ON GEOMETRIC AND ALGEBRAIC PROPERTIES OF HUMAN BRAIN FUNCTIONAL NETWORKS

by

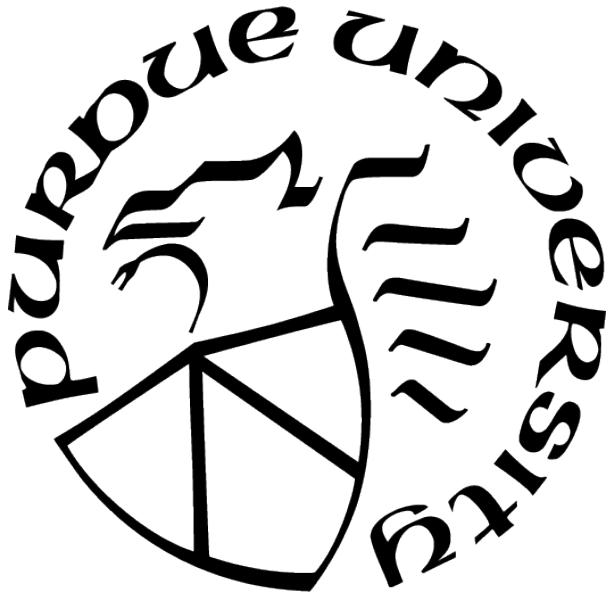
Duy Duong-Tran

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy



School of Industrial Engineering

West Lafayette, Indiana

May 2022

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Joaquín Goñi, Chair

School of Industrial Engineering

Weldon School of Biomedical Engineering

Dr. Mario Dzemidzic

Indiana University School of Medicine

Dr. Mario Ventresca

School of Industrial Engineering

Dr. Thomas Talavage

Department of Biomedical Engineering

University of Cincinnati

Dr. Juan Wachs

School of Industrial Engineering

Approved by:

Dr. Barret Caldwell

Just as the constant increase of entropy is the basic law of the universe, so it is the basic law of life to be ever more highly structured and to struggle against entropy.

Václav Havel

Dedicated to my parents, my sister, and my grandma.

ACKNOWLEDGMENTS

First and foremost, I would like to send the deepest appreciation to my advisor, my mentor, and most importantly, my friend: Dr. Joaquín Goñi. I still recalled the first day I drove from Michigan to Purdue campus and ended up driving back without being able to meet in early Fall of 2016. Joaquín is more than just a PhD advisor to me; for the past four and a half years, he has been a constant source of encouragement to the challenges I faced in my PhD journey and in life. I am deeply touched and motivated by Joaquín's stories, wisdom, and personal experience over the years. It is fair to say that Joaquín is the main source of motivation for me to pursue a career in academia.

I would also like to thank Dr. Mario Dzemidzic for his constant support during my PhD journey. Dr. Dzemidzic has always been willing to spend extra time with me to make sure that I understand concepts in neuroscience and neurology correctly. His passion and dedication are the qualities that I will instill and relay to the future generations in the next chapter of my academic career.

I would like to thank Dr. Mario Ventresca for pushing me to always think and approach any scientific problem in the most scientific way (it might sound simple but it is really not!). I will always remember our conversation about "semi-randomness" in scientific endeavor.

I would also like to express my gratitude to Dr. Thomas Talavage and Dr. Juan Wachs for their guidance on how to position myself effectively as a researcher, a scholar, and a scientific communicator.

I would like to express my sincere thanks to Lawrence Livermore National Laboratory for the incredible opportunities to intern there for two summers and to Dr. Alan David Kaplan for providing the valuable mentorship to me. I have learnt and grown so much during the time at the lab, under Alan's guidance. I will always treasure our lunch-time soccer pick-up games near the lab (and tacos lunch afterwards, of course!).

I would also like to send my sincere thanks to Dr. Li Shen. Dr. Shen has been tremendously helpful and welcoming to provide the opportunities to interact with his lab members, to brainstorm on research ideas, despite the hardship of COVID-19 pandemic. I look forward to keep growing under his guidance and advice in the future.

I would like to also thank my colleagues and friends at the CONN*plexity* Lab: Dr. Enrico Amico, Dr. Kausar Abbas, Dr. Diana Svaldi, Dr. Federico Ramrez Torano, Meenusree Rajapandian, Dr. Uttara Tipnis, Mintao Liu, Dr. Benjamin Chiem, and Angeles Tepper for providing such memorable moments together. These are the moments that I will cherish for a long time to come. I would like to send special thanks to Dr. Amico and Dr. Abbas for interesting conversations over coffee breaks. In the same vein, I would like to acknowledge the Purdue IE Admin team and staff: Dr. Patrick Brunese, Anita Park, Elizabeth Pearsons, Morgan Kuryla, Leza Dellinger for their dedication to Purdue IE students.

Next, I would like to acknowledge the people that I have met during my PhD journey that have taught me true friendship, loyalty. I would like to send my special thanks to an amazing, loving and inspiring person, who truly believes in me and my vision to the future. I am lucky to have met and learnt a lot from you. Your kindness and thoughtfulness have encouraged me to become a better person everyday.

Last but not least, I would not be able to achieve this memorable milestone without the support of my family: my parents (Nam Duong and Phuong Tran), my sister (Anh Duong-Tran). We do have our difference at times, but one thing I learn is that family is always family, no matter what. I deeply thank you for supporting me for all those years.

PREFACE

Background: Magnetic resonance imaging technologies have recently achieved high-quality levels that make comprehensive assessments of individual human brain structure and functions possible. In the mean time, there is an increasing interest in advancing our understanding in *i)* human's unique ways of thinking, reasoning, problem-solving that make up individual human brain structure and function and *ii)* personalized medicine by leveraging whole-brain functional connectivity. The latter is also referred to as clinical utility of functional connectomes[1].

Problem statement: Although there is an increasing body of literature on understanding and quantifying individual level characteristics of human brain function at the whole-brain (i.e., circuit) level, there is a lack of comprehensive understanding in subject-level properties at the functional sub-circuit level, which plays a pivotal role in understanding different phenomenon in human brain function such as consciousness. The goal of this dissertation is to provide the brain connectomics research community with key properties of human brain functional sub-circuits in the context of functional individual fingerprints. I believe that this goal also facilitates the exploration of robust individual-level biomarkers in diverse applications (both healthy controls and neurological and psychiatric disorders).

Approach: I investigated the individuality property of human brain functional sub-circuits through two separate aims:

- As the human brain reconfigures itself from a resting condition to a task (or from one task to another), we assess whether different functional sub-circuits show different levels of fingerprints. Furthermore, can these differences of functional reconfiguration across subjects be quantified?

- When applying a fixed template of an *a priori* set of functional sub-circuits to different individual functional connectomes, we investigate if there are different levels of fitness

[1] A functional connectome is a weighted network as originated by a correlation matrix, where nodes are brain regions and edges are functional couplings between brain region pairs

when mapping a fixed pre-determined functional sub-circuits across individuals. If there is, can they be quantified?

Outcomes: By investigating the first aim, I demonstrated that different individuals did, in fact, have their own signatures in re-configuring their functional sub-circuits as they switch between tasks (or from a resting condition to a task). Further, the level of configuration across resting state and tasks can be formally quantified across individuals. Through such formalism, I showed that the individual level of functional sub-circuit configuration was associated with different cognitive measures such as episodic memory, fluid/general intelligence. By investigating the second aim, I showed that there was, indeed, different levels of fitness when a pre-determined set of functional sub-circuits is mapped onto different individual functional connectomes.

Take-home message: Although the concept of identifying an individual through neuroimaging modalities is still at its infancy, evidence of unique patterns of thinking, problem-solving, and task-performing has been provided at the whole-brain level by many brain connectomics researchers. Nonetheless, whether these unique patterns are propagated to the functional sub-circuit level is not yet fully understood. This dissertation provides further understanding of different characteristics of functional sub-circuits and in turn paves the way to investigate an emerging concept of brain parcellation: *individualized* parcellation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

fMRI	Functional Magnetic Resonance Imaging
BOLD	Blood Oxygenation Level Dependent
dHb	Deoxyhemoglobin
MR	Magnetic Resonance
FC	Functional connectivity
FN	Functional networks
ROI	Region of interest
PCA	Principal component analysis
ICA	Independent component analysis
DMN	Default Mode Network
AD	Alzheimer's Disease
RSN(s)	Resting state network(s)
tp(s)	time-point(s)

ABSTRACT

It was only in the last decade that Magnetic Resonance Imaging (MRI) technologies have achieved high-quality levels that enabled comprehensive assessments of individual human brain structure and functions. One of the most important advancements put forth by Thomas Yeo and colleagues in 2011 [1] was the intrinsic functional connectivity MRI (fcMRI) networks which are highly reproducible and feature consistently across different individual brains. This dissertation aims to unravel different characteristics of human brain fcMRI networks, separately through network morphospace and collectively through stochastic block models.

The quantification of human brain functional (re-)configurations across varying cognitive demands remains an unresolved topic. Such functional reconfigurations are rather subtle at the whole-brain level. Hence, we propose a mesoscopic framework focused on functional networks (FNs) or communities to quantify functional (re-)configurations. To do so, we introduce a 2D network morphospace that relies on two novel mesoscopic metrics, Trapping Efficiency (TE) and Exit Entropy (EE). We use this framework to quantify the Network Configural Breadth across different tasks. Network configural breadth is shown to significantly predict behavioral measures, such as episodic memory, verbal episodic memory, fluid intelligence and general intelligence.

To properly estimate and assess whole-brain functional connectomes (FCs) is among one of the most challenging tasks in computational neuroscience. Among the steps in constructing large-scale brain networks, thresholding of statistically spurious edge(s) in FCs is the most critical. State-of-the-art thresholding methods are largely ad hoc. Meanwhile, a dominant proportion of the brain connectomics research relies heavily on using *a priori* set of highly-reproducible human brain functional sub-circuits (functional networks (FNs)) without properly considering whether a given FN is information-theoretically relevant with respect to a given FC. Leveraging recent theoretical developments in Stochastic block model (SBM), we first formally defined and subsequently quantified the level of information-theoretical prominence of **a priori** set of FNs across different subjects and fMRI task conditions for any given input FC. The main contribution of this work is to provide an automated thresholding method of individuals' FCs based on prior knowledge of human brain functional sub-circuitry.

1. INTRODUCTION

1.1 An *a priori* set of functional networks and human brain functional fingerprint

1.1.1 An *a priori* set of functional networks

Definition 1.1.1. *An a priori set of Functional networks (FNs) are human brain functional sub-circuits that feature consistently across individuals at a resting condition.*[1] *In network science, FNs are closely related to the concept of ground-truth communities of a given complex network.*

The use of Yeo's "functional atlas" (also known as Yeo's parcellation or Yeo's atlas, for short) quickly became a common practice, in brain connectomics research, as an *a priori* identification of functional networks for a given whole-brain functional connectivity profile (or simply Functional connectome (FC)) estimated from an individual.

Definition 1.1.2. *In the field of brain connectomics, the whole-brain large-scale human brain connectivity profile (or Functional connectome) is estimated using a weighted matrix that represents the pair-wise functional couplings between brain regions (typically computed using Pearson's correlation between Blood Oxygenation Level Dependent (BOLD) time series of two respective brain regions). A detailed treatment of brain connectomics is provided in Chapter 2 of this dissertation.*

1.1.2 Individual human brain functional fingerprint

Definition 1.1.3. ***Fingerprint*** *represents unique physical patterns of an individual through the impression of a fingertip of a human finger.*

Definition 1.1.4. *In brain connectomics domain,* ***functional fingerprint*** *[2] represents the idea that an individual's unique cognitive patterns, signatures (e.g., ways of thinking or performing a task) exist and are deeply embedded in the whole-brain functional connectivity profile of such individual.*

[1] ↑Hence, FNs are also known as Resting-State Networks (RSNs).

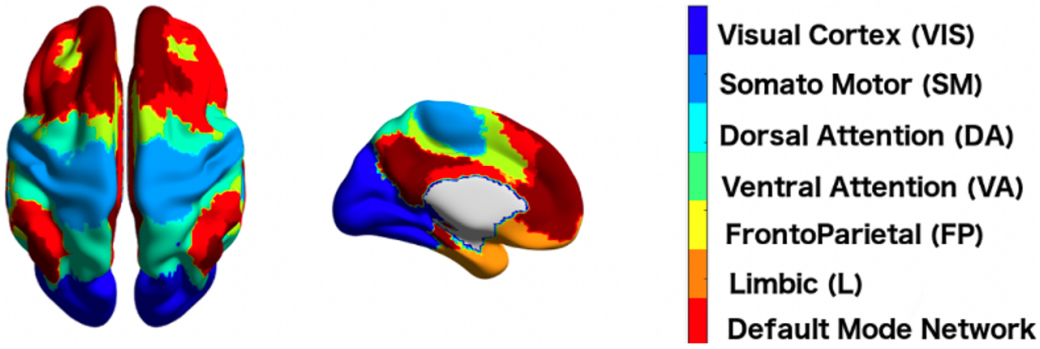


Figure 1.1. An example of Yeo's parcellation [1] of the human brain association cortex into seven distinct functional communities: Visual cortex, Somatomotor, Dorsal Attention, Ventral Attention, Fronto Partietal, Limbic, and Default Model network.

Remark 1. *Cognitive variability is a group level statement that is a consequence of individuals that have unique cognitive patterns, signatures, i.e. ways of thinking or performing a task.*

Definition 1.1.5. *In brain connectomics domain,* ***functional fingerprint identifiability*** *refers to the quest to quantify and measure unique cognitive patterns (e.g., ways of thinking or performing a task) by using individual whole-brain functional connectivity. Fingerprint identifiability relies on the hypothesis that functional fingerprint exists.*

Remark 2. *In brain connectomics research, functional fingerprint is typically assessed by a summary statistics such as identifiability rate (ID rate)*[2] *or intra-class correlation coefficient (ICC)*[3]*.*

Remark 3. *A high level of functional fingerprint, as measured by fingerprint identifiability statistics, elucidates the idea that functional fingerprint can be measured significantly above the level expected by chance.*

[2]↑An identifiability rate is computed using a ratio between correctly identified pairs of functional connectomes of the same individual (typically from two different scans (test and retest)) over the total number of individuals in the dataset.

[3]↑ICC measures the degree to which quantitative measurements within the same groups resembles themselves when they are classified into different sub-groups.

For instance, in a cohort of 100 unrelated subjects where each subject is scanned twice, there is 1% chance to correctly identify an individual. A high level of fingerprint identifiability typically ranges between 80% and 90% of correctly identified pair of the same individual in different scans, see [2]–[4] for further information.

On a large-scale functional connectivity pattern, there is evidence of functional fingerprint:

- across subjects (individual characteristics),
- across tasks (functional reconfigurations),
- across time (temporal fluctuations)

at both whole-brain (**macroscopic**) and functional edge ("**microscopic**") levels [5]. The emergent concept of human brain functional fingerprint has opened many research opportunities to personalized medicine for neurological and psychiatric disorder [6].

1.1.3 The introduction of *individualized* parcellations

A functional atlas such as Yeo's parcellation [1] is derived from a population inference approach. Specifically, parcellation approaches typically use the entire cohort's whole-brain functional connectivity profiles and look for shared properties by using different techniques to derive a **fixed template** that maps brain regions of interest to specific functional networks.

Meanwhile, functional fingerprint supports the concept that that both subject- and task-fingerprint exist at both functional edge (**microscopic**) and whole-brain (**macroscopic**) level [5]. A common, fixed atlas to different individuals engaging in the same task (or the same individual engaging in different tasks) may not optimally assess individual cognitive signatures as represented by whole-brain functional connectivity. This has motivated Salehi and colleagues to introduce the concept of "individualized parcellation" [7], [8] . Specifically, the authors argued that since there was evidence of individual functional fingerprint [2] at the whole-brain (macroscopic) large-scale functional connectivity level, one should also derive different specialized parcellations (specialized atlases) for different circumstances. For instance, one would derive a specialized parcellation for an individual subject at rest as

well as when actively engaging in a cognitive demanding task, (or even among different individuals, doing the same task).

1.2 Research Opportunity

Up to now, the practice of mapping an *a priori* set of FNs onto different whole-brain functional connectivity patterns of an individual (or different subjects performing the same fMRI task, or the same subject performing different tasks) is still much more widely accepted, compared to the practice of individualized parcellations by Salehi and colleagues [8]. Both approaches are well-defined and offer different perspectives in parcellating the human brain cortex into different functional sub-circuits. Nonetheless, due to the enormous computational requirements for computing individualized, and specialized parcellations [8], the fixed parcellation approach (e.g., Yeo's atlas) continues to maintain its popularity and wide acceptance among brain connectomics research community.

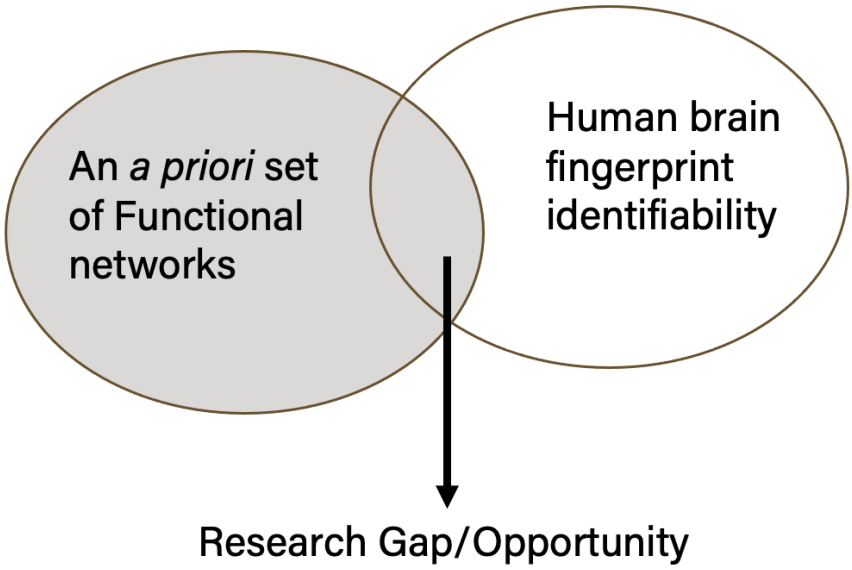


Figure 1.2. Identifying a research opportunity by exploring the individual functional fingerprint characteristics of mesoscopic structures such as FNs in the human brain.

Ever since the concept of human brain functional fingerprint and corresponding evidence of functional fingerprint at the whole-brain (macroscopic) [2], [5] or functional edges/node-

based (microscopic) [5] level were introduced, a gap was present in the brain connectomics research that required a thorough investigation of diverse characteristics of human brain fingerprint at the **mesoscopic** level (e.g. Functional networks).

This knowledge gap has become more critical to address because understanding the functional fingerprint properties of FNs allows researchers to *properly*[4] address research projects involving the mapping of a fixed parcellation (e.g., Yeo's atlas) onto different whole-brain functional connectivity for different individuals or same individual performing different fMRI tasks and resting condition.

1.3 Dissertation Philosophical Statement

This dissertation reflects the extent to which an *a priori* group-level derived set of functional networks exposes individual's functional fingerprint. Specifically, we explore the potential impacts to individual functional variations by mapping group-level derived mesoscopic structures onto individual functional connectomes, which were shown to have high level of functional fingerprint. Individual's functional network fingerprint should be identifiable and measurable at the mesoscopic (functional network) level, separately[5] by constructing a low dimensional phenotypic space and collectively[6] by measuring the level of prominence (of an *a priori* set of functional networks) using stochastic block models. This approach provides further insights into basic functional sub-circuitry behaviors. In this dissertation, the analysis of functional network fingerprint properties addresses on diverse characteristics of human brain functional sub-circuits in the context of *i)* network reconfiguration across resting condition and tasks and *ii)* determining the optimal threshold to an individual's whole-brain functional connectome.

[4]↑By "properly", I refer to the observation that although the mapping of an *a priori* set of FNs is a very common practice in brain connectomics research, the community seems to overlook the importance of such mapping and its potential impacts on whole-brain functional connectivity.
[5]↑By "separately", each FN fingerprint property is considered by itself.
[6]↑By "collectively", an *a priori* set of FNs functional fingerprint is considered together as a pre-established choice.

Remark 4. *Network reconfiguration involves cognitive changes that take place when one switches from resting state to performing a task or from one task to another. In brain connectomics research, these changes are reflected through functional connectivity.*

1.4 Broader Impacts

The quest to identify, characterize, and quantify FN (mesoscopic) fingerprint across different conditions and/or subjects contributes to the field of brain connectomics, especially when

- the practice of mapping a fixed parcellation has deeply embedded in many brain connectomic research studies [9]–[14], compared to a relatively new emergent concept of individualized parcellation as proposed by Salehi and colleagues [8];
- individualized parcellation [8], although promising, poses significant challenge in computational requirement to compute a new, specialized parcellation[7] for each condition. This is not favorable to investigate large dataset in brain connectomics.
- evidence of human brain functional fingerprint existence at both macroscopic (whole-brain) and microscopic (functional edge) levels [5] has been presented [5] but **not** at mesoscopic level such as FNs.

Therefore, this dissertation advocates that in order to support the brain connectomics research community in brain connectomics applications that involve an *a priori* identification of FNs, it is vital to investigate the functional fingerprint properties of FNs.

By accomplishing the dissertation intended impacts, I also believe that investigating FN fingerprint properties would also lay foundation knowledge in creating robust individual-level biomarkers for functional sub-circuits of the brain that are clinically useful for personalized medicine and diagnosis related to neurological and psychiatric disorders.

[7]↑there is a bigger confound than computational shortcoming: states/conditions of networks depend on lot of factors (e.g., hungry vs fed, happy vs sad, to name just a few). These effects could significantly modify specific circuits so condition/state-specific parcellations are necessary.

1.5 Overview of hypotheses, impacts, technical concepts and experiments

1.5.1 Quantifying individual's configural breadth using network morphospace

Background: Human functional brain network (re-)configuration had been shown to associate with highly individually driven cognitive measures such as general and fluid intelligence on the whole-brain (macroscopic) level by the early work of Schultz and Cole [15].

Hypothesis: Differential (re-)configuration capacity between different FNs (e.g., each FN is considered **separately**) is measurable and highly driven by individual functional fingerprint.

Impact: This research endeavor aims to better understand human brain functional subcircuit (re-)configurations and in turn provide a more comprehensive picture to brain network reconfigurations at mesoscopic level.

Technical concept:

- *Current literature:* To address and quantify the FN fingerprint in the context of brain network (re-)configuration, dimensionality reduction techniques are typically employed to collapse the high-dimensional nature of mesoscopic structures in the human brain. Studies pertaining brain network configurations have mostly leveraged techniques such as Principal Component Analysis (PCA) and Singular Values Decomposition (SVD) [16]–[19]. Nonetheless, PCA-based technique is not appropriate to study brain network reconfiguration properties of mesoscopic structures in human brain functional connectivity. In FCs, PCA-based technique requires a complete separation of sub-networks (as represented by blocks of matrix) and hence, as result, interactions among sub-networks are not considered.

- *Contribution:* An alternative approach to PCA-based technique is to use phenotypic space (also known as morphospace). Morphospace is a low dimensional technique that represents forms, shapes, and structures where each of its dimension is characterized by a hypothesized phenotype/characteristic. In order to construct a low dimensional

morphospace, the corresponding phenotypes need to be shown relevant and meaningful to the system/structure under study. The relevance of this method lies in the fact that it offers the freedom in designing the morphospace measures (space parameterization) that are relevant to the system under study (e.g., human brain functional network configuration properties). Using network morphospace, internal (within itself) and external (between itself and other sub-networks) configuration properties of functional brain networks can be considered simultaneously. Specifically, meaningful phenotypes of human brain functional network (re-)configuration need to be defined and subsequently quantified. A detailed exploration of FN fingerprint in the context of brain network (re-)configuration is presented in chapter 3 of this dissertation.

Experiment: The dataset used in this project was the 100 unrelated participants based on the Human Connectome Project (HCP), Q3 release [20]. Further information on this dataset is described in chapter 3 of this dissertation. First, I quantify two relevant phenotypes of human brain functional reconfiguration: **segregation** and **integration** as stated in the literature [16]–[19], [21]. We extend these properties to functional network level by parameterizing a two-dimensional Euclidean space by formally experimenting different designs of Module Trapping Efficiency (**TE**; to quantify FN segregation property) and Module Exit Entropy (**EE**; to quantify FN integration property). Morphospace design quality are assessed by its philosophical meaning (i.e. does it represent segregation property of FN?), its mathematical property (i.e. is it bounded? and does it have robust, well-defined normalizer to provide a fair comparisons between different FNs, which have different sizes?). After parameterizing such space with a specific geometrical design, I tested different designs to quantify and individual functional network configural breadth, which contains two components for each FN: functional reconfiguration, and functional preconfiguration. Different quantification methods are assessed based on its relevance to the neuroscientific literature. For instance, functional network reconfiguration needs to represent the coverage of different fMRI tasks represented in a 2-D morphospace whereas functional network preconfiguration needs to represent the "distance" between rest and task-positive state. Finally, based on

this particular morphospace design, I need to show that network configural breadth has a high level of individual functional fingerprint. All experiments are tested against some null models, which typically involves a random shuffle of FN memberships among the brain ROIs.

1.5.2 Thresholding a large-scale whole-brain functional connectome using stochastic block models

Background: One of the critical steps in constructing large-scale brain networks concerns the thresholding of statistically spurious edge(s) in FCs since state-of-the-art thresholding methods are largely ad hoc. A dominant proportion of brain connectomic research relies on mapping *a priori* set of FNs without properly considering whether it is relevant with respect to a given FC. The recent advancements in the understanding of individual fingerprint and personalized neuroscience/medicine, through brain connectomics research, have challenged the brain connectomics community to investigate the possible impacts of mapping group-level derived FNs to individual FC with high level of functional fingerprint.

Hypothesis: A level of FN fingerprint across different subjects in resting state condition is quantifiable[8] , as FNs are mapped on to an individual FC. As an extension to this central hypothesis, levels of FN fingerprint differ between i) resting condition and fMRI tasks (both group and individual levels); ii) different scenarios such as different scan duration, brain parcellations, and fMRI data processing pipeline.

Impact: The hypothesis was investigated in the context of thresholding the statistical spurious edges in the whole-brain functional connectivity with respect to an *a priori* set of FNs. This assessment is a step towards a comprehensive understanding of how a fixed functional atlas is mapped to an individual functional connectivity profile that are proven to exert high level of functional fingerprint at both macroscopic (whole-brain) and microscopic (functional edge) level.

[8]↑Note that FNs are also referred to as Resting-state networks (RSNs) because they were derived from group-level at resting state condition. Hence, one would expect that the level of FN fingerprint is highest at resting state condition.

Technical concept:

- *Current literature:* The FN fingerprint is investigated within the context of thresholding (eliminating) statistically spurious functional edges in a whole-brain large-scale functional connectivity. This can be viewed as a "signal-to-noise" (SNR) maximization problem where meaningful functional edges are "signal" and statistically spurious edges are "noise." A complex network (graph) can be defined and interpreted as a communicating channel [22], [23].

- *Contribution:* Hence, the problem of eliminating spurious functional edges with respect to an *a priori* set of FNs can be view as the problem of ground-truth community detectability, for a specific graph/network. The problem of eliminating statistically spurious edges in whole-brain functional connectivity can be viewed as a denoising procedure using a fixed ground-truth community as a reference across a threshold range[9]. In other words, for a thresholded whole-brain functional connectivity, there is a distinct level of signal-to-noise ratio (i.e. the ratio between meaningful and statistically spurious functional edges). These signal-to-noise ratios reflect the recoverability of a given ground-truth community structure. This is why choosing an optimal threshold for a given functional connectivity can be effectively viewed as the problem of ground-truth community recoverability for a given network (e.g., higher SNR means there is higher chance of recovering ground-truth communities in the network). In this particular case, "weak" recovery [22], [23] is argued to be the most relevant recovery criteria because there is no golden standard in functional network parcellation of human brain whole-brain functional connectivity [1], [24]–[26]. In chapter 4, a information-theoretic measure, an extended version of Signal-to-noise ratio as described in [22], [23] is proposed to measure FN fingerprint level of prominence and to act as a guiding measure to eliminate statistical spurious edges in the functional connectome. SNR is motivated by the theory of weak recovery of ground-truth community for a given complex network as proposed by [22], [23].

[9]↑In the case of brain connectomics research where functional connectivity is estimated using Pearson's correlation, the magnitude of the functional edges are between -1 to 1.

Experiment: The dataset used in this project was the 409 unrelated participants based on HCP Q3 release [20]. The goal is to address the collective behavior of an *a priori* set of FN fingerprint (and correspondingly, FN fingerprint identifiability) across different individual subjects and/or fMRI tasks at different levels of functional brain network granularity levels.[10] For a given individual whole-brain FC, we extended the SNR measure proposed in [22], [23] to the weighted network case to measure the FN level of information-theoretic prominence (FN functional fingerprint) of such individual. We tested the behavior of SNR across the thresholding range of $[0, 1]$ for a given whole-brain FC and anticipate that the SNR profile to behave non-monotonically across threshold values. We also tested SNR behavior across different tasks and resting condition, at both cohort and individual levels, for a fixed Schaefer granularity level. Finally, we verify the relevance of using SNR as a guiding measure to test the goodness of fit of an *a priori* set of FNs, across different threshold values, to the whole-brain FC by implementing different community detection algorithms and compute the adjusted mutual information between inferred (community structures recovered by community detection algorithms) and ground-truth (e.g., Yeo's atlas [1]) partition.

1.6 Overview of the Dissertation

In chapter 1, the motivation, research opportunity, scope of work and the research questions are stated. In chapter 2, an overview of brain connectomics, as an emergent research field is briefly introduced. In chapter 3, brain network configural properties are studied through the lenses of FNs. In chapter 4, a principled method to threshold whole-brain functional connectivity is proposed using stochastic block models. Finally, chapter 5 introduces an outlook on the future of brain connectomics research. All relevant mathematical notions and background are introduced in the Appendix of this dissertation.

[10]↑ Schaefer's parcellation [26] has different granularity levels with increasing number of brain regions for the same individual with a fixed fMRI task or resting condition.

2. BRAIN CONNECTIVITY: AN INTRODUCTION

2.1 Quantum mechanics of Magnetic Resonance Imaging

Preface. Nuclear Magnetic Resonance (NMR) was first introduced by the seminal work of Purcell [27] and Bloch [28] in 1946. In 1973, Lautebur [29] and Mansfield [30] leveraged NMR principles in describing a technique to determine physical structures. This has marked the beginning of Magnetic Resonance Imaging (MRI). Since 1973, MRI has been utilized in diverse contexts, including but not limited to biomedical, and engineering applications.

Quantum Mechanical Description of NMR. The spin angular momentum is effectively predicted by the quantum mechanical description of atomic nuclei [31]. The spin angular momentum is quantitatively described by the spin quantum number I. The nucleus must have the spin quantum number, denoted as I, in order to possess the property of magnetic resonance. In terms of biomedical and medical applications, the proton is the most relevant nucleus to study as it is highly abundant in nature.

Quantum mechanics of nuclei spin. Each nucleus has a fixed nuclei spin property that is either an integer or a half integer. The component to the nuclei spin, denoted as m_I, is parallel to the $z-$ axis and has $2I+1$ values $\{I, I-1, ..., -I+1, -I\}$. The magnitude of the spin angular momentum is given by:

$$|\mathbf{P}| = \hbar\sqrt{I(I+1)}$$

where $\mathbf{P}$ is a vector with the $z-$ component, when applied a long the $z-$ axis in a magnetic field, is

$$P_z = \hbar m_I$$

The nuclei have a magnetic moment μ that is proportional to $\mathbf{P}$:

$$\mu = \gamma \mathbf{P}$$

A nuclei with a constant proportionality, γ, is defined as the gyromagnetic ratio. In biomedical and medical applications, the proton has a very specific, fixed magnetogyric ratio $\gamma = 2.675 \times 10^8\ rad/s/T$.

Property of proton spin. Whenever a single proton spin is measured, it can only be measured in two orientations: parallel and anti-parallel. The quantum state of a proton can be compartmentalized into linear combinations of parallel and anti-parallel states. In the presence of a magnetic field with strength $\mathbf{B_0}$, the protons precess at Lamor frequency, which is determined by the magnetogyric ratio and the strength of field. This is also referred to as bulk magnetization due to nucleus spins precession. The angular momentum $\mathbf{P}$ precesses about the external field axis (e.g., $z-$ axis with a specific angular frequency called the Larmor frequency. The static field that is most commonly used in MRI applications causes a precession corresponding to a specific photon called the radiofrequency (RF) photon.

The parallel and anti-parallel spins of protons. The longitudinal net polarization is existence of a small proportion excess of protons that are in a lower energy state. When an RF pulse is applied, this net polarization was tipped sideway (i.e., perpendicular to the magnetic field vector) or reverse (i.e., 180^0 pulse). The protons will become in-phase with the RF and, as a direct consequence, with each other, as well.

2.2 Functional magnetic resonance imaging

The advent of neuroimaging techniques (modalities) has significant improved our understanding of brain functions and structure. Magnetic resonance imaging (MRI) is considered one of the most important imaging modalities to explore human brain both structurally and functionally. In this dissertation, the key focus is on functional MRI (fMRI) which has proven to be leading data source to unravel human brain functions. This chapter aims to provide an overview of Magnetic Resonance Imaging (MRI), and functional Magnetic Resonance Imaging (fMRI) that are relevant to construct whole-brain functional connectivity. A more thorough, detailed description of MRI physics can be found in the textbook by Hornak [32] on MRI basics.

2.2.1 A quick overview on the physics of Magnetic Resonance Imaging

According to Hornak [32], MRI is a technology that uses to produce "high quality images of the inside of the human body." As a starting point, MRI technology used tomographic imaging technique which "produced an image of the nuclear magnetic resonance signal in a thin slice through the human body", according to [32].

The MR signal measures the difference between the energy transition of an unpaired proton in a nucleus (most commonly hydrogen) from a relaxed, low-energy state (as the subject enters the scanner at static field) to the post-application of the magnetic field (excited phase). The static magnetic field strength can be controlled and measured in the unit of Tesla. The magnetic field preferentially aligns proton spins in hydrogen nuclei with the main pole of the magnet. A radio frequency pulse is introduced to excite the nuclei into a higher energy state. While returning to the ground state, a proton emits a photon, and this electromagnetic radiation is subsequently detected. The key data extracted from this process is the relaxation time, e.g. the two exponential processes describing the elapsed time between nuclei excitation to relaxation in the direction of a static B0 field (for the construction of T1 image) and perpendicular field (to construct T2 image). Both times are measured through coils (sensors) placed insider the scanner.

The key principle of MRI is the resonance equation in which the resonance frequency, denoted as v, is proportional to the static magnetic field, B_0: $v \sim B_0$. Another important concept is the gradient magnetic field generates a variation in the magnetic field with respect to each region of spins are oriented randomly [32]. For instance, in the absence of a magnetic field, hydrogen protons spin randomly. As soon as a magnetic field is introduced (e.g., B_0), proton spin tend to align with it (either parallel or anti-parallel). The net magnetization vector (NMV), which represents the collective behaviors of hydrogen protons orientations start to form an angle with B_0. Then, a radio frequency pulse is applied to excite the NMV to flip it towards the transverse plane. When the pulse is turned off, the longitudinal component (i.e., T1) recovers. The imaging contrasts emerges because different tissues will have different T1 responses (i.e., recovery times).

2.2.2 The Blood Oxygen Level Dependent Signal in functional magnetic resonance imaging

The T2-contrast images are of critical importance to fMRI data, e.g. the transverse relaxation. Ideally, homogeneous magnetic field results in an exponential decay for the transverse relaxation time. The time constant of such function is called T2. In reality, the inhomogeneity of magnetic field yields a different rate, denoted as T2*.

The T2* constant has close ties with mechanism measuring the Blood Oxygenation Level Dependent (BOLD) contrast. Specifically, it is sensitive to relative ratios of oxygenated and deoxygenated blood supply that depends on local neural activity. Deoxyhemoglobin (dHb) is para-magnetic and hence, influences both MR signals and T2* constant [34]. The neuronal activity induces into dHb changes that are quantified through **hemodynamic response function** that peaks *approximately 4-5 seconds later*. Different brain regions could bear different lags in BOLD signals. The hemodynamic response provides a reliable measure of neural inputs to relevant areas of the brain and their corresponding processes [35]. In this dissertation, whenever fMRI is mentioned and used, it implies the use of the BOLD signal. In short, BOLD fMRI demonstrates the changes in deoxyhemoglobin concentration induced by neural activity under either a "task-induced condition or spontaneous modulation of neural metabolism" according to [36].

2.2.3 Resting-state as a functional magnetic resonance imaging condition

Since the invention of fMRI, it took more than a decade for studies on co-activation of spontaneous resting-state fMRI time-series to pick up pace with pioneer work by Biswal and colleagues [37]. In essence, they showed that the brain neuronal activities at rest were not idle, but rather collectively spontaneous and highly correlated among many brain regions even without the introduction of any task-evoked condition [38].

Interestingly, low frequency resting-state fMRI BOLD time-series with frequency from $0.01 \sim 0.1\ Hz$ are shown to be a robust proxy reflecting spontaneous neural activity at a whole-brain level. This has allowed a growing proportion of neuroscience literature to study functional coupling among brain regions using resting-state BOLD signals. In the beginning

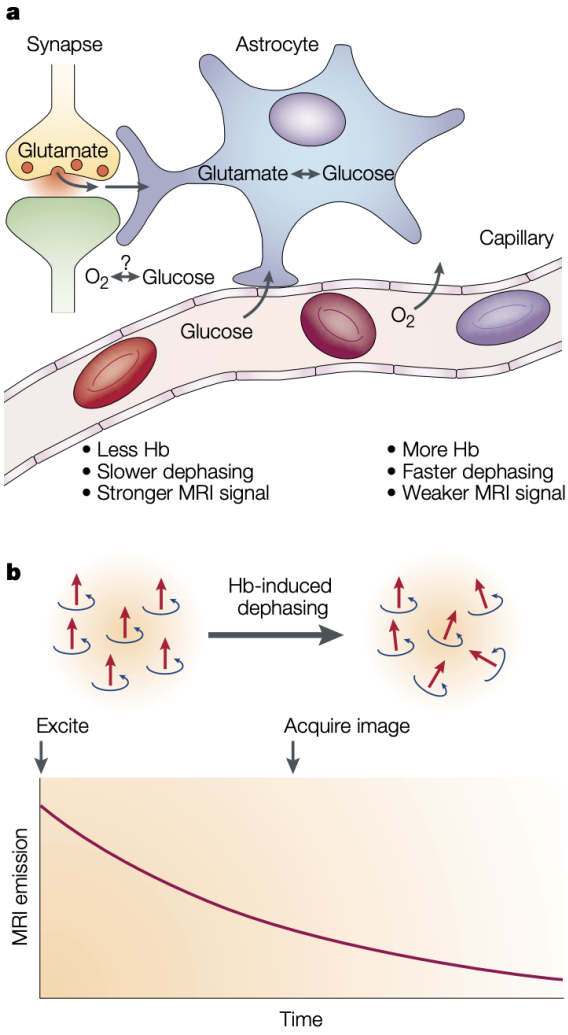


Figure 2.1. The illustration of fMRI BOLD contrast mechanism. In the presence of a magnetic field, hydrogen atom in water molecule get excited by a specific characteristic radio frequency (RF). After such excitation, hydrogen nuclei emit a similar RF until they gradually return to their equilibrium (low energy) state. BOLD signal contrasts measures changes of blood oxygenation, in part, resulting from the inhomogeneous magnetic field intensity shift. Panel (a) shows the synaptic activities neurotransmitter recycling according to some metabolic demand. Panel (b) shows deoxyhaemoglobin effects on fMRI image acquisition. Adapted from [33].

of resting-state fMRI studies, there had been ongoing debates whether resting-state fMRI time-series recorded at different brain regions resulted from physiological processes such as respiratory. Since the early work of Biswal and colleagues, a fast-paced growing body of

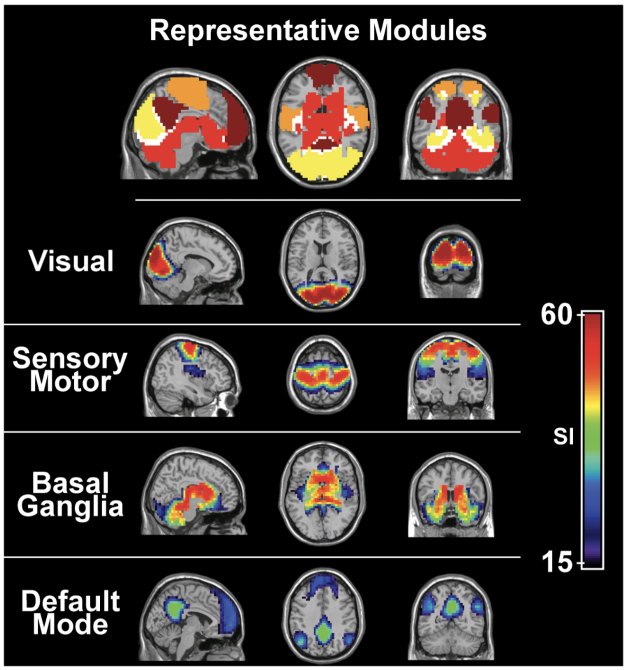


Figure 2.2. A study led by Moussa and colleagues was among one of the first manuscript identified network modules (brain functional sub-circuitry) that are highly reproducible across subjects. Specifically, four modules were identified: visual (yellow), sensory/motor (orange), basal ganglia (red) and Default Mode Network (blue/green). Figure is adapted from [39].

literature emerged supporting the evidence that resting-state fMRI measure spontaneous neural activities [38], [40]. PET studies has also confirmed the neural source of LF BOLD fluctuations. This had laid a foundation for subsequent studies to leverage BOLD fMRI signals as a robust measure of resting-state functional connectivity, at resting state, among brain regions from a whole-brain level.

