Gyrus	9	Inferior Temporal Gyrus R
5	Precentral Gyrus	10	Superior Parietal Lobule
Default			
1	Hippocampus	12	Orbital superior Frontal Gyrus
2	Frontal Pole	13	Inferior Frontal Gyrus pars triangularis
3	Temporal Pole	14	Dorsolateral superior Frontal Gyrus
4	Precuneus	15	Medial superior Frontal Gyrus
5	Frontal Orbital Cortex	16	Superior Temporal Gyrus ant. division
6	Frontal Medial Cortex	17	Middle Temporal Gyrus post. division
7	Angular Gyrus L	18	Middle Temporal Gyrus ant. Division
8	Paracingulate Gyrus	19	Superior Temporal Gyrus post. division
9	Fusiform gyrus	20	Cingulate Gyrus post. division
10	Lingual Gyrus	21	Parahippocampal Gyrus post. division
11	Middle Frontal Gyrus		

Table 3.1: Regions considered for each network. If R or L are not specified then both hemispheres are contemplated.

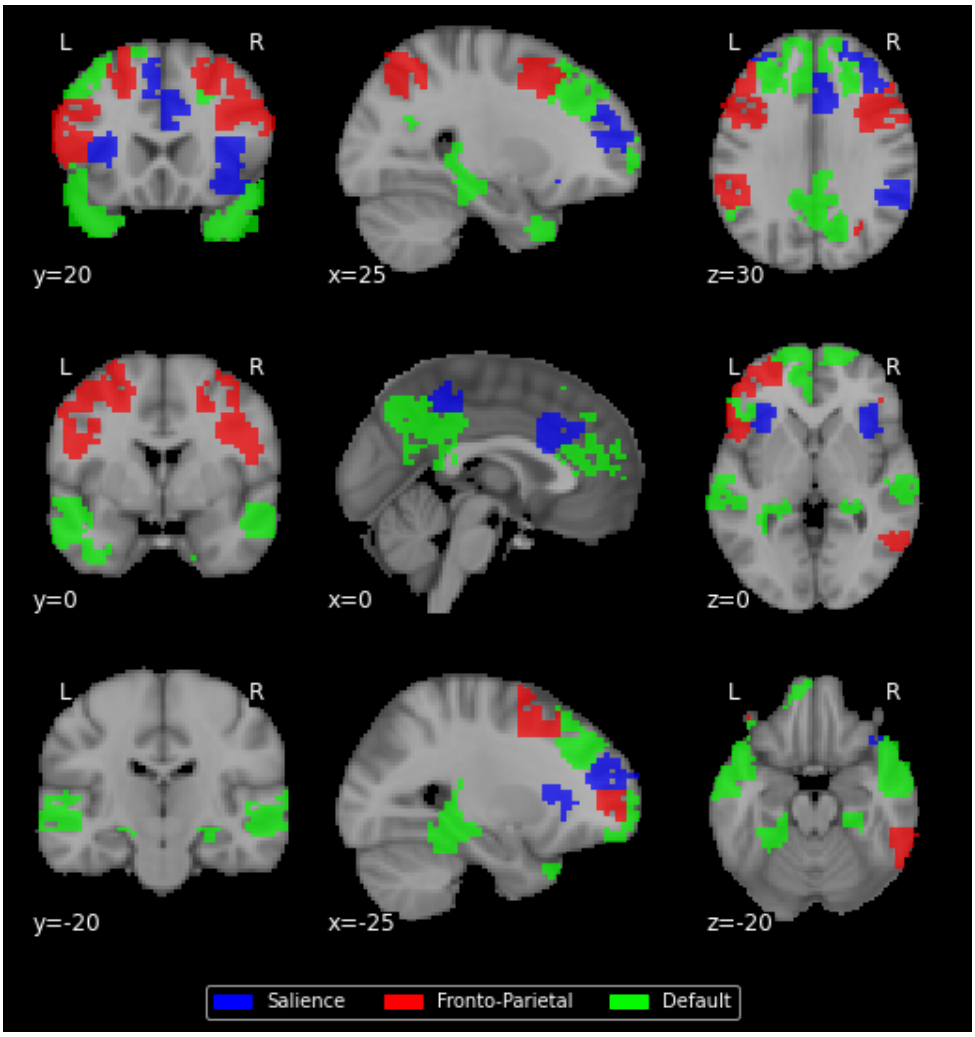


Figure 3.2: Regions of the Salience, Fronto-Parietal and Default networks depicted on the brain.

3.3 Characterization of the states

The final dataset consists on $P = 492$ patients, each with a scan of $N = 74$ regions tracked in $T = 193$ volumes (states). Let us consider only one patient, i.e., an $N \times T$ matrix representing regional activity magnitudes as a function of time. Our aim is comparing the T brain volumes to capture time-evolving changes. We can characterize each state by its pattern of activity i.e., BOLD activity for each region, by its pattern of connectivity i.e., the FC between the regions or a mixed of both approaches [50].

For example, Saggar et. al. [39] in their application of `Mapper`, considered the first characterization, thus each state is represented by a vector of N entries. On the other hand, following our hypothesis on dFC, we characterize each state by an $N \times N$ matrix capturing the FC between the regions. The simplest way to generate a set of dynamic graphs is the *sliding window method*.

3.3.1 Sliding window method

First, to asses FC between regions, we apply the *Pearson correlation*, the most common measure for functional connectivity [70]. Pearson correlation is a statistic measure of linear correlation between two sets of data. Given two samples $X = \{x_0, ..., x_n\}$ and $Y = \{y_0, ..., y_n\}$ the Pearson correlation between X and Y is defined as

$$r(X, Y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{(\sum_{i=1}^{n}(x_i - \bar{x})^2)\,(\sum_{i=1}^{n}(y_i - \bar{y})^2)}} \tag{3.1}$$

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Now, in the framework of the sliding window method (SWM), we select a window of size W. Then, we define a $N \times W$ matrix representing the regional activity within the temporal span from time $t = 1$ to time $t = W$. Subsequently, we compute the FC between each pair of regions. Afterwards, the window is shifted (slided) by a step S, and the same calculations are repeated over the time interval $[1 + S, W + S\ ]$. This process is iterated until the window spans the end part of the time series. (Figure 3.3**A** top part). This procedure yields $U = T - W + 1$ windows each with $N \times (N - 1)/2$ values, which are summarized into a connectivity matrix describing the *state* of the brain during the examined temporal interval. Let us call State i the matrix found from the window $[i, i+W]$.

When we consider all windows, we recover a set of U connectivity matrices describing the temporal evolution of whole-brain FC (Figure 3.3**A** lower part) [70, 71].

As is noted from the description of the SWM we need to choose the parameters W and S. For S, we choose 1 TR as step since it has been proved to be the optimal value to detect connectivity changes between the windows [71]. On the other hand, W is a matter of debate, since the output of the SWM can be highly dependant on it. Too short window lengths increase the risk of introducing spurious fluctuations [72], in contrast, too long windows are not capable to capture short-lived FC variations. Nevertheless, previous studies suggested that window sizes around $30 - 60$ s are able to produce robust results in dFC and in most cases, different window lengths, when chosen in this interval stabilize and do not yield substantially different results [70, 73]. Consequently, we choose $W = 18$ TR which gives an span of approximately 44/5s falling into the recommended interval as we want.

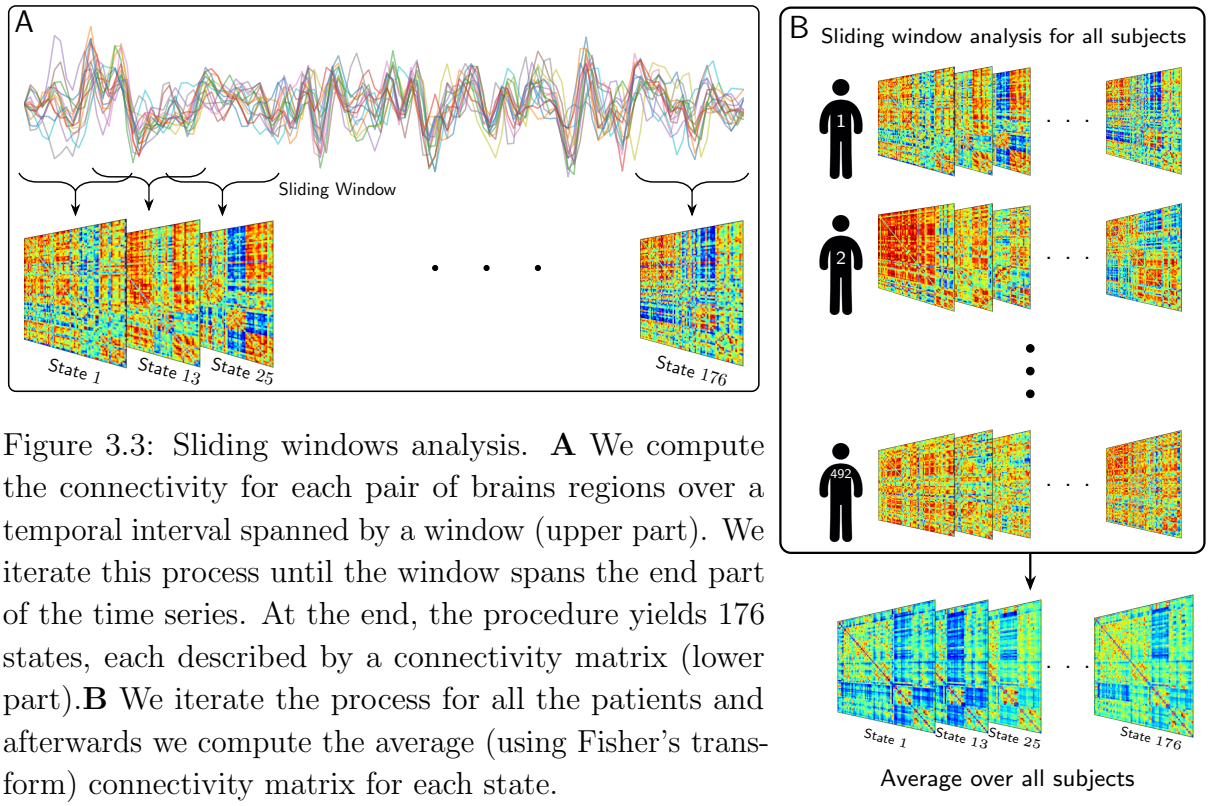


Figure 3.3: Sliding windows analysis. **A** We compute the connectivity for each pair of brains regions over a temporal interval spanned by a window (upper part). We iterate this process until the window spans the end part of the time series. At the end, the procedure yields 176 states, each described by a connectivity matrix (lower part).**B** We iterate the process for all the patients and afterwards we compute the average (using Fisher's transform) connectivity matrix for each state.

Afterwards, we use the SWM over all individuals. Since the correlation matrices among the patients may be quite variable, to smooth and stabilize the correlation coefficients between the regions [74], we take the average of all connectivity matrices for each State

(Figure 3.3**B**). For this average, we first apply Fisher's transformation to all correlations, then we average and subsequently back-transform the values. In this way, we reduce the bias induced by the bounds of correlation at -1 and 1 [75].

3.3.2 Thresholding connectivity matrices

Having estimated the connectivity matrices for each state, we need to apply a global *threshold* τ in order to dismiss negative connections and spurious or weak positive connections. Let $C = [c_{ij}]$ be a connectivity matrix. We form the matrix $\hat{C} = [\hat{c}_{ij}]$

$$\hat{c}_{ij} = \begin{cases} c_{ij} & \text{if } c_{ij} \geq \tau \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

As usual, if there is a nonzero element in the connectivity matrix, this is equivalent to saying there is a weighted and undirected edge between the respective nodes on the corresponding functional network.

Now, with the purpose of choosing a proper τ we consider the notion of connection density. We can define the connection density κ as

$$\kappa = \frac{2\epsilon}{N(N-1)} \tag{3.3}$$

where ϵ is the total number of edges which weight is equal or greater than τ.

From [76] we know brain graphs typically have connection densities below 0.5 since over this value brain graphs present random topologies. Since the density decreases when τ increases, we choose the minimum τ such that the density of all states do not surpass 0.5. Specifically, we start at $\tau = 0.3$ and we increase τ at a step of 0.1 until every State's density is below 0.5.

3.4 The Mapper construction

Before describing in detail the steps on the `Mapper` process we summarize it in Figure 3.4 as a comprehensive guide.

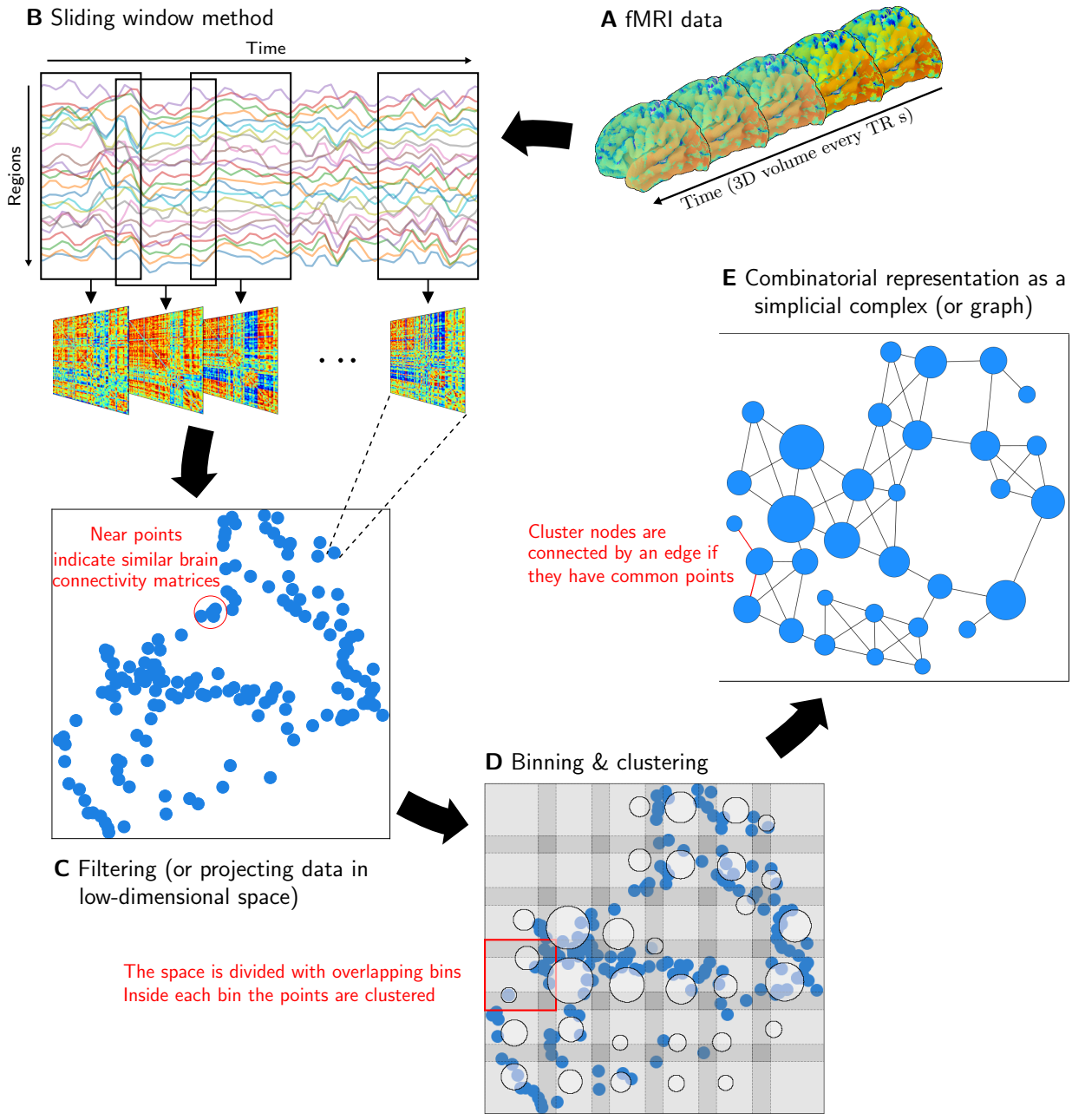


Figure 3.4: The schematic flowchart of the Mapper construction. **A** The process start with pre-processed fMRI dataset from 492 participants. **B** The sliding window method is applied to compute the functional connectivity over time. **C** In the filtering step the set of connectivity matrices are projected into a two-dimensional space. **D** The space is divided into smaller bins determined by the number in each dimension and the percent of overlap between them. Next, partial clustering is applied within each bin. **E** To create a compressed combinatorial representation (graph) each cluster is treated as a node and two nodes are connected if they share data points (connectivity matrices in this case).