One of the most monumental works on resting-state BOLD fMRI signals was published in 2011 by Thomas Yeo and colleagues [1]. In this manuscript, the authors proposed the concept of intrinsic functional connectivity MRI networks (fcMRI networks or functional network (FN) for short). These networks which are brain functional sub-circuits distributed across the cortex, are introduced explicitly in the later sections.

2.3 From functional Magnetic Resonance Imaging to Functional Connectome

2.3.1 The functional connectomes at different scales

It is widely accepted that neurons are the basics building blocks of the nervous system per neuron doctrine by Ramón y Cajal and Golgi [41], [42] and others. Nonetheless, tracking neuronal behaviors for human brain in the current stage poses tremendous challenge as acquired data is rather noisy and hence, difficult to process and analyze. As mentioned in the previous section, the increasingly rich repertoire of literature on resting-state fMRI has provide solid ground to use BOLD fMRI signals as a robust measure of human brain functions and structure. Such advancement has opened doors for investigation of functional and structural connections in the human brain, modeled as networks as a new field of Brain Connectomics (or equivalently Network Neuroscience). In brain connectomics field, the construction of functional connectome (a proxy representation of human brain's neuro-physiological activities that is approximated using network[1]) can be divided into three distinct scales:

- Microscale: each neuron is treated as a single node (vertex) in the network;
- Mesoscale: nodes are defined as functionally specialized cell assemblies or neuron populations; [43];
- Macroscale (large scale): nodes are defined using spatially connected voxels in a fMRI dataset.

The current state of neuroimaging allows the construction of macroscale functional connectomes with proven usefulness in subsequent analyses and reasonable processing time. Omitting the details of starting from a microscopic level, this dissertation uses macroscale whole-brain functional connectivity for all research endeavors. From this point, FC construction is referred to the large-scale (macroscale) estimation of functional connectivity.

[1]↑Note that the term network used in this chapter and beyond is referred to the concept of complex network introduced in Chapter 1 of this dissertation.

2.3.2 The estimation of whole-brain large scale brain networks

One of the first tasks in converting neuroimaging fMRI[2] data into network scientific data (for further analysis) involves the definition of nodes. In large-scale whole-brain network, nodes are referred to as brain regions of interest (ROIs), formed by agglomerating spatially adjacent voxels. Such process is implemented by co-registering individual brain images with an anatomically parcellated template image. The BOLD signals of individual voxels registered in the same brain ROI are averaged to form the nodal BOLD time-series. The mean of voxels' BOLD contrasts alleviates the noise, hence, provides a better estimation of neuro-physiological signals in the brain. The T1-weighted image is typically used to anatomically register brain ROIs at a voxel level. Early proposals of parcellation schemes for brain connectivity analysis are made by Desikan with 68 brain regions [44] and Destrieux with 148 ROIs [45] implemented in the Freesurfer software. Over the past decades, several

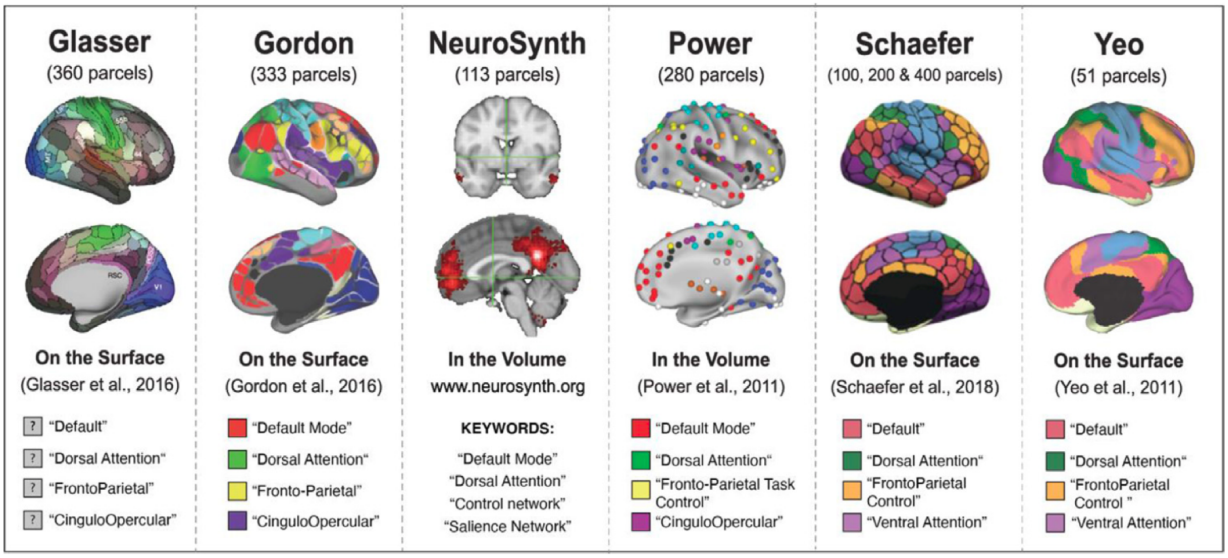


Figure 2.3. Different choice of parcellations leads to different partitions of corresponding functional networks. Note that the choice of node parcellations can be independent of functional network (FN) parcellations. For instance, one can choose Schaefer node parcellation and Yeo's set of FN. Figure is adapted from [46].

template, parcellating cortical regions of the brain, were proposed such as Glasser (360 nodes)

[2]↑Other modalities involves: EEG, MEG, NIRS.

[20], Power (280 nodes) [24], Yeo (51 nodes) [1], and most recently Schaefer with range from 100 to 1000 nodes [26]. It is important to note that the choice of a particular parcellation is critical to further divisions into brain functional networks (or equivalently communities). There is recent effort in revisiting parcellation choices in translational research because *a priori* brain parcellations to study individual differences relies on certain assumptions such as similar network properties across subjects, or that particular choice of parcellation does not impact the analysis results [46]. The key advantage of using anatomical parcellation template is that it supports direct comparisons of results to prior studies, leveraging the same or similar template.

Once the nodes are selected, the network is fully constructed by formally defined the so-called functional couplings. There are a few approaches when it comes to quantifying the coupling level of two brain regions. The most widely used (perhaps the first) measure of functional connectivity is through Pearson correlation, which estimates the level of synchronization among pairs of time-series data in which high correlation (close to positive one) indicates in-phase coupling. As much as it is widely used and accepted, the usage of Pearson correlations are not without its shortcomings. In certain cases, the usage of mutual information among two brain regions guarantees that the weighted network has no negative functional edges.

2.4 Fingerprints in functional brain networks

The explosion of publicly available neuroimaging data [47]–[49] coupled with an increasing evidence of using BOLD fMRI data as a reliable, robust measure of brain activity [50]–[52] has allowed brain connectomics researchers to positions themselves to address fundamental questions in computational neuroscience using functional and structural connectomes as a proxy for neuro activity. Among those quests lies a very fascinating question: does neuroimaging data contain unique cognitive signatures of individuals, either at resting condition or during fMRI task performance? In the early work of Schultz and Cole, the authors associated general intelligence scores using spatial correlation between FC at rest and fMRI tasks [15]. Intelligence along with other cognitive measures such as verbal or episodic mem-

ory are highly individual-driven. Hence, the work in [15] demonstrated high potential of investigating individual cognitive signatures using whole-brain functional connectivity.

Recently, FCs have been shown to exert robust and reproducible individual fingerprint [2]–[6], [53]–[58]. In 2015, Finn and colleague were among the pioneers to put forth the concept of "brain fingerprint" [2]. In this seminal work, the authors provided evidence that one can effectively recognize an individual from others using a measure as defined as "identifiability rate" (ID rate). Conceptually, ID rate is a very simple, yet effective, measure that can be computed using Pearson correlation between vectorized FCs of an individual, obtained from the scanning session (either from different or the same imaging session). Specifically, the framework started with a sample database of FCs, the authors show that, to some extent, a target FC can be effectively identified by matching itself with one of the FCs in the database by using spatial correlation measure. The success rate of identification was above 90% for resting-state and between 54% to 87% for other conditions such as rest-task or task-task session comparisons. This early work from Finn and colleagues has opened up a whole new array of opportunities for brain connectomics researchers. The work was monumental because it has paved the way to shift the field focus from population level (group average) inference to individual inference (e.g., examining how individual networks are functionally and structurally organized in unique ways). This approach lays the solid foundation to study individual phenotypes in both healthy and disease participants. The proposed ID rate, nonetheless, does have shortcomings. Firstly, the rate is heavily dependent on the number of fMRI scans (e.g., the rate drops as the number of scans increase). Secondly, it is a binary measure (e.g., if subject i test session FC_{i}^{test} has a spatial correlation of 0.5 with one FC in the database and 0.49 with 10 other FCs (in the database), the ID rate will match the target FC with the FC with highest spatial correlation). In other words, there is no continuous degree of identification. Such technical shortcoming is sensitive to the quality of fMRI data input, especially with respect to head motion, scan sites, and others [59]. Thirdly, and most importantly, the initial proposal of Finn and colleague does not suggest a method to improve the identifiability rate for a given neuroimaging dataset.

To improve identifiability in human functional connectome, the identifiability of functional fingerprint was proposed as a objective function named differential identifiability, I_{diff}.

The framework was based on a data dimensional reduction technique described in Amico et al. [5]. Specifically, the framework leverages an optimization approach and is motivated by the idea that functional connectomes of the same individual should look more similar to themselves, compared to other individuals' (others) FCs.

$$I_{diff} = I_{self} - I_{others}$$

The identifiability matrix, denoted as $\mathbf{I} \in R^{+}_{n \times n}$, is task-based for which $I_{\mathrm{ij}} \mid \mathrm{i,j} = [n]$ represents the similarity - measured by Pearson Correlation between individual i (Test) and j (Retest) sessions' vectorized upper-triangular (functional) connectome matrices (under original and reconstructed conditions). For rest or any given task, $I_{\mathrm{ii}} \mid \mathrm{i} \in [n]$ is measured by 2 visits (test and retest) for subject i. Moreover, I_{self} and I_{others} are the average of diagonal and off-diagonal entries, respectively.

$$I_{self} = \langle \mathbf{I}_{\mathrm{ii}} \rangle \forall \mathrm{i} \quad ; \quad I_{others} = \langle \mathbf{I}_{\mathrm{ij}} \rangle \forall \mathrm{i} \neq \mathrm{j}$$

The objective function is maximized, discretely, by deleting one principle component (PC) at a time, starting from the one with least explained variance and, subsequently, reconstruct the functional connectomes based on the remaining Eigen modes, denoted as PC. The original FCs estimated from the cohort is named original FC, which corresponds to a original score of $I_d\mathrm{i}ff$. The constructed FCs are estimated by removing the PC; for each time a PC is removed, a new $I_d\mathrm{i}ff$ score is computed. The cohort's FCs are then reconstructed, namely FC_{recon} as follows:

$$FC^{k}_{Recon} = \mu^{k} + \sum_{\mathrm{i}=1}^{k} w^{k}_{\mathrm{i}} PC_{\mathrm{i}}$$

where w^{k}_{i}'s are weights corresponding to PC_{i}'s. In each step, reconstructed FCs are mapped from connectome space to identifiability score space. Hence, k is found by computing $argmax_k\,[I_{diff}]$.

2.5 Resting state functional networks

The human brain is an extremely complex multi-scale system whose interactions among smallest elements such as neurons give rise to complex behavior (e.g., cognition). Besides exhibiting the hierarchical structures across different spatial scales (e.g., micro-scale where each neuron is treated as an individual node), human brain functional organizations, for a fixed scale, also display "modular" characteristic. Specifically, the human brain can be decomposed into independent, yet highly interacting modules (or communities) [60].

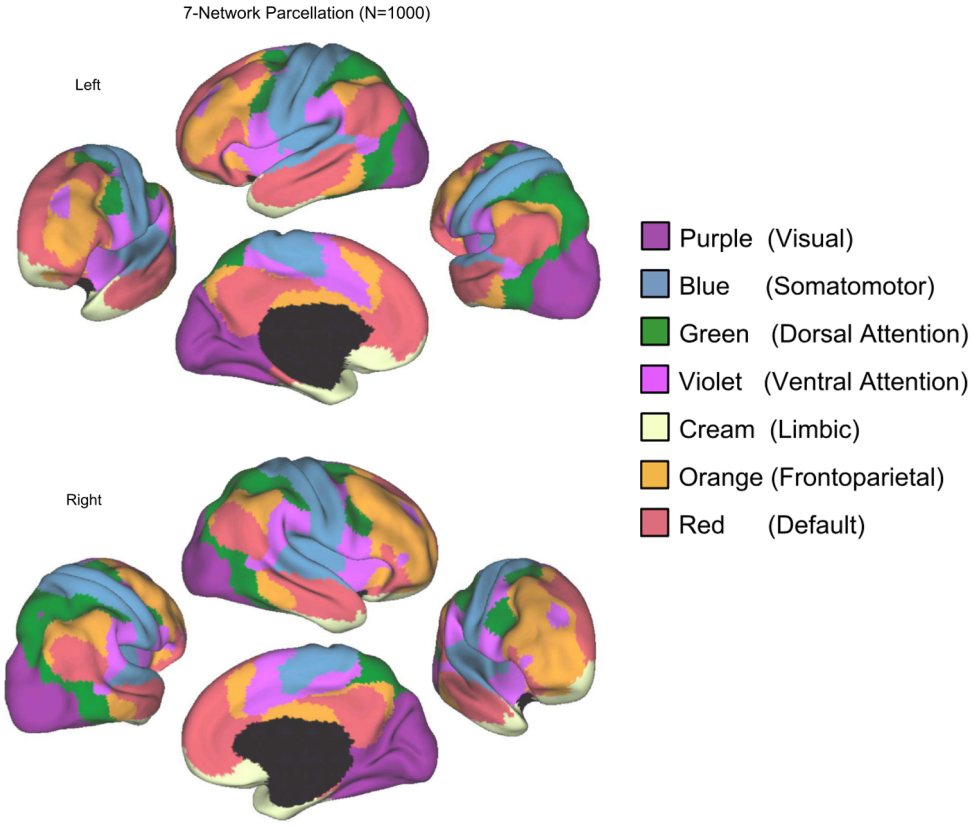


Figure 2.4. Coarse-grained parcellation of human brain cortex into seven FNs. These functional circuits are consistently reproduced across resting state fMRI scans from 1000 participants. Figure is reproduced in courtesy of [1].

The modular characteristic of human brain is one of the backbone mechanisms that allows human brain function to adapt flexibly with diverse cognitive demands. Functional brain modularity is also an important tool to explain brain complexity (e.g., cognition as an emergent property of complex systems). The modular characteristics of human brains were

also noted in a very important work by Bullmore and Sporns [61] reporting that the human brain can be sufficiently characterized into "modules" whose elements (e.g., nodes/vertices in network-scientific sense) are contributed by different distributed areas across the cortex.

Yeo and colleagues in 2011 [1], see figure 2.4, put forth a very important concept: intrinsic functional connectivity MRI (fcMRI) networks or functional networks (FNs) for short. These FNs are indeed **parallel distributed circuits** extracted from noninvasive imaging techniques such as fMRI. Specifically, Yeo and colleagues noted that: "An intriguing possibility is that the majority of the human cerebral cortex involves multiple parallel circuits that are interdigitated throughout association cortex such that each cortical lobe contains components of multiple association networks."

Some of the noted approaches to identifying different FN sets include Power et al. [24], Glasser et al. [62], Gordon et al. [25], and most recently Schaefer et al [26]. The establishments of different atlases are, in part, due to the implementations of different approaches and techniques. Some of the most common one are meta-analysis of intrinsic functional connectivity patterns [24], [63], multi-model approach [62], functional edge detection approach [25]. More exhaustive review can be found in the work by Bryce et al. [46]. An important note is that different parcellations yield different assignments of human brain regions of interest (ROIs) to one particular (or occasionally multiple) FN(s). For instance, a given brain region might belong to the default mode network in Glasser's [62] but not in Power's parcellation [24].

An *a priori* identification of human brain FNs creates a *template* to reveal different executive functional organization in cognitive, developmental, healthy or neurodegenerative disease research [46]. After applying a functional atlas (e.g., a guidance to which brain region(s) belong to which FN(s)), researchers have a **baseline reference** for the physiological, functional, individual differences of the same FN across different conditions or different FNs across the same task. Specifically, the uses of an *a priori* set of FNs allow examination of *i)* the functional differences among individuals under different cognitive conditions [9], [10]; *ii)* aging condition [9], [11], [12]; *iii)* psychopathology or neurological dysfunctions [13], [14]. A comprehensive review on the practice of *a priori* set of FN mappings can be found in the work of Bryce and colleagues [46].

Importantly, as shown in a study by Amico and Goñi [5] per figure 2.5, functional edges as sorted by Yeo's FNs also exert different level of fingerprint (task- and subject-). Specifically, some FN such as the visual cortex contains more functional edges with more task- than subject- fingerprint; other FN such as the default mode network contains more functional edges with more subject- than task- fingerprint.

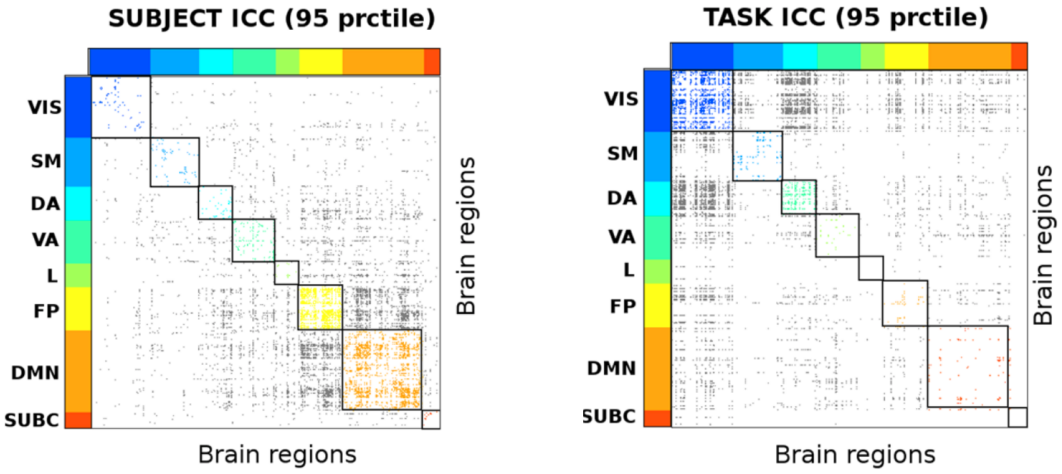


Figure 2.5. Subject Intra-class correlation (ICC) plot highlights functional edges for which there is more subject variability than task variability, i.e. functional coupling is more determined by who is the subject than what the subject is doing. Task ICC Plot highlights functional edges for which there is more task variability than subject variability.

Although FNs open many promising opportunities to explore diverse brain network functions, they are not unique. In other words, there is no "golden standards" regarding how to parcellate the association cortex into distinct parallel circuits. Nonetheless, in this dissertation, Yeo's association cortex parcellations [1] into distributed functional circuits are reviewed and utilized in the subsequent sections. These circuits are, from this point on, referred to as Yeo's functional networks (FNs) or Resting-state-networks (RSNs), see figure 2.4 for further details.

3. A MORPHOSPACE FRAMEWORK TO STUDY FUNCTIONAL BRAIN NETWORK CONFIGURATIONS

3.1 Preliminaries on Network Morphospace

3.1.1 Literature Review

In general, **morpho-** (in morphology) is the study (*-ology*) of living (biological) shape (*morph-*). The motivation of this field is to disentangle the complex taxonomic relationships between, for instance, species or biological entities, see figure 3.1 as an example of a phenotypic space.

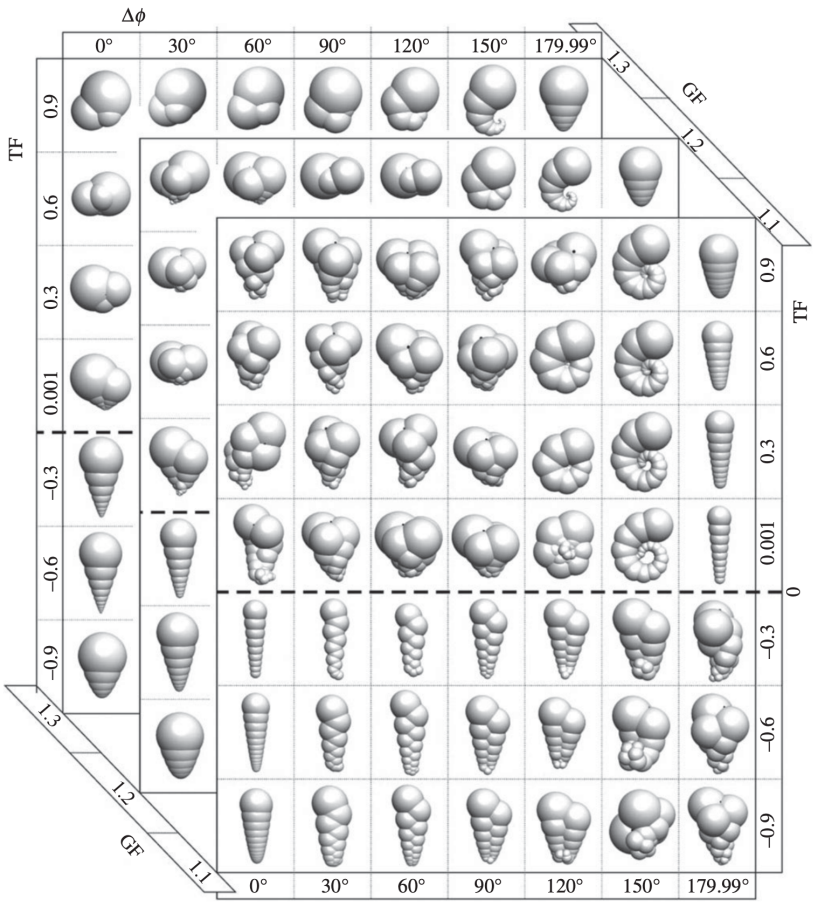


Figure 3.1. An example on how to construct a three dimensional phenotypic space, representing three distinct external features of foraminiferal. In such case, those phenotypes are deviation angle, translation factor, and growth factor. Reproduced in courtesy of [64].

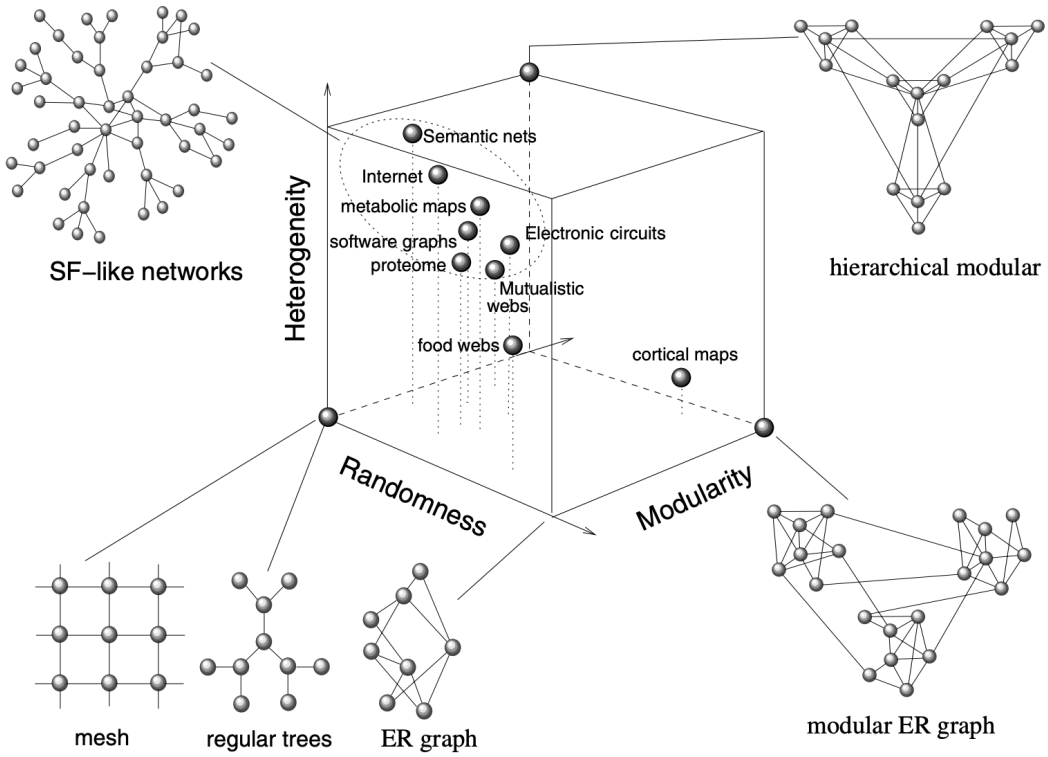


Figure 3.2. A phenotypic hyperspace is constructed to characterize different types of complex networks using three parameters: randomness, heterogeneity and modularity. Figure is adapted from [66].

McGhee first proposed the concept of theoretical morphospace in [65]. As such, it is constructed using *N-dimensional geometric hyperspaces*. An example of such space is given in the figure 3.2. In the last two decades, with the unprecedented growth pace of network science there is an emergent need to study the origins (or subtle similarity/difference in topological features) of networks (or networked systems). Such interesting interplay ultimately gave rise to the interdisciplinary field called **Network Morphospace**. Here, a new term is added, i.e. *-space* which can be understood as mathematical space, such as vector space (as introduced in the Appendix 5.5).

Network morphospace can be characterized as a comprehensive/quantitative description of complex systems, i.e. networked systems, using low (finite) dimensional space. The literature on network morphospace is active with exciting efforts from multiple disciplines such as i) language [67], ii) efficiency in communication [68], origins of hierarchy [69] for complex

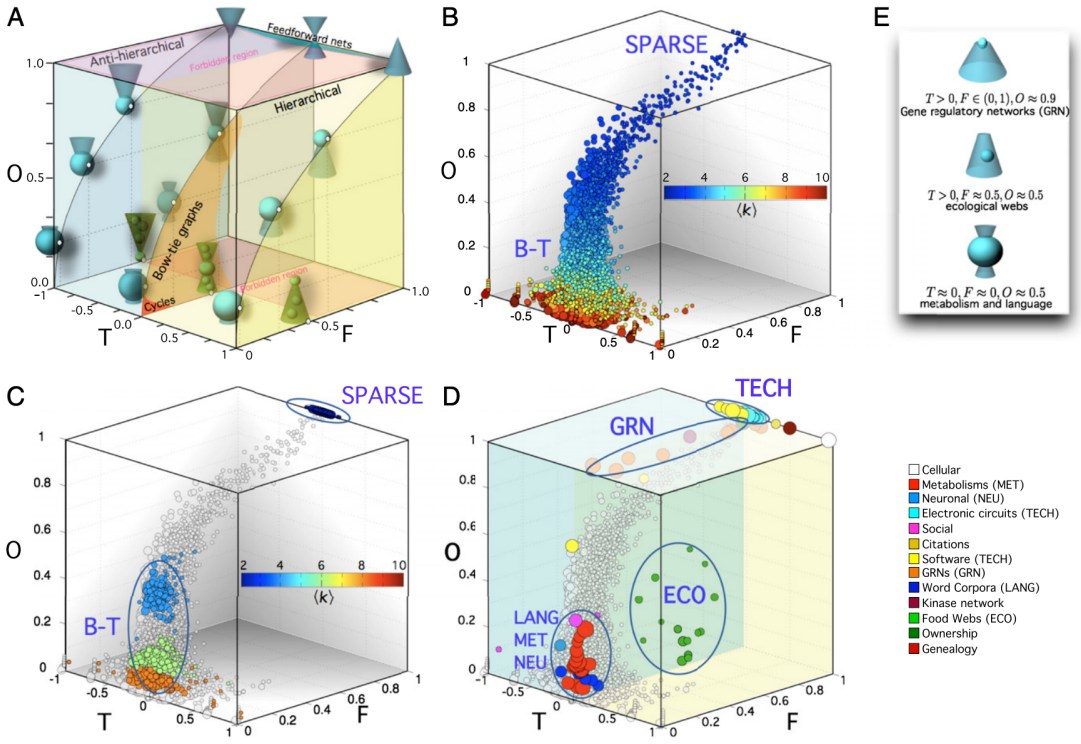


Figure 3.3. A morphospace construction to study the hierarchy properties of different types of networks. Reproduced in courtesy of [69].

networks (see figure 3.3 for more details), or ii) configuration property [70], consciousness [71], topological motifs [72] in brain networks, among many others.

Comprehensively, a principled way of looking how (networked) systems evolve and change is through phenotypic spaces, also called morphospaces, see [65], [68], [69], [72]–[76]. The concept of a morphospace can be used to analyze many other mathematical objects, including networks. When applied to networks, quantitative traits of global or local network topology are conceptualized through the Cartesian coordinates defined in this abstract space. A brain's subsystem configuration is topologically represented by a point in this multidimensional space.

3.1.2 A formalism on network morphospace

Let $G \in \mathcal{G}$ denotes a network/graph that belongs to the space of graphs/networks; $v \in \mathcal{V} \subset R^n$ be a point in a multidimensional space (over real field $\mathbf{R}$).

Definition 3.1.1. *A network morphospace (denoted as f) is an operator that maps networks (graphs) into finite dimensional phenotypic space $\mathcal{P}^d$ where $d << \infty$ is the space dimension.*

$$\mathcal{G} \xrightarrow{f} \mathcal{P}^d$$

Morphospace Complexity. As carried over by the concept of complexity and (mathematical) space, morphospace complexity can be measured by the number of geometric parameters (morphospace measures) used to efficiently describe the network phenotypes. In the above definition, the complexity is measured by d. It is important to note that, one of the advantages of defining theoretical morphospace is the freedom of designing geometric parameters without considering whether a specific form exists or not. Once mapped on to the theoretical morphospace, the robust parametric design would yield an area which non-existent forms occupy.

3.2 Brain Functional Network Configuration and Morphospace

3.2.1 Background and Motivation

A particular challenge in network neuroscience is the derivation of a comprehensive means to quantify brain network configurations across different mental states and cognitive tasks. Configurations across a collection of cognitive tasks can be conceptualized at three distinct levels of granularity. **1) Network configural breadth**: the overall extent, *across many different mental states and tasks*, to which the brain networks change in configuration. **2) Task-to-task (transitional) reconfiguration**: brain network reconfigurations that occur when transitioning from *one* specific cognitive or mental state to *a second*, different state. **3) Within-task reconfigurations**: reconfigurations that occur within one task, such as shifts from lower to higher cognitive demands, or vice versa. Formal definitions are provided in later section.

To properly study functional brain configuration properties, one must first determine at which scale network configurations can be efficiently identified at both the group level by task and at the phenotypic level by individual subject. To this end, Cole and colleagues in [77] note that task configurations are rather subtle in both macro- (whole-brain) and micro- (edge-to-edge) levels, relative to resting configuration. Hence, to efficiently disentangle task [78] and subject fingerprints [2], [5], [6], [53], [57], [58], [79]–[81] existing in brain network configurations , one needs to consider mesoscopic scale (i.e. functional networks). It has been argued that higher level of cognition emerges through interactions of subsystems [82]. Specifically, mesoscopic structures exhibit modular characteristics that can adapt to cognitive demands "*without adversely perturbing the remainder of the system*" [60]. Mesoscopic structures can be viewed as either (i) brain connectivity patterns related to unique cognitive modes [83], [84], or (ii) subsets of brain regions that sustain and/or modulate one particular function [1], [6], [8], [24], [62].

Traditionally, a mesoscopic exploration of functional brain networks would either involve the detection of functional communities [85] based on topology [86], [87] or on the information flow [88], [89]. But, both these approaches are limited in **evaluating** the dynamics of detected communities across time, tasks, and/or subject. On the other hand, we would like to have a framework that can capture the behavior of a reference set of FNs with changing mental states; a framework that can not only characterize the topology of FNs and but also the flow of information within and between the FNs. Such a formalism can help us define and quantify different types of configurations that functional brain networks can assume and re-configurations that they go through when switching between seemingly infinite number of mental states.

Two primary approaches examine mesoscopic configurations across different cognitive/mental tasks. One approach unravels newly emergent functional modules in each single task and/or each subject separately. Many interesting concepts rely on this approach, such as individualized (atlas-free) parcellation [90], [91] or task-dependent atlases [8]. An alternative is to maintain a set of baseline functional modules and, instead, monitor their properties across tasks. For a framework to assess network configural breadth, the first approach does not permit tracking changes between cognitive states that are both *subject-comparable* and

task-compatible. Subject-comparable refers to a common brain parcellation across subjects, whereas task-compatible refers to a common brain parcellation across different tasks. As stated by Cole et al., the brain's common intrinsic network is prominently reflected by its resting state architecture [77]. Similarly, Shine and colleagues hypothesize that the human brain exhibits and maintains a core integrative network that dynamically configures during cognitive demands [19]. In that context, task configurations in general can be viewed as topological perturbations departing from the resting state architecture that can be quantified by comprehensive edge-wise changes in the functional connectomes.

3.2.2 The necessity of a mesoscopic morphospace

Assessing brain network (re-)configuration requires an appropriate identification of a system scale in which changes can be efficiently detected among various cognitive tasks/mental states and/or subjects. At both the functional edge (microscale) and entire whole-brain functional connectome (macroscale) levels, these changes are rather subtle [18] and hence, insufficient to detect underlying shifts across tasks and reflect cognitive changes. This leaves mesoscopic structures (for instance, FNs in the case of brain functional networks) as a suitable scale to investigate network configural breadth.

Further, an effective measure of functional network adaptations in response to varied and changing cognitive demands should, to some degree, reflect cognitive capacity. For example, greater intelligence has been associated with reduced across-task network configurations, suggesting a degree of efficiency (smaller changes in FN connectivity across tasks; [15]).

Inspired by [15], we consider rest a reference with respect to network configural breadth. As opposed to measuring similarity (reconfiguration efficiency) in a pairwise fashion rest-to-task reconfigurations [15], we expand the concept of reconfiguration by assessing the differentiating capacity of functional connectivity through sampling the cognitive space [80], [92] using multiple fMRI tasks simultaneously and collectively measuring change. Our goal is to propose a minimal number of measure(s) that can capture functional network task and subject characteristics that strive beyond similarity measure induced from pair-wise interactions, i.e. Pearson correlations as reconfiguration efficiency measure proposed in [15].

Moreover, this would allow an assessment of complex cognitive changes induced from configural breadth. Further, large whole-brain (macroscale) configuration changes while performing tasks or across individuals might not be expected [77] and do not provide specificity on the spatial organization of those changes along the cortex. Here, we track reconfigurations at the mesoscopic functional networks level, allowing the hypothesis testing that some functional circuits reconfigure more than others while not focusing on edge-to-edge functional changes that are subtle among rest and task states.

As we look into how mesoscopic interactions changes with respect to task, at the same time we have to consider reflecting changes in both internal (within one FN) and external (between a pair of FNs) changes. Hence, straightforward pair-wise similarity would not suffice. In other words, these functional circuits cannot be studied in an isolated fashion. One needs to consider functional network configural breadth with respect to the entire cortex, e.g. both within- and between- functional networks simultaneously.

Consequently, assessing specific network configural breadth requires a mathematical mapping of mesoscopic changes into a well-defined space comprising measures that integrate interactions among sub-systems (such as functional communities in brain functional networks). To accomplish this goal, two relevant theories that can satisfy the aforementioned requirements, are (i) stochastic processes, [93], [94], and (ii) information-theory, [95]–[97]. Stochastic processes provide an appropriate tool to, metaphorically, inject a random particle that walks among connectome nodes, through their interactions (i.e. functional edges), in such a way that no particular pairwise or local interaction can fully describe its behavior. In other words, random walk theory allows us to study configural breadth based on the topological changes induced by inter- and intra- communities when a participant performs different tasks and rest. Information theory provides the quantification of uncertainty in choosing the preference of communicating channels among functional communities. Specifically, it allows a fine-grained (edge-wise) approach that is complementary with the stochastic modelling approach, as seen in [88], [89], especially in the domain of community structures in complex networks. In other words, configural breadth is investigated through information-theoretic changes induced by exiting edges between a given FN with others when a participant is involved in different tasks and rest.

3.2.3 Conceptualization of configuration morphospace

Here, we propose such a formalism to quantify network configural breadth using two distinct assessments (measures/metrics): i) "trapping efficiency" (**TE**), describing the extent to which a particular FN (e.g., the frontoparietal network) "traps" an incoming hypothetical signal, and ii) "exit entropy (**EE**)," describing the uncertainty as to where (what nodes) that same signal would exit a given FN to enter another FN. We propose that for a given FN, the relative combinations of these two measures across a comprehensive range of task and mental states defines a "mesoscopic morphospace." This 2-dimensional geometric shape formed by unique combinations of **TE** and **EE** can quantify the extent to which a specific FN changes in its behavior as *a module*. In other words, it is degeneration of modular structure or otherwise, within the repertoire of tasks situated in a "cognitive space." [80], [92].

This work aims to formally parametrize subsystem changes that occur across a broad span of cognitive and mental states. We refer to all those across-task (flexibility) changes occurring in a functional network as "network configural breadth". Furthermore, this framework also paves the way to model the nature of transitional FN reconfigurations that occur either between two different tasks ("task-to-task transitional reconfiguration") or within-task according to varying levels of difficulty and demand. The purpose of this work is, hence, to define and assess theoretical network properties of *a priori* functional modules as determined by resting state FNs, leveraging the idea that tasks modify a common intrinsic network to efficiently meet cognitive demands [77]. As noted by several authors [15,16,26], executive subsystems in the brain are consistently reproducible across many individuals at rest [1], [24], [62]. We consequently see resting-state functional communities as a common (fixed) foundation upon which modifications induced by cognitive demands from tasks occur.

To model *network configural breadth*, one then needs to map *FN configurations* into a well-defined mathematical space. On a practical level, the framework needs to estimate the minimum number of parameters required to characterize network subsystem changes, which constitute global comprehensive changes (i.e., functional connectome changes due to different cognitive/mental tasks). A principled way investigating how systems evolve and change is through phenotypic spaces, also called morphospaces, see [65], [68], [69], [72]–[76].

The concept of a morphospace could be used to analyze many other mathematical objects, including networks. When applied to networks, quantitative traits of global or local network topology can be conceptualized through the Cartesian coordinates defined in this abstract space. A brain's subsystem configuration is topologically represented by a point in this multidimensional space.

3.3 Scope of the project

The primary aim of this work is to clearly define and quantify different configurations that FNs can assume, as well as measure their nature of re-configurations switching between a large number of cognitive states. From a graph-theoretical perspective, FNs and their corresponding reconfigurations are described by two attributes: topology and communication. From a system dynamic perspective, FNs can be characterized by segregation and integration [98] properties across which the human brain reconfigures across varied cognitive demands [16]–[19], [21]. To formally capture these diverse characteristics of FNs, we constructed a mathematically well-defined and well-behaved 2D "mesoscopic morphospace" based on two novel measures defined for non-negative, undirected, weighted functional connectomes: Trapping Efficiency (**TE**) and Exit Entropy (**EE**). Trapping Efficiency captures the level of segregation/integration of a functional network embedded within the functional connectome and quantifies the extent to which a particular FN "traps" an incoming signal. Exit Entropy captures the specificity of integration of an FN with the rest of the functional connectome, and quantifies the uncertainty as to where (in terms of exit nodes) that same signal would exit the FN. In summary, this mesoscopic morphospace is a representation of the cognitive space as explored within and between cognitive states, as reflected by brain activity in fMRI. Such representation relies on FNs reconfigurations that can be tracked, at an individual level, and at different granularity levels in network (re-)configurations. All three induced sub-graphs have the same cardinality ($|\mathcal{C}| = 8$) with different number of exits (connections to $G \backslash \mathcal{C}$). Nonetheless, depending on their topological structures, the corresponding morphospace measurements (**TE** and **EE**) have rather distinct values.