3.4.1 The metric space

From our characterization of the states, we have two mathematical representations, the vector and the connectivity graph.

To calculate the distance between the states in the form of vectors, we use the Manhattan distance, in other words, we sum up the differences in the activity of the regions. On the other hand, to measure the difference in form of graphs, we sum up the differences in the functional connectivity for each pair of regions, formally let $G = (V, E)$ and $G' = (V', E')$ be the connectivity matrices of some States. We define their distance $D(G, G')$ as

$$D(G, G') = \sum_{i<j} \left| e_{ij} - e'_{ij} \right| \tag{3.4}$$

where e_{ij} is the weight of the edge connecting vertices i and j.

3.4.2 Filter function

The election of the filter function is inevitably application-specific, however some functions have been proposed [41, 77–79]. These functions rely on distance, thereby, they carry geometric information about the data even in non-euclidean spaces. The `Mapper` is highly dependent on f, thus different choices of filters could highlight distinct properties of the data.

The filter we used is an algorithm for nonlinear dimensionality reduction method called *Isometric mapping (Isomap)* that uses geodesic distances to learn the global geometry of the data set and converges asymptotically to the true structure [80]. With this filter we aim to preserve and highlight the neighboring States exploiting the local linearity of the dataset.

Previous to describing this approach, we present the following definitions:

- n is the number of empirical objects in our dataset.
- δ_{ij} is the dissimilarity between objects i and j.
- X is an $n \times m$ matrix that describes a point configuration of n points in m-dimensional Euclidean space. The elements of X are called *configurations* of the objects.

Now, we describe an outline of the algorithm:

- k-nearest neighbors graph: We construct a graph of n where nodes i and j are connected if i belongs to the k nearest neighbors of j or vice versa. We set the weight of the edge to δ_{ij}. Note that this relation is not symmetric.
- Shortest path: We compute the geodesic distance matrix Δ between all pairs of nodes. Δ_{ij} corresponds to the length of the shortest path between nodes i and j.
- MDS: We apply the classical Multidimensional Scaling (MDS) [81] to matrix Δ. In this algorithm we find the eigendecomposition of the centered matrix $B = -\frac{1}{2}J\Delta^{(2)}J$, where $J = I - \frac{1}{n}O$ is the centering matrix with I as the identity matrix and O the matrix of all 1's, both of size $n \times n$. Then we take the first m eigenvectors (corresponding to the m largest eigenvalues) as the columns of X.

3.4.3 Covering

Considering the filter function stated above, the parameter space Z is a real n-space $\mathbb{R}^m$, for simplicity m is equal to 2. The next step is finding a suitable open covering $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$.

Theoretically speaking, `Mapper` makes no assertions restricting the shape of the sets U_α, but regular coverings are preferred for example rectangles, hexagons or balls [41, 82]. Each of these shapes depends on parameters selected by the user, in particular we use rectangles uniformly distributed in the parameter space. The parameters for this covering are the number of intervals (n) that each dimension of the parameter space is divided in and the percentage of overlapping between intervals called overlap (o). Thus, we have a cover of n^2 rectangles. These parameters determined the coarseness of the `Mapper` output: low values of n would generate a small number of nodes each containing probably a large amount of data points, thus, we would have the risk of loosing structural information, on the other hand, high values of n make computational expensive `Mapper` since it would generate nodes that do not carry additional information of the underlying structure. Meanwhile, high values of o could generate spurious connections (edges) on the combinatorial object and low values of o could loss them, again, losing structural information.

As the choice of different parameters leads to different and even misleading outputs, tuning them is crucial. For this project, we are going to choose them aiming to replicate the homology of our data.

For this purpose we use another topological tool called persistent homology. The details are presented elsewhere [56]. Briefly, for a given real ϵ we create a cover of open balls of radius ϵ centered in each data point. Then, we compute the homology of the cover's nerve.

Subsequently, we repeat this algorithm for $0 < \epsilon < R$, the homological classes that *persist* trough a large range of ϵ reveal the homology of the set.

As we simply consider the 0- and 1- simplices in the `Mapper` output, we only need to compute the 0 and 1 homology of the dataset, meaning, the number of connected components and the number of 1-dimensional holes i.e., circles.

3.4.4 Clustering

Clustering is the task of grouping by some natural criterion of similarity. Even though there is not a clear definition of cluster [83] the heuristic is that the inter-point distance within each cluster would be smaller than the distance between clusters. `Mapper` does not specify a clustering algorithm, however, the desired characteristics of the clustering are: a) Is not restricted to data in Euclidean space, b) Do not require specifying the number of clusters before hand [41].

In our research, we consider the agglomerative hierarchical clustering methods since this type of clustering satisfies the mentioned characteristics. There are different types of methods depending on how we calculate the distance between two clusters. For example, in the single-link method [84], the distance between any two clusters C and D is defined as the minimum distance between the elements in C and D:

$$d(C, D) = \min_{i \in C, j \in D} d(i, j) \tag{3.5}$$

In contrast, the complete-link method defines the distance between clusters using the maximum.

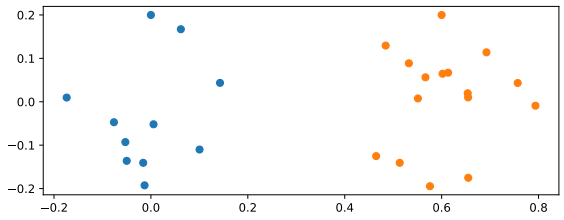
$$d(C, D) = \max_{i \in C, j \in D} d(i, j) \tag{3.6}$$

The hierarchical method works as follows:

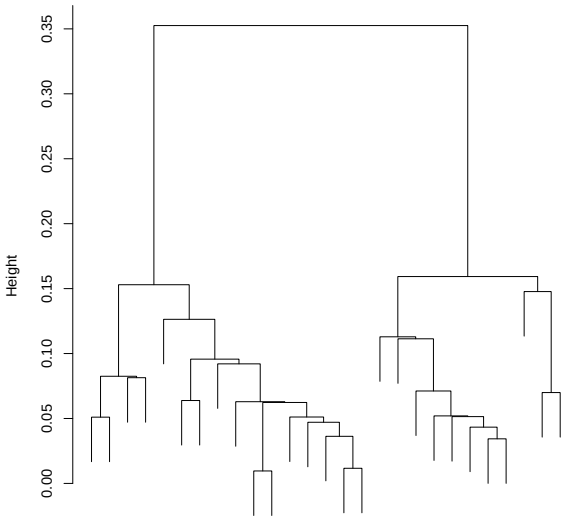
1. We start with each data point as his own cluster.
2. We calculate $\mathcal{D} = \{d(C, D) \mid C, D \text{ clusters } \}$.
3. We form a new cluster $C' = C_0 \cup D_0$ as the new cluster if $d(C_0, D_0) = h = \min \mathcal{D}$ and go back to step 2.
4. We finish when there is only one cluster.

Let us call the values h's from step 3, heights. Note that for each height we have a different set of clusters, moreover, we can extend the relation to all values between 0 and the last height (Figure 3.5b), then, it is crucial to choose which set represents an appropriated

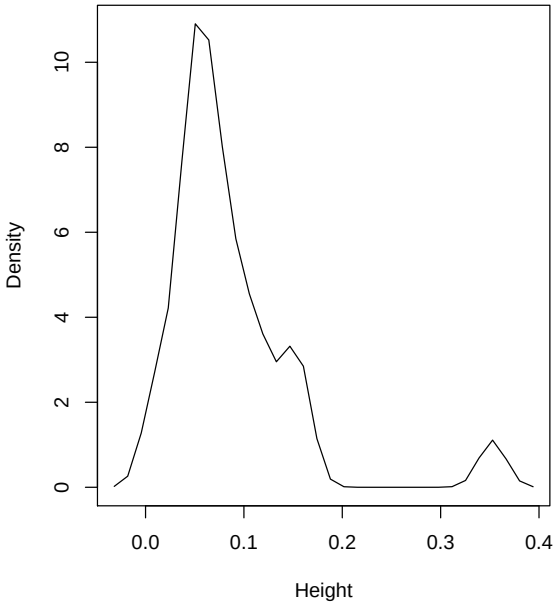
clustering of the data. We use the cut-off method proposed in the `Mapper` implementation called *continuous* to choose the proper h.



(a) Sample of 27 points divided into two groups.



(b) Dendogram representing the arrangement of the clusters produced by hierarchical clustering. Each vertical line represents a cluster, when two of them met in an horizontal line, it means the clusters merge at that height forming a new cluster.



(c) Density estimation of the vector of heights where new clusters are formed.

Figure 3.5: Example of hierarchical clustering and the method we used to determine the best clustering

Let us use a toy example (Figure 3.5a) to show how the continuous cut-off method works. It is simple: we form the vector $\mathbf{h}$ of heights from step 3 and we calculate the kernel density estimation [85] of $\mathbf{h}$ (Figure 3.5c). We use a Gaussian kernel and we have the possibility of tuning the *bandwidth* of the kernel. After that, we choose as the cut-off height the minimum height (continuous interval) where the density is smaller than certain ϵ of order 10^{-8}. If none of the heights satisfy this condition we choose the last height of

the range, meaning, we choose a single cluster. In this way, we follow the heuristic, we aim to have larger distances between clusters compared to distances inside clusters. In the example shown in Figure 3.5, after calculating the density of $\mathbf{h}$ we find that $h \approx 0.469$ is the cut-off height, this height corresponds to a clustering of two clusters and perfectly matches the initial configuration of our sample.

3.4.5 Software

For this construction, we use the implementation of `Mapper` developed by Piekenbrock et. al in R language [82,86] . Besides, for the filter we use the implementation of Isomap found in the Python's package scikit-learn [87] and Python's package ripser [88] to calculate the persistent homology.

3.5 Analysis of the Mapper graph

3.5.1 Special nodes

Let us give some definitions from graph theory. Let $G = (V, E)$ be a graph.

Definition 3.1. $G' = (V', E')$ with $V \subseteq V$ is a *connected component* of G if any two nodes in V' are path connected and the edges of E' only connects nodes of V'.

Definition 3.2. Let C be a path of G. C is a cycle if the first and last nodes are the same.

Let u be a vertex of G and let G_u be the graph after removing u.

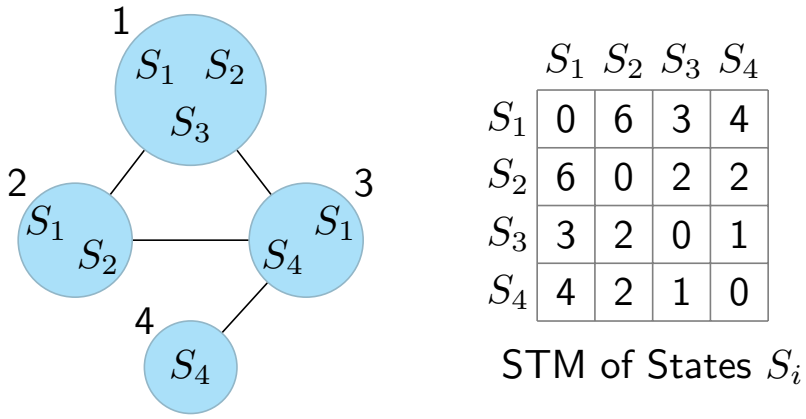
Definition 3.3. u is a *connector node* if the number of connected components of G_u and G are different.

Definition 3.4. u is a *cyclic node* if the number of cycles of G_u and G are different.

3.5.2 State transitions

To estimate the transitions of the FC in our brain networks, we convert the `Mapper`'s graph into a weighted adjacency matrix for States. Let us called this matrix State Transition Matrix (STM). For each pair of States, we count the number of times c they are contained

by the same node or the nodes containing these States are connected by an edge in the graph. After, we set c as the weight of the edge connecting the former pair. We see in Figure 3.6 a minimalist example. For instance, for S_1 and S_4 their corresponding entry is 4 since they share node 3 (+1), and are in the connected nodes 1/3, 2/3 and 3/4 (+3).



	S_1	S_2	S_3	S_4
S_1	0	6	3	4
S_2	6	0	2	2
S_3	3	2	0	1
S_4	4	2	1	0

STM of States S_i

Figure 3.6: Example of State Transition Matrix (STM) with a graph with four nodes and four states.