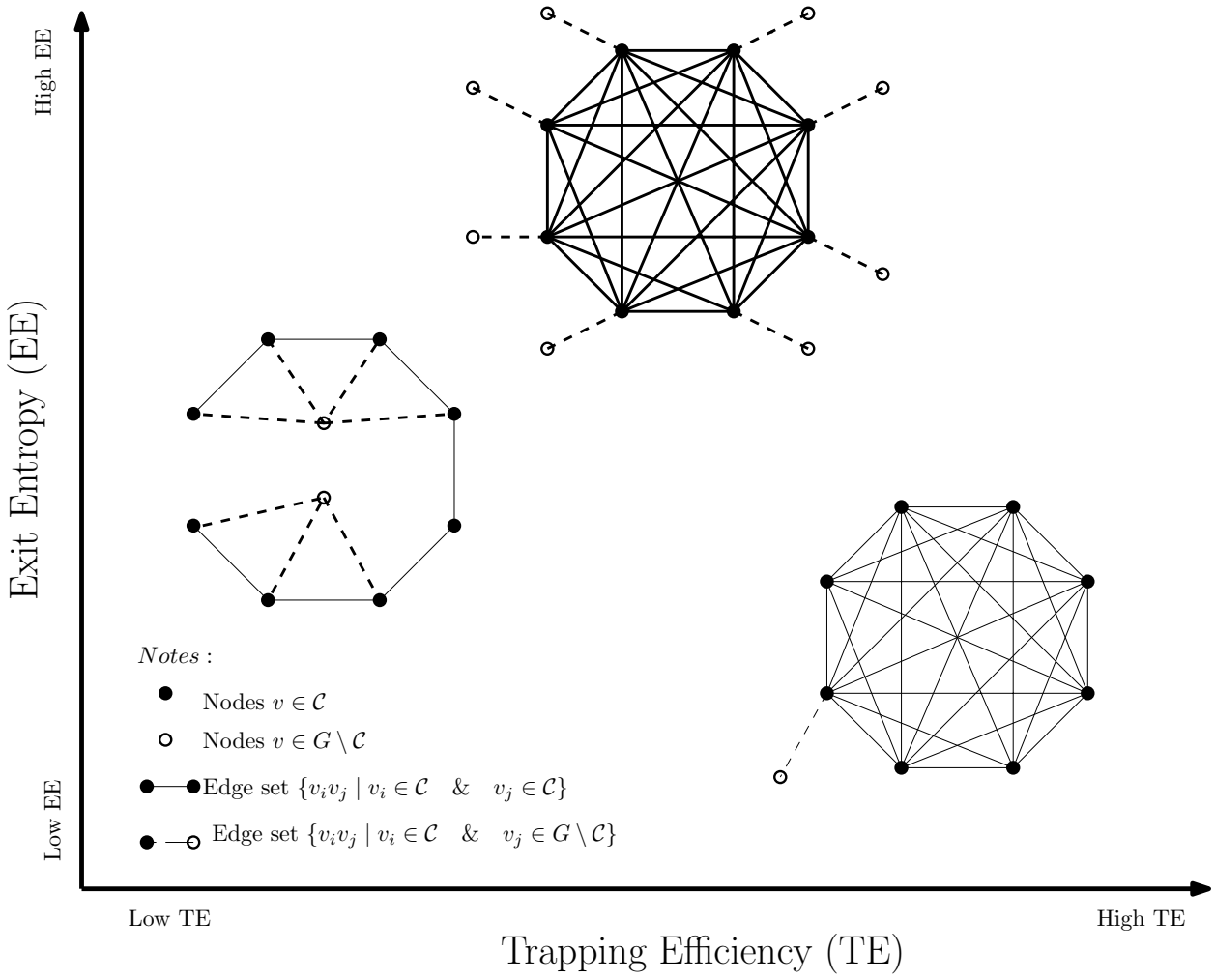


Figure 3.4. Morphospace Measurements - examples. $\mathcal{C}$ represents communities in network $\mathcal{G}$. v represents nodes and uv represents edges in the network.

By using this 2D **TE,EE**-based morphospace, we formally study Network Configural Breadth (Figure 3.4), the most global and coarse grain exploration of the cognitive space, and its subsequent functional configuration components. To that end, we formally define measures of (1) functional reconfiguration (capacity of an individual to reconfigure across widely differing cognitive operations) and (2) functional preconfiguration (efficiency of transition from resting-state to task-positive state [15], for potentially any community or FN. We thus intend the idea of a "mesoscopic morphospace" to capture brain's subsystem configurations across multiple cognitive/mental states, which in turn may relate to behavioral measures, as shown in [15], [99]. In particular, we aim to determine if FN configural properties, which

comprises of: (i) "functional preconfiguration" - functional readiness transitioning from rest to active task engagement and (ii) "functional reconfiguration" - FN transformations across mental/emotional states, relate to cognitive abilities such as intelligence.

In summary, the purpose of this work is to 1) assess and subsequently quantify functional network pre- and reconfiguration and 2) test the associations of FN configural breadth with cognitive ability. We first present the theoretical aspects of this framework and then apply it to the dataset of one hundred unrelated subjects from the Human Connectome Project (HCP) [20], [100].

3.4 Dataset and Data Processing Pipeline

In this section, we provide the details related to the dataset used to analyze the notion of configural breadth. We also provide information related to the brain atlas.

Brain atlas

The brain atlas used in this work is the based on the cortical parcellation of 360 brain regions as recently proposed by Glasser et al. [62]. Similarly to reference [5], [84], 14 sub-cortical regions were added, as provided by the HCP release (filename $Atlas\_ROI2.nii.gz$). We accomplish this by converting this file from NIFTI to CIFTI format by using the HCP workbench software http://www.humanconnectome.org/software/connectomeworkbench.html, with the command -cifti- create-label. This resulted in a brain atlas of 374 brain regions (360 cortical + 14 sub-cortical nodes).

Using Human Connectome Project Dataset, we explore the characteristics of functional networks' configural breadth by utilizing Resting State Networks (FNs), see [1], which includes seven functional networks (FNs): Visual (VIS), SomatoMotor (SM), Dorsal Attention (DA), Ventral Attention (VA), Limbic (LIM), Frontoparietal (FP), Default Mode Network (DMN); Sub-cortical (SUBC) region, as mentioned before, is added into this atlas for completeness. Thus, the parcellation used in this paper comprises of eight (8) FNs.

HCP Dataset

The fMRI dataset used in this paper is available in the Human Connectome Project (HCP) depository (http://www.humanconnectome.org/), with Released Q3. The processed functional connectomes obtained from these data and used for the current study are available from the corresponding author on reasonable request. Please refer to below detailed descriptions on the dataset and data processing.

HCP Functional Data

The fMRI data from the 100 unrelated subjects in the HCP Q3 release were employed in this study [47], [48]. Per HCP protocol, all subjects gave written informed consent to the HCP consortium. The two resting-state functional MRI acquisitions (HCP filenames: $rfMRI\_REST_1$ and $rfMRI\_REST_2$) were acquired in separate sessions on two different days, with two acquisition patterns (left to right and right to left) in each day, [20], [47], and [48] for details. This release includes also data from seven different fMRI tasks: gambling ($tfMRI\_GAMBLING$), relational or reasoning ($tfMRI\_RELATIONAL$), social ($tfMRI\_SOCIAL$), working memory ($tfMRI\_WM$), motor ($tfMRI\_MOTOR$), language ($tfMRI\_LANGUAGE$, including both a story-listening and arithmetic task), and emotion ($tfMRI\_EMOTION$). Per [20], [101], three tasks MRIs are obtained: working memory, motor, and gambling.

The local Institutional Review Board at Washington University in St. Louis approve all the protocol used during the data acquisition process. Please refer to [20], [100], [101] for further details on the HCP dataset. All tasks and resting functional MRIs were equally weighted in importance, i.e. each task is equally weighted.

Constructing functional connectomes

We used the standard HCP functional pre-processing pipeline, which includes artifact removal, motion correction and registration to standard space, as described in [20], [100] for this dataset. For the resting-state fMRI data, we also added the following steps: global

gray matter signal regression; a bandpass first-order Butterworth filter in both directions; z-scores of voxel time courses with outlier eliminations beyond the three standard deviations from first moment [102], [103]. For task fMRI data, aforementioned steps are applied, with a relaxation for bandpass filter [0.001 Hz, 0.25 Hz]. Starting from each pairs of nodal time courses, Pearson correlation coefficient is calculated to fill out the functional connectomes for all subjects at rest and seven designated tasks. This yields symmetrical connectivity matrix for all fMRI sections.

Resting State Connectomes: There are two resting scanning sections conducted in two different days. In each day, individual MRIs are obtained independently in the morning and afternoon sections. We averaged the resting functional connectome in the first day (which contains morning/afternoon scans) to obtain the "Test" connectome. By the same token, we obtained "Retest" connectome for the resting condition.

FC's matrix entries: For all considered fMRI images presented here, we removed negative correlations as the morphospace axes are built upon stochastic models; hence, numerically it is not possible to utilize negative entries. The remaining matrix are, then, squared.

Improve Individual fingerprint The framework proposed by Amico et al. [5] was used to maximize individual fingerprints where each subplot represents rest and seven tasks in HCP dataset. The optimal reconstructed number of orthogonal components is indicated by a black dot, see figure 3.5 for further details.

3.5 A formalism of brain network configurations

Human behavior arises out of a complex interplay of functional dynamics between different brain networks [60]. These interactions are reflected in functional network reconfigurations as participants perform different tasks or are at rest [77], [104], [105]. One of the network neuroscience challenges is to develop a comprehensive framework to quantify the brain network (re-)configurations across different mental states and cognitive tasks. To that end, configurations across a collection of cognitive tasks can be conceptualized at three distinct levels of granularity (see figure 3.6 for geometrical demonstration):

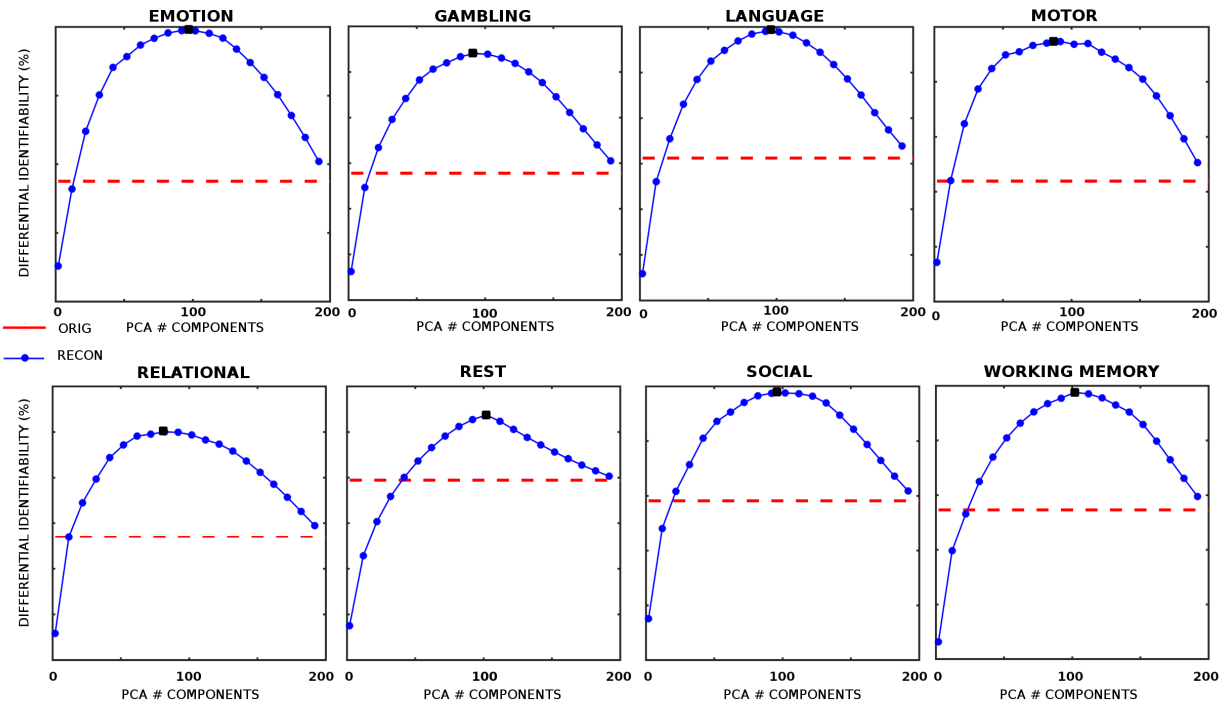


Figure 3.5. Individual fingerprints unveiled through the framework proposed by Amico et al. [5]. In this case, I_{diff} score is computed for all fRMI task and resting state. The maximum value of I_{diff} for each task is denoted by $*$ which is also indicative of the number of principal components used to reconstruct the FCs.

- **Network configural breadth** represents, for an FN, a given individual's repertoire of cognitive and emotional states through functional configurations while performing different tasks. In practice, how well the entire "cognitive space" [80], [92] is sampled depends on the number and choice of the tasks. This concept is inspired by [15].

- **Task-to-task transitional reconfiguration** represents the specific shift in network functional configuration when a subject switches between cognitive/mental tasks [78], [99]. For instance, task transitions and accompanying reconfigurations will occur when a subject transitions from quiet reflection to engage in a spatial problem solving task, or from a lexical retrieval to a decision making paradigm.

- **Within-task reconfiguration** represents specific network functional configuration changes that may occur within a single task. This phenomenon has been assessed at

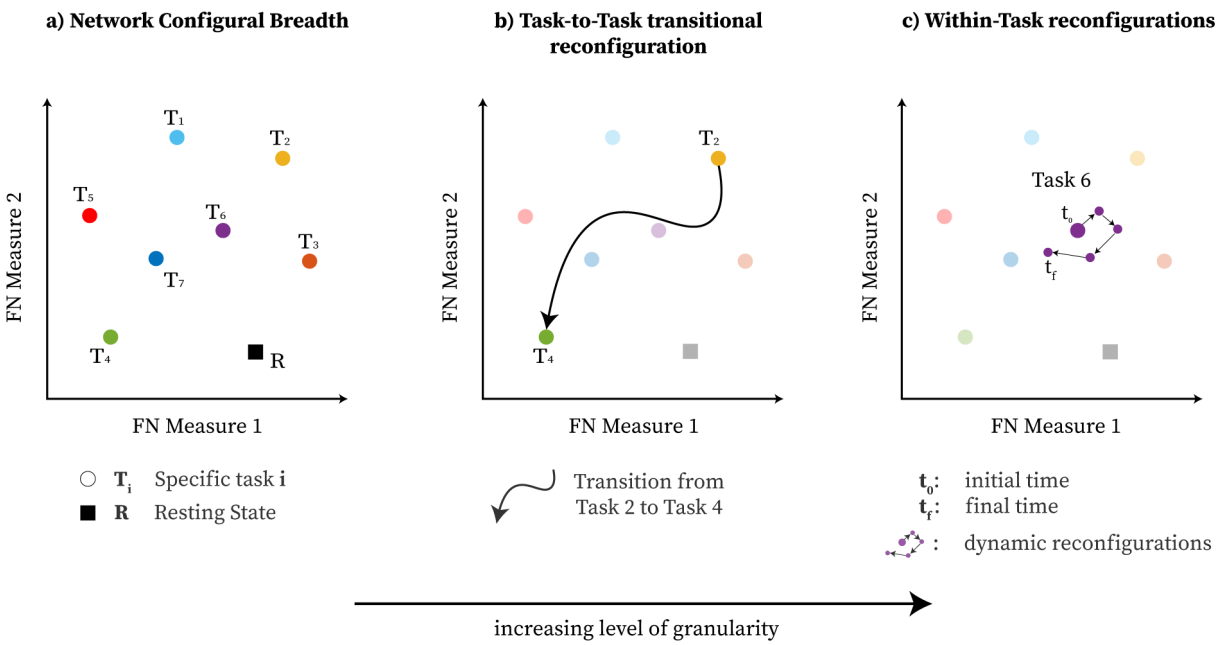


Figure 3.6. The three types of brain (re-)configurations that can be represented by a mathematical space characterized by: **TE** and **EE**.

the whole-brain level, showing the presence of distinct brain states within a task [16], [17], [19], [82], [106].

3.6 The construction of a mesoscopic morphospace to access brain functional configural breadth

The *mesoscopic morphospace* proposed here is a two dimensional space built upon Trapping Efficiency and Exit Entropy measures for assessing functional networks or communities of functional connectomes. In this framework, functional connectomes must be undirected (symmetrical) weighted graphs, with *non-negative* functional couplings. This framework allows for any *a-priori* partition into functional communities. In this work, we assess the resting-state functional networks as proposed by [1] as the *a priori* FNs. Also, we use functional connectivity (without incorporating structural connectivity information), which is a quantification of statistical dependencies between BOLD time-series of brain regions, and it can be used as a proxy of communication dynamics in the brain [43].

It is important to be aware that any generic network can be fragmented (disconnected). This is rarely the case for brain functional networks because of how we compute functional couplings, typically using Pearson Correlation Coefficient. Despite of that, *a priori* functional community induced from global thresholded adjacency structure is not guaranteed to be connected. As pointed out in [107] among others, a meaningful cluster should, at minimum, be connected. To respect the global topology of the functional networks G, i.e. connectedness, we applied a small perturbation to edges with zero weight, i.e.

$$a_{\mathrm{ij}} = 0 \rightarrow a_{\mathrm{ij}} =\mid \forall \mathrm{i}, \mathrm{j} \in \mathcal{C}$$

According to [108], if one concerns solely the connectedness property, then topological spaces are similar to graphs. Therefore, the goal is to have the connectedness property carried from the graph G to all of its induced subgraphs $\mathcal{C}$. To do so, one needs to maintain the topology defined on G, i.e. the collections of open sets $\mathcal{U}$ (which can be thought of as the edge set E defined on G).

3.6.1 Computing mechanistic components for morphospace measures

A mesoscopic morphospace is constructed to assess functional network behaviors through two focal lenses: level of segregation/integration (using graph topology), and specificity of integration (using information theory). We first define all necessary components to compute **TE** and **EE** as follows:

1. The whole-brain FC is graph-theoretically denoted by $G(V, E)$ where V is the set of vertices (represented by the regions-of-interest (ROIs)) and E is the set of edges (quantified by functional couplings between pairs of ROIs). The whole-brain FC is mathematically represented by an adjacency structure denoted as $\mathbf{A} = [w_{\mathrm{ij}}]$ where i, j are indexed over vertex set V and $w_{\mathrm{ij}} \in [0, 1]$ are functional couplings;

2. Using a pre-defined set of FNs, a functional community (graph-theoretically denoted as $G_{\mathcal{C}}(V_{\mathcal{C}}, E_{\mathcal{C}})$ or $\mathcal{C}$ for short) is defined to have the corresponding node set $V_{\mathcal{C}} \subset V$

and edge set $E_{\mathcal{C}} \subset E$ for which the union over all FNs exhaust the vertex and edge set of G such that:

$$\cup V_{\mathcal{C}} = V \quad \& \quad \cup E_{\mathcal{C}} = E$$

3. For a given functional community $\mathcal{C} \subset G$, define the set of states (or equivalently, vertices) S which contains the set of transient states (denoted as $S_{trans} = V_{\mathcal{C}}$), and absorbing states (denoted as $S_{abs} = \{\mathrm{j} \mid w_{\mathrm{ij}} > 0; \mathrm{j} \notin V_{\mathcal{C}}, \forall \mathrm{i} \in V_{\mathcal{C}}\}$) such that

$$S = S_{trans} \cup S_{abs}$$

4. We mathematically denote a whole brain FC as $\mathbf{A} = [w_{\mathrm{ij}}]$, where i and j are brain regions (from now on denoted as vertices or states) of the specified parcellation or atlas. Each matrix $\mathbf{A}$ represents a single subject, single session, single task whole-brain FC. We assess the whole-brain FC with respect to organizations into FNs, here denoted by $\mathcal{C}$. For a specific $\mathbf{A}$ and a specific $\mathcal{C}$, we obtain an induced sub-matrix $\mathbf{A}_{\mathcal{C}}$ by extracting the corresponding rows and columns of matrix $\mathbf{A}$ using only the vertices that belong to S, which results in the matrix:

$$\mathbf{A}_{\mathcal{C}} \in (0,1)^{|S| \times |S|}$$

We note that the row and column order of the states (or vertices) of $\mathbf{A}_{\mathcal{C}}$ respects the order of $S = S_{trans} \cup S_{abs}$ with transient states followed by absorbing ones which results in a blockage structure:

$$\mathbf{A}_{\mathcal{C}} = \begin{array}{c} \\ \textbf{Transient} \\ \textbf{Absorbing} \end{array} \begin{array}{c} \begin{array}{cc} \textbf{Transient} & \textbf{Absorbing} \end{array} \\ \begin{pmatrix} \mathbf{A}(S_{trans}, S_{trans}) & \mathbf{A}(S_{trans}, S_{abs}) \\ \mathbf{A}(S_{abs}, S_{trans}) & \mathbf{A}(S_{abs}, S_{abs}) \end{pmatrix} \end{array}$$

where $\mathbf{A}(S_{trans}, S_{trans})$ means that we extract the sub-matrix of $\mathbf{A}$ that corresponds to states in S_{trans} for the rows (first argument) and S_{trans} for the columns (second argument);

5. For any functional network $\mathcal{C}$, using the induced adjacency structure $\mathbf{A}_{\mathcal{C}}$ in the previous step, we define each vertex in S to be a state in the stochastic process and construct the corresponding Terminating Markov Chain by computing:

   - the normalization of $\mathbf{A}_{\mathcal{C}}$ by the nodal connectivity strength:

   $$Q = \mathbf{D}_{\mathcal{C}}^{-1}\mathbf{A}_{\mathcal{C}} \in (0,1)^{|S|\times|S|}$$

   where $\mathbf{D}_{\mathcal{C}}$ is the weighted degree sequence matrix filled with the node strength (defined by the row (or equivalently, column) sum of $\mathbf{A}_{\mathcal{C}}$) in the diagonal entries and zeros for the off-diagonal elements:

   $$\mathbf{D}_{\mathcal{C}} = [d_{\mathrm{ij}}] = \begin{cases} \sum_{\mathrm{j}=1}^{\mathrm{j}=|V_{\mathcal{C}}|} w_{\mathrm{ij}}, \forall \mathrm{i} = \mathrm{j} \\ 0, \forall \mathrm{i} \neq \mathrm{j} \end{cases}$$

   where i, j are indexed over S. Note that the order of rows and columns of Q and $\mathbf{D}_{\mathcal{C}}$ also respect the order of S;

   - the transition probability matrix of the terminating Markov Chain:

   $$\mathbf{P} = \begin{array}{c} \\ \textbf{Transient} \\ \textbf{Absorbing} \end{array} \begin{array}{c} \begin{array}{cc} \textbf{Transient} & \textbf{Absorbing} \end{array} \\ \begin{pmatrix} Q(S_{trans}, S_{trans}) & Q(S_{trans}, S_{abs}) \\ \mathbf{0}_{|S_{abs}|\times|S_{trans}|} & \mathbf{I}_{|S_{abs}|} \end{pmatrix} \end{array}$$

   where $\mathbf{0}_{|S_{abs}|\times|S_{trans}|}$ is the matrix of all zeros (size $|S_{abs}|$ rows by $|S_{trans}|$ columns); $\mathbf{I}_{|S_{abs}|}$ is identity matrix of size $|S_{abs}|$; the index $\mathcal{C}$ for Q and $\mathbf{P}$ is dropped for simplicity.

6. Using matrix $\mathbf{P}$, we extract the sub-matrix induced by states in S_{trans} (denoted by $\mathbf{P}S_{trans}$). Note that $\mathbf{P}S_{trans} = Q(S_{trans}, S_{trans})$ because rows and columns of $\mathbf{P}$ respect the order of S. We then compute the fundamental matrix (denoted as $\mathbf{Z}$) [94] which contains the mean number of steps a specific transient state in S_{trans} is visited, for any

pair of transient states in S_{trans}, before the random walker is absorbed by one of the states in S_{abs}:

$$\mathbf{Z} = (\mathbf{I}_{|S_{trans}|} - \mathbf{P}S_{trans})^{-1} \in R_{+}^{|S_{trans}|\times|S_{trans}|}$$

7. Compute the mean time to absorption (denoted as τ) which contains the mean number of steps that the random particle needs to be absorbed by one of the states in S_{abs}, given that it starts in some state in S_{trans}:

$$\tau = \mathbf{Z}\mathbf{1}_{|S_{trans}|} \in R_{+}^{|S_{trans}|\times 1}$$

where $\mathbf{1}_{|S_{trans}|}$ is the all one vector of size $|S_{trans}|$.

8. Compute the absorption probability matrix (denoted as Ψ), which contains the likelihood of being absorbed by one of the absorbing states, given that the stochastic process starts in some transient state:

$$\Psi = \mathbf{Z}\left[\mathbf{P}S_{trans}, S_{abs}\right] \in R_{+}^{|S_{trans}|\times|S_{abs}|}$$

where $\mathbf{P}S_{trans}, S_{abs}$ is the sub-transition probability matrix induced from (row) state S_{trans} and (column) state S_{abs}. Hence, $\mathbf{P}S_{trans}, S_{abs} = Q(S_{trans}, S_{abs})$.

3.6.2 Module Trapping Efficiency

Module Trapping Efficiency, denoted as **TE** (unit: $\frac{steps}{weight}$), quantifies a module's capacity to contain a random particle from leaving its local topology, i.e. $\mathcal{C}$. Specifically, through FN topology, we want to assess its level of *segregation/integration*, measured by the L_2 norm of τ (unit: *steps*), i.e. the mean time to absorption of nodes in $\mathcal{C}$, normalized by its total exiting strength (unit: *weight*), measured by

$$\mathcal{L}_{\mathcal{C}} = \sum_{\mathrm{i}\in S_{trans},\mathrm{j}\in S_{abs}} A_{\mathrm{ij}} = \mathbf{A}(S_{trans}, S_{abs})$$

Mathematically, trapping efficiency is quantified as follows:

$$\mathbf{TE} = \frac{||\tau||_2}{\mathcal{L}_{\mathcal{C}}} \tag{3.1}$$

We see that the mean time to absorption vector, τ, is dependent on both **density-based** [87], [109] and **flow-based** [88], [89], [107] modularity. The mean-time-to-absorption vector τ for which τ_{i} contains the average number of steps a random walker needs to escape the FN topology, given that it starts from node i.

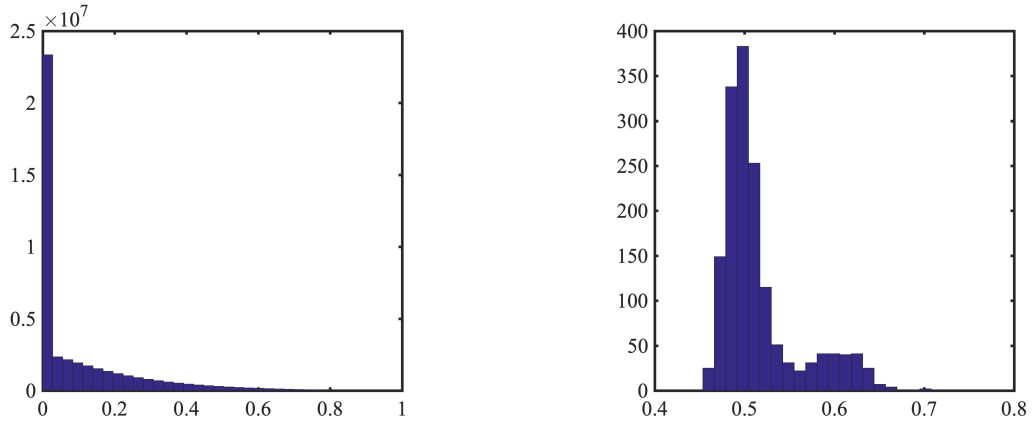
This means that the numerical values in τ are always greater than or equal to 1. We chose to use L_2 norms because it squares the input values of the vector and thus enhance our capacity to quantify FN (re-)configuration. On the other hand, the denominator $\mathcal{L}_{\mathcal{C}}$ is a simple statistical summary of the module "leakages" to the rest of the cortex. Since all the values in $\mathcal{L}_{\mathcal{C}}$ are between $(0, 1)$, L_2 norm would have diminished the differences across FNs. Hence, we chose L_1-norm for the denominator. The role of $\mathcal{L}_{\mathcal{C}}$ is to account for potential differences in trapping efficiency due to community size. Numerically, higher **TE** indicates that a module is more segregated (or equivalently, less integrated). This is because the FN topology traps the incoming signal efficiently, relatively to its exiting edges when embedded in the cortex.

Numerator τ

As claimed in the main text, **TE** is finitely bounded. There are several ways to observe this; one approach involves applying hierarchical community detection algorithm [110] and look for the first time G split into more than one subgraphs. Thus, let i be indices representing communities belong to the first hierarchical layer, then

$$M = \max_k \left[\mathbf{TE}(S_k)\right] \mid \forall k \in [l]$$

where l represents the number of communities. Such value is well-defined and finite. An



(a) Edge strength, after post-processing steps, histogram of all considered FCs.

(b) FC density, defined to be the number of non-zero functional weighted edges out of $\binom{N}{2}$ possible edges, histogram of all considered FCs.

Figure 3.7. Mean Density and majority of edge strength falls in the first bin [0,0.025] are the two major factors into the maximum value of **TE**.

alternative way to see the trivial bound of the measures is as follows: Let us consider the entire network G, we have:

$$\mathbf{TE}(\mathcal{C} \equiv G) = \frac{||\tau||_2}{\mathcal{L}_{\mathcal{C}}} = \frac{\infty}{0} = \infty$$

because there is no exits if the configurations is the entire network; moreover, there is zero leakages. Hence, any cut into G would have to be strictly less than this upper bound. Note that $\frac{\infty}{0}$ is undefined. However, in such case, we define this quantity to be unbounded which is the notion of infinity.

In the context of the data set at hand, we can, however provide a better bound. We proceed by obtaining the maximum value of **TE** when all subjects and all tasks are under consideration which yields the result (see further evidence on edge strength and FC density for further details 3.7):

$$\max_{subjects,tasks}(\mathbf{TE}) = 0.5064$$

One can relate this numerical value with two factors: Functional connectome density and edge strengths.

Normalization

Realistically, since larger communities carry more exits which is driven purely from a topological viewpoint, $\mathcal{L}_{\mathcal{C}}$ is a logical choice to normalize the magnitude of τ.

Additionally, $\mathcal{L}_{\mathcal{C}}$ is deemed to perform as $||\tau||_2$-damping. Notice that there also exists functional communities with low total exiting strength with large cardinality, theoretically. In such case, these structures are rewarded from the standpoint of **TE** as it converges to $\mathbf{TE}(\mathcal{C} \equiv G)$.

τ and modularity

In terms of notation, given the sub-system $\mathcal{C}$, we define two vectors:

$$\mathbf{R}_{in} = \left[\frac{k_1^{in}}{k_1}, \frac{k_2^{in}}{k_2}, ...\frac{k_{|C|}^{in}}{k_{|\mathcal{C}|}}\right] \quad \mathbf{R}_{out} = \left[\frac{k_1^{out}}{k_1}, \frac{k_2^{out}}{k_2}, ...\frac{k_{|\mathcal{C}|}^{out}}{k_{|\mathcal{C}|}}\right]$$

Further, given a community $\mathcal{C}$, there are three possible node types:

1. Node type 1, denoted through set I_1, are nodes with all connections belonging to $\mathcal{C}$;
2. Node type 2, denoted through set I_2, are nodes with some connections belonging to $\mathcal{C}$ and others belong to $G \setminus \mathcal{C}$;
3. Node type 3, denoted through set I_3, are nodes with all connections belonging to $G \setminus \mathcal{C}$.

We note that:

- If $||\mathbf{R}_{out}||_2 = ||\mathbf{R}||_2$ then community $\mathcal{C}$ is disconnected i.e. no external connectivities. This is impossible due to our assumption on connectedness.
- If $||\mathbf{R}_{in}||_2 = ||\mathbf{R}||_2$ then $\mathcal{C}$ resembles an empty subgraph i.e. no internal connectivities. This is impossible as we reveal our algorithm in the later section.
- $||\mathbf{R}_{in}||_2$ and $||\mathbf{R}_{out}||_2$ is well-defined as there exists no disconnected components in our working graph.
- $\tau_{i_3} = 1 \quad \forall i_3 \in I_3$.

Theorem 3.6.1. *Given a non-empty induced subgraph $\mathcal{C} \in G$, $\langle \mathbf{R}_{out}, \tau \rangle$ is $|\mathcal{C}|$.*

Proof. We begin with two vectors in $R^{|\mathcal{C}|}$. The inner product between $\mathbf{R}_{out}$ and τ τ

$$\begin{aligned}\langle \mathbf{R}_{out}^T \tau \rangle &= \mathbf{1}_{|\mathcal{C}|}^T \mathbf{I}_{|\mathcal{C}|} - \mathbf{Q}_{\mathcal{C}} \mathbf{I}_{|\mathcal{C}|} - \mathbf{Q}_{\mathcal{C}}{}^{-1} \mathbf{1}_{|\mathcal{C}|} \\ &= \mathbf{1}^T \mathbf{I}_{|\mathcal{C}|} \mathbf{1} \\ &= |\mathcal{C}|\end{aligned}$$

where $\mathbf{I}_{|C|}$ is the identity matrix of size $|\mathcal{C}|$ and $\mathbf{1}$ is the appropriate sized vector of all ones. □

Note that the aforementioned theorems and remarks hold for both binary, i.e. $a_{\mathrm{ij}} = \{0, 1\}$, and weighted, i.e. $w_{\mathrm{ij}} \in [0, 1]$, graphs. In this section, we use the graph theoretical notation for binary graph i.e. k_{i} represents the degree of node i although, in general, binary graph notations can be substituted by weighted graph ones without loss of generosity.

Based on the three types of nodes defined in the problem setting, we obtain the following remarks:

Remark 5. *The norm of $\mathbf{R}_{out}$:*

$$\begin{aligned}||\mathbf{R}_{out}||_2 &= \sqrt{\sum_{\mathrm{i}}^{|\mathcal{C}|} \left\{ \frac{k_{\mathrm{i}}^{out}}{k_{\mathrm{i}}} \right\}^2} \\ &= \sqrt{|I_3| + \sum_{\mathrm{i}_2=1}^{|I_2|} \left\{ \frac{k_{\mathrm{i}_2}^{out}}{k_{\mathrm{i}_2}} \right\}^2} \\ &= \sqrt{|I_3| + ||\mathbf{R}_{1[II]}||_2^2} = \eta_2\end{aligned}$$

in which i_2 is used to index Nodes Type II in $\mathcal{C}$; $||\mathbf{R}_{1[II]}||_2$ being the norm of external edge proportion of nodes Type II and $|I_3|$ is the cardinality of node type 3 in $\mathcal{C}$.

Remark 6. *Given a community $\mathcal{C}$,the lower bound of $||\mathbf{R}_{out}||_2$:*

$$\begin{aligned}
\eta_2 = ||\mathbf{R}_{out}||_2 &\geq ||\mathbf{R}_{out} + \mathbf{R}_{\mathrm{i}n}||_2 - ||\mathbf{R}_{\mathrm{i}n}||_2 \\
&\geq ||\mathbf{R}||_2 - ||\mathbf{R}_{\mathrm{i}n}||_2 \\
&= \sqrt{|\mathcal{C}|} - \sqrt{\sum_{\mathrm{i}}^{|\mathcal{C}|} \left\{ \frac{k_{\mathrm{i}}^{in}}{k_{\mathrm{i}}} \right\}^2} \\
&= \sqrt{|\mathcal{C}|} - \sqrt{|I_1| + \sum_{\mathrm{i}_2=1}^{|I_2|} \left\{ \frac{k_{\mathrm{i}_2}^{in}}{k_{\mathrm{i}_2}} \right\}^2} \\
&= \sqrt{|\mathcal{C}|} - \sqrt{|I_1| + ||\mathbf{R}_{\mathrm{i}n[II]}||_2^2} = \eta_1 > 0
\end{aligned}$$

in which i_2 *is used to index Node Type II in* $\mathcal{C}$*;* $||R_{2[II]}||_2$ *being the norm of internal edge proportion of nodes Type II and* $|I_1|$ *is the cardinality of node type 1 in* $\mathcal{C}$*. The first inequality is due to Result 1.*

We proceed to analytically show that $||\tau||_2$ intrinsically carries both *density-based* and *flow-based* notion of community in $\mathcal{G}$.

$$||\tau||_2^2 = \frac{\langle \tau^T \mathbf{R}_{out} \rangle}{||\mathbf{R}_{out}||_2 cos(\mathbf{R}_{out}, \tau)} \tag{3.2}$$

$$= \left[\frac{|\mathcal{C}|}{\eta_1 \eta_2} \right] \left[\frac{\alpha}{cos(\mathbf{R}_{out}, \tau)} \right] \tag{3.3}$$

where $\alpha = \frac{\eta_1}{\eta_2} \in (0, 1]$; per Remarks 3 and 4,

$$\eta_1 = \sqrt{|\mathcal{C}|} - \sqrt{|I_1| + ||\mathbf{R}_{out[II]}||_2^2} \tag{3.4}$$

and

$$\eta_2 = \sqrt{|I_3| + ||\mathbf{R}_{\mathrm{i}n[II]}||_2^2} \tag{3.5}$$

in which $||\mathbf{R}_{x[y]}||_2^2$ is the contribution of node type $y \in \{1, 2, 3\}$ to the L_2-norm of $\mathbf{R}_x \forall x \in \{\mathrm{i}n, out\}$. Combining equation 3 and 5, τ-induced modularity is scored as follows:

Hence, given an induced subgraph $\mathcal{C}$ in $\mathcal{G}$,

$$DM(\mathcal{C}) = \frac{|\mathcal{C}|}{\sqrt{(\eta_1 \eta_2)}}$$

and

$$SM(\mathcal{C}) = \frac{\sqrt{\alpha}}{cos(\mathbf{R}_{out}, \tau)}$$

represent density-based and flow-based modularity, respectively, of community $\mathcal{C}$ in $\mathcal{G}$. Note that the dependency of $\eta_1, \eta_2, \alpha, cos(\bullet, \bullet)$ on community $\mathcal{C}$ are dropped for notation simplicity.

τ & *Density-based* Communities

In this section, we show that the fitness score assigned to a community $\mathcal{C} \in \mathcal{G}$ effectively compliment the notion of dense subgraph in sparse graph, as mentioned in [111], among others. Specifically, per equation (4), density-based τ-induced modularity of $\mathcal{C}$ is scored as follow:

$$DM(\mathcal{C}) = \frac{|\mathcal{C}|}{(\eta_1 \eta_2)^{0.5}},$$

It follows that[1]

$$DM(\mathcal{C}) \propto \eta_1^{-1} \quad , \quad DM(\mathcal{C}) \propto \eta_2^{-1}$$

Using equation 3.4 and 3.5 and the fact that given any induced subgraph $\mathcal{C} \in \mathcal{G}$, $|\mathcal{C}|$ is fixed, we obtain:

$$DM(\mathcal{C}) \propto ||\mathbf{R}_{2[II]}||_2$$
$$DM(\mathcal{C}) \propto |I_1|$$
$$DM(\mathcal{C}) \propto ||\mathbf{R}_{1[II]}||_2^{-1}$$
$$DM(\mathcal{C}) \propto |I_3|^{-1}$$

It is trivial that the density-based score is increased with respect to the number of Type I nodes in $\mathcal{C}$ and the internal number of edges contributed by nodes type II. On the other

[1]↑Notation $a \propto b$ is used to denotes "a is proportional to quantity b".

hand, such score is penalized by the number of type III nodes and the external number of edges contributed by nodes type II. Collectively, τ compliments the traditional definition of a *good* community in network, i.e. particularly dense subgraph within sparse graph.

τ & *Flow-based* Communities

Mathematically, given a community $\mathcal{C}$, the $flow - based$ modular aspect of module $\mathcal{C}$ is scored as follows:

$$SM(\mathcal{C}) = \frac{\sqrt{\alpha}}{cos(\mathbf{R}_{out}, \tau)}$$

Since the three vectors, namely $\mathbf{R}_{out}$, $\mathbf{R}_{in}$, and τ, contains all non-negative elements, the angle between any two vectors are upper-bounded by 90^o which makes the denominator of $SM(\mathcal{C})$, $cos(\mathbf{R}_{out}, \tau)$, bounded in $(0, 1]$, see figure 3.8 below for further details.

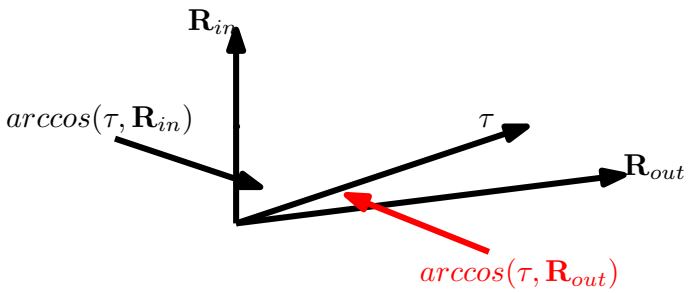


Figure 3.8. Schematic presentation of 3 vectors τ, $\mathbf{R}_{out}$ and $\mathbf{R}_{in}$ and their relationships in space $R^{|\mathcal{C}|}$.

Intuitively, one would expect that the more intra-edges $\mathcal{C}$ would "delay" random walker first visit absorbing states; and this is *generally* true; especially when we view this relationships through the lenses of τ - containing the $mfpt$ information, $\mathbf{R}_{out}$ - the "absorption" probability vector. It is important to note that $||\mathbf{R}_{in}||_2$ and $||\mathbf{R}_{out}||_2$ are competing norms i.e. $\max(||\mathbf{R}_{in}||_2) = \max(||\mathbf{R}_{out}||_2) = ||\mathbf{R}_{in} + \mathbf{R}_{out}||_2 = \sqrt{|\mathcal{C}|}$ where $||\mathbf{R}_{in}||_2 = \sqrt{|\mathcal{C}|}$ happens when module $\mathcal{C}$ is disconnected, i.e. no external connectivities and $||\mathbf{R}_{out}||_2 = \sqrt{|\mathcal{C}|}$ takes place if $\mathcal{C}$ has zero internal density. Geometrically, since $||\mathbf{R}_{out}||_2$ and $||\mathbf{R}_{in}||_2$ are competing norms, consequently, $arcos(\mathbf{R}_{out}, \tau)$ and $arcos(\mathbf{R}_{in}, \tau)$ are competing angles because their pairwise angle is, at most, 90^0. For example, if external edge(s) are deleted, while keeping

all other edges intact, then two things will follow: (1) the norm of τ and $\mathbf{R}_{in}$ increases and (2) the norm of $\mathbf{R}_{out}$ decreases, see figure 3.8 for further details.