3.5.3 Underlying connectivity patterns

With the purpose of anchoring the graph and their properties into neurophysiology, we measure the strength of connections inside and between large-scale networks to detect the changes in the FC responsible for the observed features. For this measurement, we utilized the within- and between-network connectivity weights, defined in [33] as follows:

$$W_{N_1,N_1} = \frac{2}{n_1(n_1-1)} \sum_{i,j \in N_1, i<j} c_{ij} \tag{3.7}$$

$$W_{N_1,N_2} = \frac{1}{n_1 n_2} \sum_{i \in N_1, j \in N_2} c_{ij} \tag{3.8}$$

where N_1 and N_2 are networks, n_1 and n_2 are the number of regions respectively and c_{ij} is the measure of functional connectivity between regions.

To compute dynamic changes in the network's connections, we measure all weights within and between networks for each State forming a time series. Specifically, we form six

time series corresponding to three within-networks weigths,(Salience, Fronto-Parietal and Default) and three between-network weights (Salience/Fronto-Parietal, Salience/Default and Fronto-Parietal/Default).

To evaluate the relationship between the dynamics of the networks and the level of suspense, we calculate the Pearson correlation between the level of suspense (Fig. 3.1) and the resulting six time series. Note that meanwhile the level of suspense understood as a time series has a length of $T = 193$, our time series corresponds to $U = 176$ states. Therefore, we consider the ratings of suspense between times $t = 11$ and $t = 186$.

We evaluate the statistical significance of our results through the null model explained in section 3.6.

3.6 Statistical test for dFC

To validate that our results are due to real changes in the FC of the brain we define a "statistical hypothesis test in which the null hypothesis corresponds to the correlation being static and the alternative hypothesis corresponds to the correlation being dynamic" [89]. To start we need a proper null distribution. As suggested in [89], the null hypothesis under which the distribution of the test is constructed should correspond to the sFC.

To approximate a suitable null distribution, we generate a large number of phase-randomized surrogate time series following the method introduced in [90]. In this method, given a time series $x(t)$ we take its discrete Fourier transform $X(f)$

$$X(f) = \mathcal{F}\{x(t)\} = A(f)e^{i\phi(f)}$$

where $A(f)$ is the amplitude and $\phi(f)$ is the phase. Note that $X(f)$ is evaluated at discrete frequencies. Then, we rotate the phase ϕ at each frequency f by an independent random variable ψ which is chosen uniformly in the range $[0, 2\pi)$. That is

$$\tilde{X}(f) = A(f)e^{i\phi(f)+\psi(f}$$

Finally, we obtain the surrogate time series calculating the inverse Fourier transform

$$\tilde{x}(t) = \mathcal{F}^{-1}\{X(f)e^{i\psi(f)}\}$$

By the Weiner-Khintchine theorem $\tilde{x}(t)$ has the same autocorrelation function as $x(t)$. However , for multivariate time series, not only we want the linear properties for each

time series, but we also want to reproduce the linear correlations between time series. Suppose we have m simultaneously measured variables $x_1(t), ...x_m(t)$ with zero mean and unit variance and let $X_1(f), ...X_m(f)$ denote their respective Fourier transforms. By an extension of the Weiner-Khintchine theorem, the Fourier transform of the cross-correlation function is the cross spectrum:

$$X_j^*(f)X_k(f) = A_j(f)A_k(f)e^{i[\phi_k(f)-\phi_j(f)]} \tag{3.9}$$

Since Eq. 3.9 only involves the differences of phases, to preserve all linear auto-correlations and cross-correlations, we only need to add the same phase $\psi(f)$ for all j. That is,

$$\tilde{x}_j(t) = \mathcal{F}^{-1}\{X_j(f)e^{i\psi(f)}\}$$

In our case, to form a set of surrogate data, for each patient, we have 74 time courses corresponding to our ROIs and we phase-randomized them by the same phase. We repeat this process on all 492 patients. We iterate this process a thousand times to obtain a proper null distribution.

The second step for the statistical hypothesis is picking a test capable to detect dFC. We choose *variance* as our test since it is the most widely used and the most straightforward [89].

Let c_i be the connection value for some pair regions at State i and form the time series $c_0, ..., c_U$.We can calculate the variance of the sample correlation series and we will note it by κ

$$\eta^2 = \frac{1}{U}\sum_{i=1}^{U}(c_i - \mu)^2 \tag{3.10}$$

where μ is the mean of the time series.

In this way, the absence of dFC corresponds to the null hypothesis

$$H_0 : \eta = 0 \tag{3.11}$$

and the presence of dFC corresponds to the alternative hypothesis

$$H_1 : \eta > 0 \tag{3.12}$$

Suppose we reject the null hypothesis for our data i.e., $\eta_0 > 0$ To decide whether η_0 is statistically significant or not, we calculate the variance for each of the surrogate

copies. Since we chose a proper null distribution, we should end with a distribution approximating the unknown null distribution of η. Then, we calculate the percentile of η_0 in the distribution of η, if η_0 falls within the 5% highest values or lowest values (if the test is one-sided) we can conclude it is statistically significant [89]. As usual we note the significance level as p (percentile) and we choose 0.05 as our threshold. In this way, we reject η_0 if $p > 0.05$.

From the test, we can determine the pair of regions that present a significant dynamic functional connectivity and exclude from our analysis the pairs which do not have a positive variance.

4

Results

4.1 Characterization of the states

We used the SWM in all subjects and took the average overall resulting in a matrix of dimensions $176 \times 74 \times 74$ corresponding to 176 States represented in connectivity matrices of 74 regions. In Figure 4.1 we show the connectivity matrices for some States.

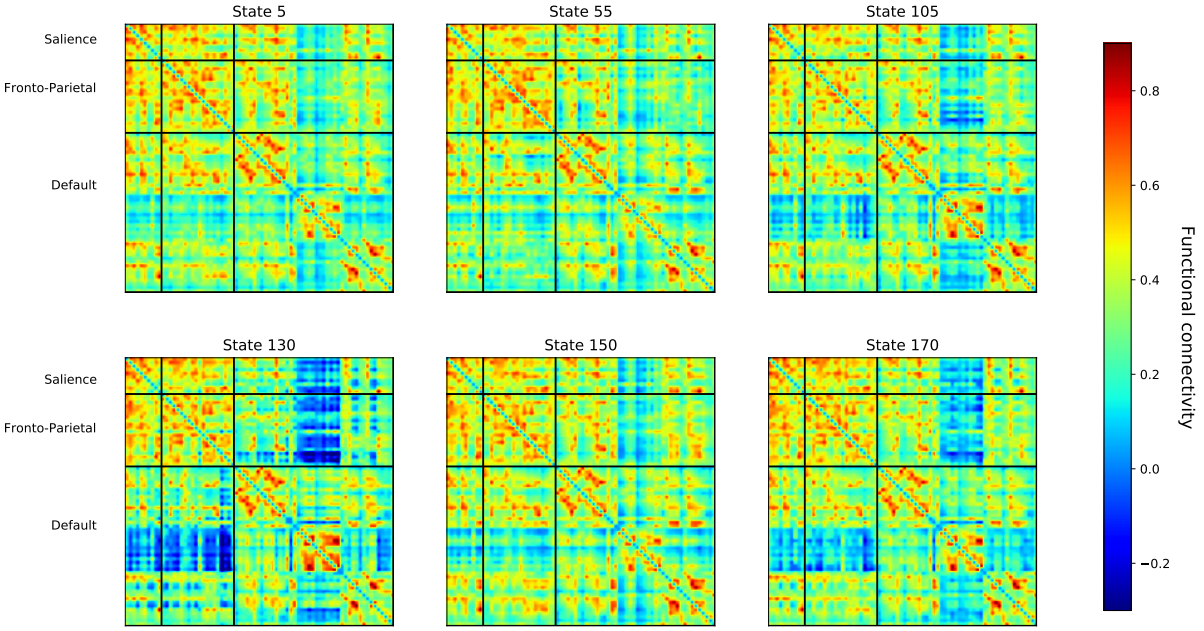


Figure 4.1: Snapshots of correlation matrices of some States. The regions are ordered by networks (Salience, Fronto-Parietal and Default) and each network by lobes (Prefrontal, MotorStrip, Insula, Parietal,Temporal, and Limbic).

Subsequently, we calculated τ to threshold all connectivity matrices (Eq. 3.2). We found that $\tau = 0.37$ satisfies the condition we gave. We can see in Figure 4.2 an histogram of the densities (Eq. 3.3) of the resulting matrices. The density values are between 0.325

and 0.475, having a range $R = 0.15$. At the same time, all resulting graphs (after thresholding) are connected i.e., there is a path between any pair of nodes we choose. On the other hand, when we did the SWM in our null distribution we found their τ values are either 0.34 or 0.35 and generate States which ranges of their density values R are between 0.0307 and 0.0925. In this way, both values, τ and R, are at the top of their null distributions.

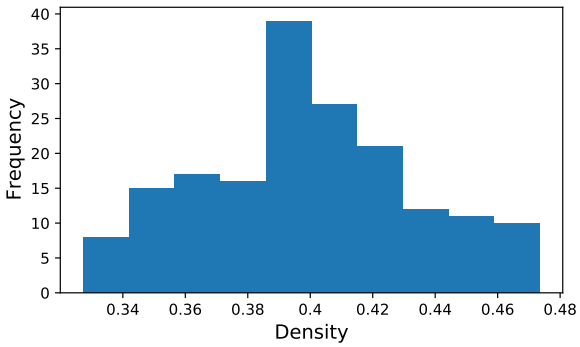


Figure 4.2: Histogram of the densities of the States' connectivity graphs.

Nonetheless, we reran the following `Mapper` construction without considering the density and just taking $\tau = 0$. The results from this analysis are presented in Appendix C.

4.2 Statistical test for dFC

We calculated the variance (Eq. 3.10) of the FC through time, to do so, we formed a matrix η of 74×74 where η_{ij} is the variance of FC between regions i and j trough time. On the other hand, we computed the p-value for each variance, then, we formed matrix p of same dimensions as η where p_{ij} is the p-value of η_{ij}. In Figure 4.3 we see a matrix divided in upper and lower triangular parts. The lower part corresponds to η meanwhile the upper part corresponds to p.

The 64% of the pairs of regions presented a variance greater than 0, in addition, approximately 93% have a p-value less than 0.05. On the other hand, the pairs that had zero variance were the result of zero-constant time series due to to the fact that we threshold the connectivity matrices.

From these results, we could proceed with our analysis without omitting the pair of regions that present a null variance since their time series were null and did not affect the distance calculation nor the pairs that had a non-zero variance but have a high p-value since they only represented the 7% of the total.

In addition, we can conclude the data shows indeed a dynamical behavior in the FC since its variance is greater than 0, thus, rejecting the null hypothesis.

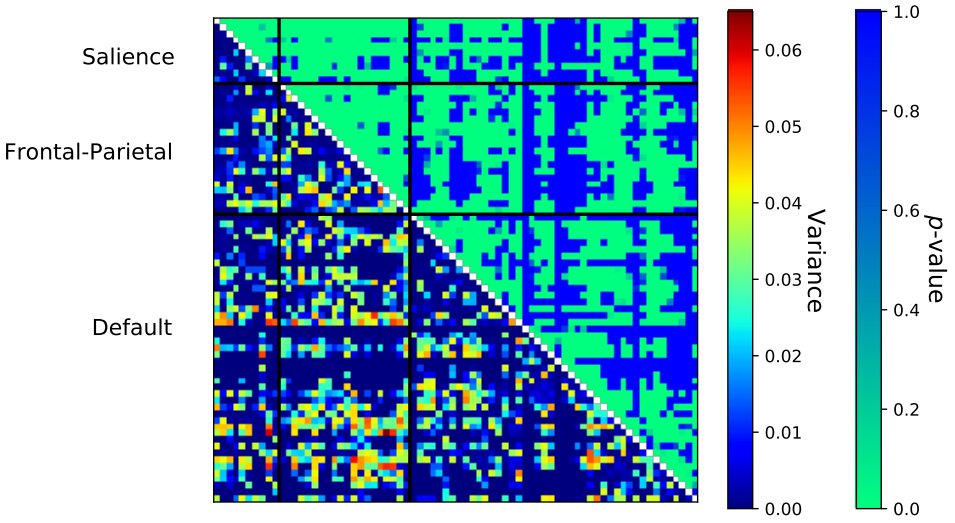


Figure 4.3: Variance of the FC trough time for each pair of regions.

4.3 The Mapper construction

We calculated the distance between the states in the two characterizations we presented. For the vector space, the distance matrix is presented in Appendix B in Figure B1**A** and for the graph space, the matrix is in Figure 4.4. On the other hand, as mentioned in Section 3.4 we calculated the 0- and 1-homology of our dataset using persistent homology. The results are presented in the persistence diagram (Fig. 4.5).

In the diagram, each dot represents a homological class, for example, each blue dot is a connected component. The x coordinate of the dot corresponds to the ϵ value when the homology class appears (Birth) and the y the *epsilon* value when the homology class disappears (Death). The dotted line represents the infinity. Thus, the farther the dot is from the diagonal, the more persistent is. Let us call *persistence* the difference between the Death and Birth. With this term, we have one connected component with infinity persistence and one circle with persistence 97 (from 125 to 222) with the rest of circles not surpassing 25 in persistence. Thus, mathematically speaking, we could infer the objective homology is $H_0 = 1$ and $H_1 = 1$, i.e., a *shape homologically equivalent to a circle*.

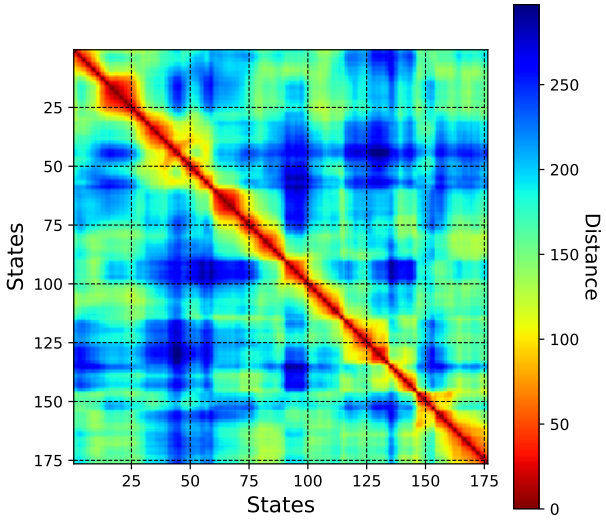


Figure 4.4: Distance matrix of States

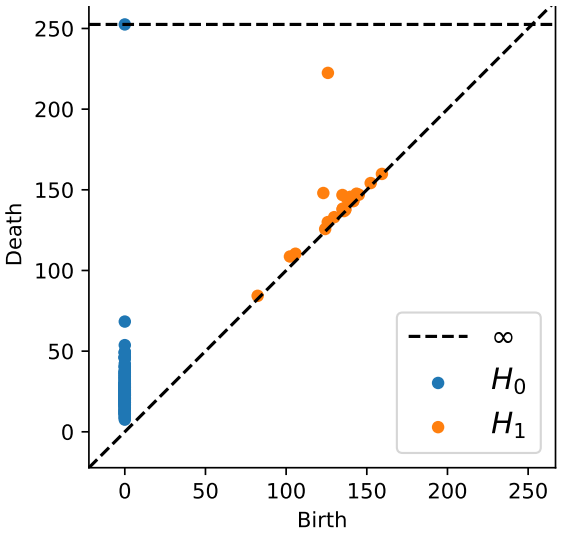


Figure 4.5: Persistence diagram presenting the homology of the set of States. H_i stands for i-homology.