τ and topological sensitivity

In terms of sensitivity (to topological perturbations), since $\tau = \mathbf{I}_{|\mathcal{C}|} - {\mathbf{Q}_{\mathcal{C}}}^{-1}\mathbf{1}_{|\mathcal{C}|}$, it is trivial to see that τ is unique and specific to $\mathbf{Q}_{\mathcal{C}}$. In other words, it is intolerant of any changes to the local adjacency structure (ultimately the graph topology). To illustrate this point, one can relate a graph with fixed n and m and perturb the current adjacency structure by randomly removing an edge and subsequently adding another edge, it is very likely that τ would be altered. For realistic network where topological symmetries are rare, the L_2–norm of τ would definitely depend on the amount of perturbation one makes to the original graph i.e. the number of edge swaps. In addition, we also provide a toy example of two induced substructures with, essentially, the same number of internal and external edges that is completely unrecognizable under Newman-Girvan modularity notion, [87] but under **TE**, these two configurations are much different, see figure 3.9 for details.

Final Remarks on τ and modularity

We collect key remarks and observations on time-to-absorption τ in this section: (i) τ access both $density-based$ and $flow-based$ modularity of a given community; (ii) Community with dense internal edge density might or might not embrace $flow-based$; (iii) Communities with the same internal and external edges might or might not have the same overall scores in terms of module escaping efficiency due to topological sensitivity offered by τ; (iv) there exists a trade-off relationship between *density-based* and *flow-based* modularity in which higher score in one aspect does not suggest high score in the other. For instance, when considering the configuration of clique size n and n exits, if an edge that connects a nodes with degree $n-1$ to the sole node with degree $2n-1$ is deleted then $DM(\mathcal{C})$ decreases while $SM(\mathcal{C})$ increases. Internally, both configurations are formed by a clique (of size 8), i.e. $\mathcal{K}_8$. Intuitively, one would expect two graphs with the same number of nodes and edges but different topological structures, i.e. topological edge arrangement, would have drastically

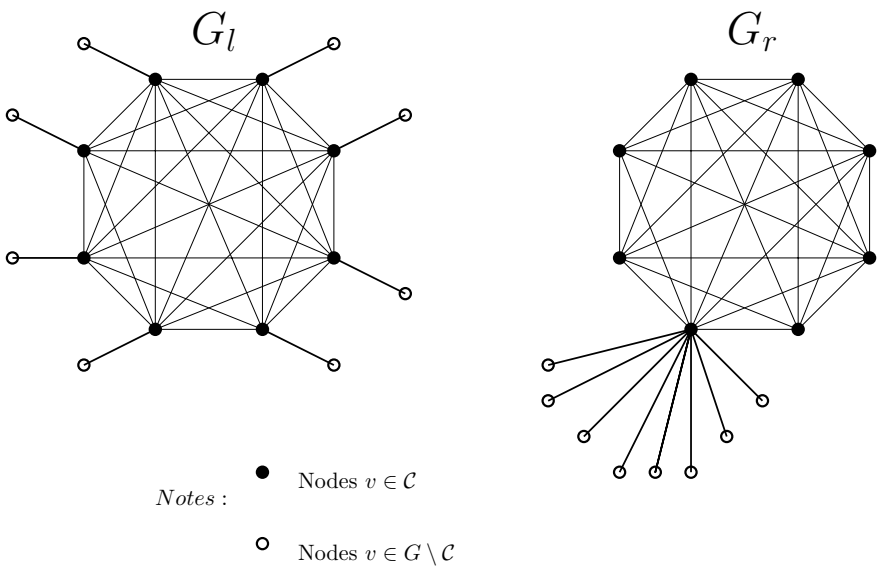


Figure 3.9. Schematic presentation between two *binary* graphs with the same number of internal and external edges.

different structural dynamics which is effectively measured through τ.Externally, the left configuration, denoted as G_l, has evenly-distributed exits while the right one, denoted as G_r, has congested/bottle necke- exit. Firstly, $SM(G_l) = 1$ because $\alpha = cos(tau, \mathbf{R}_{out}) = 1$. On the other hand, analogously, $SM(G_r) = \frac{\sqrt{\alpha}}{cos(\tau, \mathbf{R}_{out})} \approx 1.40$. Mathematically, the only non-zero entry in $\mathbf{R}_{out}(G_r)$ locates in node with bottle-necked exits which diminishes the denominator of $SM(G_r)$, i.e. $cos(\mathbf{R}_{out}, \tau)$ significantly. Hence, having the maximal numerator does not help the overall SM score for G_l. Another important result is that having the maximal internal subgraph like the clique structure does not necessarily help the notion of "trapped" random walker. As the matter of fact, some times, it carries side-effects. This is where the measure pushes beyond the "particular density within sparsity" community notion such as Newman-Girvan modularity.

3.6.3 Module Exit Entropy

Module Exit Entropy (denoted as **EE**, and in the range $\mathbf{EE} \in (0, 1]$ and unitless) assesses the normalized level of uncertainty in selecting an exiting node in S_{abs} of a random particle that starts in $\mathcal{C}$. The exit entropy, denoted as $\mathcal{H}_{\mathrm{e}}$, measures the level of uncertainty exiting

node j $\in S_{abs}$ (outside of the module) is preferred. Module exit entropy is mathematically formalized as:

$$\mathbf{EE} = \frac{\mathcal{H}_{\mathrm{e}}}{\mathcal{N}_{\mathcal{C}}} = \frac{-\sum_{\mathrm{i}=1}^{|S_{abs}|} \psi_{\mathrm{i}} \log(\psi_{\mathrm{i}})}{\log(|S_{abs}|)} \tag{3.6}$$

where preferential exit probability is the probability vector that contains $|S_{abs}|$ entries that represents the likelihood of an exit signal selects a specific exiting state j $\in S_{abs}$ such that $\sum_{\mathrm{j}\in S_{abs}} \psi_{\mathrm{j}} = 1$.

The numerator of $\mathbf{EE}(\mathcal{C})$, i.e. $-\sum_{\mathrm{i}=1}^{|S_{abs}|} \psi_{\mathrm{i}} log(\psi_{\mathrm{i}})$, measures the degree to which channels of communication between nodes in S_{trans} and S_{abs} are preferred for a fixed task/subject. It is noteworthy that **EE** is not influenced by the (cumulative) magnitudes (of functional connectivity values) that connect nodes from within the FN to outside (exiting) nodes. It is only affected by the distribution of such values. In particular, homogeneous distributions display high entropy levels and uneven distributions favoring certain exiting node(s) display low entropy. To demonstrate this point, an example is provided in **SI** under section **C.3**. The normalizer, $\mathcal{N}_{\mathcal{C}} = log(|S_{abs}|)$, is the maximum entropy obtained from a module in which all exit nodes have the same absorption rate. Numerically, a high **EE** would denote the homogeneous integration within the rest of the system whereas a low **EE** would indicate a preferential communication or integration of the module with the rest of the system. In terms of functional brain networks, module exit entropy facilitates the understanding of collective behavior from $\mathcal{C}$ to other FNs through its outreach channels (edges formed by nodes in $\mathcal{C}$ and exiting nodes in $G \setminus \mathcal{C}$). This is because entropy measures the level of uncertainty in communication; hence, lower entropy means higher specificity in communication between the FN with the rest of the cortex. **EE** value ranges are bounded between 0 and 1.

Numerator

The numerator of $\mathbf{EE}(\mathcal{C})$, i.e. $-\sum_{\mathrm{i}=1}^{|S_{trans}|} \psi_{\mathrm{i}} log(\psi_{\mathrm{i}})$, measures the extent to which specified channels of communications, under finest scale (i.e. node/edge-level), is established between nodes in $\mathcal{C}$ with nodes that belong to other functional communities in G. Therefore, $-\sum_{\mathrm{i}=1}^{|S_{trans}|} \psi_{\mathrm{i}} log(\psi_{\mathrm{i}})$ captures the properties of the distribution of w_{ij} as a whole, represented by $w_{\mathrm{ij}} \forall \mathrm{i} \in \mathcal{C}, \mathrm{j} \in \mathcal{J}$.

For example, let us say that we have two communities with the same state set $S_{trans} = 1, 2, 3$ and $S_{abs} = a, b, c, d$. In community 1, $w_{\text{ij}} = 0.01, \forall \text{i} \in S_{trans}, \text{j} \in S_{abs}$; and community 2, $w_{\text{ij}} = 0.9, \forall \text{i} \in S_{trans}, \text{j} \in S_{abs}$. Once we compute the entropy, for both cases, we see that they both have no communication preference, hence numerator is one for both cases.

Normalization

This is the coordinate where normalization is possible. Note that since entropy is normalized by its maximum value (i.e. $log(|S_{trans}|)$), the number of exits $|S_{trans}|$ impacts is, consequently, neutralized. Thus, one does not need to concern about the cardinality of a community with respect to its number of exits as, in reality, a typically larger community usually carries more exits.

3.6.4 TE, EE behavior across thresholds

In this section, we explore some further characterization of these new metrics used in the morphospace. Specifically, we explore how **TE** and **EE** depend on FC density, across different thresholding values. To do so, we performed the analysis of morphospace measure **TE** and **EE** with different thresholding values 0 to 0.2 at an increment of 0.05. Specifically, for each subject, each task, each FN, and each threshold value, we compute **TE** and **EE**. We then average across subjects, and tasks to obtain **TE** and **EE** for each threshold value.

From figure 3.10, we observe that the values of **TE** and **EE** are stable only for very small magnitudes of the threshold, up to $\tau = 0.075$. As threshold increases, and more functional edges with low-to-intermediate values are removed, the **TE** increases and **EE** decreases. This happens because these edges with low-to-intermediate values, are most likely inter-network connections (i.e. functional edges connecting different FNs) , and their removal makes those networks more segregated, ultimately producing and impact in both **TE** and **EE** measures.

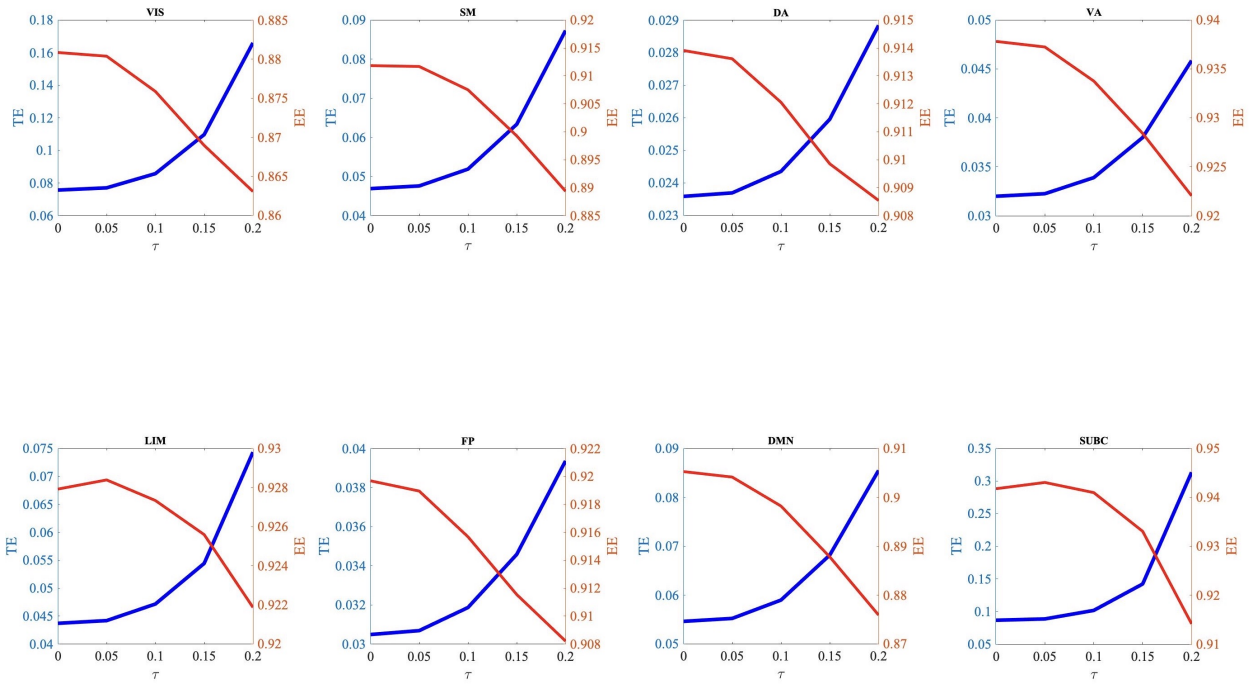


Figure 3.10. Morphospace Metric values (**TE** and **EE**) at different thresholds of τ for each functional network within the range $[0, 0.2]$.

The definition of the Mesoscopic Morphospace Ω

The two distinct features of each FN in brain graphs are addressed by a point $\mathbf{u}(\mathcal{C})$ in $\Omega \subset (0, M) \times [0, 1] \subset R^2$ as follows:

$$\mathbf{u}(\mathcal{C}) = (\mathbf{TE}(\mathcal{C}), \mathbf{EE}(\mathcal{C})) \in \Omega \tag{3.7}$$

where $M < \infty$. for a given subject and task, a functional brain network G is obtained with a pre-defined parcellation that results in l induced subgraph $\mathcal{C} \subset G$, we can obtain l points $\mathbf{u}(\mathcal{C})$ corresponding to l FNs in network G.

In general, trapping efficiency, $\mathbf{TE}(\mathcal{C})$ is finitely bounded by construction (see more details in **Section D.6** under **SI**). However, a better bound is possible for the HCP dataset used for this study. This is due to two driving factors: connectome sparsity and edge weights

[64]. We address the upper bound for **TE** as: $\max(\mathbf{TE}(\mathcal{C})) = M = 1$. In terms of $\mathbf{EE}(\mathcal{C})$, its numerical range $\mathbf{EE}(\mathcal{C}) \in (0, 1]$. Hence, $\Omega \subset (0, 1) \times [0, 1]$ for this dataset.

3.6.5 Morphospace null Model

Randomization of $G(V, E)$

Given a weighted network $A = [w_{\mathrm{ij}}]$, we apply randomized algorithm Xswap, see [112] with number of desired changes that are set to be $[2, 2^3, 2^5..., 2^{19}]$ (with exponent increment of 2) and maximum iterations set at 100 times the corresponding changes. This algorithm **preserves network basic topological characteristics** such as size, density and degree sequence. As the desired number of changes increases, the difference between the original matrix, denoted as $\mathbf{A}_{orig}$, and the randomized counterpart, denoted as $\mathbf{A}_{rand}$, also increases. The difference between two graphs can be quantified as follows:

$$Diss = \frac{\sum_{\mathrm{i,j}=1}^{n} |\mathbf{A}_{rand}(\mathrm{ij}) - \mathbf{A}_{orig}(\mathrm{ij})|}{\sum_{\mathrm{i,j}=1}^{n} \mathbf{A}_{rand}(\mathrm{ij})}$$

where n is graph's size and $Diss \in [0, 1]$. It is important to note that the difference between two graphs saturates after a certain number of changes and each graph topology saturates at different values (not necessarily 1).

Hence, getting $Diss$ to arbitrarily close to saturation with the smallest number of changes is genuinely the target for this procedure. In our case, we pick participant 100307 and run the randomization procedure for all available tasks and rest. We first found that the acceptable $Diss$ occurs at 2^{15} desired changes at resting state. We note that the $Diss$ saturates at 0.6 because the (sub)graphs we are dealing with are very dense and some links will be repeated in force, leading to a non-zero overlap between the links of the graphs in the random ensemble and the original one. We then used the same number of changes for the investigated tasks in this dataset, see figure 3.11 for further details.

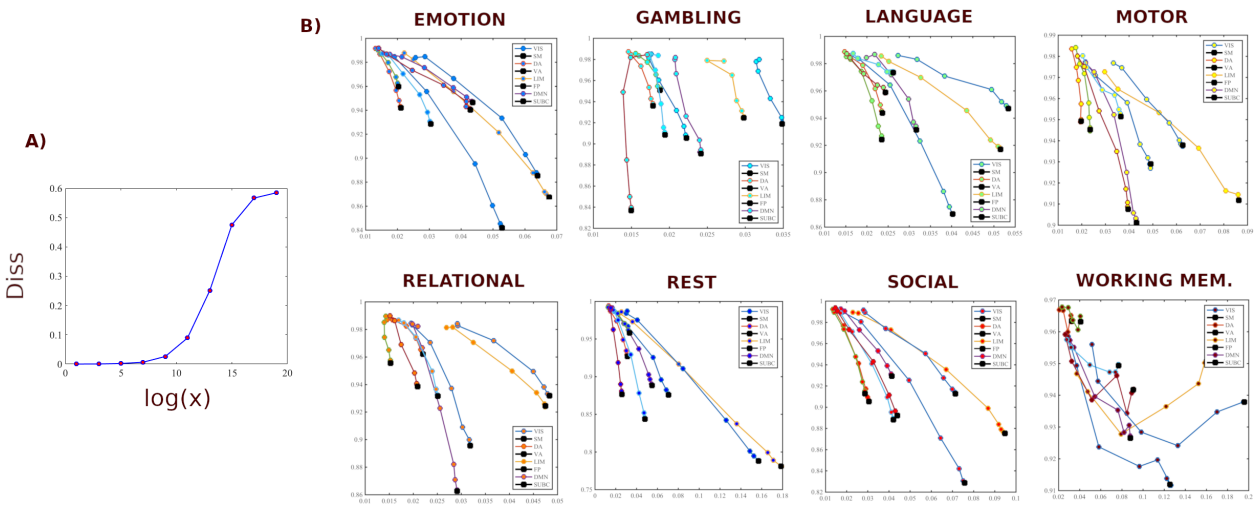


Figure 3.11. Morphospace Null Model for Subject 100307 across resting state and 7 other fMRI tasks.

The Null Model

The main drive for studying trajectories, through randomization, is because it could provide further evidence of the robustness in design of the metrics. Specifically, if done correctly, measurements should highlight unique characteristics of functional communities (and not the randomized counterpart). As the randomized graph (with topological preserved features) get assigned the same partition into functional networks (e.g. Yeo's parcellation) as the original one, any destruction of such topology, at global scale G, would also be carried over (hence, identified) by the morphospace itself. One common theme emerges is that regardless of which functional network and task, as the dissimilarity increases with the desired number of changes, all functional communities are pushed towards to top left corner. This regime of the morphospace represent random exiting strategy from module $\mathcal{C}$ (high value of **EE** - high level of uncertainty in communication preference) and high degree of non-assortative community (low **TE** - low level of segregation).

In panel **A)** in the above figure, Participant 100307 with resting state functional connectome is randomized using Xswap procedure. Panel **B)** represents the application of the selected number of changes to all other tasks. This is an important results to see that functional networks' topology is truly well-defined and highly reproducible across subject

domain. Note that black square dot denote functional community **TE** and **EE** with no randomization.

3.7 The network configural breadth formalism

Studying the manifold topology defined in this 2D mesoscopic morphospace theoretically requires an infinite amount of points. In finite domain with discrete sampling of the morphospace, polytope theory, a mathematical branch that studies object geometry, allows us to create a reasonable scaffold presentation with well-defined properties to formally define and quantify configural components of the functional networks.

Given a set of points in this space, $W = \left\{\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{x}_{|W|}\right\}$, a convex hull formed by W is represented by

$$\mathbf{Conv}(W) = \left\{ \sum_{\mathrm{j}=1}^{|W|} \alpha_{\mathrm{j}} \mathbf{x}_{\mathrm{j}} \mid \sum_{\mathrm{j}=1}^{|W|} \alpha_{\mathrm{j}} = 1, \alpha_{\mathrm{j}} \geq 0 \right\}$$

One can compute the notion of volume of the convex hull enclosed by $\mathbf{Conv}(W)$, denoted as $Vol(\mathbf{Conv}(W))$. Given that the morphospace is 2D, the manifold dimension can be from 0 up to 2.

The functional network configural breadth, for the i^{th} subject, is compartmentalized into two components:

- FN (task) reconfiguration and
- FN rest-to-[task-positive] preconfiguration.

We then propose a mathematical relation between network configural breadth with FN reconfiguration and preconfiguration as follows:

$$\mathcal{F}_{\mathrm{i}} = f(\mathcal{R}_{\mathrm{i}}^{FN}, \mathcal{P}_{\mathrm{i}}^{FN}) \tag{3.8}$$

where $\mathcal{F}_{\mathrm{i}}$ represents configural breadth for subject i^{th}. Here, we provide directly the measures that quantify (functional) reconfiguration and preconfiguration of FNs for i^{th} subject's configural breadth, see figure 3.12 for further details. Tasks are assigned the same level of importance and hence, no task is weighted more than others. **Functional Network con-**

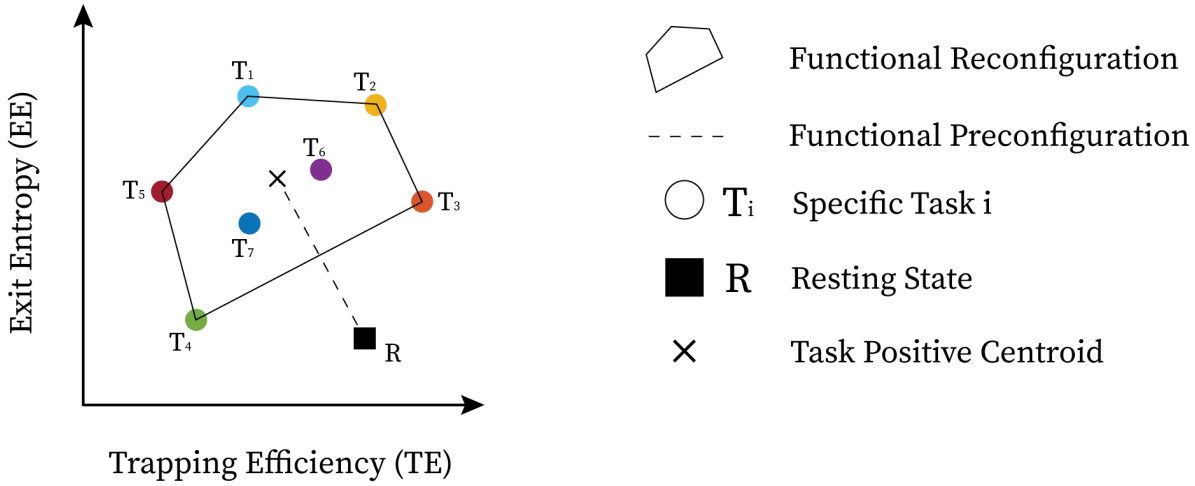


Figure 3.12. Geometrically presentation of functional pre- and re-configuration.

figural breadth is geometrically represented using two predefined morphospace measures. Specifically, for mesoscopic structures such as communities in functional brain networks, the first measure is Trapping Efficiency (**TE**) while the second is is Exit Entropy (**EE**). In this case, tasks T1 to T5 belong to the convex hull (e.g. Pareto front - further details are discussed in the first chapter of this dissertation) while T6 and T7 is in the interior enclosed by the convex hull.

3.7.1 Functional Reconfiguration

Definition 3.7.1. *Functional reconfiguration in this work is represented by a 2-dimensional spatial volume derived from given FN's **EE** and **TE** coordinate values across different cognitive tasks. As such, it represents an example of "cognitive space" [80], [92] within a functional domain that spans a variety of network states under various task-evoked conditions. We quantify this as*

$$\mathcal{R}_{\mathrm{i}}^{FN} = Vol(\mathbf{Conv}(W_{\mathrm{i}}^{FN})) \tag{3.9}$$

where W_{i}^{FN} *represents the set containing all investigated task coordinates of subject* i*'s FN;* $\mathbf{Vol}(Conv(W_{\mathrm{i}}^{FN}))$ *is the convex hull volume induced by points in* W_{i}^{FN}*.*

For a given subject ith's FN, note that $\mathbf{Conv}(W_{\mathrm{i}}^{FN})$ represents the broad span (breadth) of task configurations for a given functional community. Subsequently, $\mathcal{R}_{\mathrm{i}}^{FN}$ represents the amount of breadth as measured by the volume of $\mathbf{Conv}(W)$. Functional reconfiguration for a given subject's FN, denoted as $\mathcal{R}_{\mathrm{i}}^{FN}$, is geometrically depicted in the above figure.

Furthermore, since we can only obtain finite number of tasks (hence, points in this space), we see that convex hull notion is logical to represent distinct points (FN tasks) that constitute the Pareto front (hull boundary). To measure the notion of capacity (potential to shift), one needs to measure the notion of spreading given finite number of points in the hull. If we use first order measurements such as distance among two points in the hull, we face the following problems:

- inability to capture the reservoir defined by the interior of the convex hull;
- assumption of linearity between task points

The notion of distance does not cover the space of possibility [64] parameterized by **TE** and **EE**. Hence, second order measurement, i.e. volume (or area in this case), is more appealing.

3.7.2 Functional Preconfiguration

Definition 3.7.2. *Functional preconfiguration reflects the topologically distributed equipotentiality that is theoretically designed to enable an efficient switch from a resting state configuration to a task-positive state [15], and is quantified as follows*

$$\mathcal{P}_{\mathrm{i}}^{FN} = ||Rest_{\mathrm{i}}^{FN} - \eta_{W_{\mathrm{i}}^{FN}}||_2 \tag{3.10}$$

where $\eta_{W_{\mathrm{i}}^{FN}}$ *is the geometrical centroid of* W_{i}^{FN}*;* $\mathcal{P}$ *measures the distance between rest to task-general position (represented by* $\eta_{W_{\mathrm{i}}^{FN}}$*). It is defined with the selected metric space, in this case is the* 2*-norm in Euclidean space.*

Analogously, once the points are well-defined in this space, in order to effectively measure the notion of functional preconfiguration, we need to highlight the functional readiness, from a cognition standpoint, to switch between resting configuration to a generic task. Here, we first provide the formula proposed in main-text for functional preconfiguration:

$$\mathcal{P}_{\mathrm{i}}^{FN} = ||Rest_{\mathrm{i}}^{FN} - \eta_{W_{\mathrm{i}}^{FN}}||_2$$

where $Rest_{\mathrm{i}}^{FN}$ and $\eta_{W_{\mathrm{i}}^{FN}}$ represent FN coordinate at rest, and geometric centroid considering all FN tasks.

Firstly, the geometric centroid of all FN task coordinates might or might not be cognitively possible, i.e. there might not be a connectome that result in FN task centroid being numerically exact. However, that is not the purpose of using this notion. If the goal is to reflect the degree of functionally readiness between resting and task-engagement, the notion of distance, in this case, is meaningful. Here, complexity of trajectory between rest and task-evoked condition is irrelevant to consider.

Note that functional preconfiguration can be viewed as $Vol(\mathbf{Conv}(W))$ where the convex hull is defined solely by two points: FN's rest and FN's geometrical centroid of task convex hull, i.e. $W = \left\{R_{\mathrm{i}}^{FN}, \eta_{W_{\mathrm{i}}^{FN}}\right\}$. In such regards, the notion of $Vol(\mathbf{Conv}(W))$ is also suitable to describe the configural breadth between rest and task positive location. Functional preconfiguration is geometrically depicted in the above figure.

3.8 Individual Fingerprints and Behavioral measures

3.8.1 Subject Sensitivity

To quantify subject sensitivity (through network configural breadth), for each subject/scan, we obtain one measure. We then concatenate the data into a 100 by 2 matrix and run intra-class correlation (ICC) analysis. To test subject sensitivity result robustness, for each functional network's preconfiguration or reconfiguration, we keep one column of ICC input intact (Test) and shuffle the second column (Retest) and measure ICC for each permutation. The same procedure is repeated 10,000 times and the 95%-ile is reported in the main text.

3.8.2 Iterative Multi-Linear Regression Model (MLM) Description

We apply iteratively multi-linear correlation models (MLM) to correlate $\mathcal{F}_{\mathrm{i}} = f(\mathcal{R}_{\mathrm{i}}^{FN}, \mathcal{P}_{\mathrm{i}}^{FN})$ with various behavioral measures, m_{i}. We hypothesize there exists a high level of functional subject fingerprint. Iteratively, we start by using only 1 predictor ($\mathcal{P}^{FP}$); in every subsequent step, we append one extra predictor to the existing predictor(s). At the end of iterative process, we consequently obtain 16 MLMs.

In order to pick the best MLM (and their corresponding number of linear descriptors in the model), we use the model with smallest p-value among all 16 MLMs.

Model Specificity (MS)

Constructing the MLM to infer the intrinsic relationship is a necessary but not sufficient if the ultimate goal is to discover if there is a truly robust relationship between network configural breadth and behavioral measures. Such robustness exists, then there has to be a certain degree of specificity in these models such that only significant correlations are observed when linear predictors are correlated with the true behavioral measures. Specifically, network configural breadth, as mathematically formulated using linear descriptors, must show that it is strongly correlated with a designated measures and not anything else, say a randomized vector.

Model Description

We further test the strength of our hypothesis by splitting available data into two subsets: test and validation set. Specifically, we first extract the optimal number of predictors by applying the procedure described in the main article. We then proceed with the model specificity by creating 2000 simulations; for each simulation - indexed by $\mathrm{j} = \{1, 2, 3..., J = 2000\}$ - we first find a randomized order of indices from 1 to 100, denoted as $\vec{d}$, and divide them into five batches (indexed by $\mathrm{i} = \{1, 2, ..., I = 5\}$) of 20 subjects. In other words, each batch of 20 randomly picked subjects, indexed by the set Q_{i}, are used to validate the authenticity of the coefficients proposed by utilizing the remaining 80 unpicked subjects. We see that

we recover the permutation of the randomized order vector as follows: $\vec{d} = Q = \cup Q_{\mathrm{i}}$. It is important to note that we use this procedure because it minimizes the chance of picking the same (or highly overlapped) batch of 20 subjects. For each simulation j, in each batch i, the remaining 80 subjects are then used to acquire multi-linear correlation model's parameters, denoted as $\vec{\beta} \in R^{[*]}$ where $[\,*\,]$ denotes the optimal MLM driven by procedure described in previous section. These corresponding coefficients are then used to predict the remaining 20 unused data points, indexed by $w \in W_{\mathrm{i}}$, denoted as $\hat{y}$.

$$\hat{y}_w = \vec{\beta}_0 + \left[\left\{\mathcal{P}_w^{FN}, \mathcal{R}_w^{FN}\right\}^{[*]}\right] \vec{\beta}$$

where $\left\{\mathcal{P}_w^{FN}, \mathcal{R}_w^{FN}\right\}^{[*]} \in R^{+,[*]}$ is the $[\,*\,]$-tupled vector representing functional preconfiguration, reconfiguration, obeying the descending order of concatenated subject sensitivity. Next, for each batch, we compute the correlation between actual values, y_w with predicted ones, $\hat{y}_w$ and record the correlating result, denoted as $R_{\mathrm{i}}, \quad \forall \mathrm{i} = 1, 2, ..., I = 5$. Consequently, at each simulation, we obtain 5 values of R_{i} corresponding to 5 batches. Lastly, for each simulation j, the mean and standard deviation of 5 validation models R_{i}'s is obtained

$$R_{\mathrm{j}} = \sum_{\mathrm{i}=1}^{I} R_{\mathrm{ij}} = \langle R_{:,\mathrm{j}} \rangle \quad \& \quad \sigma_{\mathrm{j}} = \sqrt{\frac{\sum_{\mathrm{i}=1}^{I} (R_{:,\mathrm{j}} - R_{\mathrm{j}})^2}{I}}$$

Per Central Limit Theorem, the statistic $R_{\mathrm{j}} \mid \forall \mathrm{j} = \{1, 2, ..., J = 2000\}$ is normally distributed, i.e. $R_{\mathrm{j}} \sim N(\mu_0, \sigma_0)$. This would create an empirically normal distribution $R_{\mathrm{j}} \sim N(\mu_0, \sigma_0)$ such that

$$\mu_0 = \frac{\sum_{\mathrm{j}} \sum_{\mathrm{i}} R_{\mathrm{ij}}}{I \times J} \quad \& \quad \sigma_0 = \sqrt{\frac{\sum_{\mathrm{j}=1}^{J} \sigma_{\mathrm{j}}^2}{J}}$$

The null model and paired t-test

Similarly to the MLMs, we want to test the authenticity of selected models by testing it against artifacts such as random vectors. The same procedure is applied for the random vector to populate the null model's empirically normal distribution (its means is notated as μ_1): $R_{\mathrm{j}}^{rand} \sim N(\mu_1, \sigma_1)$. Finally, paired t-tests are applied between the two aforemen-

tioned distributions, i.e. R_{j} and R_{j}^{rand}, to test the capacity of configural breadth predictors towards behavioral measures. Interestingly, given the investigated behavioral measures, all null model empirical distributions have very similar first and second moments, independently of behavioral measures.

3.9 Results

The mesoscopic morphospace formalized in previous sections is used to assess network configural breadth in terms of functional preconfiguration and reconfiguration for the one hundred unrelated subjects of HCP 900 subjects data release [47], [48]. This dataset includes (test and retest) sessions for resting state and seven fMRI tasks: gambling (GAM), relational (REL), social (SOC), working memory (WM), language processing (LANG), emotion (EMOT), and motor (MOT). Whole-brain functional connectomes estimated from this fMRI dataset include 360 cortical brain regions [62] and 14 subcortical regions. The functional communities evaluated in the morphospace include seven cortical resting state FNs from [1]: visual (VIS), somatomotor (SM), dorsal attention (DA), ventral attention (VA), frontoparietal (FP), limbic (LIM), default mode (DMN) and one comprised of subcortical regions (SUBC).

3.9.1 Task- and subject-sensitivity

Within- and between-subject task sensitivity

We first evaluate the capacity of module trapping efficiency and exit entropy to differentiate between tasks **within** subject in below figure. For both test and retest sessions of each subject, we compute the **TE** and **EE** metrics for each FN. We compute these values for all 8 fMRI conditions. We compute the intraclass correlation coefficient (ICC), with test and retest (per subject) being the repeated measurements and task being the class variable (**TE** (top panels) and **EE** (bottom panels), respectively, where each ICC is computed using a 2 (test, retest) by 7(tasks) design, and the ICC reflects task within-subject sensitivity). For most subjects, ICC values in all FNs are high and positive values. **EE** displays a higher within-subject task sensitivity than **TE**. Specifically, **TE** in VIS, DA and DMN most distin-

guished between the cognitive tasks, whereas **EE** in VA and FP was best at distinguishing the within-subject task-based configural changes. The ICC values for both coordinates were the lowest for LIM. In figure 3.13 Panel A is Within-subject task sensitivity of Module trap-

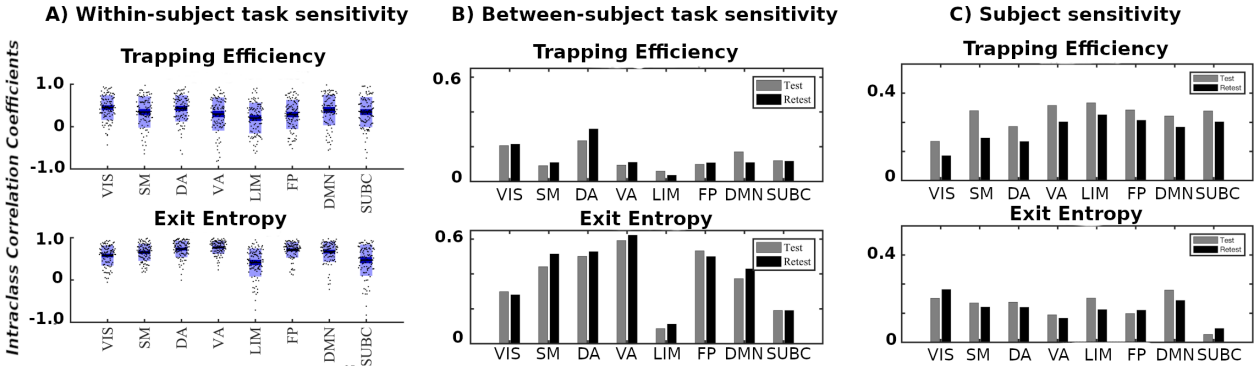


Figure 3.13. Morphospace measures and their task- and subject- sensitivity measured by intra-class correlation coefficients for each functional network.

ping efficiency (**TE**) and exit entropy (**EE**) for each FN per subject . figure 3.13 Panel B is Between-subject task sensitivity of **TE** (top) and **EE** (bottom). Finally, figure 3.13 - Panel C represents Subject-sensitivity ICC of **TE** (top) and **EE** (bottom).

We then evaluate the degree to which morphospace metrics capture cohort-level configural changes. To test this, for each morphospace metrics (**TE** or **EE**), we compute ICC with each FN where subjects as the repeated measures and task the class variable. We performed the evaluation separately for test and retest sessions as denoted by gray and dark bars, respectively. **EE** captures cohort-level task-based signatures as ICC values are consistently higher than those of **TE**. Interestingly, LIM has the lowest cohort-level task-based sensitivity for both morphospace metrics.

Subject sensitivity across tasks

Here, we compute ICC considering the tasks (fMRI conditions) the repeated measurements and subjects as the class variable. It is note-worthy that **TE** is superior in uncovering subject fingerprints, compared to **EE**, for the majority of FNs. This is complementary to **EE** being more task-sensitive.

TE and EE are disjoint features

Results in these sections suggests that **TE** and **EE** have the differentiating capacity to highlight non-overlapping characteristics of objects under consideration, i.e. task- and subject- based FNs. First of all, for within-subject task differentiation, FNs with high ICC values in one measure do not necessarily show a similar tendency in the other. For instance, VA has the third lowest mean **TE** value in characterizing within-subject tasks differentiation but it has the highest mean **EE** score. Similarly, FP has second lowest average **TE** score while third highest **EE** score indicating that each of the two measure captures unique aspects of a given FN. Secondly, evidence of disjoint features is shown through the ICC results in cohort-level task-sensitivity and subject-sensitivity configural changes. Indeed, **TE** is superior in detecting subject fingerprints while **EE** is better in unraveling task fingerprints. The idea is that, for a given studied object (i.e. task-based FNs), configurations are shown to "stretch" in exclusive/disjoint directions (subject-sensitive trapping efficiency and task-sensitive exit entropy).

3.9.2 Quantifying network configural breadth on functional networks

The mesoscopic morphospace allows the quantification of network configural breadth. For a given functional community, we compute functional reconfiguration (degree of configurations across tasks) and preconfiguration (distance from rest to task positive state).

Group-Average Results

The group average behavior of functional communities is shown in figure 3.14. Functional reconfiguration of FNs are shown as filled convex hulls whereas preconfiguration of FNs are shown as dashed lines from rest to the corresponding task hull geometric centroid.

In figure 3.14, functional reconfiguration and preconfiguration for all FNs are represented using group average of individual subjects' coordinates. Task coordinates in this space are represented by either asterisk (*) or the plus (+) symbol. The asterisk symbol is used for those tasks that are part of the Pareto front of the convex hull; the plus symbol represents

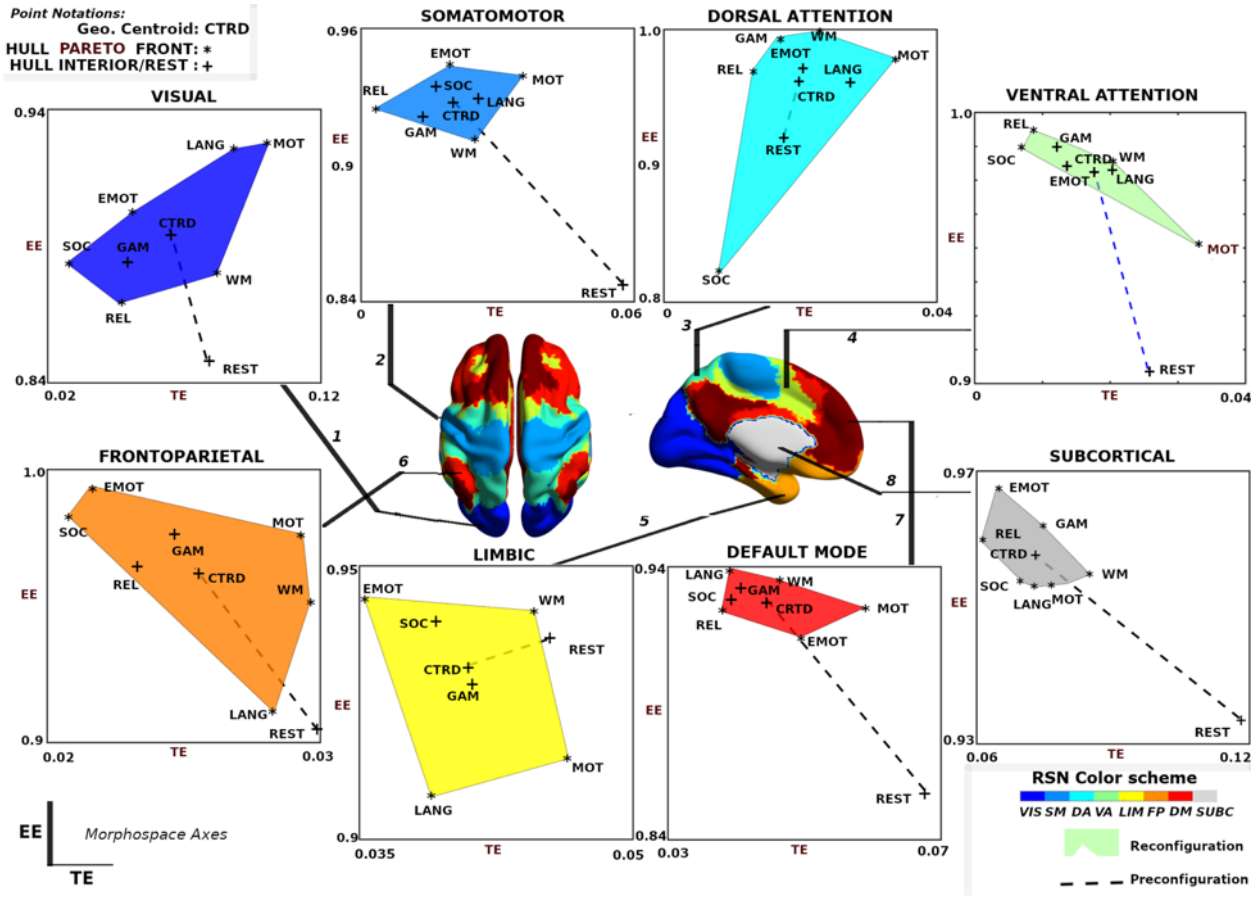


Figure 3.14. Visualization of Differential Network Configural breadth among different fcMRI functional networks distributed across the association cortex.

either the resting state or task that belongs to the interior of the convex hull. Note that $x-$ and $y-$ axis are purposely not scaled in the same range so that the full range of values for all tasks, task-centroid, and rest can be more easily visualized.