Besides, we used Isomap to filter our data to $\mathbb{R}^2$. Note that Isomap has a parameter k corresponding to the number of nearest neighbors we have to considered. However, the previous computation gave us a suitable $k = 30$ since we used geodesic distance to highlight the homological properties. The result of the filtering is shown in Figure 4.6, each dot represents a State, for qualitative examination, we have colored the States by their level of suspense and join the dots with a line indicating the flow of time. In Figure 4.6**B** we show the level of suspense for the States as a guidance to better recognize the corresponding State in the filtering.

Following, we generated 25 different covers of the parameter space varying the intervals from 4 to 8 (steps of 1) and the overlap from 25 to 45 (steps of 5). The results are shown in Figure B3 on the Appendix. Then, we localized the `Mapper` graphs that resemble the objective homology with the minimum amount of nodes and multi-color nodes. At the end, the best `Mapper` graph is generated by the rectangular cover of parameters $n = 6$ and $o = 35$ (Appendix Fig. B2**A**).

The final graph (Figure 4.7), posses 52 nodes (as opposed to 176 States) and 100 edges. For qualitative analysis, the size and color of the nodes are based on the States they contain. In this way, we can observe the flow of time and its relation with the suspense.

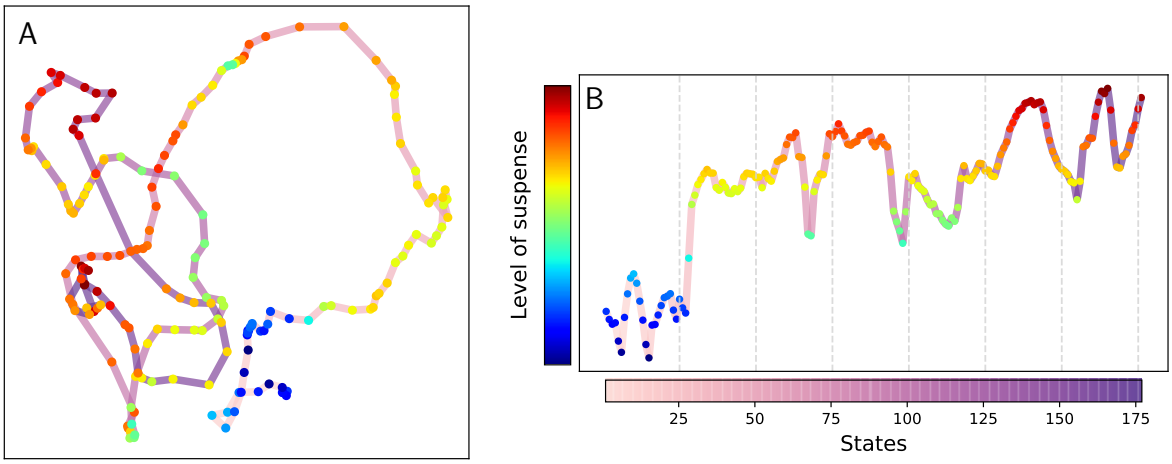


Figure 4.6: **A** Isomap filtering. The colored States are joined by a degraded pink-purple line representing the flow of time i.e., State i is joined with a segment with State $i-1$ and a slightly darker segment with State $i+1$. **B** Reported level of suspense for the States (see Fig. 3.1)

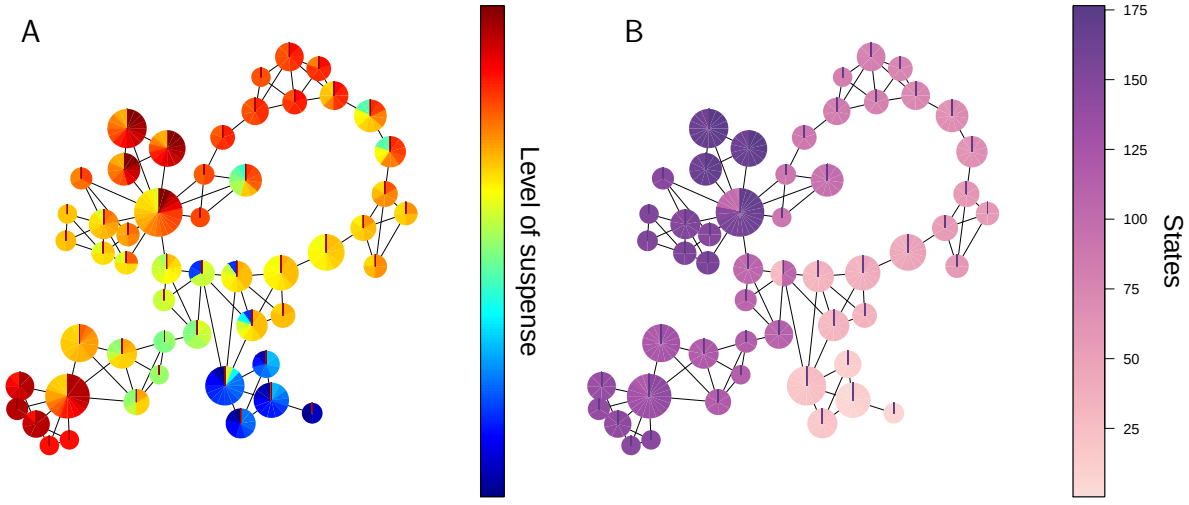
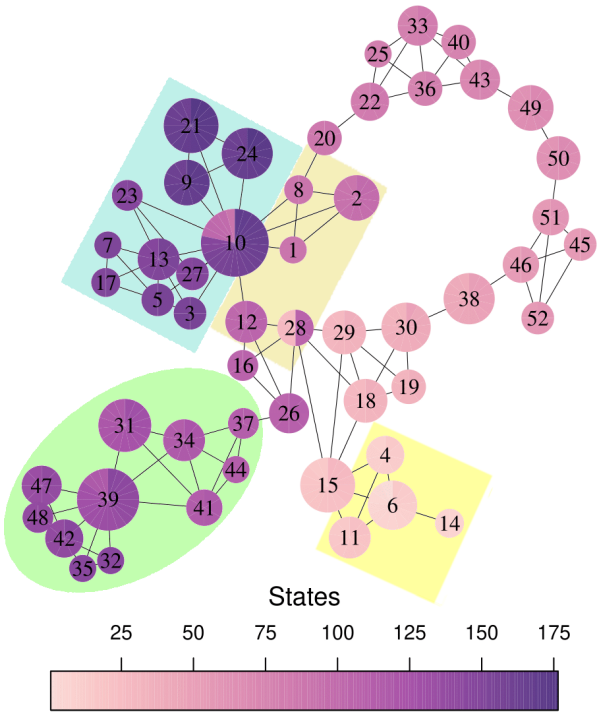


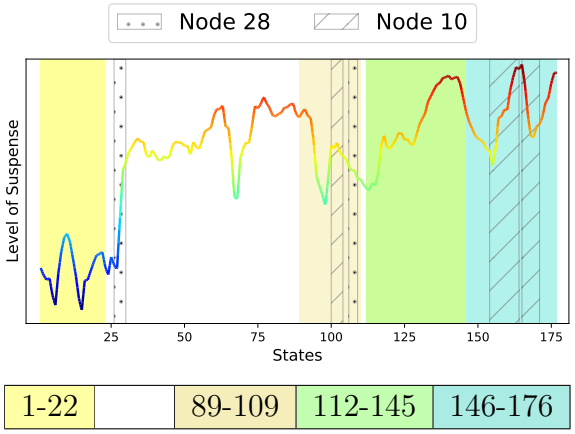
Figure 4.7: Final `Mapper` graph. Each node is sized in proportion to the number of contained States. The nodes are colored using a pie chart denoting **A** the level of suspense **B** the index of the States

4.4 Analysis of the Mapper graph

Let us consider the graph in Figure 4.8a. In this graph, we can define three groups of nodes which we enclosed with colored patches: *yellow, green and blue.*



(a) `Mapper` graph. The nodes are numbered and colored denoting the indices of States. In yellow, green and blue are enclose noteworthy groups of nodes.



(b) Level of suspense of the States. The vertical colored rectangles indicate the intervals associated with the groups of nodes from the `Mapper` graph. Besides, the contained States of nodes 10 and 28 are indicated by patterned rectangles.

Figure 4.8: Analysis of `Mapper` graph. We show the relation between the graph and the States including their level suspense.

The groups are defined by two common properties:

- In a process similar to the Girvan–Newman algorithm [91] for community detection, we can define groups as the ones left behind after removing connector nodes. Specifically, the groups we considered are defined by the connector nodes 15, 26 and 10, respectively.
- They are mono-colored groups. This stems from the fact that the States contained inside all nodes in a group form an interval of States. For instance, the yellow group

has all States between 1 and 22, it is not the case, one State inside the interval is missed. Thus, we can associate intervals of States to each group. In Figure 4.8b we identify the level of suspense of the related intervals.

In particular, the yellow interval corresponds to the States of low suspense. The gray interval, explained later, possesses States of decreasing suspense even though it has a peak. The green interval encloses States of increasing suspense including a peak of suspense, meanwhile the blue interval has one peak and two troughs of suspense.

Furthermore, when we focused on the cyclic nodes, two of them stand out since they are bicolored:

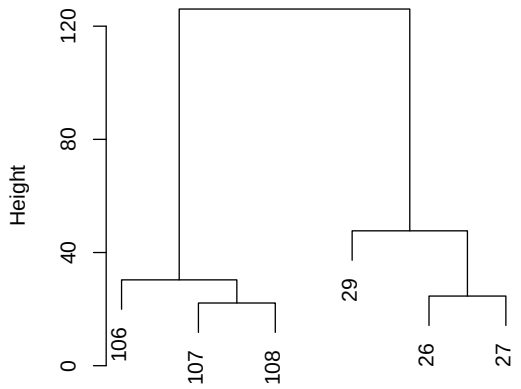


Figure 4.9: Cluster corresponding to node 28

- Node 10 connects the blue group to the rest of the graph. It contains States from different intervals, specifically it contains the set {90, 100-103, 147,148, 154-163, 165-170} of States. In Figure 4.8b we indicate these States with a diagonal pattern.
- Node 28 connects the yellow group to the graph. As before, it connects two different intervals of States {26, 27, 29, 106, 107, 108}. We indicate these States with a dotted pattern in the Figure 4.8b. In addition, it is connected to the connector nodes 15 and 26.

Besides, we defined an extra group that contains the nodes connected to node 10 that are not part of the blue group. We enclose this new group with a gray patch. Furthermore, the gray patch also includes part of the States contained in Node 28.

Additionally, we calculate the State Transition Matrix (STM) for our `Mapper` graph (Fig. 4.10). The connections in the matrix are easily explained by the features present in the graph:

- The two well-defined sub-matrices that we find in the inferior part of the diagonal in descendent order correspond to the connections inside the groups green and blue, respectively.
- The general absence of connections of the interval 112-145 with other intervals derives from the fact that the green group is connected to the graph only through node 26.

- There are two blocks of connections outside the diagonals. The upper one derives from the connections of node 28 and the one below to node 10 corresponding to the gray group.

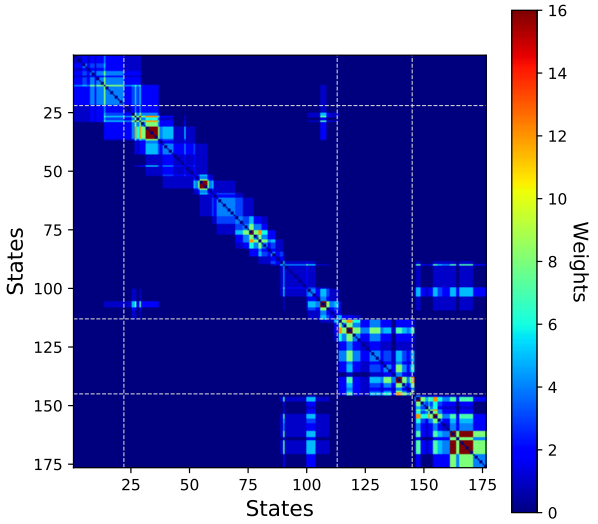


Figure 4.10: State Transition Matrix (STM)

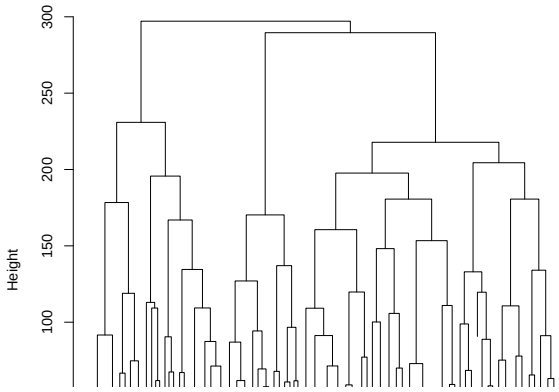


Figure 4.11: Upper part of the dendogram representing the clustering of the rows of the distance matrix.

Along with this analysis, we wanted to contrast the groups found with the `Mapper` graph against clustering from the distance matrix directly with the complete-link method (Eq. 3.6). In the Figure 4.11 we see the steps of the clustering. Our cut-off method detects a single cluster, however we chose as cut-off height 250 to obtain the three clusters shown in Table 4.1

Cluster 1	**Cluster 2**	**Cluster 3**
60-79	32-59	1-31
115-145		80-114
		146-176

Table 4.1: Table of intervals of States forming the clusters

Next, we computed the weights within- and between networks between all pairs of networks. The corresponding time series are shown in Figure 4.12. We colored the intervals associated with the groups of nodes from the `Mapper` graph. Within these intervals, we can describe some particular features.

To begin with, all series present a peak and a subsequent descending in the yellow interval. Second, the Salience & Default and Fronto-Parietal & Default series present a very steep increasing transition between green and blue intervals, giving a great difference between the averages. This difference in averages is also present in Salience & Fronto-Parietal and

Default & Default series. Finally, the Salience & Default and Fronto-Parietal & Default series also present a difference of the average connection between gray and green intervals.