To facilitate comparing network configural breadth across all functional networks, these same convex hulls are shown in figure 3.15 panel A with the same $x-$ and $y-$ axis values. VIS network polytope, representing group-average behavior, is lower in **EE** relative to other FNs.

With the exception of VIS and SUBC, all other FNs cluster in a similar, high **EE**/ low **TE** area of the morphospace (see figure 3.15 panel A). It should be noted that different tasks and subject populations (e.g., older or clinical groups) might cluster FNs differently.

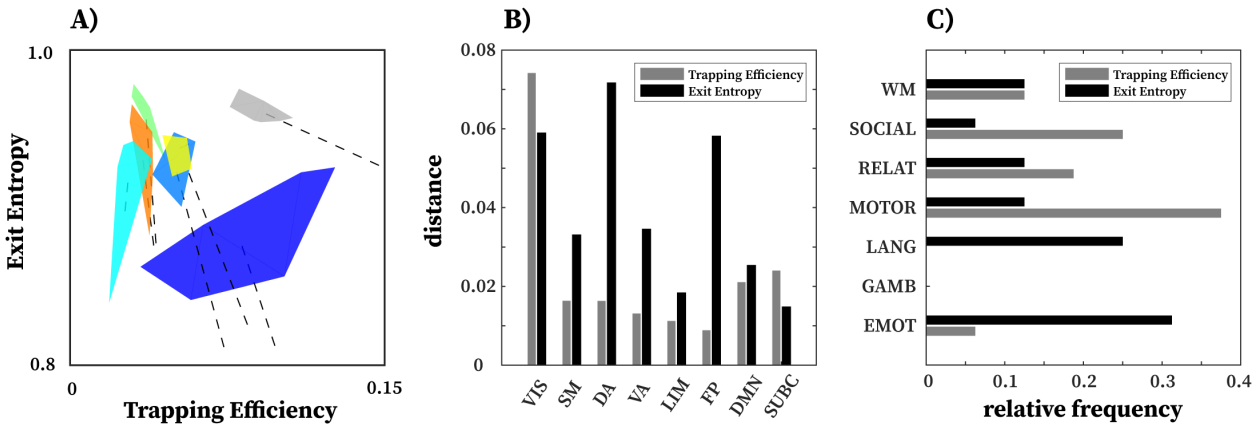


Figure 3.15. Network Configural Breadth insights are unravelled through mesoscopic morphospace. The polytope colors are the same as 3.14.

We also note that the subcortical polytope is relatively high in exit entropy. However, the subcortical parcellation might not optimally reflect the functional and/or structural makeup of various subcortical regions (e.g., role of the basal ganglia in the motor system) so these results should be interpreted cautiously.

One observation drawn from such a presentation is that the morphospace framework reconfirms, quantitatively, that functional dichotomy of the brain between task-positive and rest state [113]. Specifically, the default mode network acts more as a segregated module with high level of integration specificity at rest - as seen in the lower right regime with high **TE**, low **EE** values - as opposed to under task-evoked conditions - as seen in the top left corner with low **TE**, high **EE** values (see figure 3.14 *Default Mode*) [40], [113].

Another observation is that in terms of segregation level measured by **TE**, the lower bound of subcortical convex hull is, approximately, the upper bound of other FNs, with the exception of the visual network. The above figures also summarize functional reconfiguration and preconfiguration respectively, for test and retest fMRI sessions in all subjects and FNs. Here, the VIS system displays the largest functional reconfiguration, functional preconfigurations display a more comparable magnitude among all FNs.

Further evidence of disjoint feature is also displayed in the below figures. Maximal distance is computed using pairwise distances for two given tasks for a specific FN. The result shows that for a given FN, the two measures complement each other and in many

cases, stretch the cognitive space in one direction or the other. For instance, in case of DA and FP, the maximal distance in **EE** is very high but low for **TE** whereas in VIS and SUBC, **TE** maximal distance is higher than that of **EE**. Furthermore, only specific tasks (e.g. Motor and Emotion) that push the cognitive space in particular direction (which is captured by maximal distance computation). Evidence of disjoint features is also illustrated by the relative frequency of Motor and Emotion tasks for which **TE** and **EE** are complimentary.

Figure 3.14- panel **(A)** illustrates network configural breadth for all functional communities where *polytope colors are analogous to the ones scheme shown in previous figure*. For each functional community, the dash line represents the amount of functional preconfiguration whereas the polytope volume represents the amount of functional reconfiguration. Figure 3.14-Panel **(B)** represents maximal distance is computed using the maximum pairwise distance between two tasks for a given functional network. Finally, Figure 3.14-Panel **(C)** is the relative frequency with which a task appears in the maximal distance normalized by 16 (8 FNs and 2 task per FN).

Subject specificity of pre- and reconfiguration of functional networks

The formulation of network configural breadth (in terms of preconfiguration and reconfiguration) enables us to assess these properties at the subject level.

Figure 3.16 - Panels **1** and **2** show functional reconfiguration and preconfiguration, respectively, from both magnitude and subject-sensitivity viewpoints. For each functional network, the **A** Panels of figure 3.16 report subject's preconfiguration and reconfiguration values whereas the **B** Panels quantify subject sensitivity. Reconfiguration and preconfiguration measures are displayed in blue and red, respectively. Panel **C** of figure 3.16 merges all 16 configural breadth terms in descending order of subject sensitivity.

In the above figure, we use ICC to analyze the ability of morphospace measures (in the form or reconfiguration (panels 1 of figure 3.16) and preconfiguration (panels 2 of figure 3.16) to reflect subject identity within each FN. For all FNs from [1], the ICCs suggest that subjects can be differentiated from each other when contrasted against a corresponding null model. We see that subject sensitivity scores of all eight FNs for both pre- and re-

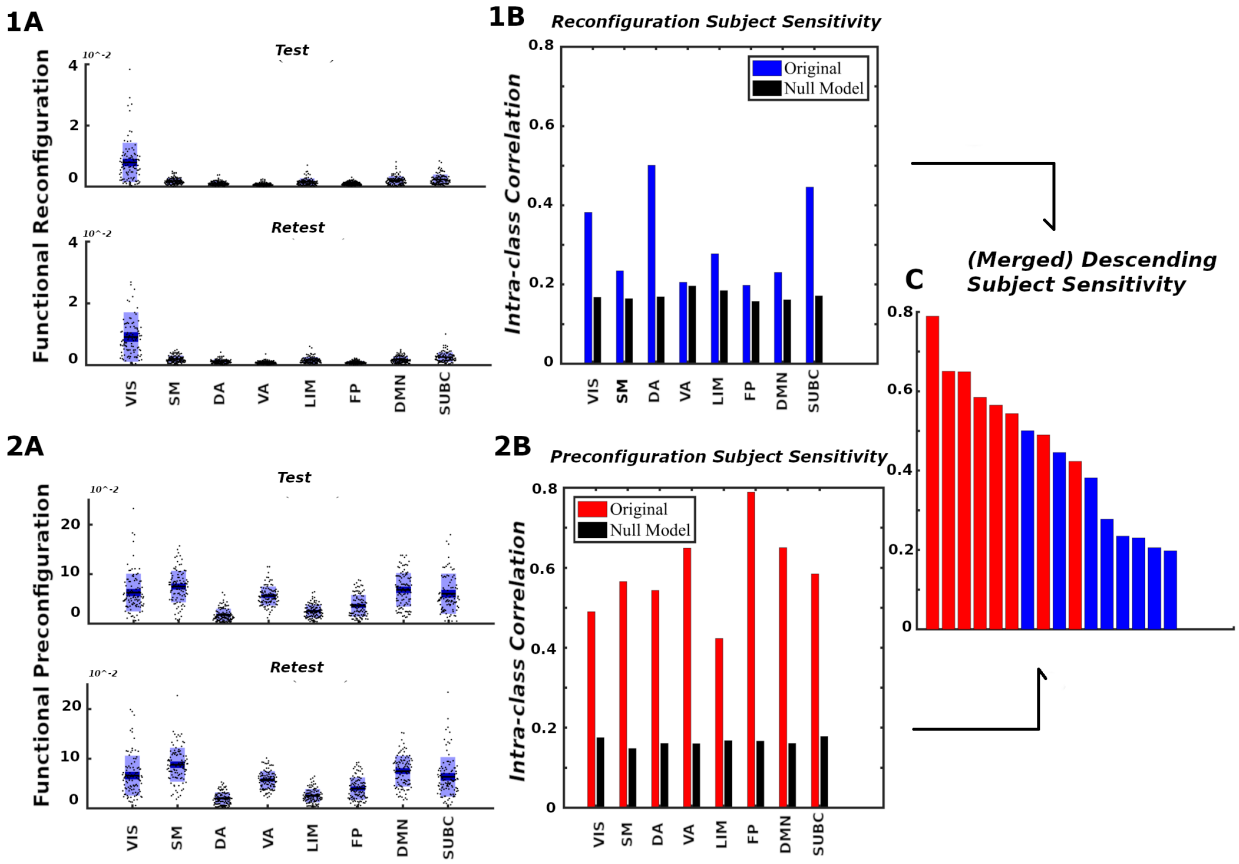


Figure 3.16. Network Configural breadth - subject specificity analysis.

configurations are higher than their corresponding null models. Finally, for a fixed FN, functional preconfigurations dominated the subject sensitivity ranking, as illustrated by panel C of figure 3.16. Furthermore, FP, DMN and VA preconfigurations are among the FNs with highest subject fingerprints in overall subject sensitivity ranking.

3.9.3 Network configural breadth and behavior

Network configural breadth, compartmentalized into FN reconfiguration $\mathcal{R}^{FN}$ and preconfiguration $\mathcal{P}^{FN}$, shows high level of subject sensitivity. This allows us to assume that $\mathcal{F}_{\mathrm{i}}$ is associated with an individual's behavioral measures (denoted as m_{i} for subject i^{th}). Several studies reported that FP and DMN networks are associated with memory and intelligence [15], [114], [115]. Therefore, we evaluated if the outlined framework reflects four

widely studied cognitive/behavioral measures, related to memory and intelligence: episodic memory, verbal episodic memory (Verb. Epi. Mem.), fluid intelligence gF, and general intelligence g. While fluid intelligence reflects subject capacity to solve novel problems, general intelligence, g, reflects not only fluid intelligence, gF, traits but also crystallized (i.e. acquired) knowledge ([116] and typically denoted as gC). The early notion of general intelligence is conceptualized by Spearman's positive manifold [117] that cannot be fully described using a single task. Quantification of g can be accomplished using subspace extraction techniques such as explanatory factor analysis ([118]) or principal component analysis (PCA [15]). In this work, we quantified g using the PCA approach described in Schultz and Cole [15]. Mathematically, we propose the following composite relationship:

$$m_{\mathrm{i}} = \Upsilon(\mathcal{R}_{\mathrm{i}}^{FN}, \mathcal{P}_{\mathrm{i}}^{FN}) \tag{3.11}$$

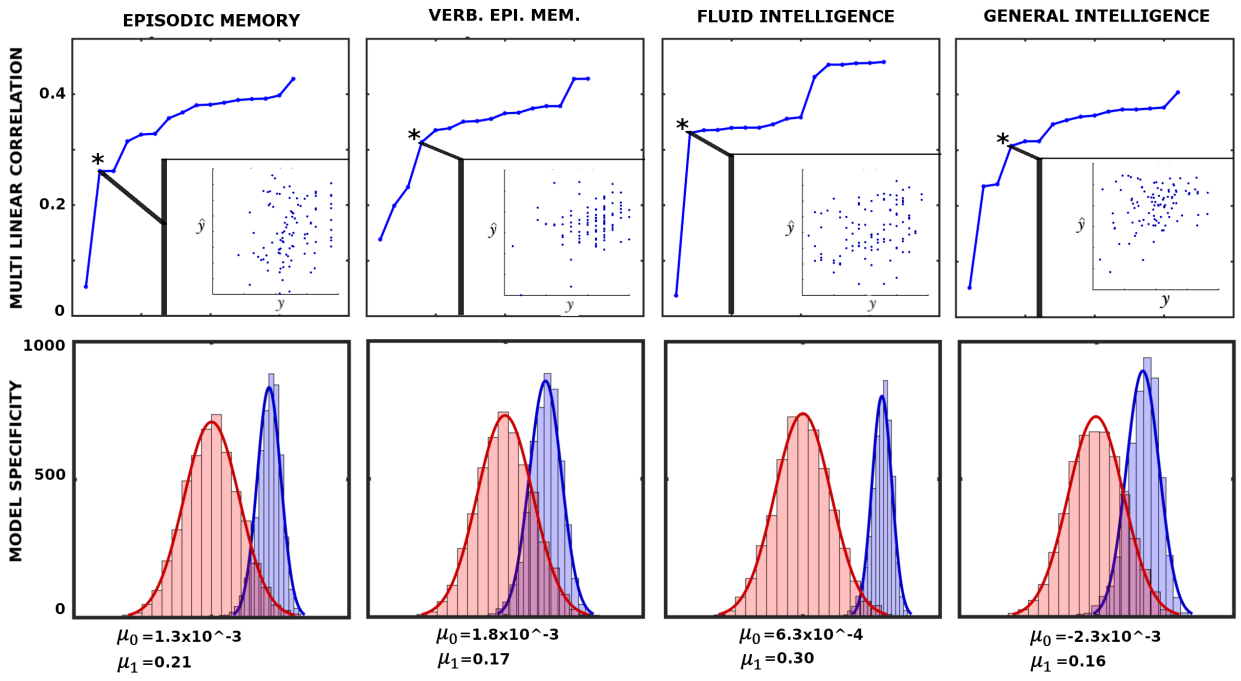


Figure 3.17. Associations between network configural breadth and behavior.

Having established a plausible connection between behavioral measures and $\mathcal{P}^{FN}$, $\mathcal{R}^{FN}$, the above equation can be viewed as a multi-linear model (MLM) using FN preconfigura-

tion and reconfiguration as independent variables (or predictors). The MLM is constructed iteratively, starting with the descriptor with the highest individual fingerprints in panel C of figure 3.16. In each iteration, the subsequently ranked descriptor (according to figure 3.16 - panel C) is appended to the existing ones. The best MLM (denoted with an asterisk in figure 3.17), which determines the number of linear descriptors included the model, is selected based on the model p-value.

The x-axis represents functional network preconfiguration and reconfiguration terms, i.e. $\mathcal{P}_{\mathrm{i}}^{FN}$ and $\mathcal{R}_{\mathrm{i}}^{FN}$, ordered in decreasing subject fingerprints (as shown in previous figure). The *top panels* illustrate iterative multilinear regression model (MLM) while *bottom panels* show model specificity (MS) for corresponding behavioral measures. Asterisk represents the optimal MLM with lowest p-value. The top panels in the above figure show that as more linear descriptors (FN's functional pre- and re-configurations) are added to iterative MLMs, variance associating with behavioral/cognitive performance measures decreases with linear descriptors that bear less subject sensitivity. This result highlights the importance of appending linear predictors in descending order with respect to the subject sensitivity. Specifically, as individual specificity reduces from left to right (see figure 3.16 panel (C) for further details), the differential correlations, i.e. the difference between two consecutive correlation values, decreases.

To test the level of specificity in the model, we performed 2000 simulations of $k-$ fold cross validation where $k = 5$ between the selected MLM and the corresponding behavioral measure. Specifically, for each cross validation (per simulation), we obtain a correlation between the 20 left-out values (y) with the predicted values ($\hat{y}$). Hence, in each simulation we obtained five correlation distribution (one of each cognitive measures) and their corresponding mean values. It can be shown that those means follows a normal distribution. Lastly, to provide the level of specificity of linear descriptors, we present a corresponding null model where the same descriptors are evaluated to predict random vectors of appropriate size. To test our model and its ability to predict the behavioral measures, we rely completely on network configural breadth predictors ranked in descending order of subject specificity.

Table 3.1. MLM's linear constant terms for all statistically significant terms and overall model.

MLM terms, coefficients	Constant	$\mathcal{P}^{FP}$	$\mathcal{P}^{DMN}$	$\mathcal{P}^{VA}$	$\mathcal{P}^{SUBC}$
	β_0	β_1	β_2	β_3	β_4
Episodic Memory	0.6	2.9	-9.3		
Verbal Episodic Memory	0.5	11.8	-1.1	-8.8	-6.1
gF	0.7	5.1	-12		
g	0.8	3.9	-5.5	-3.6	-5.7

In the above table, multi-linear regression models with corresponding standardized β coefficients. Dependent variables for each model are: episodic memory, verbal episodic memory, fluid intelligence (gF) and general intelligence (g).

Table 3.2. MLM's p-value results for all statistically significant terms and overall model.

MLM terms, p-values	Constant	$\mathcal{P}^{FP}$	$\mathcal{P}^{DMN}$	$\mathcal{P}^{VA}$	$\mathcal{P}^{SUBC}$	Entire
	p_0	p_1	p_2	p_3	p_4	**Model**
Episodic Memory	0	0.57	0.01			0.03
Verbal Episodic Memory	0	0.02	0.77	0.17	0.03	0.04
gF	0	0.3	9×10^{-4}			0.004
g	0.03	0.44	0.16	0.57	0.05	0.05

In the above table, multi-linear models with corresponding p-values. Note that we do not use step-wise linear model which discards descriptors that are not statistically significant. Column entire model shows the significance of the entire model.

3.10 Discussion

In this work, we fill an existing gap in the field of network neuroscience by proposing a mathematical framework that captures the extent to which subject-level functional networks, as estimated by fMRI, reconfigure across diverse mental/emotional states. This proposed framework can also be potentially useful in application to a clinical population (AD; drug abuse; ...). We first propose that brain networks can undergo three different types of (re-)configurations: **i)** Network Configural Breadth, **ii)** Task-to-Task transitional reconfiguration, and **iii)** Within-Task reconfiguration. Unlike other existing frameworks [15], [17], [19], [21], the framework presented here can be applied to all three reconfiguration types. As a first

step, we focus on assessing the broadest aspect of reconfiguration, i.e. Network Configural Breadth. We postulate, based on previous literature [77], that macro-scale (whole-brain) and micro-scale (edge-level) reconfigurations of brain networks are subtle, and hence difficult to disentangle. At the same time, mesoscopic structures in the brain (e.g. functional networks (FNs)) reconfigure substantially across different mental/emotional states as elicited by different tasks [119]. The framework presented here constitutes the first attempt to formalize such (re)configurations of mesoscopic structures of the brain, and quantify the behavior of a reference set of FNs with changing mental states. We set forth a mathematically well-defined and well-behaved 2D network morphospace using novel mesoscopic metrics of Trapping Efficiency (**TE**) and Exit Entropy (**EE**). This morphospace not only characterizes the topology of FNs but also the flow of information within and between FNs. We show that this morphospace is sensitive to FNs, tasks, subjects, and the levels of cognitive performance. We show that both of these measures are highly subject-sensitive for some FNs, while preconfiguration is highly subject-sensitive for all of them. Lastly, we also formalize and quantify the concepts of functional reconfiguration (the extent to which an FN has the capacity to reconfigure across different tasks) and functional preconfiguration (amount of transition from resting-state to a task-positive centroid). We thus construct a formalism that can explore FN changes across different cognitive states in a comprehensive manner and at different levels of granularity.

Ideally, a morphospace framework [64], [65], [68], [69], [72]–[76] would have a minimal complexity and, in this particular case, capture distinct features of functional network changes. As discussed in [64], metrics parametrizing a given morphospace should be disjoint. We see that, for any specific FN, high within-subject task sensitivity of **TE** does not necessarily imply a high value in **EE** and vice versa (e.g. VA and FP in 3.13). In addition, we see that both **TE** and **EE** offer their unique insights in capturing non-overlapping features with **TE** is more subject-sensitive and **EE** more task-sensitive at the cohort level. Our result also highlights the disjoint nature of the two metrics as well, where we compute maximal distance per FN polytope in the TE and the EE axes separately. Results show that corresponding **TE** and **EE** maximal distances are disjoint and FN dependent. In other words, for a specific FN, the polytope is "stretched" in a particular task direction, where each morphospace measurement (**TE** or **EE**) unravels distinct properties. In 3.15, we further see

that a subset of tasks dominantly contribute to the maximal distance computation, such as Motion, Language, and Social tasks. Interestingly, we see that Motion and Language tasks can be considered "orthogonal" tasks with respect to **TE** and **EE**.

Interestingly, the limbic network possesses the lowest ability to distinguish between tasks. This might be because tasks used are far from optimal for assessing the functional fingerprint property of LIM in configuration scenarios. Further, LIM may act as a relay station, transferring information across the cortex. Another possibility is that LIM might consider these tasks irrelevant. Similar behavior has been observed in [120] when using Jensen-Shannon divergence as a distance metric of functional connectivity. In addition, the limbic network seems to work as a "relay" in brain communication [104]. One potential explanation for this unique behavior is that the limbic network maintains a minimal cognitive load across various tasks, most of which comprises relaying information from one part of the brain to the others; it thus does not reconfigure as much across different mental states.

Brain network configuration is typically studied considering a specific task at multiple spatial and temporal scales, see [16]–[19], [21], [82], [106], [119]. Previous investigations have mainly focused on the mechanism of how the brain traverses between high/low cognitive demands [19], [21], [73], [98], [120], [121], or on periods of integration and segregation at rest [17]–[19], [21], defined in this paper as within-task reconfigurations. On the other hand, whole-brain configurations have also been investigated across different tasks (one configuration per task) with respect to rest, which led to the concept of general efficiency [15]. This approach would belong to a wider category that we formally generalize as the Network Configural Breadth. The idea of general efficiency in [15] relied on whole-brain FC correlations between task(s) and rest. While intuitive in quantifying similarity/distance between a single task and rest, quantification across multiple tasks becomes a challenge. Specifically, note that in [15], general efficiency is quantified using the first eigenmode, which explains most of the variance, after measuring the correlation between resting FC and three distinct task FCs. As more and more tasks are included, using the first eigenmode would become less and less representative of the task-related variations present in the data (in this paper summarized as the Network Configural Breadth). The proposed network morphospace overcomes these limitations and can be used to study brain network (re-)configurations across any number

of tasks. It allows us to study different types of brain network (re-)configurations, as mentioned in above, using one comprehensive mathematical framework, which also facilitates a meaningful comparison between these seemingly disparate kinds of (re-)configurations. Schultz and Cole proposed that configurations can be compartmentalized into two differentiated concepts: functional reconfiguration and preconfiguration [15]. Note that although the term **reconfiguration** is also used in [15], it is not referring to the action of switching among multiple mental/emotional states, i.e. as represented by task-to-task transitional reconfiguration or within-task reconfiguration. Rather, it refers to the overall competence in exploring the total repertoire of task space of each subject given its resting configuration. That is why when we translate the corresponding idea into the mesoscopic morphospace, we call it the network configural breadth. We have also incorporated the two concepts of functional pre- and re-configurations into a well-defined mathematical space, which solves some of the technical difficulties and generalizes these concepts to mesoscopic structures.

Brain network within-task reconfigurations have been almost exclusively qualitatively assessed. For instance, [16] show that the whole-brain functional connectome traverses segregated and integrated states as it reconfigures while performing a task. They also found that integrated states are associated with faster, more effective performance. Our formalism of within-task reconfigurations permits assessing such reconfigurations in a quantitative manner. Potentially, such within-task reconfigurations could also be used to assess cognitive fatigue, effort or learning across time.

[77] have shown that the resting architecture network modifies itself to fit task requirements through subtle changes in functional edges. Numerically, small changes constituted by functional edges between rest and task-based connectivity might not be statistically significant when looking at edge level. Moreover, we also observe that while such changes might be negligible on a whole-brain global scale, they are more evident when looking at subsystems or functional brain networks, as clearly observed in the VIS network, relative to others. For functional preconfiguration, this effect is observable in all the FNs. In essence, we are postulating that a mesoscopic explorations of changes in brain network configurations with changing mental states is more informative than a macroscopic or microscopic exploration.

A key feature of this morphospace is that, in order to, to study brain network (re-)configuration, an FN is not removed from the overall network for exploration. On the contrary, both metrics that define the morphospace, namely **TE** and **EE**, account for a particular FN's place embedded within the overall functional brain network, both in terms of topological structure and flow of information. That is why it is important to begin with a reference set of FNs (e.g., RSNs), so as to study how these FNs adapt to changing mental states within the context of the overall network.

Another benefit of a mesoscopic framework is that we can compare individual cognitive traits in each FN, instead of the whole brain. Specifically, after quantifying reconfiguration and preconfiguration for all FNs, we determine if these quantities incorporate information about individual traits. We observe different levels of subject fingerprint in different FNs for both re- and pre-configuration measures. This subject fingerprint heterogeneity across different FNs is consistent with previous literature on functional connectome fingerprinting, [2], [5]. Interestingly, functional pre-configuration (amount of transition from a resting-state to a task-positive state) displayed greater subject fingerprint than functional reconfiguration for all FNs. Based on this observation, we argue that to have better subject differentiability, we need to design tasks where the subject transitions from a stable resting-state to a task-positive state and/or vice versa [105]. This could be a significant step forward in precision psychiatry [122], where we can identify regional brain dysfunction more precisely as a function of the type and degree of cognitive or emotional load.

Subject-sensitivity of the proposed network morphospace framework is also supported by significant associations of the frontoparietal and default mode networks with fluid intelligence. Specifically, as pointed out by [115], high fluid intelligence is associated with a greater frontoparietal network activation, which is also consistent with findings from a three-back working memory task ([114]). In the domain of network configural breadth, we observe a higher reconfiguration as represented by a positive frontoparietal functional preconfiguration coefficient (see table 3.1).

This work has several limitations. The framework was tested specifically on the Human Connectome Project dataset and using a single whole-brain parcellation. Alternative parcellations [26], [123], additional fMRI tasks to better sample the cognitive space, and

other datasets might offer further insights about the mesoscopic network morphospace, see [64], [69]. In addition, we did not perform a sensitivity analysis on how small fluctuations in functional connectomes affect mapping into the network morphospace. Due to the nature of module trapping efficiency and exit entropy metrics, negative functional couplings were not considered and hence, were set to zero. In future work, other combinations of L1 and L2 norms, or even other norm choices, should be evaluated when defining trapping efficiency. This would impact not only the magnitude of the morphospace measure but also the differentiating capacity of configuration across different functional networks.

Future studies should incorporate a sensitivity study of the behavior of this network morphospace with respect to small fluctuations in the input functional connectomes. Further studies could also incorporate structural connectivity information to inform both **TE** and **EE** measures when assessing the morphospace coordinates of functional reconfiguration. Additional exploration of different aspects of this morphospace could provide further insights. For example, location of the polytopes in the morphospace might improve individual fingerprint. An important aspect of the proposed mesoscopic network morphospace is that it allows for an exhaustive and continuous exploration of network reconfigurations, including those that are continuous in time [21], [99]. For example, if the subject performs several tasks within the same scanning session, including extended resting-state periods (such as the fMRI experiment done at [124]. This would allow us to fully explore the cognitive space and gain a valuable insight into how different subjects adapt to different levels of cognitive demands. One can also study the trajectory of changing mental states using dynamic functional connectivity [78], which can easily be mapped to this morphospace for additional insights. Another potential avenue could be the application of this framework to characterize and understand different brain disorders.

In summary, this mesoscopic network morphospace is our first attempt to create a mathematically well-defined framework to explore an individual's cognitive space at different levels of granularity. It allows us to characterize the structure and dynamics of specific subsystems in the brain. This type of framework can be extremely helpful in characterizing brain dynamics at individual-level, in healthy and clinical population, which in turn would pave the way for the development of personalized medicine for brain disorders.

4. A PRINCIPLED METHOD TO THRESHOLD THE WHOLE-BRAIN FUNCTIONAL CONNECTIVITY BASED ON STOCHASTIC BLOCK MODELS

4.1 Backgrounds and motivations

Graph-theory has played an increasingly important role in analyzing large-scale brain networks (e.g. functional connectomes). Nonetheless, to properly addressing any research questions in brain connectomic area, it is absolutely vital to obtain the most possibly well-defined representation of functional connectivity. Evidently, having proper brain network representations plays a critical role in multi modalities of neuroimaging research such as fMRI [125]–[127], MEG [128] and EEG-based [129] connectomes. In functional connectivity domain, the connectome is constructed by computing a statistical dependency measure (e.g. Pearson correlation coefficient) for all given pairs of brain's regions of interest (ROIs) using the aggregated voxel level blood-oxygen-level-dependent (BOLD) signals. In fact, functional connectome construction, induced from BOLD signals with activation delays due to inhibitory-excitatory nature and negative-valued correlations among ROIs, could cause potentially severe impacts in estimating population-level functional connectome [125], functional brain network topological features (e.g nodes' centrality [130], global network measures [126]) and geometry-topology relation [131], to name a few. Recent efforts have focused on improving functional connectome construction, taking into account neural signal activation delays [126] and negative correlations [127].

In the domain of network neuroscience, thresholding (or more generally, the process of eliminating statistically spurious functional edges) in large-scale functional brain networks, has become such an important step to properly address all subsequent research inquiries with tremendous applications in not only healthy control studies, but also schizophrenia [132] and unipolar depression & bipolar disorder [133]. Thresholding, if not done correctly, could cause negative impacts in subsequent analyses such as parametric statistical testing [134] and random network characterization [135]. Often, thresholding technique is deployed to preserve certain desired properties of the original weighted network such as proportional

thresholding (to maintain the absolute number of edges across different subjects and tasks by Van den Heuvel et al. [135]), modular similarity [133], percolation (to preserve original weighted graph's topological features [136]). The process of eliminating spurious edges is typically accomplished using a wide spectrum of methods such as wavelet-based methods [132], mixture modeling [137], topological data analysis through persistent homology [138], [139], branch-and-bound based algorithm (to study cognitive activity [129]), and orthogonal minimal spanning trees for dynamical functional brain networks [140]. Alternatives to thresholding treatment to functional connectomes is also proposed using hierarchical Bayesian mixture model [141].

Human brain functional sub-circuits (e.g. [1]) and their modularity characteristics [121], [142]–[144] are the fundamental building blocks to understand brain complexity [60], differential configural properties [144], modular structures [85], [142], information processing [104], [120], among many others. In fact, human brain modular organizations are of monumental importance to study neuro pathological applications such as aging [145], [146] and schizophrenia [147], among many others. As noted by several authors, executive subsystems in the brain are consistently reproducible across many individuals at rest (such as [1], [24]). The usage of brain functional sub-circuits, especially at rest (i.e. resting state networks (RSNs) or, equivalently, functional networks (FNs)) is very common in control studies [148] and pathological conditions [149] and predicting individual differences [150]. Nonetheless, to-date, there have been no studies assessing a particular set of a priori FNs with respect to some functional connectivity representation, i.e. functional connectomes (FCs). In other words, common practice in FC processing involves some initial representation of functional connectivity that is thresholded, based on some arbitrary rules or research hypotheses as mentioned above. After such step, subsequent analyses pre-dominantly involve the mapping of a priori fixed set of FNs onto constructed FCs, across different subjects and fMRI task conditions, without examining whether those mappings is information-theoretically relevant to the constructed FCs.

Besides many decisions that need to be made along the processing of neuroimaging data to obtain the eventual whole-brain estimates of functional connectivity, e.g. functional connectomes, the choice of brain parcellations, i.e. how to define the notion of node in functional

brain networks, is undoubtedly one of the most critical steps [8], [26], [62]. This is because it defines the topology of the input networks that feed into subsequent analyses. Recent research has shown that different levels of parcellation granularity can impact the discovery of subject fingerprints [2], [3]. In an effort to register the unprocessed neuroimaging data into a sequence of increasing granularity, Schaefer and colleagues has recently published a scheme of atlases that increase in network sizes. Additionally, Schaefer parcellations with increasing granularity levels subdivide the highly reproducible set of RSNs proposed by Yeo et al. [1]. In the light of such developments, brain connectivity community can now investigate many interesting characteristics of sequential functional brain networks, coupled with corresponding a priori set of FNs.

In this work, our aims are two-fold: i) to formalize and subsequently quantify the level of information prominence of a given fixed set of FNs across different subjects, and tasks; ii) use the level of prominence as guidance to eliminate spurious functional edges in whole-brain FCs. To do so, we utilize Schaefer parcellations [26] with nine distinct granularity levels from 100 to 900 nodes, in an increment of 100 nodes. We first present some theoretical relevance of stochastic block models in exploring our quest in Section 2 of this chapter. We then proceed to propose the reconstruction pipeline in Section 3. We wrap up with Results (Section 4) and Discussion (Section 5). Our framework can be generalized to any given pair of FN partition and parcellations (e.g. [20], [123]).

Recoverability of ground-truth partition depends on the degree regime of the network (indicated by the degree scaling factor s_n). For instance, weak recovery only requires the necessary condition for limiting graph ($n \to \infty$) to be in constant degree regime, i.e. $O(\frac{1}{n})$. On the other hand, exact recovery requires the necessary condition (for limiting graph) to be asymptotically connected, i.e. in the degree regime of logarithmic $O(\frac{log(n)}{n})$. The sufficient condition for all recovery criteria is stated in the respective theorems with different proposed measures with sharp phase transitions, see [22].If a measure (say for weak or exact recovery) is below a certain algebraic threshold (stated in the respective theorems), recovery is not possible although necessary condition is satisfied.

Here, we chose weak-recovery as a guidance for whole-brain functional connectivity estimation because of four reasons:

1. Although Schaefer parcellations with increasing number of nodes allows us to empirically project some insights onto its degree regime, a rigorous theoretical argument on degree regime is not possible for any empirical graph sequence. Hence, exact recovery of a priori unique ground-truth partition is non-relevant in the case of brain functional connectomes;

2. Even in the empirical domain, we observe that both group-average and individual FCs disconnected (e.g. the number of connected components is more than one) after a relatively small threshold value in the interval $\tau \in [0.2, 0.3]$ (further details is available in figure 4.3). Theoretically, graph sequence is required to be connected, asymptotically, in order to consider exact recovery. On the other hand, weak-recovery (detection of mesoscopic structures) offers a more realistic.relaxed criteria in this particular application to estimate a whole-brain FC that is most suitable for a priori set of FNs without evaluating the number of connected components of the thresholded FC.

3. Most (if not all) mesoscopic studies of brain functional sub-circuits such as [151], [152] are based on pre-defined hypotheses (e.g. the brain functional sub-circuits involve a more diverse classes of community than just assortative ones [151]). Such assumption leads to the appropriate usage of different community detection algorithms such as Weighted Stochastic Block Models (WSBM) in the case of [151], [152]. As mentioned above, weak-recovery is equivalent to detection in theoretical SBM literature;

4. No set of functional sub-circuits is universally agreed and uniquely identified as ground-truth communities. Hence, all proposed brain functional sub-circuit parcellations (e.g. [1]) are relative.

4.2 Data & Atlases

Neuroimaging Data Acquisitions

The fMRI dataset used in this paper is available in the Human Connectome Project (HCP) depository (http://www.humanconnectome.org/), with Released Q3. The processed functional connectomes obtained from these data and used for the current study are avail-

able from the corresponding author on reasonable request. Please refer to below detailed descriptions on the dataset and data processing.

We first describe the acquisitions of unprocessed neuroimaging data from 409 Unrelated Participants chosen from the list of 1200 participants by Essen et al. [47], [48] in the Human Connectome Project (HCP) release. This subset of participants ensures that no two participants have any family relations, i.e. sharing a parents or being siblings. This selection is particularly critical to avoid any confounding effects in our subsequent analyses, such as group average analysis, due to family structures.

Per HCP protocol, all subjects gave written informed consent to the HCP consortium. The two resting-state functional MRI acquisitions (HCP filenames: $rfMRI\_REST_1$ and $rfMRI\_REST_2$) were acquired in separate sessions on two different days, with two distinct scanning patterns (left to right and right to left) in each day, [20], [47], and [48] for details. This release includes also data from seven different fMRI tasks: gambling ($tfMRI\_GAMBLING$), relational or reasoning ($tfMRI\_RELATIONAL$), social ($tfMRI\_SOCIAL$), working memory ($tfMRI\_WM$), motor ($tfMRI\_MOTOR$), language ($tfMRI\_LANGUAGE$, including both a story-listening and arithmetic task), and emotion ($tfMRI\_EMOTION$). Per [20], [101], three tasks MRIs are obtained: working memory, motor, and gambling. The local Institutional Review Board at Washington University in St. Louis approve all the protocol used during the data acquisition process. Please refer to [20], [100], [101] for further details on the HCP dataset.

Constructing functional connectomes

We used the standard HCP functional pre-processing pipeline, which includes artifact removal, motion correction and registration to standard space, as described in [20], [100] for this dataset. For the resting-state fMRI data, we also added the following steps: global gray matter signal regression; a bandpass first-order Butterworth filter in both directions; z-scores of voxel time courses with outlier eliminations beyond the three standard deviations from first moment [102], [103].

For task fMRI data, aforementioned steps are applied, with a relaxation for bandpass filter [0.001 Hz, 0.25 Hz]. Starting from each pairs of nodal time courses, Pearson correlation is used to fill out the functional connectomes for all subjects at rest and seven designated tasks. This would yield symmetrical connectivity matrix for all fMRI sessions.

Brain Atlases

The brain atlases used in this work is sequential, in the sense that its granularity increases, ranging from 100 nodes to 900 nodes (increment of 100 node each time), registered on the cortical surface of the brain. This sequential atlases are made possible thanks to the work of Schaefer and colleagues [26]. Similarly to reference [5], [84], 14 sub-cortical regions were added, as provided by the HCP release (filename $Atlas\_ROI2.nii.gz$). We accomplish this by converting this file from NIFTI to CIFTI format by using the HCP workbench software [(http://www.humanconnectome.org/software/connectomeworkbench.html, with the command -cifti- create-label]. The resultant sizes of ROI-based connectome are, hence, the number of brain ROIs are $100, 200, ..., 900$ nodes for rest and any given fMRI tasks. Mathematically, we denote the Schaefer parcellations as a sequence of graphs G_{n_t} where $t \in [9]$ and $n_t = [100, 200, ..., 900]$.

Moreover, Schaefer parcellation are also coupled nicely with further subdivisions of Yeo's functional networks [1]. For a fixed Schaefer granularity (indexed n_t), we denote the corresponding Yeo's RSNs to be σ_{n_t}. For a fixed granularity, the Yeo's FN partition applied to a Schaefer parcellation has a one-to-one relationship. Specifically, one brain region in a Schaefer-parcellated functional connectome belongs only to one Yeo's FN.

4.3 SBM Inference and extended usage

To infer SBM parameters, the basis is to reverse engineer using maximum likelihood principle. Specifically, since both G and σ (subsequently, the number of communities in an *a priori* set of FNs is denoted as $k = \max_{u \in [n]} \sigma_u$) are priors, in expectation, we can infer $SBM(P, W)$ using Bayesian approach as follows:

1. $P = \frac{\Omega}{n} = [p_{\mathrm{i}}] = \left[\frac{|\Omega_{\mathrm{i}}|}{n}\right]$

2. Infer $W_{bin} = \frac{C^{bin}}{C^{max}}$

3. Compute $W_{\text{wei}} = \frac{C^{\text{wei}}}{C^{max}}$

4. $Q = nW$ as $s_n = 1$ for weak-recovery

5. Compute PQ (Matrix Multiplication)

where C_{bin} is a simple edge count of M_τ between/within blocks of communities whereas C^{wei} is the sum of weighted edges of FC_τ (also between/within communities). Specifically,

$$C_{bin} = \sum_{u,v \in [n]} \mathbf{1}_{\sigma_u = \sigma_v}$$

$$C_{\text{wei}} = \sum_{u,v \in [n]} |\mathbf{a}_{uv}|, \sigma_u = \sigma_v$$

and

$$C_{max} = \Omega\Omega^T$$

The inference of matrix P are based on the law of large numbers [22]. The inference of W_{bin} is a entry-wise divisions between matrix C_{bin} and C_{max} which infers the Bernoulli random variable parameter p, representing the number of successes running independent Bernoulli trials of edge existence among all pairs of stochastically equivalent nodes in or between communities. In the case of C_{wei}, note that we use the term computing instead of inference because we have extended the usage of SNR to be mesoscopic prominence measure. We use the absolute values $|\mathbf{a}_{uv}|$ to only consider the overall magnitude (and not the sign) of functional couplings within/between FNs.