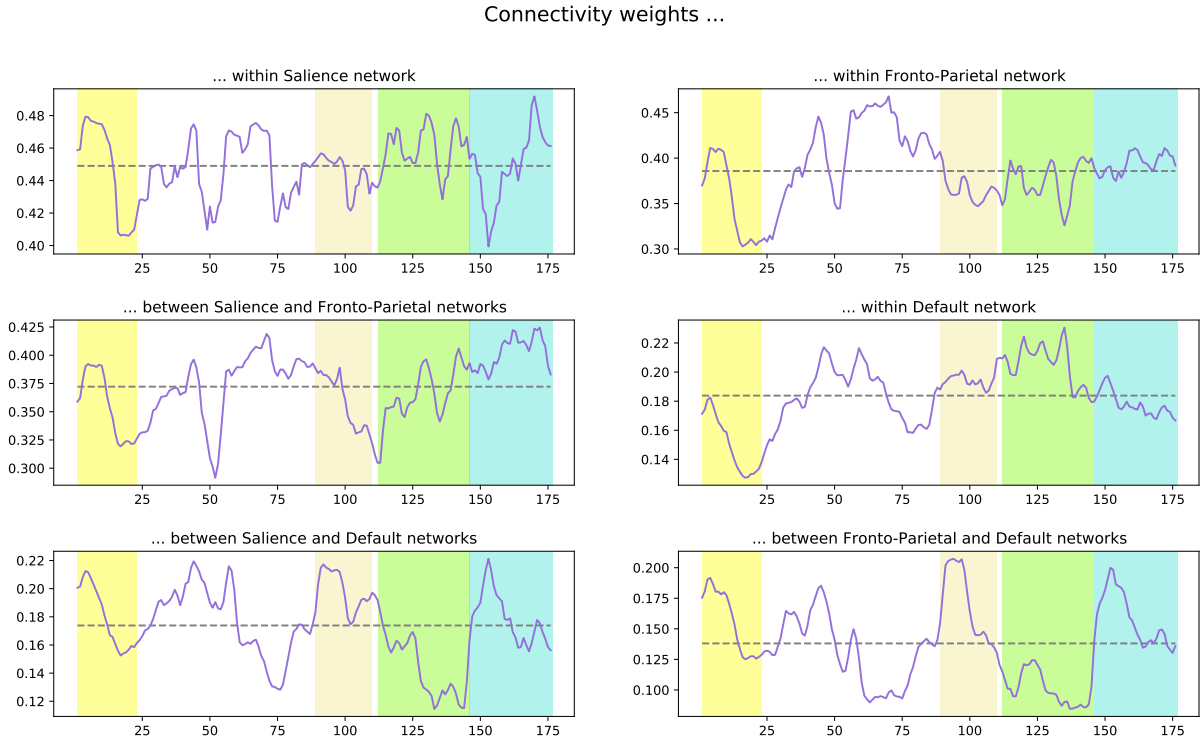


Figure 4.12: Time series of connections' weights within- and between-networks (purple) and the average value (dashed gray). The vertical colored rectangles indicate the intervals associated with the groups of nodes from the `Mapper` graph in Figure 4.8a.

Additionally, for each series we calculated the Pearson correlation with the level of suspense cutting from different States as starting points, i.e., we did not consider the States before the starting States. All the correlation values are presented in Figure B4 in Appendix B. As an interesting remark, the correlation values from State 1 to 29 have a similar tendency expected from the similarity of the yellow interval pointed out before, and the values started to increase (in absolute value) from State 68, coinciding with the trough of suspense right before the first suspense peak associated with a close-up of the gun aiming an innocent person (Fig. 3.1). Thus, we chose 29 as the starting State where we considered the correlation values, since we focus in the section of medium and high suspense. The values presented in Table 4.2.

Networks	Correlation value
Salience & Salience	0.09
Fronto-Parietal & Fronto-Parietal	0.15
Salience & Fronto-Parietal	0.44
Default & Default	-0.29
Salience & Default	-0.55
Fronto-Parietal & Default	-0.29

Table 4.2: Correlation values between the level of suspense and the series of weights within- and between-networks. The number of State indicates the position from where we considered the series.

We observed a positive correlation for the connectivity weights within Salience and Fronto-Parietal networks as well as between the networks. However, the correlation value between them drops when we omit the first States. In addition, for the Salience & Fronto-Parietal series in the gray interval descends and increases in the green interval. This is similar to the behavior of the suspense's level.

In contrast, we found a negative correlation for the connectivity weights within the Default network, however, when we admit the first States the value turns positive. In contrast, the correlation values for the connections between the Default and the other two networks are negative. In these series, it is evidently an opposite tendency to suspense, especially in the green interval where the connections' weights descend. In addition, all series with the Default network have lower values of connections in comparison with the other three pairs since their values are between 0.1 and 0.22, meanwhile the minimum for the other pairs is 0.3.

Finally, none of the samples in the null distribution present correlation values with the same sign and close absolute values e.g., r values for the pair of Salience and Fronto-Parietal networks are less than 0.45 in samples that coincide with the signs.

5

Discussion

In the present study, we employed `Mapper`, a topological based tool to investigate, for the first time, the experience of suspense using a network-based approach. We found that the functional connectivity within and between Salience, Fronto-Parietal and Default networks changed dynamically as the level of suspense varied. Additionally, we generated a graphical representation of the brain's transitions through different functional connectivity States exhibiting the relationship of the States between them and the level of suspense.

5.1 The Mapper construction

Let us begin the discussion unraveling the result in each step of the Mapper construction.

Starting with the filter (Fig. 4.6) we can identify some characteristics of the States and relate them with the distance matrix and the `Mapper` graph. To begin with, the degraded line in general has not large trails between States, indicating that neighboring States in time are neighbors in distance. This is also clear in the reddish band around the diagonal in the distance matrix. The feature is derived from the SWM, since State i and State $i+1$ record the correlation of pairs of time series that only differ in one entry.

In the graph, this characteristic is evidenced in the general uniformity in the color of the nodes, indicating that neighboring States remain in the same clusters. Besides, this fact is confirmed by the State Transition Matrix, where most connections are present on the diagonal.

Another feature we can recover from the filter is the fact that States between 30 and 70 are far from the rest reflected in the absence of crossings in the right upper part of the filter. In the same way, we can verify the distances of these States with other are above 200. We also confirm the `Mapper` graph also retains this characteristic in the STM, since we can appreciate the absence of connections in the interval between 37 and 89 with other States. From the difference of intervals before (30-70 / 37-89), we can point the key of this

topological construction. In the filter, close to the center, States around 75 (dark orange color) and around 117 (green color) intersect each other's lines. When we apply clustering to the open set containing them (See Appendix Figure B2), the algorithm recognizes, using the original distances of the States, they are in different clusters, thus, not being part of the same node. In this way, the filter, which offers a first good approximation of the data, can create artifacts derived from the projection, nonetheless, from this initial information, `Mapper` tries to detect the real structure of the data.

In the search for a good approximation of the data, we generated different graphs (Fig. B3) varying the parameters. We perceive not all graphs have homologies $H_0 = 1$ nor $H_1 = 1$, showing the mentioned need of tuning. In particular, for H_1, we can describe the reason of this variation. Let us explain using our final graph.

As mentioned, node 28 is the key to form the cycle, its connection between the initial section to the middle part of States guarantees the formation of the loop. Yet, when we see the corresponding clustering (Fig. 4.9) it looks like it should be broken into two disjointed nodes, then breaking the loop. The reason is simple, from the persistent homology we know the circular shape appears from the distance 125, evidenced in the merge height of the node 28 found in 126. Then, if we let all open sets to be cut in heights less than 125 we would never recover the objective homology. Thus, we allowed this counter-intuitive clustering, not only to respect the cut-off method but to recover the homology of the data, and we chose our `Mapper` graph allowing the minimum of these nodes or as we call them, multicolor nodes.

Before continuing, we want to highlight the results from the same analysis without setting a threshold presented in Appendix C. We find the results are quite similar, both have an homology of the circle (Fig. C2), the filters have the same shape (Fig. C3) and the `Mapper` graph (after similar selection) presented common characteristics like maintaining the homology, general uniformity in the color of the nodes and detection of akin groups in the graph. However, the multicolor nodes are bigger and more numerous, some of the times, due to a spurious clustering detection. Thus, we can conclude the set threshold maintains the structure of the data, reducing the possible noise of the data. On the other hand, the density values we found after thresholding were between 0.325 and 0.475. The range is similar to the density values found in [92] after a similar process.

Then, moving to the final graph, we recognized different groups of nodes when we identified the connector and cyclic nodes, which are the key to preserving the homology. As we could associate the groups with intervals of States we detected a peculiarity in the flow of time: the blue and green groups are neighboring intervals, but in the graph, the blue group is completely separated from the green one and instead is connected with the gray group by

the node 10, which is separated in terms of intervals. Given this behavior by intervals, we did a clustering directly from our distance matrix (Figure 4.11). In particular, the third cluster is pretty interesting, the first interval is comparable with the yellow group, the second with the gray group, and the third is exactly the blue group. Their belonging to the cluster denotes a similarity given precisely by nodes 10 and 28 in our graph. In addition, we find an interval analogous to the green group in the first cluster, as expected, a different cluster from the groups before.

Having said that, some studies have used this approach of hierarchical clustering [93, 94] or k-means [95, 96] to summarize the obtained States into smaller sets. The `Mapper` approach has the same ability to recover the set's information without constraining the number of sets or collapsing the data into big groups. In addition, it generates a combinatorial object (graph) that shows the relationship between the sets, indicates the States that are key in the formation of the groups, and remarkably reveals how the brain navigates through different States during the scan.

5.2 Dynamic brain networks in suspense

Let us follow the discussion with the the within- and between-network connectivity weights, to ultimately link the results with the graph's shape.

The results from the series of connectivity's weights within and between networks revealed that all networks exhibit functional connectivity changes during the scan. In addition, we related the variations of connectivity to changes of the level of suspense via the Pearson correlation. Interestingly, the correlation values varied if we took as starting points for the computation States related to the increase of suspense (Figure B4), showing that the networks demonstrate a differentiated behaviour in the presence of the stimulus of suspense.

The functional connections within the Salience network have a small correlation value with the suspense's level. This result is not in line with the one reported by Najafi et. al. [33] where the connections within the network presented the strongest change, increasing in periods of anxious anticipation and decreasing when the threat creating anxiety retreated. Even though we did not find a paired behavior with the level of suspense, the network presents the highest values of connections, with a interval between 0.4 and 0.5, thus, having a sustained connection during suspense. A potential reason is that the movie stimulus is motivationally salient, peaks of suspense signaled potential dangerous events, meanwhile, troughs of suspense signaled safety and hopeful events.

Contrasting, the behavior of the connections between the Salience and Fronto-Parietal networks are similar to Najafi's work, in fact, with a larger and significant correlation, indicating the communication between the two networks increased in the presence of the stimulus of suspense. Meanwhile, the connections within the Fronto-Parietal network do not possess a high positive correlation with suspense, however, the connections' weights decreased as the level of suspense does, this remark corresponds to the gray patch and the connections' weights stayed the same when the suspense increased, as we see in the green patch. The behaviour is similar to Najafi's results.

On the other hand, the connections' weights between the Default network with the Salience and Fronto-Parietal networks had a large and moderate anti-correlation with the level of suspense, respectively. The results are again in line with Najafi's work and reveal that the increasing suspense is associated with the segregation of the Default network from the Salience and Fronto-Parietal network, at the same time, the decrease of the suspense prompts again the communication between the networks.

Conversely, the correlation of the suspense with the connections within the Default network is negative, which differs from the result in [33] where is positive. However, it can be argued that likewise the connections within the Fronto-Parietal network, the series presented a positive large correlation value considering the series from State 1. This result derives from the fact that the starting sections of both connectivity weight series are below the average. This behaviour suggests that the large sudden rise in suspense increased the connection within the series, but the following smaller increments did not have the same effect.

Besides the negative correlation values, the regions in the Default network have connectivity weights in the interval between 0.2 and 0.3. These values are low in comparison to the other series. The thresholding is partially responsible for the low values, since most of the values that were reduced to 0 belong to the Default network, then the average values of connections diminished, however, this fact indicates that the lower values on the matrices are precisely in the Default network. We consider that it indicates that the communication within and between the Default network during the scan is not strong and could be related to the low activity of the Default network during externally oriented tasks [97]. In particular, the regions in the temporal lobe have the lowest connectivity values, especially with the Salience and Fronto-Parietal regions, as we observe in Figure 4.1, it is consistent with the fact that temporal regions are not usually consider inside the Default network given the small connection with the other regions within the network.

Then, going back to the graph, one of the most significant features is the separation between the green and the blue group. This disconnection emerges from the transition between the green and blue intervals, especially in the Salience & Default and Fronto-Parietal &

Default networks where the series present a steep rising slope. Furthermore, the change is associated with a transition from a period of increasing suspense to a descend. Hence, the graph captures one the most significant transition between different States that is associated with the level of suspense.

On the other hand, we have nodes 10 and 28 that connect separate intervals. We speculate that the similarity of the States in these intervals is related to steep changes in the suspense. For example, both gray and blue groups are peaks between two local minima within a difference less than 15 TR. In the same way, the transition between low and high suspense happens in 5 TR. This result could even imply the functional connectivity of the networks are similar in steep variations of suspense, independent of its direction (upwards or downwards). Nevertheless, this relationship is hard to determine given the transitions are fast and less than our window width of 18 TR.

As a final remark, when we tried to compute the `Mapper` graph with same approach as Saggar et al. [39], computing the distance between the States without computing the FC (Figure B1) the results were defective, since the data did not show any structure related to time or suspense. Importantly, this outcome reveals that the experience of suspense relies on the dynamic connectivity within and between networks and not only on the variation of activation of the regions inside each network. This illustrates an important principle, differences in activation can be disassociated from differences in co-activation.

5.3 Limitations and future directions

The results presented here must be considered in the context of experimental and methodological limitations. To begin with, the movie we used as stimuli was originally designed to evoke suspense [61], but not to contrast levels of suspense, in contrast with [33] where the experiment was designed to have defined periods of anxiety and relief. In the movie, we had a short period of low suspense followed by a period of sustained high suspense, thus, we had a transition from low to high suspense but not the opposite. Besides, in the second part we have difficulty to identify the transitions between peaks and trough of the level suspense. This fact in combination with the low temporal resolution of fMRI pose an upper limit for our ability to resolve the neural effects of suspense in time. Our results suggest future researches to design experiments with multiple blocks of low and high suspense, including periods of transition from both states.