Technically, this inference is a less challenging compared to traditional inference problems where σ is also a latent variable in the model and graph ensemble G is the only observable ensemble available. Specifically,

$$(G, n, \sigma, k) \sim SBM(P, W)$$

where G and σ are priors.

4.4 Whole-brain functional connectivity estimation pipeline

procedure RECONFC(Γ, n, k, σ, τ, P, W FC,M)
  for $\forall$ Schaefer Granularity level and fMRI tasks **do**
    $\rightarrow$ Vetting Step: Step 1 and 2
    Use all individual FC (FC_γ) for a given Schaefer parcellation and fMRI task, compute group-average FC

$$FC^{avg} = \frac{\sum_{\gamma=1}^{\Gamma} FC_\gamma}{\Gamma}$$

    for $\tau \in [0,1]$ ***do***
      Compute *Masked, thresholded group-average FC* $M_t = \begin{cases} 1, & |FC^{avg}| \geq \tau \\ 0 & |FC^{avg}| < \tau \end{cases}$
      Infer SBM parameters and apply Theorem 1 to compute $SNR[M_\tau]$
    end for
    Determine the weak-recoverability sub-interval (a_w, b_w) by

$$a_w = argInf_{\tau \in [0,1]}(SNR[M_\tau] > 1)$$

$$b_w = argSup_{\tau \in [0,1]}(SNR[M_\tau] < 1)$$

    $\rightarrow$ Compute Individual FC *mesoscopic prominence measure*: Step 3 and 4
    for all $\gamma \in [\Gamma]$ **do**
      for $\tau \in [0,1]$ **do**
        Compute individual thresholded weighted FC: $FC_{\gamma,\tau} = \begin{cases} |FC_\gamma|, & |FC_\gamma| \geq \tau \\ 0 & |FC_\gamma| < \tau \end{cases}$
        Compute *mesoscopic prominence measure* $SNR[FC_{\gamma,\tau}] = \frac{\lambda_2^2}{\lambda_1}(FC_{\gamma,\tau})$
      end for
    end for
      $\rightarrow$ Step 5
      Obtain $\tau_{opt} = argmax(SNR[FC_\tau], \tau)$
      Check if $\tau_{opt} \in (a_w, b_w)$
    end for
  end procedure

Figure 4.1. Pseudo-code for *reconFC* routine using the number of individual FCs Γ, Schaefer granularity n, number of functional networks k, a priori partition σ, threshold range τ, community assignment likelihood P and connectivity pattern matrix W.

The reconFC pipeline, see figure 4.1, describes the process to compute the optimal threshold for a given fMRI condition, Schaefer granularity, cohort for two particular cases:

- individually driven threshold τ_{opt}^{i};
- constant (cohort-driven) threshold τ_{opt}^{GA}.

The pipeline contains five distinct steps:

Step 1: For each Schaefer granularity level and task, compute the binarized (masked) group-average FC (denoted as M_τ) using the entry-average of individual FCs (the number of individual FC is denoted as n_{FC})

Step 2 (Vetting Step):

For each threshold value $\tau \in [0, 1]$, infer the Stochastic Block Model (SBM) parameters for the compute Signal-to-noise ratio (SNR) of M_τ:

$$SNR[M_\tau^{GA}] = \frac{\lambda_2^2}{\lambda_1} \left\{ [PQ]_{bin}^{GA} \right\}$$

Repeat this computation for all threshold values, apply Theorem 1 to determine the weak-recoverability sub-interval $(a_w, b_w)[0, 1]$ for the group-average FC, i.e. M_τ

Step 3: For a given individual FC and threshold value τ, compute the associated thresholded FC (FC_τ^{i}) ; compute the Stochastic Block Model (SBM) parameters for FC_τ^{i}. Extend the usage of SNR as a mesoscopic prominence measure:

$$SNR[FC_\tau^{\mathrm{i}}] = \frac{\lambda_2^2}{\lambda_1} \left\{ [PQ]_{w\mathrm{ei}}^{\mathrm{i}} \right\}$$

Analogously, we can also compute SNR, using group-average FC (FC^{GA}) as follows:

$$SNR[FC_\tau^{GA}] = \frac{\lambda_2^2}{\lambda_1} \left\{ [PQ]_{w\mathrm{ei}}^{GA} \right\}$$

Step 4: Repeat steps 3 for all threshold values $\tau \in [0, 1]$ for all individual FCs for a given fixed Schaefer parcellation and fMRI task pair; Step 5:

1. Obtain the threshold value that maximizes SNR of the thresholded FC and the corresponding optimally reconstructed whole-brain FC;

$$\tau_{opt}^{\mathrm{i}} = argmax(SNR[FC_\tau], \tau)$$

Note that if group-average FC (FC^{GA}) is use in Step 3 then:

$$\tau_{opt}^{GA} = argmax(SNR[FC_{\tau}^{GA}], \tau)$$

2. Check if τ_{opt} is in the weak-recoverability sub-interval (Step2)

$$\tau_{opt} \in (a_w, b_w)$$

Note that one needs to check the optimal threshold against the weak-recovery sub-interval, whether it is individualized (τ_{opt}^{i}) or group-average threshold (τ_{opt}^{GA}).

In figure 4.2, a demonstrated computation of SBM inference and extended usage is shown. Note that the "for" loop indicated by $(*)$ is used to find individualized optimal threshold for each subject, τ_{opt}^{i}. One can substitute this "for" loop by finding one unique cohort optimal threshold, τ_{opt}^{GA} using group-average FC directly, in which case, we do not need such "for" loop.

4.5 Granularity Analysis of the Schaefer parcellations in Resting State

4.5.1 Number of connected components

In this section, we investigate the topological features of Schaefer FC graph sequence across the entire threshold period $\tau \in [0, 1]$. Specifically, we look into the number of components across threshold range and Schaefer granularity levels.

We use all nine available Schaefer parcellations with $n = [100, 200, ..., 900]$ and their corresponding mappings of seven Yeo's RSNs [1] for each granularity level. Besides the individual level FC, group-average FC (denoted as FC_{group}) is computed using entry-wise mean across the individual FCs (denoted as FC):

$$FC_{group} = \frac{\sum_{\gamma=1}^{\gamma=\Gamma} FC_{\gamma}}{\Gamma}$$

where Γ denotes the number of participants and $\gamma \in [\Gamma]$.

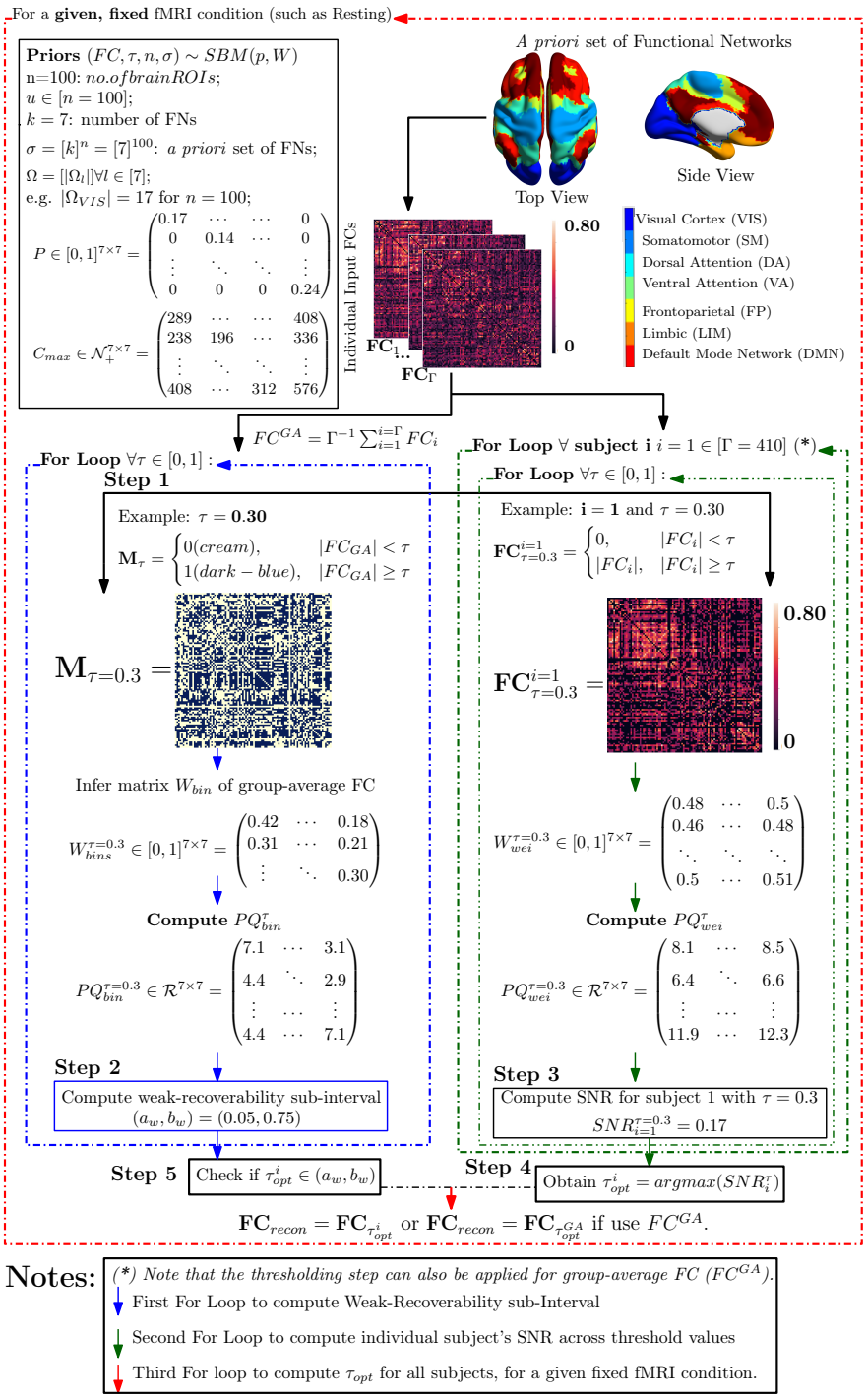


Figure 4.2. Example of FC reconstruction routine based on Schaefer granularity level of 100 nodes and Resting state fMRI with scanning pattern *LR*.

Specifically, for each Schaefer granularity and threshold combination, we compute the number of components for each individual and group-average FC.

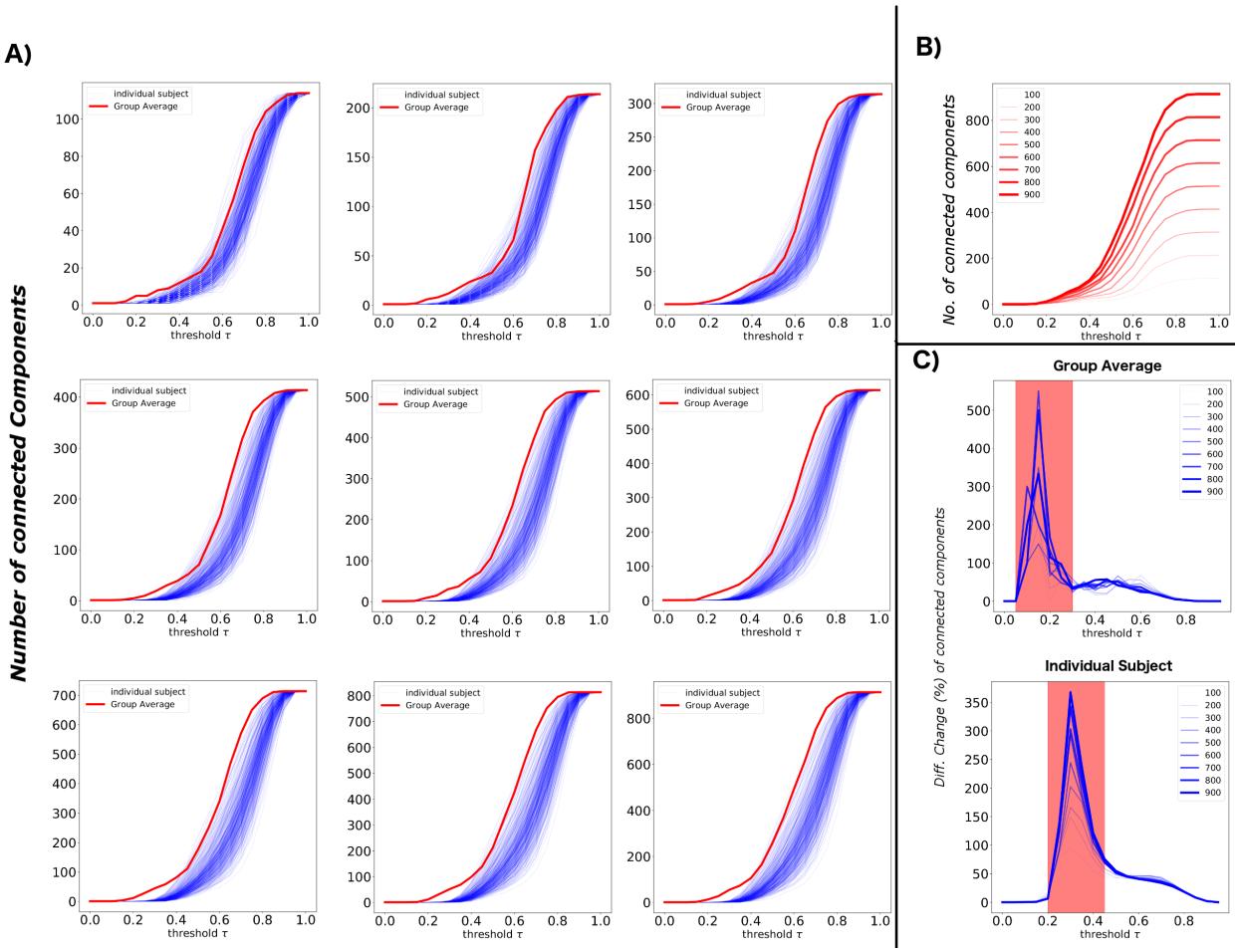


Figure 4.3. Panel (A) represents the number of connected components, for each Schaefer parcellation (from 100 to 900 nodes with an increment of 100 nodes each time), across the pre-defined thresholding range $\tau \in [0, 1]$. Panel (B) represents the overlap number of components of the group-average FC, for each Schaefer parcellation. Panel (C) is the differential change (in %) between two consecutive number of component statistics across τ for group-average FCs (*top*) and **mean** of individual subject FCs (*bottom*).

To study this characteristic, we pick the resting state fMRI data, e.g. $rsfMRI$. Without loss of generality, we pick the first resting scan, i.e. $REST_1$, with phase encoding LR. It is important to note that the connectivity (through computing number of connected components) of the thresholded FC (where the absolute values of functional edges are set to zero - only apply step (a) above) is analogous to its binarized thresholded counterpart (where the surviving functional edges are set to one - applying both step (a) and (b) for any

given threshold and Schaefer parcellation choice). The number of components is computed using the python package networkx after converting the FC matrix to a graph object.

First, for all considered Schaefer parcellations, we see that the group-average FC fragments (e.g. splits into more than one connected components) earlier than the individual subjects' ones. This is because the normalization of functional edges across cohort domain, which neutralizes individual differences and zeroes out relatively faster across the thresholding range (Panel A). Moreover, it is also expected that the group-average number of connected components increases proportionally with the parcellation sizes and that for a fixed threshold value, the number of connected component statistics of a coarser parcellation is always smaller than that of a finer one (Panel B).

We also compute the differential change (in percentage) between two consecutive C across threshold range as follows:

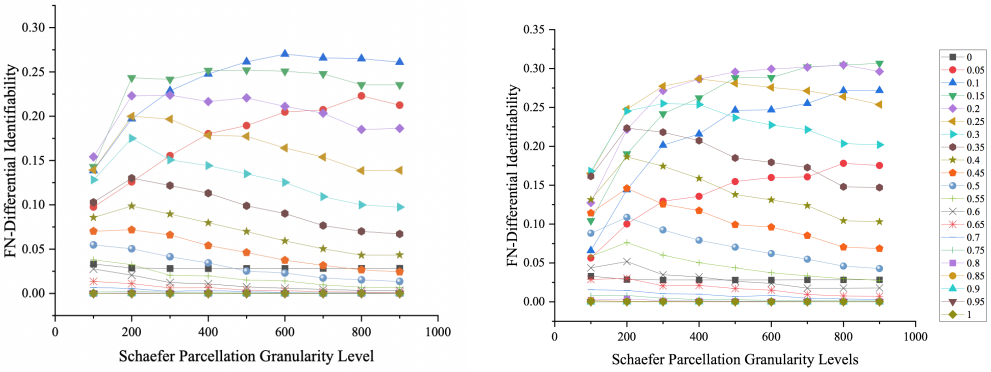
$$\Delta C_l(\%) = \frac{|C_{l+1} - C_l|}{C_l} * 100$$

where l is indexed over threshold range. Firstly, we see that both group-average (Panel C-*top*) and individual level (Panel C-*bottom*) show an empirical phase transition in the number of connected components. Such phase transition happens at sub-interval $(0.05, 0.30)$ and $(0.20, 0.45)$ for group-average and individual level, respectively. Although there is an numerical overlap between the two phases, group-average transitions earlier than individual one.

4.5.2 FN-Differential Identifiability

In this section, we investigate the behavior of matrix W_{bin} of group-average FCs, using Yeo's 7 RSNs [1]. To make some empirical observations about Schaefer FC sequence degree regime, we look at the group-average masked FC, M^{GA}, across all nine granularity levels and threshold interval $\tau \in [0, 1]$ with an increment of 0.05. The reason we look at only the masked (binarized) FCs is because

- The Sandon et al. [22] theorem on weak-recovery is written for binary graphs. Hence, the recoverability requirement on degree-regime is only applied to the binary scaffold.



(a) FN-Differential Identifiability I_{diff}^{FNs} score for Language fMRI task (LR scanning pattern).

(b) FN-Differential Identifiability I_{diff}^{FNs} for resting state (LR scanning pattern).

Figure 4.4. FN-Differential Identifiability I_{diff} for resting state and one particular fMRI task across all threshold and Schaefer Granularity Level combinations.

- We see that looking at weighted graphs is not appropriate in this case as the row (or column) sum of the FC matrix would yield connectivity strength of a node, not degree.

Here, we investigate empirical degree regime of Schaefer graph sequence based on the behavior of W_{bin}. For all studied Schaefer granularity level and threshold combinations, to infer W_{bin}, we simply use the maximum likelihood rule as mentioned in the main text. Recall that since matrix $W_{bin} = [w_{\text{ij}}]$ where w_{ij} contains the probability that a node u in community i is connected (e.g. $a_{uv} = 1$) or not-connected (e.g. $a_{uv} = 0$) to another node v in community j. Its entries are bounded between 0 and 1. Also, recall that in previous section on degree regime, graph sequence is in constant degree regime if the corresponding matrix W does not scale with n, e.g. $s_n = 1$.

Here, we look at the behavior of degree regime through a propose measure, called FN-differential identifiability, inspired by Amico et al. [5], as follows:

$$I_{diff}^{FNs} = I_{self}^{FNs} - I_{others}^{FNs} \tag{4.1}$$

$$= \langle W_{\text{ii}} \rangle - \langle W_{\text{ij}} \rangle \tag{4.2}$$

where $\mathrm{i}, \mathrm{j} \in [k]$ and $k = 7$ in our study. Moreover, $\langle W_{\mathrm{ii}} \rangle$ and $\langle W_{\mathrm{ij}} \rangle$ are average of diagonal and off-diagonal entries of matrix W, respectively. We formally define $\langle W_{\mathrm{ii}} \rangle$ and $\langle W_{\mathrm{ii}} \rangle$ to be the differential identifiability within (e.g. I_{self}^{FNs}) and between (e.g. I_{others}^{FNs}) FNs, see Chapter 2 for further details on identifiability function definition.

4.6 Results

In this section, we investigate the level of information-theoretical prominence of *a priori* set of FNs with respect to different FCs (both group-average and individual subject levels) across thresholding values using weak recovery criteria. We provide further insights on using SNR as a information-theoretic prominence of *a priori* set of FNs. The dataset used in this paper contains the 410 unrelated participants (HCP, Q3 release). This includes test and retest sessions for resting state and seven fMRI tasks: gambling (GAM), relational (REL), social (SOC), working memory (WM), language processing (LANG), emotion (EMOT), and motor (MOT). Whole-brain functional connectomes estimated from this fMRI dataset include 9 distinct Schaefer granularity levels that parcellate the cortical regions into $n = 100$ to $n = 900$ nodes, with a 100 nodes increment for each parcellation. The functional communities evaluated in this framework include seven cortical resting state FNs from [1]: visual (VIS), somatomotor (SM), dorsal attention (DA), ventral attention (VA), frontoparietal (FP), limbic (LIM), default mode (DMN). For each Schaefer granularity, there is a corresponding Yeo's FN parcellation.

4.6.1 Weak-recoverability sub-interval (a_w, b_w)

Based on panel (A) of 4.5, we see that for most Schaefer granularity levels (with the exception of $n = 100$), the lower and upper bound of theoretically guaranteed sub-interval of weak-recovery stay fairly stable: $\tau = [0.05, 0.8]$. The lower bound a_w stabilizes faster than the upper bound b_w, across Schaefer parcellations. With the exception of low resolution parcellation $n = 100$, the weak-recovery valid range is respectfully stable and parcellation-independent. This could elucidate that the information-theoretical relevance of *a priori* set of FNs is, to some extent, parcellation-free. In other words, for all investigated granularity

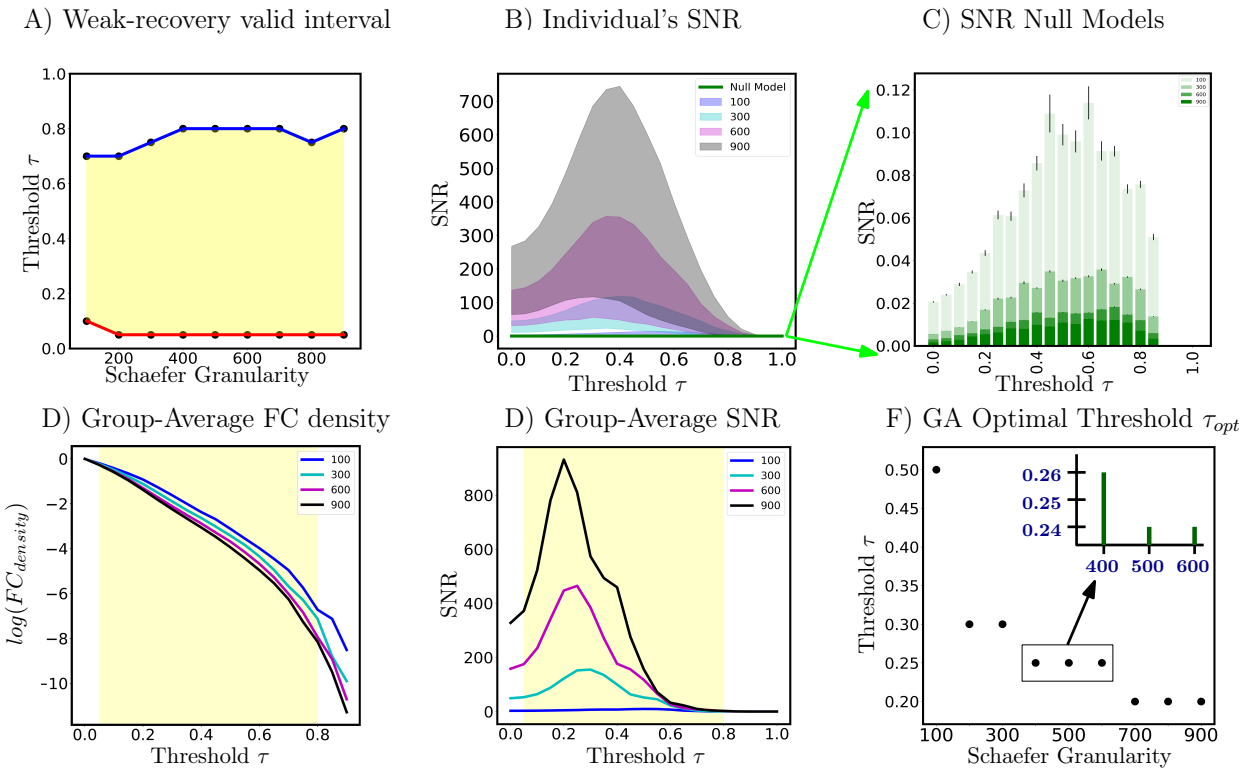


Figure 4.5. Panel (A) is the weak-recoverability sub-interval of $\tau \in (a_w, b_w)[0,1]$ (**Step 2**). Panel (B) is the 5- and 95- percentile of individual subjects' SNR for four distinct Schaefer parcellations $n = 100, 300, 600, 900$. Panel (C) is the SNR null models. Panel (D) is the FC density, on logarithmic scale, across the same 4 granularity levels. Panel (E) is SNR profiles computed on group-average FCs (again, over the same granularity levels). Finally Panel (F) reports the optimal threshold τ_{opt} computed based on maximum SNR of group-average FCs. Note that in panel (D) and (E) the weak-recoverability sub-interval using the maximum and minimum values for upper and lower bound, respectively, across Schaefer parcellations.

levels, the thresholded graphs are in weak-recoverability regime, except for the complete graph($\tau = [0, 0.05]$) or empty graph extreme ($\tau = (0.8, 1]$), see figure 4.5 panel **D** for further details on FC density. This is rather interesting because at those two extremes, networks contains either too much noise (complete graphs) or too little signal (empty graph) for any highly putative partitions to be information-theoretically relevant.

4.6.2 Resting State: Group-Average versus Individuals

Based on panel (B) and (E) of figure 4.5, it is evident that all SNR profiles (including the group average and individual levels) behave non-monotonically across thresholding range. There exists a threshold value such that SNR is maximized in the investigated range $\tau \in [0, 1]$. In addition, all optimal threshold values (for both group-average and individual FCs) are within the weak-recoverability sub-interval (a_w, b_w) for all investigated Schaefer granularity levels. Secondly, we see that both group-average and individual SNR profiles scales with

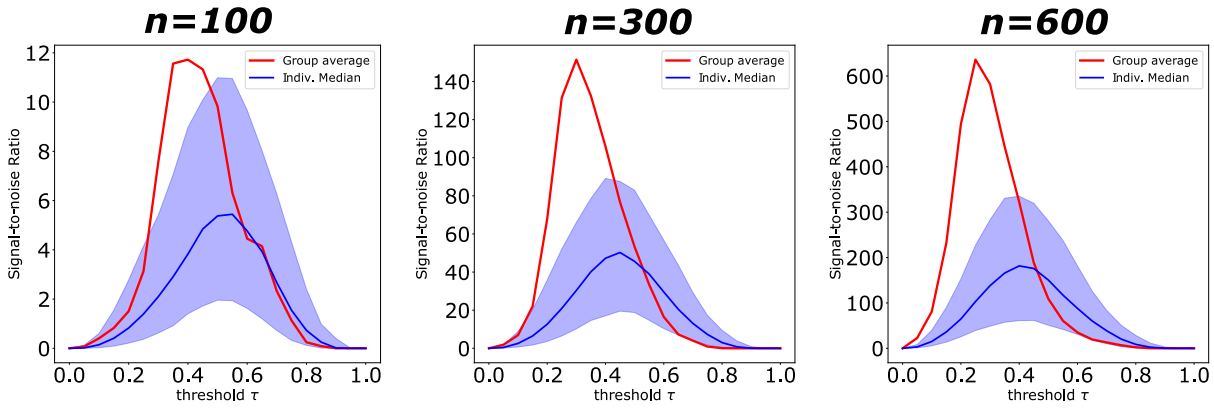


Figure 4.6. Overlaid SNR profiles between group-average and individual SNR for $n = [100, 300, 600]$ of scanning pattern LR.

n. This is because the scaling factor s_n for the Schaefer FC sequence is not constant. In other words, as the graph size gets larger, one can expect the community profile matrix PQ, whose entries $[PQ]_{\text{ij}}, \forall \text{i}, \text{j} \in [k]$ represents the number of expected "friends" between FN i and j (e.g. between DMN and LIM) gets larger numerically.

Third, we see that for a fixed Schaefer granularity level, group-average SNR peaks **higher and earlier** (across investigated threshold range) than that of individual subject's SNR. Interestingly, the topological property of *connected components* for both individual and group-average FCs, across all Schaefer parcellations, also yield a similar trend. Specifically, individual FC fragments (the number of connected component is larger than one) earlier, compared to the corresponding group-average FCs, for a fixed granularity level.

Topologically and numerically, this can be explained as averaging FC entries across subject domain damps down the individual fingerprints (which could be presented as high magnitude Pearson correlation values in FCs). This results in a smaller (magnitude-wise) functional connectivity entries in FCs which get eliminated by smaller threshold value τ. On the other hand, using the same analogy, one can see that it takes higher threshold value for individual FC entries to be annihilated. Further, the group-average SNR peaks not only higher but earlier compared to the individual SNR, for a fixed parcellation. Results are shown in figure 4.6.

4.6.3 Null models

Null model is assessed by feeding a randomized partition that respects Yeo's FNs sizes. The number of simulations is 100 and the scanning session is LR. Results on empirical distribution of randomized SNR scores are shown in figure 4.7.

4.6.4 Individualized optimal thresholds

As one can observe from figure 4.8, individualized optimal thresholds vary across different individuals which demonstrates there exists FN functional fingerprint evidence across subjects exists. In addition, the average of these individualized thresholds, for a given parcellation granularity, is roughly equal to the group-average optimal threshold.

4.6.5 Group-average: Resting State vs. fMRI Task Analysis

Next, we investigate the prominence of Yeo's RSNs with respect to the fMRI conditions (including 7 tasks and resting state) through SNR properties using group-average FCs, across all Schaefer granularity levels and thresholds. Using resting state SNR profile as a baseline, we compare all task responses in two particular scenarios:

- constructing FCs with the maximum number of time-points (tp) available for each fMRI conditions;

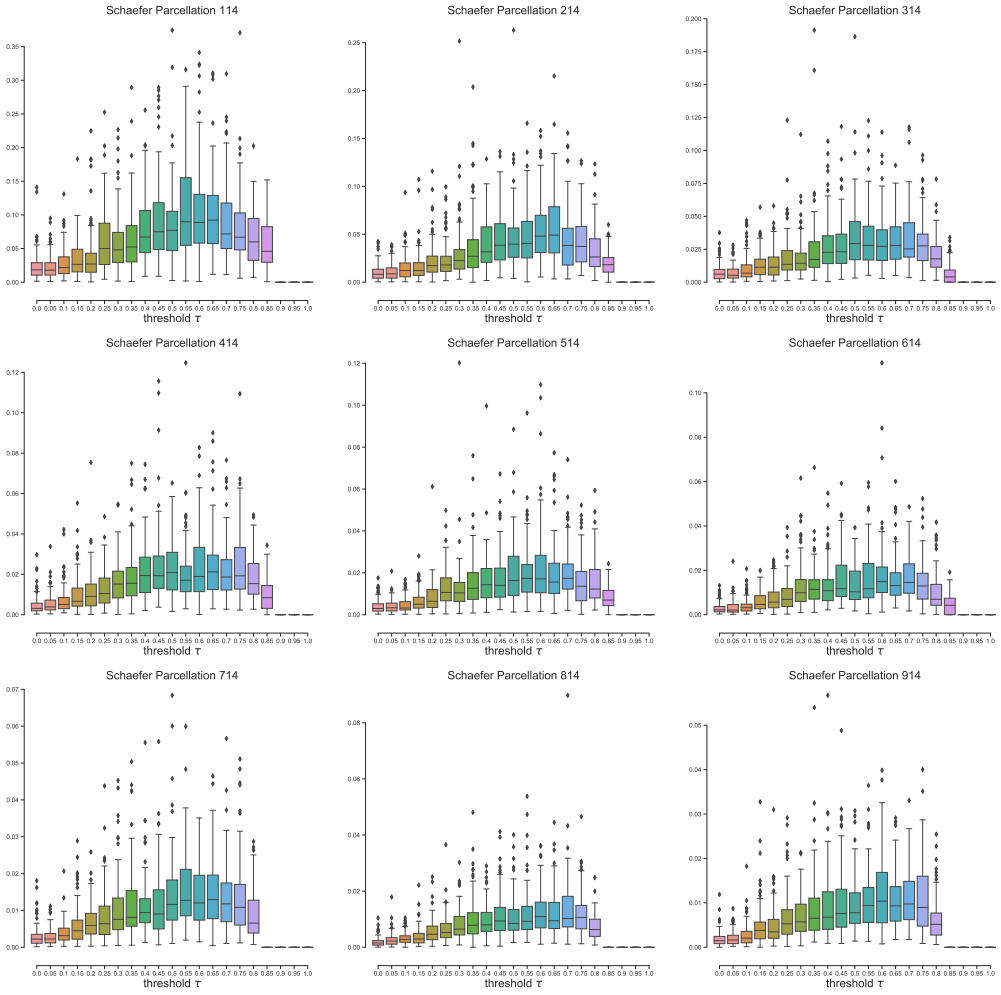
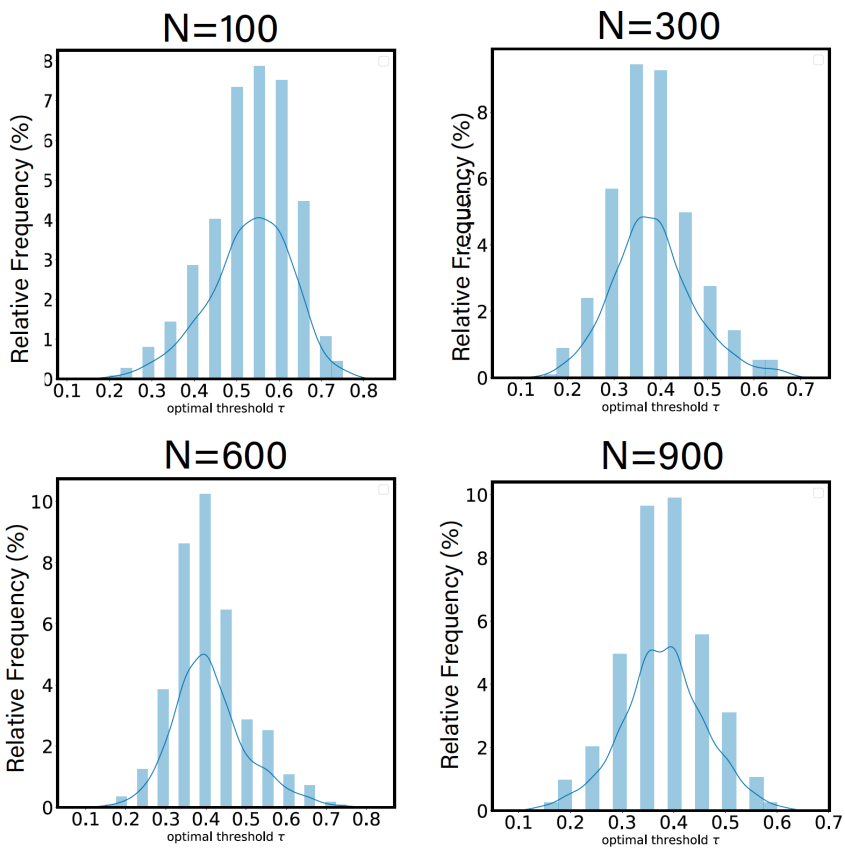


Figure 4.7. Each subplot represents the SNR profiles corresponding to 100 randomized parcellations for each thresholding value $\tau \in [0, 1]$ for all nine Schaefer parcellations.

- constructing resting state FCs using fMRI task scanning length and compute the corresponding SNR.

Firstly, for both scenarios, the **maximum** SNR values for all studied tasks are above the hard threshold $SNR = 1$ for weak recoverability. Moreover, the associated τ_{opt} falls



Note: *relative frequency is measured in percentage.*

Figure 4.8. Individualized optimal threshold is derived using the SNR behavior of each individual for 4 distinct Schaefer's parcellations $n = \{100, 300, 600, 900\}$.

also within (a_w, b_w). Trivially, resting state SNR dominates all available tasks, and across all parcellation levels. This is expected because the selected set of FNs are Yeo's RSNs. Secondly, we see that Working Memory (WM) task responds fairly consistently across all granularity levels, for both scenarios. Information-theoretically, EMOT is the closest task to resting state, with respect to Yeo's RSNs.

Thirdly, in the maximum tp case, with the exception of $n = 100$ parcellation, most task SNR profiles are, at most, approximately half magnitude, compared to resting-state SNR. In addition, group-average task FCs seems to have an earlier SNR-peak, relatively to resting-state, for all investigated Schaefer parcellations. Further details are indicated in 4.9 - Panel **A**.

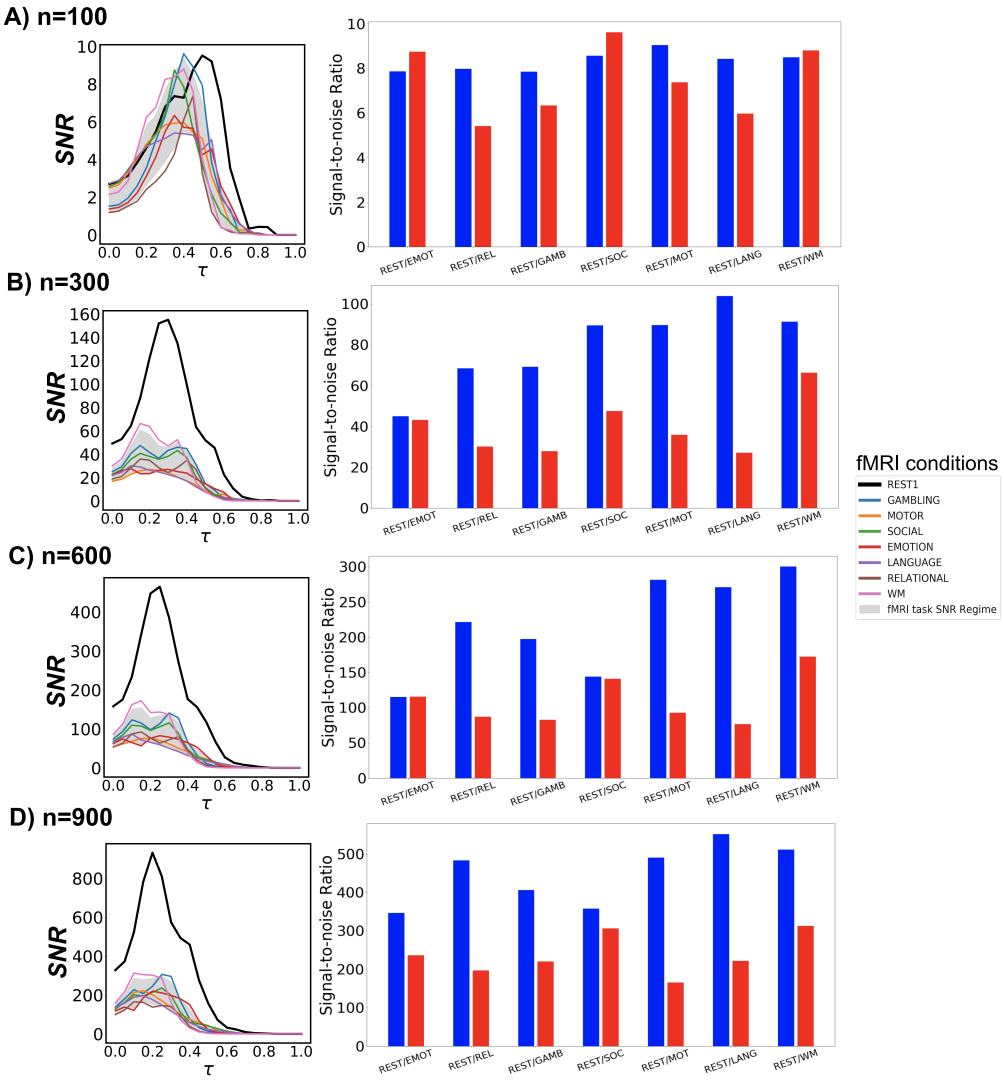


Figure 4.9. fMRI task and rest SNR profiles. Each panel (A,B,C,D) represent the SNR between rest and fMRI task using i) maximum scanning length (**left plots**) and ii) fMRI-task scanning lengths to estimate resting state FCs (**right plots**). Results are presented for 4 Schaefer parcellation levels: $n = [100, 300, 600, 900]$.

In the second scenario (when the minimum number of tp is used across all fMRI conditions), the SNR magnitude gap between resting state and fMRI task get reduced significantly, although resting-state SNR still dominates fMRI tasks. Further details are indicated in 4.9 - Panel **B**. Note that the gray shaded area indicates the 5- and 95- percentile of SNR responses among all fMRI tasks.