Furthermore, our results are circumscribed to Shen Atlas, used to define the brain regions and the selection of brain regions for each network based on the work of Power et al. [45].

However, we hypothesize the results will be similar for a different selection of atlas based on functional connectivity and network separation since we did not consider roles for particular regions but general behaviors of connected regions. Nonetheless, an interesting path for future research could be considering the hypothesis of overlapping networks, where specific regions belong to several intersecting networks and its participation at a given time is context dependant [47].

We should also consider that observed dFC might be related to time-varying noise (e.g., subject motion and variable respiratory and cardiac rhythms), nevertheless, the filtering and processing of the data diminished the effects of this noise.

Additionally, we inherited the limitations of the SWM such as the lack of a gold standard for the choice of the window length, the window step, the type of window and other parameters [71]. As suggested in [72] one promising way to overcome this parameter dependency is the use of time frequency analysis like the wavelet transform coherence (WTC). Furthermore, other methods to asses the FC, alternative to the Pearson correlation have been proposed, like applying the the regularization strategy to the inverse of the covariance matrix [95] or using magnitude squared coherence or power spectrum coherence [98, 99].

Besides, future research should investigate the contributions of the subcortical regions linked to emotional experiences [5] like the amygdala, PAG, BNST, hypothalamus and thalamus in the dynamic interactions of the networks in the presence of suspense. Due to time limitations, in the present investigation it was not possible to extend the analysis to these structures. In addition, the work could be extended to recognize the contribution of the regions inside each network, e.g., the Frontal Orbital Cortex inside the Salience network.

Finally, another venue for future research could be applied the proposed methodology used in the thesis to the study of the dynamic processes related to emotions like anger, happiness, sadness among others. Looking forward, it can be argued, we could characterize each emotion by the temporal transitions between different connectivity states [4].

6

Conclusions

Inspired by the psychological model of suspense by Lehne and Koelsch [1] and its consistency with a network-based approach of the emotional brain [20, 42], in our study, we examined the hypothesis that the emotional experience of suspense depends on dynamic interactions within and between large-scale networks. For this purpose, we use `Mapper`, a topological based tool revealed as a "novel method to distill brain dynamics" [39].

Our study appears to be the first to find dynamic functional changes within and between the the Salience, Fronto-Parietal and Default networks associated with the variation of the level of suspense. In particular, the functional connectivity between the Salience and Fronto Parietal networks has a positive correlation with the level of suspense. Conversely, we found a negative correlation with the functional connectivity between the Salience and Fronto-Parietal networks with the Default network. These findings add substantially to our understanding of the neural processes underlying the emotional experience of suspense and demonstrate the potential of understanding emotional experiences as dynamic interactions of domain-general neural systems.

Finally, through the use of `Mapper`, we obtained a graph that recovers the shape of the data in a low-dimensional representation and depicts how the brain evolved across different functional States in time, revealing the brain's dynamical organization across variations of the level of suspense. Furthermore, the results of this thesis highlight the topological-based methods as novel and powerful alternatives to carry out future researchs on dynamic networks.

Bibliography

[1] Lehne, M., and Koelsch, S., *Toward a general psychological model of tension and suspense*, Frontiers in Psychology **6** (2015).

[2] Schmälzle, R., and Grall, C., *The coupled brains of captivated audiences*, Journal of Media Psychology (2020).

[3] LeDoux, J. E., *Emotion circuits in the brain*, Annual Review of Neuroscience **23** (2000), no. 1.

[4] Celeghin, A., Diano, M., Bagnis, A., Viola, M., and Tamietto, M., *Basic emotions in human neuroscience: neuroimaging and beyond*, Frontiers in psychology **8** (2017).

[5] Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., and Barrett, L. F., *The brain basis of emotion: A meta-analytic review*, Behavioral and Brain Sciences **35** (2012), no. 3.

[6] Knobloch-Westerwick, S., David, P., Eastin, M. S., Tamborini, R., and Greenwood, D., *Sports spectators' suspense: Affect and uncertainty in sports entertainment*, Journal of Communication **59** (2009), no. 4.

[7] Knobloch-Westerwick, S., and Keplinger, C., *Thrilling news: Factors generating suspense during news exposure*, Media Psychology **9** (2007), no. 1.

[8] Löker, A., *Film and suspense*, Trafford Publishing, 2005.

[9] Smith, G. M., *Film structure and the emotion system*, Cambridge University Press, 2003.

[10] Prieto-Pablos, J. A., *The paradox of suspense*, Poetics **26** (1998), no. 2.

[11] Smuts, A., *The paradox of suspense*, 2009, https://plato.stanford.edu/archives/fall2009/entries/paradox-suspense/.

[12] Vorderer, P., Wulff, H. J., and Friedrichsen, M., *Suspense: Conceptualizations, theoretical analyses, and empirical explorations*, Routledge, 1996.

[13] Bezdek, M. A., Gerrig, R. J., Wenzel, W. G., Shin, J., Revill, K. P., and Schumacher, E. H., *Neural evidence that suspense narrows attentional focus*, Neuroscience **303** (2015).

[14] Bezdek, M. A., Wenzel, W. G., and Schumacher, E. H., *The effect of visual and musical suspense on brain activation and memory during naturalistic viewing*, Biological Psychology **129** (2017).

[15] Lehne, M., *Emotional experiences of tension and suspense: psychological mechanisms and neural correlates*, Ph.D. thesis, Fachbereich Erziehungswissenschaft und Psychologie der Freien Universität Berlin, 2014.

[16] Lehne, M., Engel, P., Rohrmeier, M., Menninghaus, W., Jacobs, A. M., and Koelsch, S., *Reading a suspenseful literary text activates brain areas related to social cognition and predictive inference*, PLoS One **10** (2015), no. 5.

[17] Lehne, M., Rohrmeier, M., and Koelsch, S., *Tension-related activity in the orbitofrontal cortex and amygdala: an fMRI study with music*, Social Cognitive and Affective Neuroscience **9** (2013), no. 10.

[18] NORDEN, M. F., *Toward a theory of audience response to suspenseful films*, Journal of the University Film Association **32** (1980), no. 1/2.

[19] Steinbeis, N., and Koelsch, S., *Shared Neural Resources between Music and Language Indicate Semantic Processing of Musical Tension-Resolution Patterns*, Cerebral Cortex **18** (2007), no. 5.

[20] Pessoa, L., *A network model of the emotional brain*, Trends in cognitive sciences **21** (2017), no. 5.

[21] Bassett, D. S., and Gazzaniga, M. S., *Understanding complexity in the human brain*, Trends in cognitive sciences **15** (2011), no. 5.

[22] Wang, H. E., Bénar, C. G., Quilichini, P. P., Friston, K. J., Jirsa, V. K., and Bernard, C., *A systematic framework for functional connectivity measures*, Frontiers in Neuroscience **8** (2014).

[23] Telesford, Q. K., Lynall, M.-E., Vettel, J., Miller, M. B., Grafton, S. T., and Bassett, D. S., *Detection of functional brain network reconfiguration during task-driven cognitive states*, NeuroImage **142** (2016).

[24] Kinnison, J., Padmala, S., Choi, J.-M., and Pessoa, L., *Network analysis reveals increased integration during emotional and motivational processing*, Journal of Neuroscience **32** (2012), no. 24.

[25] Raz, G., Winetraub, Y., Jacob, Y., Kinreich, S., Maron-Katz, A., Shaham, G., Podlipsky, I., Gilam, G., Soreq, E., and Hendler, T., *Portraying emotions at their unfolding: A multilayered approach for probing dynamics of neural networks*, NeuroImage **60** (2012), no. 2.

[26] Hermans, E. J., Henckens, M. J., Joëls, M., and Fernández, G., *Dynamic adaptation of large-scale brain networks in response to acute stressors*, Trends in Neurosciences **37** (2014), no. 6.

[27] McMenamin, B. W., Langeslag, S. J. E., Sirbu, M., Padmala, S., and Pessoa, L., *Network organization unfolds over time during periods of anxious anticipation*, Journal of Neuroscience **34** (2014), no. 34.

[28] Touroutoglou, A., Bickart, K. C., Barrett, L. F., and Dickerson, B. C., *Amygdala task-evoked activity and task-free connectivity independently contribute to feelings of arousal*, Human Brain Mapping **35** (2014), no. 10.

[29] Touroutoglou, A., Lindquist, K. A., Dickerson, B. C., and Barrett, L. F., *Intrinsic connectivity in the human brain does not reveal networks for 'basic'emotions*, Social cognitive and affective neuroscience **10** (2015), no. 9.

[30] Wilson-Mendenhall, C. D., Barrett, L. F., and Barsalou, L. W., *Variety in emotional life: within-category typicality of emotional experiences is associated with neural activity in large-scale brain networks*, Social cognitive and affective neuroscience **10** (2015), no. 1.

[31] Raz, G., Touroutoglou, A., Wilson-Mendenhall, C., Gilam, G., Lin, T., Gonen, T., Jacob, Y., Atzil, S., Admon, R., Bleich-Cohen, M., et al., *Functional connectivity dynamics during film viewing reveal common networks for different emotional experiences*, Cognitive, Affective, & Behavioral Neuroscience **16** (2016), no. 4.

[32] Diano, M., Tamietto, M., Celeghin, A., Weiskrantz, L., Tatu, M.-K., Bagnis, A., Duca, S., Geminiani, G., Cauda, F., and Costa, T., *Dynamic changes in amygdala psychophysiological connectivity reveal distinct neural networks for facial expressions of basic emotions*, Scientific Reports **7** (2017), no. 1.

[33] Najafi, M., Kinnison, J., and Pessoa, L., *Dynamics of intersubject brain networks during anxious anticipation*, Frontiers in human neuroscience **11** (2017).

[34] Satpute, A. B., and Lindquist, K. A., *The default mode network's role in discrete emotion*, Trends in Cognitive Sciences **23** (2019), no. 10.

[35] Dabaghian, Y., Mémoli, F., Frank, L., and Carlsson, G., *A topological paradigm for hippocampal spatial map formation using persistent homology*, PLoS computational biology **8** (2012), no. 8.

[36] Reimann, M. W., Nolte, M., Scolamiero, M., Turner, K., Perin, R., Chindemi, G., Dłotko, P., Levi, R., Hess, K., and Markram, H., *Cliques of neurons bound into cavities provide a missing link between structure and function*, Frontiers in computational neuroscience **11** (2017).

[37] Sizemore, A. E., Giusti, C., Kahn, A., Vettel, J. M., Betzel, R. F., and Bassett, D. S., *Cliques and cavities in the human connectome*, Journal of computational neuroscience **44** (2018), no. 1.

[38] Yoo, J., Kim, E. Y., Ahn, Y. M., and Ye, J. C., *Topological persistence vineyard for dynamic functional brain connectivity during resting and gaming stages*, Journal of neuroscience methods **267** (2016).

[39] Saggar, M., Sporns, O., Gonzalez-Castillo, J., Bandettini, P. A., Carlsson, G., Glover, G., and Reiss, A. L., *Towards a new approach to reveal dynamical organization of the brain using topological data analysis*, Nature communications **9** (2018), no. 1.

[40] Chazal, F., and Michel, B., *An introduction to topological data analysis: fundamental and practical aspects for data scientists*, ArXiv **abs/1710.04019** (2017).

[41] Singh, G., Memoli, F., and Carlsson, G., *Topological Methods for the Analysis of High Dimensional Data Sets and 3D Object Recognition*, Eurographics Symposium on Point-Based Graphics (Botsch, M., Pajarola, R., Chen, B., and Zwicker, M., eds.), The Eurographics Association, 2007.

[42] Lindquist, K., and Barrett, L., *A functional architecture of the human brain: Emerging insights from the science of emotion*, Trends in cognitive sciences **16** (2012).

[43] Bressler, S. L., and Menon, V., *Large-scale brain networks in cognition: emerging methods and principles*, Trends in cognitive sciences **14** (2010), no. 6.

[44] Uddin, L. Q., Yeo, B. T. T., and Spreng, R. N., *Towards a universal taxonomy of macro-scale functional human brain networks*, Brain Topography **32** (2019), no. 6.

[45] Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., Vogel, A. C., Laumann, T. O., Miezin, F. M., Schlaggar, B. L., et al., *Functional network organization of the human brain*, Neuron **72** (2011), no. 4.

[46] Barrett, L. F., and Satpute, A. B., *Large-scale brain networks in affective and social neuroscience: towards an integrative functional architecture of the brain*, Current opinion in neurobiology **23** (2013), no. 3.

[47] Pessoa, L., *Understanding emotion with brain networks*, Current opinion in behavioral sciences **19** (2018).

[48] Pessoa, L., and McMenamin, B., *Dynamic networks in the emotional brain*, The Neuroscientist **23** (2017), no. 4.

[49] Braun, U., Schäfer, A., Walter, H., Erk, S., Romanczuk-Seiferth, N., Haddad, L., Schweiger, J. I., Grimm, O., Heinz, A., Tost, H., et al., *Dynamic reconfiguration of frontal brain networks during executive cognition in humans*, Proceedings of the National Academy of Sciences **112** (2015), no. 37.

[50] Khambhati, A. N., Sizemore, A. E., Betzel, R. F., and Bassett, D. S., *Modeling and interpreting mesoscale network dynamics*, NeuroImage **180** (2018).

[51] Huang, X., Yao, Y., Bowman, G. R., Sun, J., Guibas, L. J., Carlsson, G. E., and Pande, V. S., *Constructing multi-resolution markov state models (msms) to elucidate rna hairpin folding mechanisms*, Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing (2010).

[52] Yao, Y., Sun, J., Huang, X., Bowman, G. R., Singh, G., Lesnick, M., Guibas, L. J., Pande, V. S., and Carlsson, G., *Topological methods for exploring low-density states in biomolecular folding pathways*, The Journal of Chemical Physics **130** (2009), no. 14.

[53] Nicolau, M., Levine, A. J., and Carlsson, G., *Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival*, Proceedings of the National Academy of Sciences **108** (2011), no. 17.

[54] Nielson, J. L., Paquette, J., Liu, A. W., Guandique, C. F., Tovar, C. A., Inoue, T., Irvine, K.-A., Gensel, J. C., Kloke, J., Petrossian, T. C., et al., *Topological data analysis for discovery in preclinical spinal cord injury and traumatic brain injury*, Nature communications **6** (2015), no. 1.

[55] Feged-Rivadeneira, A., Ángel, A., González-Casabianca, F., and Rivera, C., *Malaria intensity in colombia by regions and populations*, PLOS ONE **13** (2018), no. 9.