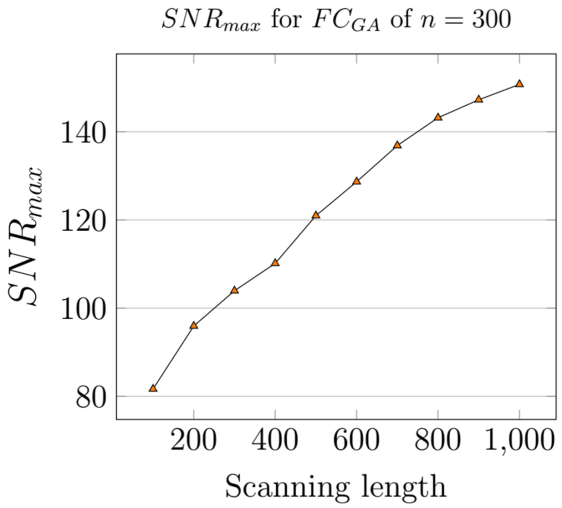


Figure 4.10. Maximum SNR computed for Resting state of Schaefer Group-average FC with $n = 300$ for increasing scanning lengths, starting at 100 to 1000 time-points, increments of 100 each.

4.6.6 The SNR-driven inequality

It is important to check if SNR is robust against randomness (or equivalently, whether it is a good marker driving the thresholding decision). To do so, we randomly shuffle the Yeo's RSNs and essentially recompute SNR response. We repeat the random shuffling procedure 100 times, and record the result for all nine Schaefer parcellations' group-average FC for $REST_1$ with scanning pattern LR.

Collectively, the null model SNR profiles are uniformly lower than all subjects, across the thresholding range, for a fixed Schaefer parcellation granularity. Further, the null model values do not exceed the hard threshold posed by weak recovery criteria, i.e. $SNR = 1$. This observation applies across all investigated Schaefer parcellations, as seen in panel **F** of figure 4.5. Another interesting observation is that the SNR gets uniformly smaller as Schaefer parcellation granularity increases, as seen also in Panel **F**.

Collectively, given the SNR results obtained at rest and task and null models, we can empirically form an inequality relation between fMRI resting and task conditions with respect

to SNR response and the corresponding level of prominence of Yeo's RSNs across different fMRI conditions:

$$0 < SNR_{null} < SNR_{task} < SNR_{rest} \tag{4.3}$$

This general order of SNR response is observed at the threshold τ that maximizes the objective function SNR, in the weak recovery case. At optimal threshold values, all fMRI task SNR profiles are in weak recoverability region while still magnitude-wise smaller than SNR at rest (when comparing maximized SNR values). Together, these inequalities constitute an empirical lower-bound and upper-bound for SNR_{task}, at least for all investigated tasks in our study.

4.6.7 Maximum SNR and threshold relationship

As the granularity of Schaefer parcellation gets finer, the corresponding group-average SNR profiles gets larger due to natural scaling for the community profile matrix PQ. This observation applies for the majority of threshold range. Moreover, per figure 4.5 Panel **F**, we see that optimal thresholds, e.g. τ_{opt}, tend to decrease as the granularity level increases which suggests that larger Schaefer FCs do not need to be thresholded as much. Another interesting observation is that with the exceptions of $n = \{100, 200, 900\}$, all other investigated granularity levels accept a very stable optimal threshold $\tau_{opt}^{GA} = 0.25$. Being a computation pipeline that relies on discretized line search on threshold τ (of increments 0.05 for $\tau = [0, 1]$), yielding this level of consistence of optimal value is unexpected.

4.6.8 Highly putative partition comparisons

The theory of weak recovery and its extended usage proposed here allowed us to argue the relevance of using SNR as a measure that guides the estimation of functional connectivity (through thresholding) with respect to *a priori* set of FNs. In this section, we would like to compare the practicality of using SNR as a driving measure, compared to other objective-function community detection methods. To do so, we use Newman's modularity Q-score [86], [87], [153]. Essentially, the Q score measures the statistical differences between a network

and its corresponding null model with some similarity topological property such as degree sequence. It can be computed as follows:

$$Q = \sum_{u,v}(A_{uv} - \alpha P_{uv})\delta(\sigma_u, \sigma_v)$$

where $\delta(\bullet, \bullet)$ and α is the Kronecker delta and tuning parameter (which by default is set to $\alpha = 1$), respectively. In network neuroscience, the majority of studies examining mesoscopic

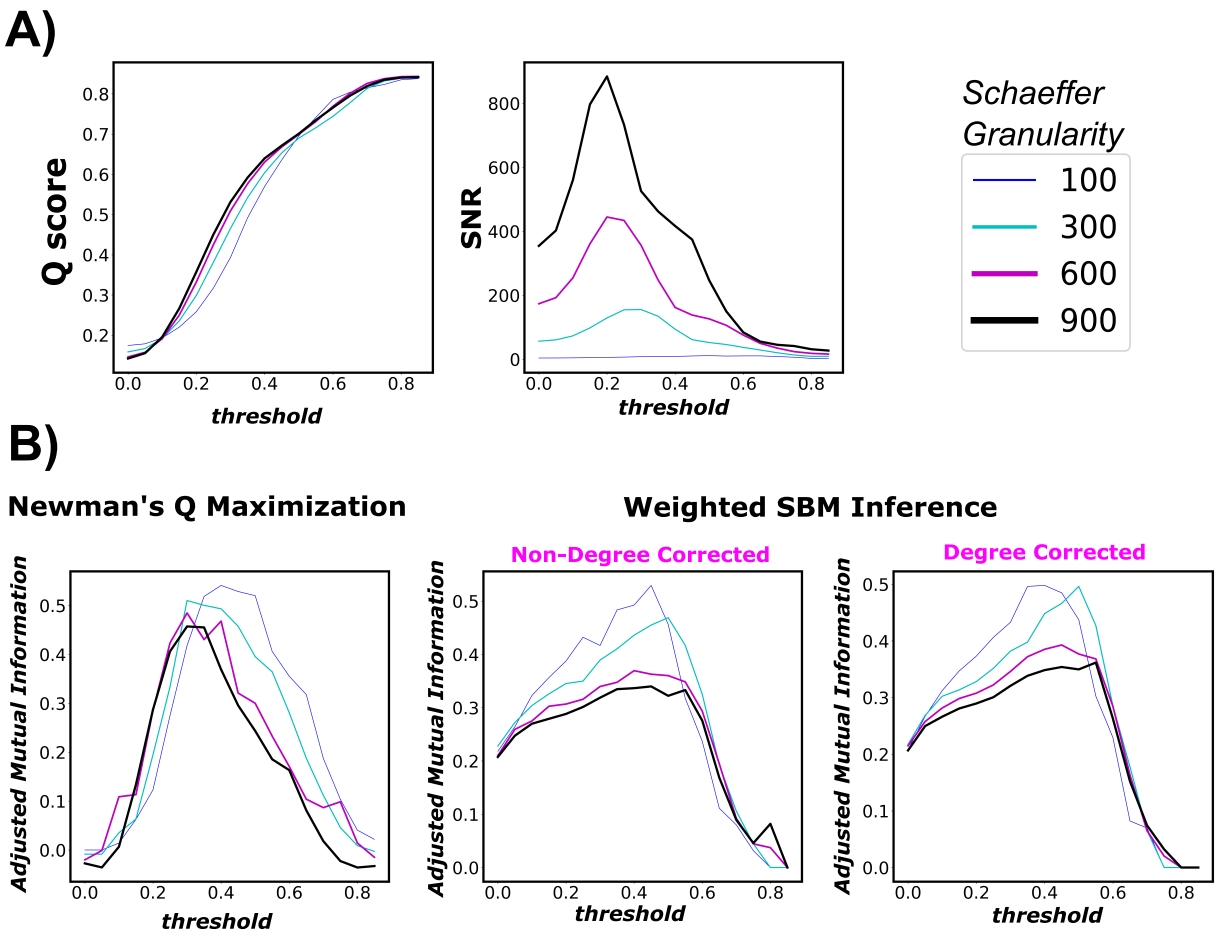


Figure 4.11. Panel (A) - left figure represents the modularity score of a thresholded group-average FC across threshold range $\tau \in [0, 0.85]$. Panel (A) - right figure reports the normalized mutual information between the inferred partition (using Q-score maximization heuristics) and the Yeo's FN partition. The same order goes to Panel (B). Panel (B) represents the results for SNR approach. Note that the full threshold range is not necessary because in the sub-interval $\tau \in [0.9, 1.0]$, the thresholded graph is almost (if not) empty graph. The displayed result is for the group-average FCs, over four Schaefer granularity levels $n = [100, 300, 600, 900]$.

structures of brain functions heavily leverage the maximization of modularity score (hereby denoted as Q_{max}), which unravels predominantly assortative communities (mesoscopic structures with denser internal edge density than external one). SBM inference method is, in principle, uncovering a more diverse types of communities, beyond assortative ones such as dis-assortative or core-periphery, [151]. Because of such distinct difference in principle between the two approaches (e.g. Q_{max} and $WSBM$ inference), modularity would provide a good benchmark test for robustness of SNR against a variety of community detection approaches. Note that for WSBM inference, we assume a Poisson distribution for the weighted graph [154]. Although other model assumptions are possible, our goal in this work is not about selecting the most fitting model assumption but rather investigating the behavior differences in communities detected using two physiologically different approaches. In other words, we are not looking to see if Q score or SNR picks up the exact threshold where the inferred partition is information-theoretically agreed with Yeo's FNs but rather either one of those two measures captures the threshold interval where the two partitions meet with relatively competitive degree of agreement. To measure information-theoretic agreement between inferred and ground-truth partitions, we use adjusted mutual information (AMI) - a measure that is adjusted for chance.

First, per figure 4.11 right panels, we see that both community detection methods (e.g. modularity score maximization or Weighted SBM inference) yield a very similar trends. Specifically, both AMI profiles go up and down crossing the threshold range. Further, AMI gets smaller as n gets larger which is expected for graphs with increasing number of nodes. Interestingly, the threshold values that maximizes modularity AMI tends to shift left as n increases. We see this particular behavior with SNR in earlier result section (Panel F of figure 4.5).

Secondly, modularity score keeps a fairly steady rise in magnitude across threshold range. Further, it does not appear that Q score is parcellation dependent; this is expected because the measure is normalized by $2m$. Moreover, Q-score peaks and plateaus at very high threshold range $\tau \in [0.6, 0.8]$. In such range, the thresholded FC is highly fragmented, see figure 4.3, extremely low edge density, see figure 4.5. and has no interesting topological insights remained for further analysis.

Lastly, we see that SNR driven curves (with *a priori* set of FNs) behave very similarly to AMI profiles of both approaches, e.g. the Newman-Girvan Q maximization heuristic and the Weighted SBM inference method. On the other hand, Q score keeps rising and plateau across threshold range which suggest no actual usage of picking a threshold that is useful for *a priori* partition such as Yeo's FNs. Collectively, our results show, once again, that SNR computation on weighted, thresholded FC provide excellent guidance to reconstruct a graph with the most information-theoretical relevance to a particular fixed set of FNs.

4.7 Discussion

In recent years, network neuroscience field strives forward with many exciting discoveries that are becoming more and more relevant to clinical applications and personalized medicine. In network neuroscience, this urges the need to improve a very important proxy of brain function: functional connectivity. Having the most proper, state-of-the-art mathematical representation of distributed brain circuits allows more accurate and confident inferences. In this work, we put forth a simple framework that allows an improvement of the mathematical representation of brain functions, given that *a priori* set of functional networks are to be utilized subsequently in a research project using FCs. Thresholding, which belongs to post-FC processing step, is usually overlooked as standard practice involves arbitrary elimination of statistically spurious edges. This step has become more critical as an increasing body of clinical research involves FC thresholding in construction pipeline.

First and foremost, there is no constant threshold value that is optimal across different parcellation granularity levels. This observation is reproducible across all Schaefer parcellation granularity levels. In particular, from coarser to finer grain of Schaefer granularity, optimal threshold value decreases. This result is partially observable in the behavior of matrix W across all studied Schaefer parcellation. Furthermore, we also see that for a fixed threshold value, as Schaefer granularity increases, Yeo's functional networks behaves more in an assortative manner (denser internal edge density and sparser external one). We see that through a brighter diagonal and a darker off-diagonal regime of matrix W, across Schaefer parcellations with fixed threshold value τ. Information-theoretically, it means that larger

graph (in size) tends to contain more relevant information to unravel the ground-truth partition (in our study, seven Yeo's RSNs); hence, we do not need to threshold the FCs as deep at the lower granularity parcellations such as $n = 100$. This result also suggests that as FC size is proportional to the level of prominence/fitness of *a priori* set of FNs. Nonetheless, it is unknown if this behavior will reach a plateau threshold even if the granularity increases.

Moreover, when using SNR as a goodness of fit measure when fitting an *a priori* set of FNs onto FC, there are distinct difference observed between resting state and fMRI task conditions, except for the low-resolution parcellation of $n = 100$. There are two ways of interpreting this result: i) there exists a degeneration in the fitness level of *a priori* set of FNs when subjects are at rest, compared to when they are in task-engaged mode; ii) There is an intrinsic shift of functional networks at the individual level between rest and task-engaged mode. Furthermore, there is also strong evidence suggesting a wide variance in individualized thresholds across all Schaefer parcellation granularity levels. In the same vein, our results also support the concept of *individualized parcellation* suggested by the work of Salehi and colleagues [8]. Individualized parcellation across subjects and tasks are intuitive and insightful but computationally demanding. To that end, our work offers a well-defined tool to examine the level of relevance a particular set of functional networks can be mapped on to individual FCs at different conditions. In other words, it allows us to, for the first time, quantify the individual difference (through information-theoretical gap) when the same atlas is mapped across cohort and/or tasks domain. It also open doors to proposed another alternative to elaborate further upon this current framework to build a task- or subject-dependent parcellation, besides the method suggested in [8].

In this study, we put forth the extended usage of weak recovery theorem (using the goodness measure SNR). Specifically, our results suggest that for the majority of threshold values, the masked binarized FCs are in the regime of week recovery. Nonetheless, there is still an open question on whether the sequence of FCs (as parcellated by Schaefer atlas, for a fixed, given individual and fMRI condition) is in exact recovery regime. Future studies need to address the information-theoretical gap between weak and exact recoverability requirements that is reflected by two measure: Signal-to-noise ratio (weak recovery) and Chernoff-Hellinger distance (exact recovery). Although exact recovery is a stronger requirement, if the Schae-

fer graph sequence falls within the exact recovery degree regime, the mutual information between inferred partition (through network inference and objective-based community detection methods) and ground-truth partition (e.g., Yeo's parcellations) will be theoretically higher.

Moreover, further studies need to address the voxels' spatial and temporal resolution limitations in HCP (and other fMRI) data and the corresponding Schaefer parcellation. Specifically, in spatial domain, further investigation should be done on the effect of voxel size when acquiring the fMRI data and its impact in the SNR result when fitting Yeo's functional networks. For the HCP data set, the fMRI voxel size is 2 mm isotropic [47]. In temporal domain, further investigation should be done on the effect of TR (time interval between two consecutive readings of BOLD time-series) when acquiring the fMRI data and its impact in the SNR result when fitting Yeo's functional networks. For the HCP data set, the TR is 720 ms [47]. One should also explore the potential impacts of SNR in a smaller dataset. Future studies need to also address the reliability of an *a priori* set of FNs (e.g., parcellation). Specifically, an in-depth analysis is needed to understand the difference and stability of an template-based alignment of FNs versus surface-based Freesurfer parcellation. Furthermore, one can also consider studying how structural connectivity play a role in shaping FNs. This study would shed light to a possible structure-function hybrid parcellation.

Our result suggests that we need to pay extra attention when apply a common/fixed atlas to individual FCs as we know that brain fingerprints [2], [5] exist. Furthermore, we show that FC post-processing step of thresholding FC matrices are not only intuitive (e.g. to arbitrarily eliminate statistically spurious edges) but also necessary if we would like to use such FCs, coupled with *a priori* set of FNs, to make quantitative statements in brain connectomics. Our research indicates a whole new direction of individualized and task-driven parcellation as a powerful alternative to using fixed parcellations such as Yeo's atlas [1] in brain connectomics research.

5. FUTURE OUTLOOK

The recent advent of neuroimaging modalities such as functional Magnetic Resonance Imaging (fMRI) has allowed a revolutionary perspective on investigating the brain as a networked system, which gave rise to the emerging field of Network Neuroscience. This has uniquely positioned network scientists to make meaningful contribution to push the frontier of neuroscience, both theoretically and clinically. The majority of my theoretical and applied research contributions exploits three distinct avenues: a) low dimensional description of human brain dynamics, b) towards a well-defined estimation of large scale whole-brain functional connectivity, and c) network anomaly detection in pathophysiological brain dysfunction.

5.1 Low Dimensional Description of human brain's functional dynamics

Human brain function involves a complex intertwined dynamics between neural elements of different scales that give rise to cognition. The high-dimensional nature of fMRI data has challenged researchers in network neuroscience field to unravel brain's key functional components using diverse technique of dimensionality reduction approaches. Many interesting insights on brain function dynamics have been unraveled using standard techniques, e.g. Principal/Independent Component Analysis. While traditional low-dimensionality reduction techniques show potentials in understanding brain dynamics at the global scales, they poses simultaneous technical shortcomings in exploring functions at mesoscale, i.e. brain sub-circuits. In the meantime, it has also been shown that human brain functional reconfigurations (switching between tasks and cognitive states) are rather subtle at macroscale, this has motivated a different low-dimensional description of brain functional dynamics through its sub-circuits' interactions. To accomplish this, a network morphospace is constructed using two key phenotypic characteristics of brain functional reconfigurations: segregation and integration. To parametrize this two dimensional space, further techniques in stochastic processes were employed to quantify how information traverses within (segregation) and between (integration) functional networks. This research projects contributes to the advanced understanding of brain functional configural properties, both theoretical (through the development of phenotypic space) and applied (through the association with individual's cognitive

measures such as fluid intelligence and working memory) aspect. Since a morphospace formalism was proposed [70], it is promising to explore human brain functional reconfiguration properties at finer scales such as task-to-task and within-task reconfiguration using dynamics functional connectivity.

5.2 Towards a well-defined estimation and efficient computation of large scale whole-brain functional connectivity

Well-defined human whole-brain functional connectivity estimation

The interdisciplinary field of network neuroscience has grown at an unprecedented pace in the last decade with, yet, many unrealized, unexplored opportunities. One of the (if not the) most critical tasks in this line of research is to properly estimate a so-called functional connectome, e.g. patterns of functional connectivity between brain neural elements, leveraging fMRI data. Mathematically, a functional connectome (FC) is often represented by a matrix pertaining pairwise interactions among given pairs of brain region of interest. The estimation of FC is of monumental importance before progressing to other subsequent analyses. This is also a critical step bridging between clinical data and theoretical studies. The human brain FC is known to be sparse (many brain regions are not connected, either functionally or structurally). This has raised the question on how to eliminate statistically spurious edges in FC in order to better represent the whole-brain functional connectivity pattern. Meanwhile, a dominant proportion of brain connectomic research relies heavily on using *a priori* set of highly-reproducible human brain functional sub-circuits (functional networks (FNs)) without properly considering whether it is information-theoretically relevant with respect to a given FC. Leveraging recent theoretical developments in Stochastic block model (SBM), we first formally defined and subsequently quantified the level of information-theoretical prominence of *a priori* set of FNs across different subjects, fMRI task conditions for any given input FC. As an extension to the first aim, the main contribution of this work is to provide an automated thresholding method of FCs based on prior knowledge of human brain functional sub-circuitry. In this study, FCs are constructed according to the Schaefer Atlas scheme, which has multiple levels of resolutions. Comprehensively, the initial results paved

the way to the proper usage of *a priori* set of FNs across different subjects, fMRI task conditions and, in turns, shed light for further studies in individualized parcellations. In future development, a well-defined FC does not need to be localized at just the thresholding step in constructing a whole-brain functional connectivity. Specifically, different combinations in pre-processing pipeline such as global signal regression, DiCER, head-motion correction etc. can significantly impact the resulting time-series and FC construction.

Towards efficient computations of whole-brain connectivity estimation

As our field striving towards high-quality big data era, the day of neuron-based connectomic is imminent. With that, a new opportunity in efficient computation emerges. For instance, how do we construct functional connectomes with billions of nodes, or perform low-dimensional studies of brain functional reservoir in a reasonable amount of time?

Graph-theoretically, FC is a weighted, complete graph (represented by a 2D matrix) of order of the number of brain regions as parcellated according to an anatomical atlas. As pointed out in the proposal "Brain Connectomics: Opportunities for High-Performance Computing", estimating the whole-brain functional connectome is an expensive computing job, from computing T1 image to defining brain regions for further statistical processing steps to correct noise, head motion and other artifacts. Each functional connectome might take up to the order of hours from start to finish. In the case of the HCP dataset (1200 participants), each subject yields 54 distinct functional connectomes which results in an enormous processing time, which could occupy a significant amount of computing resource. In order to have any realistic clinical utility, functional or structural connectome computations ought to be faster, more efficient and resource-mindful.

Recent developments in randomized numerical linear algebra (RandNLA) have opened tremendous opportunities to perform many computations for massive matrix or computationally expensive combinatorial problems such as sparse Principle Component Analysis (PCA), k-means etc. At the heart of RandNLA reside randomized algorithms. By carefully sampling rows/columns/elements of a matrix, one can construct a new, smaller matrices that are close to the original matrix (in terms of matrix norms) with proven theoretical guarantees (see

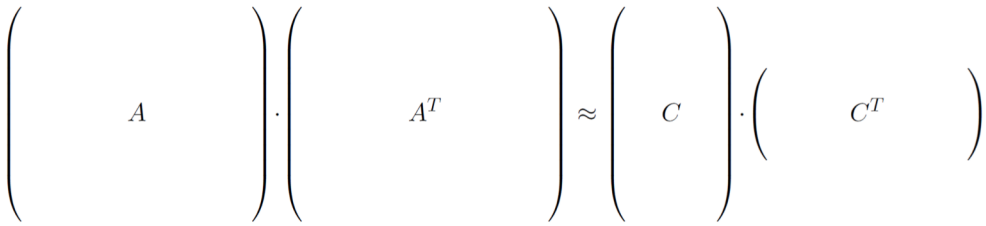


Figure 5.1. A computation demonstration of matrix multiplication between the original matrix (with the matrix transpose) on the left and its corresponding approximated sampled counterparts on the right.

figure 5.1). The resulting smaller matrices behave in similar fashion with the original matrix in terms of singular/eigen values and vectors) due to the matrix norm inequality bounds. In figure 5.1, matrix A could be viewed as the matrix containing the aggregated BOLD signals for brain parcellated regions according to some anatomical atlas. Specifically, matrix A's rows represent the brain regions while the columns represent the number of time-points (processed BOLD time series acquired at a specific time for a specific brain region).

Despite recent advancements in randomized algorithm research, there still exists a rich repertoire of unexplored opportunities in applying RandNLA to brain imaging, especially in network neuroscience domain. Specifically, opportunity for efficient computations lies across different computing steps before and/or after obtaining the human connectome. For instance, to improve the human brain fingerprints through functional connectome, Amico and Goñi [5] introduced a PCA-based framework applied to the original set of vectorized individual FCs where each subject's BOLD signals had been acquired at least twice per resting state and fMRI tasks. In this particular case, RandNLA could be used in sparse PCA procedures to speed up computing time.

5.3 Towards geometrically-aware computations of human brain whole-brain functional/structural connectivity

Nodes are the most basic element in complex networks; embedding nodes onto some geometrical structure is, often, one of the first steps to acquire further knowledge about how these elements interact [109],[155],[156]. Once these finest-scale elements are mapped

onto some host space, the notion of dimension (in that space) emerges. In a more general domain such as data analytic, when data are input as rows (representing distinct elements in the systems) and columns (representing distinct features of each elements), the notion of dimension of data-set is implicitly integrated [157],[158]. Despite the notion of data-set dimension, the pre-establishment of geometrical structure, such as Euclidean, is explicitly stated. Many techniques derived in the data analytic domain such has principle component analysis [157],[158] maintains a standardized setting of the embedded geometrical structure: Euclid. However, as we look deeper into this setting, one question arises: can we actually consider other embedded space and would that alter the course of analysis of the networks (in network science), or data-set (in data analytic domain)? Such question is addressed in [159]. Yet, before diving into whether hyperbolic geometry is an appropriate framework to unravel network features, let us briefly provide some basic concepts in hyperbolic geometry.

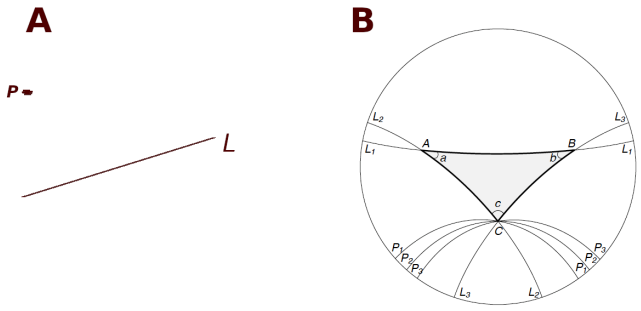


Figure 5.2. In panel **A**, Euclid's fifth postulate of parallelism, in two dimensional Euclidean space, is depicted pictorially. Specifically, given a point **P** and a line L that is not intercept **P**. There exists one and only one line L that is parallel to L and intercept **P**. In panel **B**, we consider mapping three points A, B and C on to hyperbolic space in which the triangle ABC is formed by three lines: L_1, L_2, L_3. Considering line segment AB and point C. Notice how there are more than one lines, i.e. $P_{1,2,3}$, that go through point C and parallel to line segment AB. In fact, there are infinitely many lines that have the same property as $P_{1,2,3}$ in panel **B** in when points are mapped into hyperbolic space. This example shows a fundamental difference between mapping points (or nodes in networks), into different geometrical structures. Panel **B** is re-used with permission from [159].

In terms of network science, Krioukov et al., [159], challenge the notion of Euclid embedding of nodes by proposing another embedding space, called hyperbolic geometry, in which

all classical components such as nodes, lines, distance, angles etc. are also well-defined. The work is motivated by many historical instances where significant findings were unravelled through the lenses of hyperbolic geometry such as Minkowski space-time relativity as stated in [159]. The below table summarises some of the major highlights of this paper in which many exciting features of complex networks emerges naturally from hyperbolic geometry.

Table 5.1. In this table, some features of networks embedded onto hyperbolic space are presented. Second column indicates the highlighted network properties while third column indicates the hyperbolic geometry of these properties.

Numbers	**Network Properties**	**Hyperbolic Underpinnings**	Notes
1	Topological heterogeneity - expressed by power-law degree distribution	Geometrical hyperbolicity - expressed by the notion of hyperbolic distance between two nodes in network.	This is an 'if and only if' relation, meaning: heterogeneity of degree sequence implies geometrical hyperbolicity of embedded data, and vice versa.
2	Network Modelling - generating networks with pre-determined features such as power-law exponent and average degree.	Nodes can be embedded in hyperbolic space with corresponding parameters (curvature and radius).	the notion of curvature are also provided in their previous paper [160].
3	Random graph and configuration models	Relaxing the connection probability in hyperbolic-based model to relate such model with statistical mechanics.	Connection links can be viewed under Fermi-Dirac Probability Distribution; hyperbolic distance can be interpreted as Fermi-Dirac links and corresponding parameter β as inverse temperature.

Hyperbolic geometry provides not only an appropriate space to embed network data but also an alternative viewpoint of looking at the same object (in this case, network). Recent movement in such direction has witnessed many extensions of such view into dynamical setting of scale-free networks [161] or clusters [162].

In brain connectomics domain, there is an increasing research focus in deriving well-defined frameworks in recent years, especially in the domain of underlying geometry in human brain functional networks. Recent studies have shown that whole-brain FCs, which are semi-positive definite matrices, are subsumed in Riemann's geometry (as opposed to Euclid's).

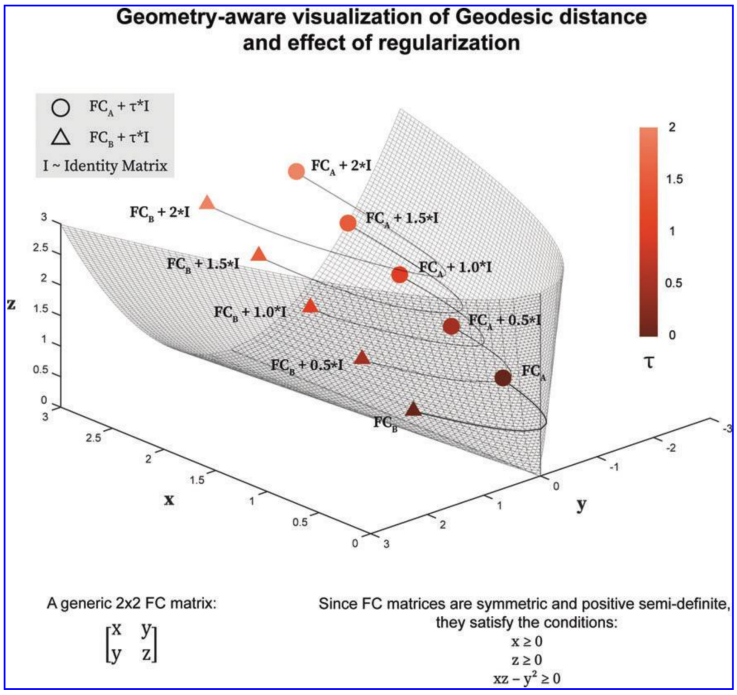


Figure 5.3. An example of 2×2 semi-positive definite matrices subsume the underlying Riemann's geometry represented by a semi-positive definite cones. In this particular case, the regularization terms τ dictate the size of the cone. Figure is adapted from [4].

5.4 Network anomaly detection in pathophysiological brain dysfunction

One of the biggest challenges in the field of network neuroscience is also about bridging the gap with a clinical domain. Specifically, can functional and structural connectome provide helpful insights for neurodegenerative disease such as Alzheimer's? Specifically, can we construct a robust statistical technique that can inform clinicians on possible structural/functional disruptions, using brain connectomics, across a given disease pathology? In network neuroscience, disease pathology can be thought as network anomaly. For instance, the healthy control FCs will be statistically different than Alzheimer's FCs, given some statistics of interest. Statistical process control (SPC) is a classical line of research developed in industrial engineering discipline for quality control applications. Recent advancement in SPC research has shown tremendous promise in networked system anomaly detection (also known as network surveillance).

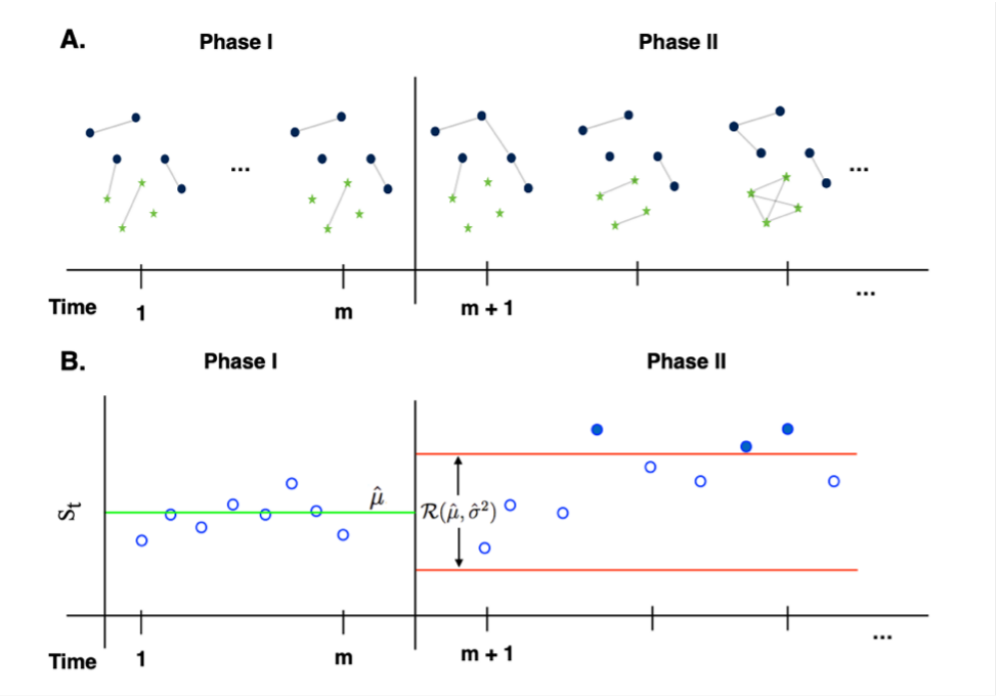


Figure 5.4. An example of using SPC to detect anomalous behaviors in dynamical networks. In a typical control chart design, there are two phases: characterization (Phase I) and monitoring (Phase II). Panel (A) represents a particular network behavioral dynamic across time. Panel (B) indicates the induced statistics of interest representing such behavior (also across time). Figure is reproduced at the courtesy of [163].

Recently, network surveillance tools are in even greater demands due to the emerging necessity to alert anomalous activities in social network applications, ideally as soon as they take place, see figure 5.4. As such, SPC has been shown to be a momentously useful framework. Nonetheless, the usage of SPC in detecting human brain's aberrant dynamics via functional/structural connectomes is completely unexploited thus far. This has opened tremendous opportunities in constructing a relevant SPC model with capability to detect early subtle structural/functional disruptions caused by different neurodegenerative diseases. Hence, building relevant SPC models could potentially aid clinicians in identifying early onset asymptomatic disruption that might, otherwise, be non-feasible. In the long run, SPC tools could prove useful in identify subtle Alzheimer's cognitive impairments which can often show no observable symptoms until it becomes too advanced.

5.5 Towards higher-order coordination patterns of human brain functions and structure

High-order topological invariant structures refer to non-trivial structures capturing interactions beyond pairwise fashion, see figure 5.5. Traditional methods in network science can only be used to study network local properties such as node degree, edge weights, or local motifs. Statistical methods quickly become cumbersome in identifying non-local mesoscopic structures within the weighted network fabrics. Since such structures cannot be reduced to known local network properties, they yield the so-called higher order coordination patterns. In this regards, topological data analysis (TDA) has also shown great potentials in identifying higher order (beyond pairwise interactions) sub-structures, i.e. functional sub-circuits, that, otherwise, are fully intact in healthy controls. Specifically, leveraging tools in algebraic topology such as persistent homology computations opens a whole new array of opportunities to identify potential non-local mesoscopic structures that distinguish between healthy controls and neurodegenerative groups. These structures could also be extremely useful bio-markers in tracking neurodegenerative disease pathology.

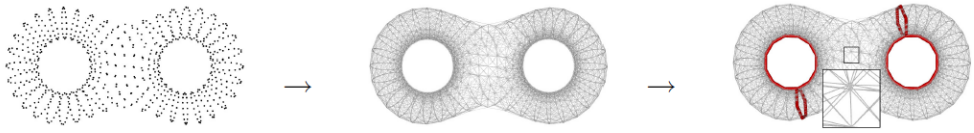


Figure 5.5. An example of TDA approach to high-dimensional dataset. In left panel, point-cloud data is sampled from a torus-shaped manifold. In the middle panel, Rip complex is constructed to create a topological scaffolding representation of the input data based on a chosen radius. In the right panel, based on the constructed Rip complex, homological computation is performed to unravel distinct topological features such as high dimensional holes (cycles). Figure is adapted from [164].

REFERENCES

[1] B. T. Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zöllei, J. R. Polimeni, *et al.*, "The organization of the human cerebral cortex estimated by intrinsic functional connectivity," *Journal of neurophysiology*, vol. 106, no. 3, pp. 1125–1165, 2011.

[2] E. S. Finn, X. Shen, D. Scheinost, M. D. Rosenberg, J. Huang, M. M. Chun, X. Papademetris, and R. T. Constable, "Functional connectome fingerprinting: Identifying individuals using patterns of brain connectivity," *Nature neuroscience*, vol. 18, no. 11, p. 1664, 2015.

[3] K. Abbas, M. Liu, M. Venkatesh, E. Amico, J. Harezlak, A. D. Kaplan, M. Ventresca, L. Pessoa, and J. Goñi, "Regularization of functional connectomes and its impact on geodesic distance and fingerprinting," *arXiv preprint arXiv:2003.05393*, 2020.

[4] K. Abbas, M. Liu, M. Venkatesh, E. Amico, A. D. Kaplan, M. Ventresca, L. Pessoa, J. Harezlak, and J. Goñi, "Geodesic distance on optimally regularized functional connectomes uncovers individual fingerprints," *Brain connectivity*, vol. 11, no. 5, pp. 333–348, 2021.

[5] E. Amico and J. Goñi, "The quest for identifiability in human functional connectomes," *Scientific reports*, vol. 8, no. 1, p. 8254, 2018.

[6] T. D. Satterthwaite, C. H. Xia, and D. S. Bassett, "Personalized neuroscience: Common and individual-specific features in functional brain networks," *Neuron*, vol. 98, no. 2, pp. 243–245, 2018.

[7] N. A. Saleh, M. A. Mahmoud, L. A. Jones-Farmer, I. Zwetsloot, and W. H. Woodall, "Another look at the ewma control chart with estimated parameters," *Journal of Quality Technology*, vol. 47, no. 4, p. 363, 2015.

[8] M. Salehi, A. S. Greene, A. Karbasi, X. Shen, D. Scheinost, and R. T. Constable, "There is no single functional atlas even for a single individual: Parcellation of the human brain is state dependent," *bioRxiv*, p. 431 833, 2018.

[9] K. C. Lopez, S. Kandala, S. Marek, and D. M. Barch, "Development of network topology and functional connectivity of the prefrontal cortex," *Cerebral Cortex*, vol. 30, no. 4, pp. 2489–2505, 2020.

[10] A. C. Murphy, M. A. Bertolero, L. Papadopoulos, D. M. Lydon-Staley, and D. S. Bassett, "Multimodal network dynamics underpinning working memory," *Nature communications*, vol. 11, no. 1, pp. 1–13, 2020.

[11] M. Jalbrzikowski, F. Liu, W. Foran, F. J. Calabro, K. Roeder, B. Devlin, and B. Luna, "Cognitive and default mode networks support developmental stability in functional connectome fingerprinting through adolescence," *BioRxiv*, p. 812 719, 2019.

[12] T. D. Satterthwaite, M. A. Elliott, R. T. Gerraty, K. Ruparel, J. Loughead, M. E. Calkins, S. B. Eickhoff, H. Hakonarson, R. C. Gur, R. E. Gur, *et al.*, "An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data," *Neuroimage*, vol. 64, pp. 240–256, 2013.

[13] J. Fan, I. F. Tso, D. F. Maixner, T. Abagis, L. Hernandez-Garcia, and S. F. Taylor, "Segregation of salience network predicts treatment response of depression to repetitive transcranial magnetic stimulation," *NeuroImage: Clinical*, vol. 22, p. 101 719, 2019.

[14] D. Lydon-Staley, C. Kuehner, V. Zamoscik, S. Huffziger, P. Kirsch, and D. Bassett, "Repetitive negative thinking in daily life and functional connectivity among default mode, fronto-parietal, and salience networks," *Translational psychiatry*, vol. 9, no. 1, pp. 1–12, 2019.

[15] D. H. Schultz and M. W. Cole, "Higher intelligence is associated with less task-related brain network reconfiguration," *Journal of neuroscience*, vol. 36, no. 33, pp. 8551–8561, 2016.

[16] J. M. Shine, P. G. Bissett, P. T. Bell, O. Koyejo, J. H. Balsters, K. J. Gorgolewski, C. A. Moodie, and R. A. Poldrack, "The dynamics of functional brain networks: Integrated network states during cognitive task performance," *Neuron*, vol. 92, no. 2, pp. 544–554, 2016.

[17] J. M. Shine and R. A. Poldrack, "Principles of dynamic network reconfiguration across diverse brain states," *NeuroImage*, 2017.

[18] J. Shine, M. Breakspear, P. Bell, K. M. Ehgoetz, R. Shine, O. Koyejo, O. Sporns, and R. Poldrack, "The dynamic basis of cognition: An integrative core under the control of the ascending neuromodulatory system," 2018.

[19] J. M. Shine, M. Breakspear, P. T. Bell, K. A. E. Martens, R. Shine, O. Koyejo, O. Sporns, and R. A. Poldrack, "Human cognition involves the dynamic integration of neural activity and neuromodulatory systems," *Nature neuroscience*, vol. 22, no. 2, p. 289, 2019.

[20] M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, *et al.*, "The minimal preprocessing pipelines for the human connectome project," *Neuroimage*, vol. 80, pp. 105–124, 2013.

[21] J. M. Shine, M. Breakspear, P. Bell, K. E. Martens, R. Shine, O. Koyejo, O. Sporns, and R. Poldrack, "Human cognition involves the dynamic integration of neural activity and neuromodulatory systems," *Nature Neuroscience*, pp. –, 2019.

[22] E. Abbe, "Community detection and stochastic block models: Recent developments," *arXiv preprint arXiv:1703.10146*, 2017.

[23] E. Abbe and C. Sandon, "Proof of the achievability conjectures for the general stochastic block model," *Communications on Pure and Applied Mathematics*, vol. 71, no. 7, pp. 1334–1406, 2018.