[56] Edelsbrunner, H., and Harer, J., *Computational topology: an introduction*, American Mathematical Soc., 2010.

[57] Bakken Stovner, R., *On the mapper algorithm*, Ph.D. thesis, Norwegian University of Science and Technolog, 2012.

[58] Campbell, K. L., Shafto, M. A., Wright, P., Tsvetanov, K. A., Geerligs, L., Cusack, R., Tyler, L. K., and ..., *Idiosyncratic responding during movie-watching predicted by age differences in attentional control*, Neurobiology of Aging **36** (2015), no. 11.

[59] Shafto, M. A., , Tyler, L. K., Dixon, M., Taylor, J. R., Rowe, J. B., Cusack, R., Calder, A. J., Marslen-Wilson, W. D., Duncan, J., Dalgleish, T., Henson, R. N., Brayne, C., and Matthews, F. E., *The cambridge centre for ageing and neuroscience (cam-CAN) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing*, BMC Neurology **14** (2014), no. 1.

[60] Gabert-Quillen, C. A., Bartolini, E. E., Abravanel, B. T., and Sanislow, C. A., *Ratings for emotion film clips*, Behavior research methods **47** (2015), no. 3.

[61] Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., and Heeger, D. J., *Neurocinematics: The neuroscience of film*, Projections **2** (2008), no. 1.

[62] Biocca, F., David, P., and West, M., *Continuous response measurement (crm): A computerized tool for research on the cognitive processing of media messages*, A. Lang (Ed.) (1993).

[63] Nummenmaa, L., Glerean, E., Viinikainen, M., Jääskeläinen, I. P., Hari, R., and Sams, M., *Emotions promote social interaction by synchronizing brain activity across individuals*, Proceedings of the National Academy of Sciences **109** (2012), no. 24.

[64] Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., and Ghosh, S. S., *Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python*, Frontiers in neuroinformatics **5** (2011).

[65] Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., and Varoquaux, G., *Machine learning for neuroimaging with scikit-learn*, Frontiers in neuroinformatics **8** (2014).

[66] Eickhoff, S. B., Yeo, B. T., and Genon, S., *Imaging-based parcellations of the human brain*, Nature Reviews Neuroscience **19** (2018), no. 11.

[67] Shen, X., Tokoglu, F., Papademetris, X., and Constable, R. T., *Groupwise whole-brain parcellation from resting-state fmri data for network node identification*, Neuroimage **82** (2013).

[68] Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., Papademetris, X., and Constable, R. T., *Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity*, Nature Neuroscience **18** (2015), no. 11.

[69] Papademetris, X., *Bioimage suite web*, https://github.com/bioimagesuiteweb/bisweb, GitHub. Retrieved February 2, 2021.

[70] Preti, M. G., Bolton, T. A., and Van De Ville, D., *The dynamic functional connectome: State-of-the-art and perspectives*, Neuroimage **160** (2017).

[71] Shakil, S., Lee, C.-H., and Keilholz, S. D., *Evaluation of sliding window correlation performance for characterizing dynamic functional connectivity and brain states*, Neuroimage **133** (2016).

[72] Leonardi, N., and Van De Ville, D., *On spurious and real fluctuations of dynamic functional connectivity during rest*, Neuroimage **104** (2015).

[73] Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., Della Penna, S., Duyn, J. H., Glover, G. H., Gonzalez-Castillo, J., et al., *Dynamic functional connectivity: promise, issues, and interpretations*, Neuroimage **80** (2013).

[74] Dunlap, W. P., Jones, M. B., and Bittner, A. C., *Average correlations vs. correlated averages*, Bulletin of the Psychonomic Society **21** (1983), no. 3.

[75] Silver, N. C., and Dunlap, W. P., *Averaging correlation coefficients: should fisher's z transformation be used?*, Journal of applied psychology **72** (1987), no. 1.

[76] Bullmore, E. T., and Bassett, D. S., *Brain graphs: graphical models of the human brain connectome*, Annual review of clinical psychology **7** (2011).

[77] Carriere, M., Michel, B., and Oudot, S., *Statistical analysis and parameter selection for mapper*, The Journal of Machine Learning Research **19** (2018), no. 1.

[78] Hajij, M., Wang, B., and Rosen, P., *Mog: Mapper on graphs for relationship preserving clustering*, 2018, arXiv preprint arXiv:1804.11242.

[79] Lum, P. Y., Singh, G., Lehman, A., Ishkanov, T., Vejdemo-Johansson, M., Alagappan, M., Carlsson, J., and Carlsson, G., *Extracting insights from the shape of complex data using topology*, Scientific reports **3** (2013).

[80] Tenenbaum, J. B., De Silva, V., and Langford, J. C., *A global geometric framework for nonlinear dimensionality reduction*, science **290** (2000), no. 5500.

[81] Borg, I., and Groenen, P. J., *Modern multidimensional scaling: Theory and applications*, Springer Science & Business Media, 2005.

[82] Piekenbrock, M., Doran, D., and Kramer, R., *Efficient multi-scale simplicial complex generation for mapper*, 2018.

[83] Estivill-Castro, V., *Why so many clustering algorithms: a position paper*, SIGKDD Explorations **4** (2002).

[84] Gan, G., Ma, C., and Wu, J., *Data clustering: theory, algorithms, and applications*, vol. 20, Siam, 2007.

[85] Silverman, B. W., *Density estimation for statistics and data analysis*, vol. 26, CRC press, 1986.

[86] Piekenbrock, M., Doran, D., and Kramer, R., *Mapper*, 2019, `https://github.com/peekxc/Mapper`.

[87] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E., *Scikit-learn: Machine learning in Python*, Journal of Machine Learning Research **12** (2011).

[88] Tralie, C., Saul, N., and Bar-On, R., *Ripser.py: A lean persistent homology library for python*, The Journal of Open Source Software **3** (2018), no. 29.

[89] Hindriks, R., Adhikari, M., Murayama, Y., Ganzetti, M., Mantini, D., Logothetis, N., and Deco, G., *Can sliding-window correlations reveal dynamic functional connectivity in resting-state fMRI?*, NeuroImage **127** (2016).

[90] Prichard, D., and Theiler, J., *Generating surrogate data for time series with several simultaneously measured variables*, Physical Review Letters **73** (1994), no. 7.

[91] Girvan, M., and Newman, M. E. J., *Community structure in social and biological networks*, Proceedings of the National Academy of Sciences **99** (2002), no. 12.

[92] Chiang, S., Cassese, A., Guindani, M., Vannucci, M., Yeh, H. J., Haneef, Z., and Stern, J. M., *Time-dependence of graph theory metrics in functional connectivity analysis*, NeuroImage **125** (2016).

[93] Ou, J., Xie, L., Jin, C., Li, X., Zhu, D., Jiang, R., Chen, Y., Zhang, J., Li, L., and Liu, T., *Characterizing and differentiating brain state dynamics via hidden markov models*, Brain Topography **28** (2014), no. 5.

[94] Yang, Z., Craddock, R. C., Margulies, D. S., Yan, C.-G., and Milham, M. P., *Common intrinsic connectivity states among posteromedial cortex subdivisions: Insights from analysis of temporal dynamics*, NeuroImage **93** (2014).

[95] Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., and Calhoun, V. D., *Tracking Whole-Brain Connectivity Dynamics in the Resting State*, Cerebral Cortex **24** (2012), no. 3.

[96] Hutchison, R. M., and Morton, J. B., *Tracking the brain's functional coupling dynamics over development*, Journal of Neuroscience **35** (2015), no. 17.

[97] Uddin, L. Q., Clare Kelly, A., Biswal, B. B., Xavier Castellanos, F., and Milham, M. P., *Functional connectivity of default mode network components: correlation, anticorrelation, and causality*, Human brain mapping **30** (2009), no. 2.

[98] Sun, F. T., Miller, L. M., and D'Esposito, M., *Measuring interregional functional connectivity using coherence and partial coherence analyses of fmri data*, NeuroImage **21** (2004), no. 2.

[99] Bastos, A. M., and Schoffelen, J.-M., *A tutorial review of functional connectivity analysis methods and their interpretational pitfalls*, Frontiers in systems neuroscience **9** (2016).

Availability of data and materials

The raw fMRI files are available in the Cam-CAN repository `https://www.mrc-cbu.cam.ac.uk/datasets/camcan` [59].

The processed data and the continuous response measurement are available on request from the corresponding authors [2].

Code to reproduce and document the analyses is accessible online at `https://github.com/aaolaveh/TDA_suspense`

Appendix A

Network	Lobe	Region	x	y	z	Network	Lobe	Region	x	y	z
Salience	Prefrontal	9	28.81	51.16	18.73	Fronto-Parietal	Motorstrip	30	25.25	12.39	49.41
		11	37.63	35.44	31.1			31	39.82	3.48	34.07
		15	6.81	21.45	31.41			164	-23.21	10.67	53.52
		20	37.05	20.8	6.0			165	-45.89	-0.51	49.24
		144	-28.91	50.13	21.65		Parietal	43	31.58	-60.74	49.28
		150	-4.84	17.76	46.21			184	-53.42	-43.52	38.81
		155	-32.47	22.17	5.88		Temporal	69	55.26	-56.28	-4.8
	Insula	36	37.63	21.21	-10.07			70	60.8	-43.38	-17.65
	Parietal	47	54.18	-45.25	37.04	Default	Prefrontal	1	14.08	56.69	-16.68
	Limbic	91	8.38	-39.78	48.1			5	8.28	46.06	-1.67
Fronto-Parietal	Prefrontal	7	30.42	54.8	-3.33			6	14.58	64.75	3.68
		8	44.47	46.1	-4.92			10	8.52	53.28	23.89
		14	40.59	14.59	48.28			12	14.58	36.74	49.11
		19	48.17	35.81	15.18			13	24.03	30.71	36.4
		21	55.34	9.6	22.05			138	-6.76	48.16	-5.73
		22	39.93	17.65	29.11			139	-18.23	56.99	-14.28
		142	-29.28	54.22	2.5			140	-5.81	48.04	11.79
		143	-42.74	47.29	-6.89			141	-11.72	65.14	4.18
		147	-46.09	28.21	26.77			145	-10.03	55.63	30.35
		154	-43.08	41.93	11.01			146	-27.39	34.15	36.31
		156	-53.02	18.47	10.64			148	-11.11	34.29	51.56
		157	-46.21	7.92	28.52			149	-39.36	17.15	46.74

Network	**Lobe**	**Region**	**x**	**y**	**z**
Default	Prefrontal	151	-46.17	28.27	-7.22
	Parietal	42	14.65	-68.45	35.0
		183	-51.32	-56.27	20.49
	Temporal	51	27.18	11.62	-39.15
		52	40.02	18.98	-34.26
		53	52.88	10.85	-21.83
		63	61.87	-23.62	-2.74
		64	56.48	-8.54	-14.32
		185	-37.83	6.06	-37.72
		186	-34.72	18.69	-32.33
		187	-49.52	11.07	-30.45
		188	-49.85	6.4	-15.1
		190	-57.61	-6.41	-22.62
		191	-58.94	-29.97	3.52
		197	-56.94	-14.55	-6.83
		198	-26.65	-42.8	-16.04
	Limbic	83	8.04	34.71	17.2
		85	5.36	-39.11	26.89
		86	12.37	-57.2	18.01
		90	6.33	-57.19	38.09
		93	28.75	-36.86	-0.1
		95	28.02	-28.48	-13.67
		219	-5.79	34.17	26.29
		222	-8.46	-58.91	17.62
		225	-6.46	-54.26	37.38
		226	-8.73	-42.78	50.25
		227	-7.46	-42.13	13.4
		229	-21.52	-36.99	5.77
		230	-32.13	-40.14	-3.96
		233	-20.73	-30.76	-11.18

Table A1: MNI coordinates of the regions from Shen atlas. The regions are numbered between 1 and 268 according to the Shen atlas and are organized by network and lobes.

Appendix B

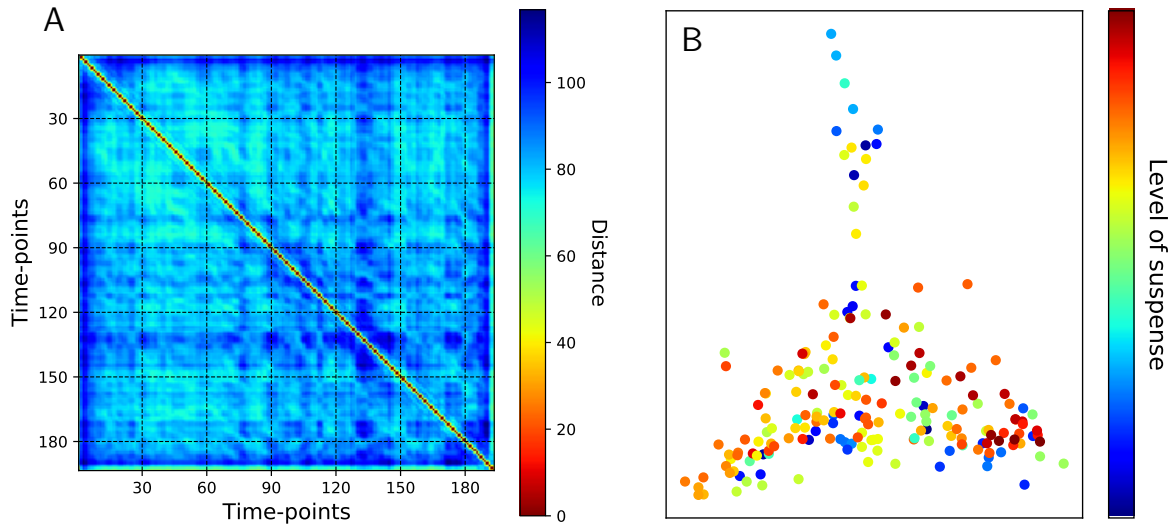


Figure B1: Results for the time-points analysis. **A** Distance matrix of time-points. **B** Isomap filtering. The timepoints are colored with the level of suspense.

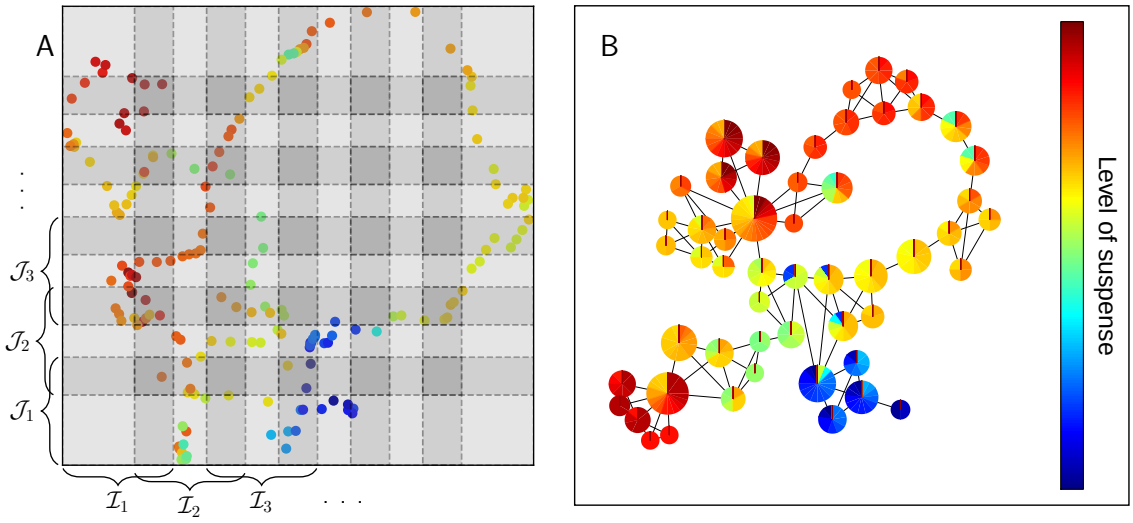


Figure B2: Final mapper graph and its generating cover. **A** Covering of the parameter space. Each dimension is divided in 6 intervals with an overlap of 35% forming a regular cover of rectangles. **B** Final mapper graph. Each node is sized in proportion to the number of contained States and colored using a pie chart by the level of suspense of these States.

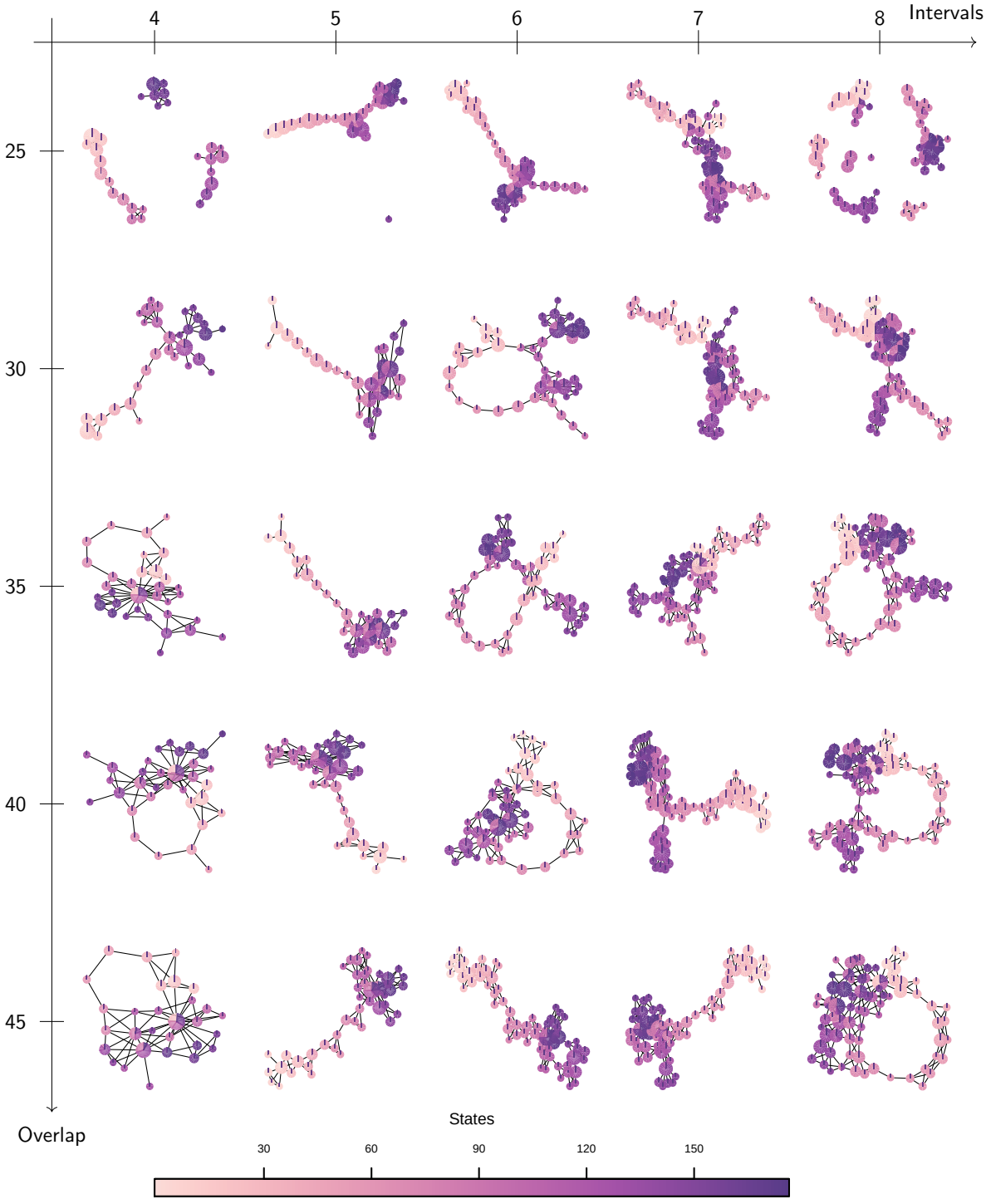


Figure B3: Perturbation of parameters and its effect on shape graphs. We depict graphs for 25 different combinations of the two parameters: number of interval and number of parameters

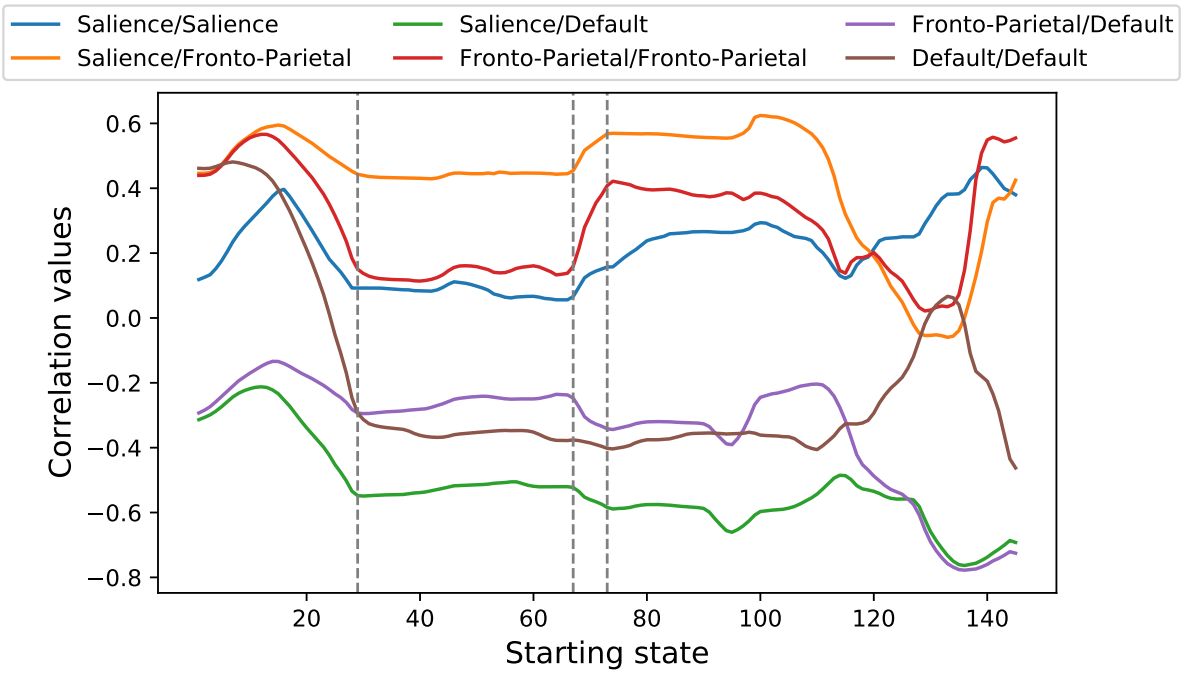


Figure B4: Correlation values of the time series of weights within- and between-networks with the suspense. We took the series from the starting state (ss) i.e we did not consider the States from 1 to ss+1. In addition we marked with a gray dashed line the States 29, 67 and 73

Appendix C

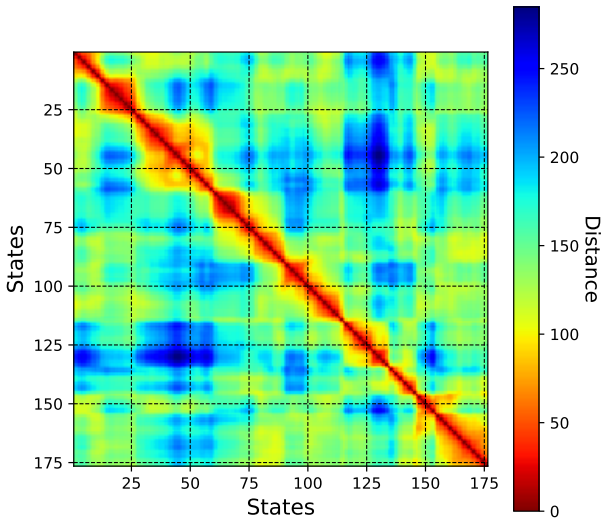


Figure C1: Distance matrix of States

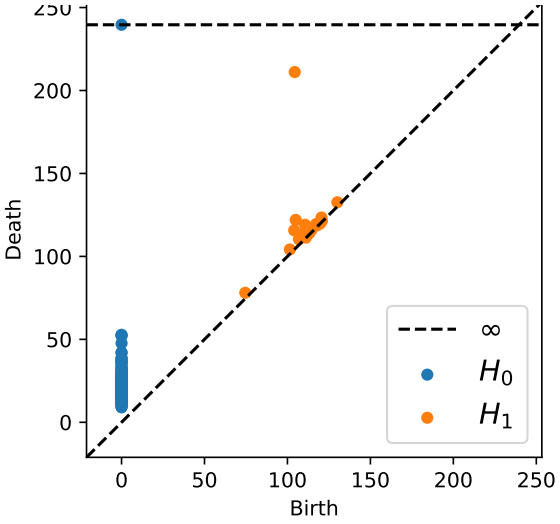


Figure C2: Persistence diagram presenting the homology of the set of States. H_i stands for i-homology.

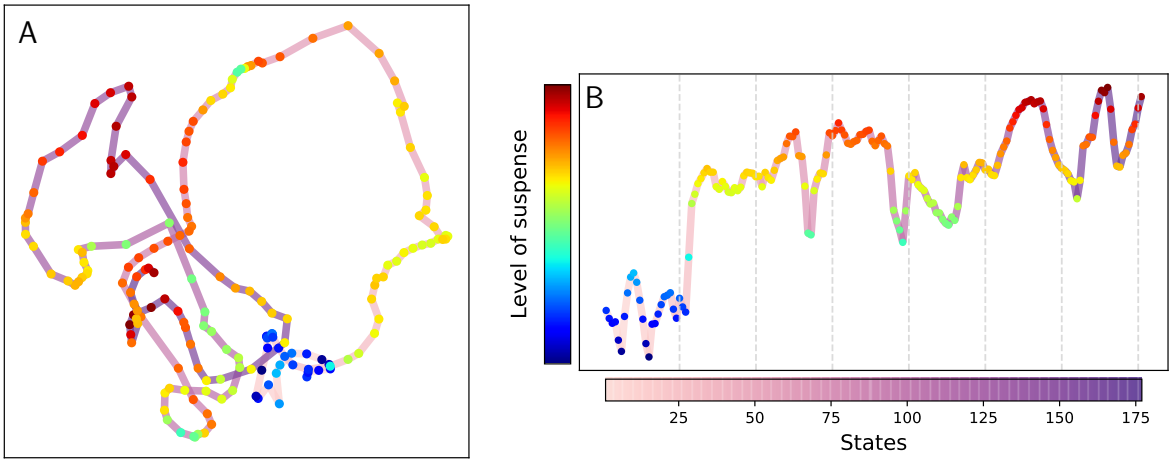


Figure C3: **A** Isomap filtering. The colored States are joined by a degraded pink-purple line representing the flow of time i.e. State i is joined with a segment with State $i-1$ and a slightly darker segment with State $i+1$. **B** Reported level of suspense for the States (see Fig. 3.1)

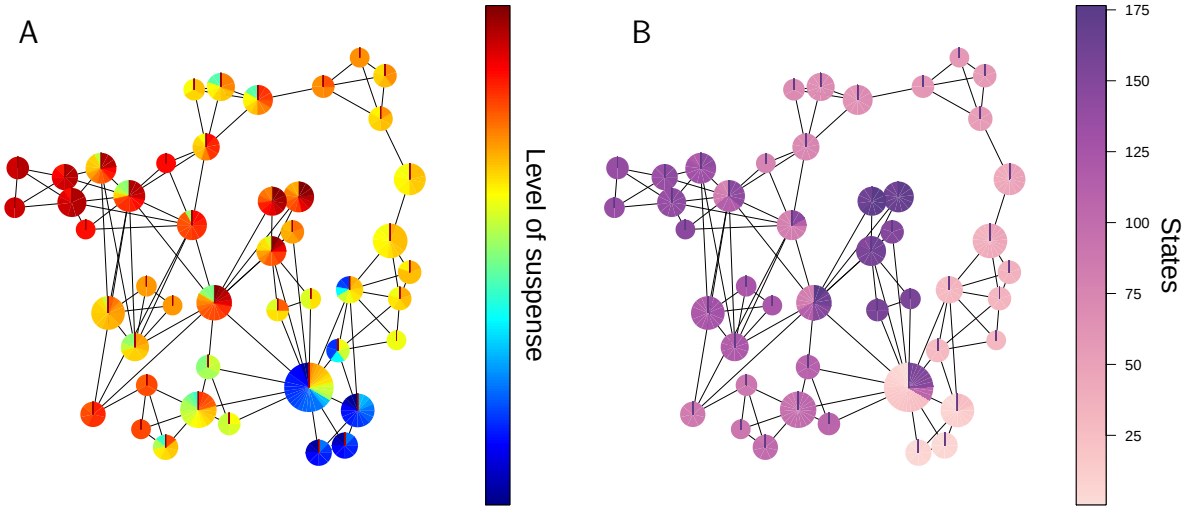


Figure C4: Final `Mapper` graph. Each node is sized in proportion to the number of contained States. The nodes are colored using a pie chart denoting **A** the level of suspense **B** the index of the States