[24] J. D. Power, A. L. Cohen, S. M. Nelson, G. S. Wig, K. A. Barnes, J. A. Church, A. C. Vogel, T. O. Laumann, F. M. Miezin, B. L. Schlaggar, *et al.*, "Functional network organization of the human brain," *Neuron*, vol. 72, no. 4, pp. 665–678, 2011.

[25] E. M. Gordon, T. O. Laumann, B. Adeyemo, J. F. Huckins, W. M. Kelley, and S. E. Petersen, "Generation and evaluation of a cortical area parcellation from resting-state correlations," *Cerebral cortex*, vol. 26, no. 1, pp. 288–303, 2016.

[26] A. Schaefer, R. Kong, E. M. Gordon, T. O. Laumann, X.-N. Zuo, A. J. Holmes, S. B. Eickhoff, and B. T. Yeo, "Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri," *Cerebral cortex*, vol. 28, no. 9, pp. 3095–3114, 2018.

[27] E. M. Purcell, H. C. Torrey, and R. V. Pound, "Resonance absorption by nuclear magnetic moments in a solid," *Physical review*, vol. 69, no. 1-2, p. 37, 1946.

[28] F. Bloch, "Nuclear induction," *Physical review*, vol. 70, no. 7-8, p. 460, 1946.

[29] P. C. Lauterbur, "Image formation by induced local interactions: Examples employing nuclear magnetic resonance," *nature*, vol. 242, no. 5394, pp. 190–191, 1973.

[30] P. Mansfield and P. K. Grannell, "Nmr'diffraction'in solids?" *Journal of Physics C: solid state physics*, vol. 6, no. 22, p. L422, 1973.

[31] P. A. Dirac, "On the annihilation of electrons and protons," in *Mathematical Proceedings of the Cambridge Philosophical Society*, Cambridge University Press, vol. 26, 1930, pp. 361–375.

[32] J. P. Hornak, "The basics of mri," *http://www. cis. rit. edu/htbooks/mri*, 2006.

[33] D. J. Heeger and D. Ress, "What does fmri tell us about neuronal activity?" *Nature Reviews Neuroscience*, vol. 3, no. 2, pp. 142–151, 2002.

[34] R. A. Brooks, J. H. Battocletti, A. Sances, S. J. Larson, R. L. Bowman, and V. Kudravcev, "Nuclear magnetic relaxation in blood," *IEEE Transactions on Biomedical Engineering*, no. 1, pp. 12–18, 1975.

[35] N. K. Logothetis, J. Pauls, M. Augath, T. Trinath, and A. Oeltermann, "Neurophysiological investigation of the basis of the fmri signal," *nature*, vol. 412, no. 6843, pp. 150–157, 2001.

[36] G. H. Glover, "Overview of functional magnetic resonance imaging," *Neurosurgery Clinics*, vol. 22, no. 2, pp. 133–139, 2011.

[37] B. B. Biswal, J. V. Kylen, and J. S. Hyde, "Simultaneous assessment of flow and bold signals in resting-state functional connectivity maps," *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In Vivo*, vol. 10, no. 4-5, pp. 165–170, 1997.

[38] R. L. Buckner and J. L. Vincent, "Unrest at rest: Default activity and spontaneous network correlations," *Neuroimage*, vol. 37, no. 4, pp. 1091–1096, 2007.

[39] M. N. Moussa, M. R. Steen, P. J. Laurienti, and S. Hayasaka, "Consistency of network modules in resting-state fmri connectome data," 2012.

[40] M. D. Greicius, B. Krasnow, A. L. Reiss, and V. Menon, "Functional connectivity in the resting brain: A network analysis of the default mode hypothesis," *Proceedings of the National Academy of Sciences*, vol. 100, no. 1, pp. 253–258, 2003.

[41] C. Golgi, "The neuron doctrine: Theory and facts," *Nobel lecture*, vol. 1921, pp. 190–217, 1906.

[42] E. Jones, "The neuron doctrine 1891," *Journal of the History of the Neurosciences*, vol. 3, no. 1, pp. 3–20, 1994.

[43] A. Fornito, A. Zalesky, and E. Bullmore, *Fundamentals of brain network analysis*. Academic Press, 2016.

[44] R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman, *et al.*, "An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest," *Neuroimage*, vol. 31, no. 3, pp. 968–980, 2006.

[45] C. Destrieux, B. Fischl, A. Dale, and E. Halgren, "Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature," *Neuroimage*, vol. 53, no. 1, pp. 1–15, 2010.

[46] N. Bryce, J. Flournoy, J. F. G. Moreira, M. L. Rosen, K. A. Sambook, P. Mair, and K. A. McLaughlin, "Brain parcellation selection: An overlooked decision point with meaningful effects on individual differences in resting-state functional connectivity," *NeuroImage*, p. 118 487, 2021.

[47] D. C. Van Essen, K. Ugurbil, E. Auerbach, D. Barch, T. E. Behrens, R. Bucholz, A. Chang, L. Chen, M. Corbetta, S. W. Curtiss, *et al.*, "The human connectome project: A data acquisition perspective," *Neuroimage*, vol. 62, no. 4, pp. 2222–2231, 2012.

[48] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, K. Ugurbil, W.-M. H. Consortium, *et al.*, "The wu-minn human connectome project: An overview," *Neuroimage*, vol. 80, pp. 62–79, 2013.

[49] M. W. Weiner, D. P. Veitch, P. S. Aisen, L. A. Beckett, N. J. Cairns, R. C. Green, D. Harvey, C. R. Jack Jr, W. Jagust, J. C. Morris, *et al.*, "The alzheimer's disease neuroimaging initiative 3: Continued innovation for clinical trial improvement," *Alzheimer's & Dementia*, vol. 13, no. 5, pp. 561–571, 2017.

[50] P. A. Bandettini, E. C. Wong, R. S. Hinks, R. S. Tikofsky, and J. S. Hyde, "Time course epi of human brain function during task activation," *Magnetic resonance in medicine*, vol. 25, no. 2, pp. 390–397, 1992.

[51] J. Frahm, H. Bruhn, K.-D. Merboldt, and W. Hänicke, "Dynamic mr imaging of human brain oxygenation during rest and photic stimulation," *Journal of Magnetic Resonance Imaging*, vol. 2, no. 5, pp. 501–505, 1992.

[52] K. K. Kwong, J. W. Belliveau, D. A. Chesler, I. E. Goldberg, R. M. Weisskoff, B. P. Poncelet, D. N. Kennedy, B. E. Hoppel, M. S. Cohen, and R. Turner, "Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation.," *Proceedings of the National Academy of Sciences*, vol. 89, no. 12, pp. 5675–5679, 1992.

[53] C. Gratton, T. O. Laumann, A. N. Nielsen, D. J. Greene, E. M. Gordon, A. W. Gilmore, S. M. Nelson, R. S. Coalson, A. Z. Snyder, B. L. Schlaggar, *et al.*, "Functional brain networks are dominated by stable group and individual factors, not cognitive or daily variation," *Neuron*, vol. 98, no. 2, pp. 439–452, 2018.

[54] R. B. Mars, R. E. Passingham, and S. Jbabdi, "Connectivity fingerprints: From areal descriptions to abstract spaces," *Trends in cognitive sciences*, vol. 22, no. 11, pp. 1026–1037, 2018.

[55] V. Pallarés, A. Insabato, A. Sanjuán, S. Kühn, D. Mantini, G. Deco, and M. Gilson, "Extracting orthogonal subject-and condition-specific signatures from fmri data using whole-brain effective connectivity," *Neuroimage*, vol. 178, pp. 238–254, 2018.

[56] M. Rajapandian, E. Amico, K. Abbas, M. Ventresca, and J. Goñi, "Uncovering differential identifiability in network properties of human brain functional connectomes," *Network Neuroscience*, vol. 4, no. 3, pp. 698–713, 2020.

[57] B. A. Seitzman, C. Gratton, T. O. Laumann, E. M. Gordon, B. Adeyemo, A. Dworetsky, B. T. Kraus, A. W. Gilmore, J. J. Berg, M. Ortega, *et al.*, "Trait-like variants in human functional brain networks," *Proceedings of the National Academy of Sciences*, vol. 116, no. 45, pp. 22 851–22 861, 2019.

[58] M. Venkatesh, J. Jaja, *et al.*, "Comparing functional connectivity matrices: A geometry-aware approach applied to participant identification," *bioRxiv*, p. 687 830, 2019.

[59] S. Bari, E. Amico, N. Vike, T. M. Talavage, and J. Goñi, "Uncovering multi-site identifiability based on resting-state functional connectomes," *NeuroImage*, vol. 202, p. 115 967, 2019.

[60] D. S. Bassett and M. S. Gazzaniga, "Understanding complexity in the human brain," *Trends in cognitive sciences*, vol. 15, no. 5, pp. 200–209, 2011.

[61] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.

[62] M. F. Glasser, T. S. Coalson, E. C. Robinson, C. D. Hacker, J. Harwell, E. Yacoub, K. Ugurbil, J. Andersson, C. F. Beckmann, M. Jenkinson, *et al.*, "A multi-modal parcellation of human cerebral cortex," *Nature*, vol. 536, no. 7615, pp. 171–178, 2016.

[63] S. B. Eickhoff, D. Bzdok, A. R. Laird, C. Roski, S. Caspers, K. Zilles, and P. T. Fox, "Co-activation patterns distinguish cortical modules, their connectivity and functional differentiation," *Neuroimage*, vol. 57, no. 3, pp. 938–949, 2011.

[64] A. Avena-Koenigsberger, J. Goñi, R. Solé, and O. Sporns, "Network morphospace," *Journal of the Royal Society Interface*, vol. 12, no. 103, p. 20 140 881, 2015.

[65] G. R. McGhee, *Theoretical morphology: the concept and its applications*. Columbia University Press, 1999.

[66] R. V. Solé and S. Valverde, "Information theory of complex networks: On evolution and architectural constraints," in *Complex networks*, Springer, 2004, pp. 189–207.

[67] L. F. Seoane and R. Solé, "The morphospace of language networks," *Scientific reports*, vol. 8, no. 1, pp. 1–14, 2018.

[68] J. Goñi, A. Avena-Koenigsberger, N. V. de Mendizabal, M. P. van den Heuvel, R. F. Betzel, and O. Sporns, "Exploring the morphospace of communication efficiency in complex networks," *PLoS One*, vol. 8, no. 3, e58070, 2013.

[69] B. Corominas-Murtra, J. Goñi, R. V. Solé, and C. Rodríguez-Caso, "On the origins of hierarchy in complex networks," *Proceedings of the National Academy of Sciences*, vol. 110, no. 33, pp. 13 316–13 321, 2013.

[70] D. Duong-Tran, A. Kausar, E. Amico, B. Corominas-Murtra, M. Dzemidzic, D. Kareken, M. Ventresca, and J. Goñi, "A morphospace of functional configuration to assess configural breadth based on brain functional networks," *Network Neuroscience*, pp. 1–36, 2021.

[71] X. D. Arsiwalla, R. Sole, C. Moulin-Frier, I. Herreros, M. Sanchez-Fibla, and P. Verschure, "The morphospace of consciousness," *arXiv preprint arXiv:1705.11190*, 2017.

[72] S. E. Morgan, S. Achard, M. Termenon, E. T. Bullmore, and P. E. Vértes, "Low-dimensional morphospace of topological motifs in human fmri brain networks," *Network Neuroscience*, vol. 2, no. 02, pp. 285–302, 2018.

[73] A. Avena-Koenigsberger, B. Misic, and O. Sporns, "Communication dynamics in complex brain networks," *Nature Reviews Neuroscience*, vol. 19, no. 1, p. 17, 2018.

[74] R. Schuetz, N. Zamboni, M. Zampieri, M. Heinemann, and U. Sauer, "Multidimensional optimality of microbial metabolism," *Science*, vol. 336, no. 6081, pp. 601–604, 2012.

[75] O. Shoval, H. Sheftel, G. Shinar, Y. Hart, O. Ramote, A. Mayo, E. Dekel, K. Kavanagh, and U. Alon, "Evolutionary trade-offs, pareto optimality, and the geometry of phenotype space," *Science*, p. 1 217 405, 2012.

[76] R. Thomas, R. M. Shearman, and G. W. Stewart, "Evolutionary exploitation of design options by the first animals with hard skeletons," *Science*, vol. 288, no. 5469, pp. 1239–1242, 2000.

[77] M. W. Cole, D. S. Bassett, J. D. Power, T. S. Braver, and S. E. Petersen, "Intrinsic and task-evoked network architectures of the human brain," *Neuron*, vol. 83, no. 1, pp. 238–251, 2014.

[78] J. Gonzalez-Castillo, C. W. Hoy, D. A. Handwerker, M. E. Robinson, L. C. Buchanan, Z. S. Saad, and P. A. Bandettini, "Tracking ongoing cognition in individuals using brief, whole-brain functional connectivity patterns," *Proceedings of the National Academy of Sciences*, vol. 112, no. 28, pp. 8762–8767, 2015.

[79] E. S. Finn, D. Scheinost, D. M. Finn, X. Shen, X. Papademetris, and R. T. Constable, "Can brain state be manipulated to emphasize individual differences in functional connectivity?" *Neuroimage*, vol. 160, pp. 140–151, 2017.

[80] P. Varona and M. I. Rabinovich, "Hierarchical dynamics of informational patterns and decision-making," *Proceedings of the Royal Society B: Biological Sciences*, vol. 283, no. 1832, p. 20 160 475, 2016.

[81] R. B. Mars, R. E. Passingham, and S. Jbabdi, "Connectivity fingerprints: From areal descriptions to abstract spaces," *Trends in cognitive sciences*, vol. 22, no. 11, pp. 1026–1037, 2018.

[82] D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, J. M. Carlson, and S. T. Grafton, "Dynamic reconfiguration of human brain networks during learning," *Proceedings of the National Academy of Sciences*, vol. 108, no. 18, pp. 7641–7646, 2011.

[83] E. Amico, D. Marinazzo, C. Di Perri, L. Heine, J. Annen, C. Martial, M. Dzemidzic, M. Kirsch, V. Bonhomme, S. Laureys, *et al.*, "Mapping the functional connectome traits of levels of consciousness," *NeuroImage*, vol. 148, pp. 201–211, 2017.

[84] E. Amico and J. Goñi, "Mapping hybrid functional-structural connectivity traits in the human connectome," *Network Neuroscience*, pp. 1–17, 2018.

[85] O. Sporns and R. F. Betzel, "Modular brain networks," *Annual review of psychology*, vol. 67, pp. 613–640, 2016.

[86] M. E. Newman, "Finding community structure in networks using the eigenvectors of matrices," *Physical review E*, vol. 74, no. 3, p. 036 104, 2006.

[87] M. E. Newman, "Modularity and community structure in networks," *Proceedings of the national academy of sciences*, vol. 103, no. 23, pp. 8577–8582, 2006.

[88] M. Rosvall and C. T. Bergstrom, "Maps of random walks on complex networks reveal community structure," *Proceedings of the National Academy of Sciences*, vol. 105, no. 4, pp. 1118–1123, 2008.

[89] M. Rosvall, D. Axelsson, and C. T. Bergstrom, "The map equation," *The European Physical Journal-Special Topics*, vol. 178, no. 1, pp. 13–23, 2009.

[90] M. Milano, P. H. Guzzi, and M. Cannataro, "Using multi network alignment for analysis of connectomes," *Procedia Computer Science*, vol. 108, pp. 1155–1164, 2017.

[91] O. Tymofiyeva, E. Ziv, A. J. Barkovich, C. P. Hess, and D. Xu, "Brain without anatomy: Construction and comparison of fully network-driven structural mri connectomes," *PloS one*, vol. 9, no. 5, e96196, 2014.

[92] G. Varoquaux, Y. Schwartz, R. A. Poldrack, B. Gauthier, D. Bzdok, J.-B. Poline, and B. Thirion, "Atlases of cognition with large-scale human brain mapping," *PLoS computational biology*, vol. 14, no. 11, e1006565, 2018.

[93] J. L. Doob and J. L. Doob, *Stochastic processes*, 2. Wiley New York, 1953, vol. 7.

[94] J. G. Kemeny, J. L. Snell, *et al.*, *Finite markov chains*. van Nostrand Princeton, NJ, 1960, vol. 356.

[95] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[96] S. Kullback, *Information theory and statistics*. Courier Corporation, 1997.

[97] C. E. Shannon, "A mathematical theory of communication," *Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[98] O. Sporns, "Network attributes for segregation and integration in the human brain," *Current opinion in neurobiology*, vol. 23, no. 2, pp. 162–171, 2013.

[99] L. Douw, D. G. Wakeman, N. Tanaka, H. Liu, and S. M. Stufflebeam, "State-dependent variability of dynamic functional connectivity between frontoparietal and default networks relates to cognitive flexibility," *Neuroscience*, vol. 339, pp. 12–21, 2016.

[100] S. M. Smith, C. F. Beckmann, J. Andersson, E. J. Auerbach, J. Bijsterbosch, G. Douaud, E. Duff, D. A. Feinberg, L. Griffanti, M. P. Harms, *et al.*, "Resting-state fmri in the human connectome project," *Neuroimage*, vol. 80, pp. 144–168, 2013.

[101] D. M. Barch, G. C. Burgess, M. P. Harms, S. E. Petersen, B. L. Schlaggar, M. Corbetta, M. F. Glasser, S. Curtiss, S. Dixit, C. Feldt, *et al.*, "Function in the human connectome: Task-fmri and individual differences in behavior," *Neuroimage*, vol. 80, pp. 169–189, 2013.

[102] D. Marcus, J. Harwell, T. Olsen, M. Hodge, M. Glasser, F. Prior, M. Jenkinson, T. Laumann, S. Curtiss, and D. Van Essen, "Informatics and data mining tools and strategies for the human connectome project," *Frontiers in neuroinformatics*, vol. 5, p. 4, 2011.

[103] J. D. Power, A. Mitra, T. O. Laumann, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen, "Methods to detect, characterize, and remove motion artifact in resting state fmri," *Neuroimage*, vol. 84, pp. 320–341, 2014.

[104] E. Amico, K. Abbas, D. A. Duong-Tran, U. Tipnis, M. Rajapandian, E. Chumin, M. Ventresca, J. Harezlak, and J. Goñi, "Towards a mathematical theory of communication for the human connectome," *arXiv preprint arXiv:1911.02601*, 2019.

[105] E. Amico, M. Dzemidzic, B. G. Oberlin, C. R. Carron, J. Harezlak, J. Goñi, and D. A. Kareken, "The disengaging brain: Dynamic transitions from cognitive engagement and alcoholism risk," *NeuroImage*, vol. 209, p. 116 515, 2020.

[106] R. F. Betzel, T. D. Satterthwaite, J. I. Gold, and D. S. Bassett, "Positive affect, surprise, and fatigue are correlates of network flexibility," *Scientific reports*, vol. 7, no. 1, p. 520, 2017.

[107] F. D. Malliaros and M. Vazirgiannis, "Clustering and community detection in directed networks: A survey," *Physics Reports*, vol. 533, no. 4, pp. 95–142, 2013.

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

[109] S. Fortunato, "Community detection in graphs," *Physics reports*, vol. 486, no. 3, pp. 75–174, 2010.

[110] S. Fortunato and D. Hric, "Community detection in networks: A user guide," *Physics Reports*, vol. 659, pp. 1–44, 2016.

[111] P. Pons and M. Latapy, "Computing communities in large networks using random walks," in *ISCIS*, vol. 3733, 2005, pp. 284–293.

[112] S. Hanhijärvi, G. C. Garriga, and K. Puolamäki, "Randomization techniques for graphs," in *Proceedings of the 2009 SIAM International Conference on Data Mining*, SIAM, 2009, pp. 780–791.

[113] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle, "The human brain is intrinsically organized into dynamic, anticorrelated functional networks," *Proceedings of the National Academy of Sciences*, vol. 102, no. 27, pp. 9673–9678, 2005.

[114] J. R. Gray, C. F. Chabris, and T. S. Braver, "Neural mechanisms of general fluid intelligence," *Nature neuroscience*, vol. 6, no. 3, p. 316, 2003.

[115] N. Tschentscher, D. Mitchell, and J. Duncan, "Fluid intelligence predicts novel rule implementation in a distributed frontoparietal control network," *Journal of Neuroscience*, pp. 2478–16, 2017.

[116] R. B. Cattell, "Theory of fluid and crystallized intelligence: A critical experiment.," *Journal of educational psychology*, vol. 54, no. 1, p. 1, 1963.

[117] C. Spearman, "" general intelligence," objectively determined and measured," *The American Journal of Psychology*, vol. 15, no. 2, pp. 201–292, 1904.

[118] J. Dubois, P. Galdi, L. K. Paul, and R. Adolphs, "A distributed brain network predicts general intelligence from resting-state human neuroimaging data," 2018.

[119] H. Mohr, U. Wolfensteller, R. F. Betzel, B. Mišić, O. Sporns, J. Richiardi, and H. Ruge, "Integration and segregation of large-scale brain networks during short-term task automatization," *Nature communications*, vol. 7, no. 1, pp. 1–12, 2016.

[120] E. Amico, A. Arenas, and J. Goñi, "Centralized and distributed cognitive task processing in the human connectome," *Network Neuroscience*, vol. 3, no. 2, pp. 455–474, 2019.

[121] M. A. Bertolero, B. T. Yeo, and M. Dâ Esposito, "The modular and integrative functional architecture of the human brain," *Proceedings of the National Academy of Sciences*, vol. 112, no. 49, E6798–E6807, 2015.

[122] D. Fraguas, C. M. Díaz-Caneja, L. Pina-Camacho, J. Janssen, and C. Arango, "Progressive brain changes in children and adolescents with early-onset psychosis: A meta-analysis of longitudinal mri studies," *Schizophrenia research*, vol. 173, no. 3, pp. 132–139, 2016.

[123] Y. Tian, D. S. Margulies, M. Breakspear, and A. Zalesky, "Hierarchical organization of the human subcortex unveiled with functional connectivity gradients," *bioRxiv*, 2020.

[124] A. Barnes, E. T. Bullmore, and J. Suckling, "Endogenous human brain dynamics recover slowly following cognitive effort," *PloS one*, vol. 4, no. 8, e6626, 2009.

[125] R. Tang, M. Ketcha, A. Badea, E. D. Calabrese, D. S. Margulies, J. T. Vogelstein, C. E. Priebe, and D. L. Sussman, "Connectome smoothing via low-rank approximations," *IEEE transactions on medical imaging*, vol. 38, no. 6, pp. 1446–1456, 2018.

[126] M. Mijalkov, J. B. Pereira, and G. Volpe, "Delayed correlations improve the reconstruction of the brain connectome," *PloS one*, vol. 15, no. 2, e0228334, 2020.

[127] L. Zhan, L. M. Jenkins, O. E. Wolfson, J. J. GadElkarim, K. Nocito, P. M. Thompson, O. A. Ajilore, M. K. Chung, and A. D. Leow, "The significance of negative correlations in brain connectivity," *Journal of Comparative Neurology*, vol. 525, no. 15, pp. 3251–3265, 2017.

[128] G. L. Colclough, M. J. Brookes, S. M. Smith, and M. W. Woolrich, "A symmetric multivariate leakage correction for meg connectomes," *Neuroimage*, vol. 117, pp. 439–448, 2015.

[129] M. Thilaga, R. Vijayalakshmi, R. Nadarajan, D. Nandagopal, B. Cocks, C. Archana, and N. Dahal, "A heuristic branch-and-bound based thresholding algorithm for unveiling cognitive activity from eeg data," *Neurocomputing*, vol. 170, pp. 32–46, 2015.

[130] T. Alakörkkö, H. Saarimäki, E. Glerean, J. Saramäki, and O. Korhonen, "Effects of spatial smoothing on functional brain networks," *European Journal of Neuroscience*, vol. 46, no. 9, pp. 2471–2480, 2017.

[131] J. A. Roberts, A. Perry, A. R. Lord, G. Roberts, P. B. Mitchell, R. E. Smith, F. Calamante, and M. Breakspear, "The contribution of geometry to the human connectome," *Neuroimage*, vol. 124, pp. 379–393, 2016.

[132] F. Váša, E. T. Bullmore, and A. X. Patel, "Probabilistic thresholding of functional connectomes: Application to schizophrenia," *Neuroimage*, vol. 172, pp. 326–340, 2018.

[133] Z. Yu, J. Qin, X. Xiong, F. Xu, J. Wang, F. Hou, and A. Yang, "Abnormal topology of brain functional networks in unipolar depression and bipolar disorder using optimal graph thresholding," *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 96, p. 109 758, 2020.

[134] N. Langer, A. Pedroni, and L. Jäncke, "The problem of thresholding in small-world network analysis," *PloS one*, vol. 8, no. 1, e53199, 2013.

[135] M. P. van den Heuvel, S. C. de Lange, A. Zalesky, C. Seguin, B. T. Yeo, and R. Schmidt, "Proportional thresholding in resting-state fmri functional connectivity networks and consequences for patient-control connectome studies: Issues and recommendations," *Neuroimage*, vol. 152, pp. 437–449, 2017.

[136] F. Z. Esfahlani and H. Sayama, "A percolation-based thresholding method with applications in functional connectivity analysis," in *International Workshop on Complex Networks*, Springer, 2018, pp. 221–231.

[137] N. Z. Bielczyk, F. Walocha, P. W. Ebel, K. V. Haak, A. Llera, J. K. Buitelaar, J. C. Glennon, and C. F. Beckmann, "Thresholding functional connectomes by means of mixture modeling," *NeuroImage*, vol. 171, pp. 402–414, 2018.

[138] H. Lee, M. K. Chung, H. Kang, B.-N. Kim, and D. S. Lee, "Discriminative persistent homology of brain networks," in *2011 IEEE international symposium on biomedical imaging: from nano to macro*, IEEE, 2011, pp. 841–844.

[139] H. Lee, H. Kang, M. K. Chung, B.-N. Kim, and D. S. Lee, "Persistent brain network homology from the perspective of dendrogram," *IEEE transactions on medical imaging*, vol. 31, no. 12, pp. 2267–2277, 2012.

[140] S. I. Dimitriadis, C. Salis, I. Tarnanas, and D. E. Linden, "Topological filtering of dynamic functional brain networks unfolds informative chronnectomics: A novel data-driven thresholding scheme based on orthogonal minimal spanning trees (omsts)," *Frontiers in neuroinformatics*, vol. 11, p. 28, 2017.

[141] T. Gorbach, A. Lundquist, X. de Luna, L. Nyberg, and A. Salami, "A hierarchical bayesian mixture modeling approach for analysis of resting-state functional brain connectivity: An alternative to thresholding," *Brain Connectivity*, no. ja, 2020.

[142] D. Meunier, R. Lambiotte, and E. T. Bullmore, "Modular and hierarchically modular organization of brain networks," *Frontiers in neuroscience*, vol. 4, p. 200, 2010.

[143] C. J. Stam, A. Hillebrand, H. Wang, and P. Van Mieghem, "Emergence of modular structure in a large-scale brain network with interactions between dynamics and connectivity," *Frontiers in computational neuroscience*, vol. 4, p. 133, 2010.

[144] D. Duong-Tran, E. Amico, B. Corominas-Murtra, K. Abbas, M. Dzemidzic, D. Kareken, M. Ventresca, and J. Goñi, "A morphospace framework to assess configural breadth based on brain functional networks," *arXiv preprint arXiv:1901.10962*, 2019.

[145] D. Meunier, S. Achard, A. Morcom, and E. Bullmore, "Age-related changes in modular organization of human brain functional networks," *Neuroimage*, vol. 44, no. 3, pp. 715–723, 2009.

[146] R. F. Betzel, L. Byrge, Y. He, J. Goñi, X.-N. Zuo, and O. Sporns, "Changes in structural and functional connectivity among resting-state networks across the human lifespan," *Neuroimage*, vol. 102, pp. 345–357, 2014.

[147] A. Alexander-Bloch, R. Lambiotte, B. Roberts, J. Giedd, N. Gogtay, and E. Bullmore, "The discovery of population differences in network community structure: New methods and applications to brain functional networks in schizophrenia," *Neuroimage*, vol. 59, no. 4, pp. 3889–3900, 2012.

[148] M. H. Lee, C. D. Hacker, A. Z. Snyder, M. Corbetta, D. Zhang, E. C. Leuthardt, and J. S. Shimony, "Clustering of resting state networks," *PloS one*, vol. 7, no. 7, e40370, 2012.

[149] M. Y. Chan, F. H. Alhazmi, D. C. Park, N. K. Savalia, and G. S. Wig, "Resting-state network topology differentiates task signals across the adult life span," *Journal of Neuroscience*, vol. 37, no. 10, pp. 2734–2745, 2017.

[150] A. E. Reineberg, J. R. Andrews-Hanna, B. E. Depue, N. P. Friedman, and M. T. Banich, "Resting-state networks predict individual differences in common and specific aspects of executive function," *Neuroimage*, vol. 104, pp. 69–78, 2015.

[151] R. F. Betzel, J. D. Medaglia, and D. S. Bassett, "Diversity of meso-scale architecture in human and non-human connectomes," *Nature Communications*, vol. 9, no. 1, p. 346, 2018.

[152] J. Faskowitz, X. Yan, X.-N. Zuo, and O. Sporns, "Weighted stochastic block models of the human connectome across the life span," *Scientific reports*, vol. 8, no. 1, pp. 1–16, 2018.

[153] M. E. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical review E*, vol. 69, no. 2, p. 026 113, 2004.

[154] B. Karrer and M. E. Newman, "Stochastic blockmodels and community structure in networks," *Physical review E*, vol. 83, no. 1, p. 016 107, 2011.

[155] M. E. Newman, "The structure and function of complex networks," *SIAM review*, vol. 45, no. 2, pp. 167–256, 2003.

[156] M. Newman, *Networks: an introduction*. Oxford university press, 2010.

[157] I. Jolliffe, "Principal component analysis," in *International encyclopedia of statistical science*, Springer, 2011, pp. 1094–1096.

[158] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1-3, pp. 37–52, 1987.

[159] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguná, "Hyperbolic geometry of complex networks," *Physical Review E*, vol. 82, no. 3, p. 036 106, 2010.

[160] D. Krioukov, F. Papadopoulos, A. Vahdat, and M. Boguñá, "Curvature and temperature of complex networks," *Physical Review E*, vol. 80, no. 3, p. 035 101, 2009.

[161] F. Papadopoulos, D. Krioukov, M. Boguñá, and A. Vahdat, "Greedy forwarding in dynamic scale-free networks embedded in hyperbolic metric spaces," in *INFOCOM, 2010 Proceedings IEEE*, IEEE, 2010, pp. 1–9.

[162] E. Candellero and N. Fountoulakis, "Clustering and the hyperbolic geometry of complex networks," in *International Workshop on Algorithms and Models for the Web-Graph*, Springer, 2014, pp. 1–12.

[163] J. D. Wilson, N. T. Stevens, and W. H. Woodall, "Modeling and detecting change in temporal networks via the degree corrected stochastic block model," *Quality and Reliability Engineering International*, vol. 35, no. 5, pp. 1363–1378, 2019.

[164] T. K. Dey, J. Sun, and Y. Wang, "Approximating loops in a shortest homology basis from point data," in *Proceedings of the twenty-sixth annual symposium on Computational geometry*, 2010, pp. 166–175.

[165] C. D. Meyer, *Matrix analysis and applied linear algebra*. Siam, 2000, vol. 2.

[166] G. M. Ziegler, *Lectures on polytopes*. Springer Science & Business Media, 2012, vol. 152.

[167] T. G. C. H. Page. (1995). Qhull the geometry center home page, [Online]. Available: http://qhull@qhull.org.

[168] M. E. Dyer and A. M. Frieze, "On the complexity of computing the volume of a polyhedron," *SIAM Journal on Computing*, vol. 17, no. 5, pp. 967–974, 1988.

[169] L. Khachiyan, "Chapter iv. complexity of polytope volume computation," *New trends in discrete and computational geometry*, vol. 10, p. 91, 1993.

[170] M. Girvan and M. E. Newman, "Community structure in social and biological networks," *Proceedings of the national academy of sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.

[171] T. Coolen, A. Annibale, and E. Roberts, *Generating random networks and graphs*. Oxford university press, 2017.

[172] A. Frieze and M. Karoński, *Introduction to random graphs*. Cambridge University Press, 2016.

[173] P. Erdos and A. Rényi, "On random graphs i," *Publ. Math. Inst. Hung. Acad. Sci*, 1959.

[174] F. R. Chung and L. Lu, *Complex graphs and networks*, 107. American Mathematical Soc., 2006.

[175] P. Erdos and A. Rényi, "On the evolution of random graphs," *Publ. Math. Inst. Hung. Acad. Sci*, vol. 5, no. 1, pp. 17–60, 1960.

[176] A. Frieze and M. Karoński, *Introduction to random graphs*. Cambridge University Press, 2015.

[177] P. W. Holland and S. Leinhardt, "An Exponential Family of Probability Distributions for Directed Graphs," *Journal of the American Statistical Association*, vol. 76, no. 373, pp. 33–50, 1981, ISSN: 0162-1459. DOI: 10.2307/2287037.

[178] S. Wasserman and K. Faust, *Social network analysis: Methods and applications*. Cambridge university press, 1994, vol. 8.

[179] M. E. Newman, "Fast algorithm for detecting community structure in networks," *Physical review E*, vol. 69, no. 6, p. 066 133, 2004.

[180] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of statistical mechanics: theory and experiment*, vol. 2008, no. 10, P10008, 2008.

[181] M. Rosvall, A. Trusina, P. Minnhagen, and K. Sneppen, "Networks and cities: An information perspective," *Physical Review Letters*, vol. 94, no. 2, p. 028 701, 2005.

[182] J.-C. Delvenne, S. N. Yaliraki, and M. Barahona, "Stability of graph communities across time scales," *Proceedings of the National Academy of Sciences*, 2010.

[183] J.-C. Delvenne, M. T. Schaub, S. N. Yaliraki, and M. Barahona, "The stability of a graph partition: A dynamics-based framework for community detection," in *Dynamics On and Of Complex Networks, Volume 2*, Springer, 2013, pp. 221–242.

[184] T. P. Peixoto, "Nonparametric weighted stochastic block models," *Physical Review E*, vol. 97, no. 1, p. 012 306, 2018.

[185] N. X. Vinh, J. Epps, and J. Bailey, "Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance," *The Journal of Machine Learning Research*, vol. 11, pp. 2837–2854, 2010.

VITA

Duy Duong-Tran is currently a Ph.D. Candidate in the School of Industrial Engineering (IE), Purdue University. His current research area is at the crossroads between applied mathematics, data science and computational neuroscience. Duong-Tran proudly joined Purdue IE because of the school's unmatched vision in approaching and solving newly emerging complex challenges in the new era. During his time at Purdue University, he had received a Ross Research Fellowship in 2017, a Magoon Teaching Excellence Award in 2019, and recently, a Dissertation Fellowship in 2021 to support his final doctoral year.

Before coming to Purdue University, he earned Bachelor's degrees in Mathematics and Industrial/Entrepreneurial Engineering, and a Master's degree in Industrial Engineering, at Western Michigan University. After obtaining his Bachelor's degrees, he worked as a planning engineer and an internal Sales/Program Manager in automobile industry. After joining Purdue, he had opportunities to work in the Data Science Summer Institute (jointly with the Machine Learning Group in 2019 and the Computing Engineering Division in 2020) at Lawrence Livermore National Laboratory, located in Livermore, California, U.S.

APPENDIX: BASIC MATHEMATICS NOTIONS

Preface. In this chapter, some of the relevant, fundamental building-blocks, spanning different fields from theory to applied are introduced. It is worth to note that most (if not all) of these fields are well-established and that only a small portion of relevant knowledge to this dissertation is referenced and cited.

.A Linear Algebra

Matrices plays an ubiquitous role in practically all aspects of science, including statistics, computer science, and applied mathematics. Often, a discrete matrix describes, quantitatively, the characteristics of m objects by n features, by an $m \times n$ matrix. In this chapter, this dissertation provides a comprehensive treatment of necessary theory in Linear Algebra that are relevant to the current and future stage of computational neuroscience. The main source of reference for fundamental Linear algebra is from the book by Carl Meyer [165].

Set theory Basics

Notations

Capital letters such as A, B are used to denote sets, while lowercase letter, i.e. a, b, are denoted as elements or objects belonging those sets, e.g. $a \in A$. Likewise, we denote not-belonging notion as $b \notin B$. Furthermore, given any set S, its cardinality is denoted as $|S|$.

We define set B to be the subset of set A if and only if all elements of B are also elements of set A, i.e. BA (also known as proper inclusion relation). Otherwise the relaxed inclusion relation is $B \subset A$.

Two sets can actually be identical, which we denote as $A = B$. Further, if $B \subset A$ and B is not identical as A then we say that B is the **proper** subset of set A. There are two conventions to express a set elements: exhaustive list, e.g. $A = \{a, b, c\}$ or imposing a rule for a generic element to belong in a set, e.g. $A = \{x \mid x = 2l\}$ where $l \in N^+$ (where N^+ is non-negative integers). It is worth noting that a set can be understood as a logic statement.

For example, $A = \{x \mid x = 2l\}$ where $l \in N^+$ can be expressed in a logical statement as *set A is a set containing all positive even numbers.*

Set Operations

The union of two sets, say A and B, are a set that contains element belonging to either set A or B:

$$A \cup B = \{x \mid x \in A \quad or \quad x \in B\}$$

On the other hand, the intersection of two sets, say A and B, is denoted as $A \cap B$. If such intersection is empty then $A \cap B = \emptyset$; in such case, we say that those two sets are disjoint.

The difference of two sets, i.e. $A - B$, are the set comprising elements that belong to A and not belong to B. Formally,

$$A - B = \{x \mid x \in A, x \notin B\}$$

Axiomatic Set Theory

Basic set operations carry over from elementary algebra such as the rule of associative and distribution. For example, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$. The most important rule/axioms on set are *De Morgan's laws* which basically boils down to two important rules:

1. The complement of unions is the intersections of complements;
2. The complement of intersection is the union of complements.

Sets/Collections of set

Another important concept in set theory that will be useful for topology construction is the collection of set notion. Specifically, given any generic set, one can define a set whose elements are subsets (proper or not) of the original set. Formally, given a set A, all subsets of A live in a space $\mathcal{P}$. For example, if $A = \{1, 2, 3\}$ then the following statements follows:

- $a \in A$;

- $\{a\} \in \mathcal{P}(A)$

Note that $a \in \mathcal{P}(A)$ is not a correct statement. Note that this allows us to formally define the notion of arbitrary intersection and union as $\cap_{A \in \mathcal{P}(A)}$ and $\cup_{A \in \mathcal{P}(A)}$.

Cartesian Products

One of the critical way to form a new set, from old ones, is Cartesian products. It is an important mechanics, in elementary geometry, to define, for example, points in a $k-$ dimensional plane.

Given a set A and B, the Cartesian product formed by the two sets are the set of all ordered pairs in which the first element of the pair is an element belonging to A (likewise, the second element in the pair is an element belonging to B). Formally,

$$A \times B = \{(a, b) \mid a \in A; b \in B\}$$

Computational Basics

In this section, we establish some of the key mathematical notations and linear algebraic theory used throughout this dissertation. Specifically, scalar is italicized, a. A vector is denoted as bold letter, $\mathbf{a}$ (default form is in column format). A matrix is notated as capitalized letter, A. If $r \in [q]$ where $q \in N^+$, it means that r takes on integer values from 1 up to, including, q. Given any two vectors $\mathbf{a}, \mathbf{b} \in R^n$, $\langle \mathbf{a}, \mathbf{b} \rangle$ denote inner product.

A generic matrix $\mathbf{A}$ over the field F (typically over real numbers $F = R$), represented by a continuous interval $[x, y]$, z_1 rows and z_2 columns is denoted as $\mathbf{A} \in [x, y]^{z_1 \times z_2}$. The notation $\mathbf{A}_{\mathrm{i}*}$ and $\mathbf{A}_{*\mathrm{j}}$ are used to denote the i^{th} row and j^{th} column of matrix $\mathbf{A}$, respectively. Further, if we want to induce a sub-matrix from the original matrix $\mathbf{A}$ based on a specific collection of rows, denoted as S_{rows}, and columns, denoted as $S_{columns}$, we use the notation: $\mathbf{A}(S_{rows}, S_{columns})$. If both rows and columns are matched (both denoted as S_w), then we will ease notation by using $\mathbf{A}(S_{rows})$.

Matrix Inverse. A matrix $\mathbf{A} \in R^{n \times n}$ is nonsingular (or invertible) if and only if there exists a matrix $\mathbf{A}^{-1} \in R^{n \times n}$ such that

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_{n \times n} = \mathbf{A}^{-1}\mathbf{A}$$

Note that in this case, matrix $\mathbf{A}$ is said to have both left and right inverse. Note that this also implies that all the columns (and all the rows) of $\mathbf{A}$ must be linearly independent (full rank). Equivalently, if there exists a vector $\mathbf{x}$ such that $\mathbf{A}\mathbf{x} = \mathbf{0}_n$ then $\mathbf{x} = \mathbf{0}$.

Inverse and Transpose Operations. Typically, inverse and transpose operator is interchangeable, i.e. $(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}$. Nonetheless, $(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

The building blocks of Spectral Decomposition

Linear Function

Note that although function is defined in the context of linear algebra; in later sections, an extended (yet, shortened) version of function will also be briefly define. In the context of linear algebra, a function is simply a mappings between a source space, say $\mathbf{V}$, to a target space, say $\mathbf{U}$ where both of them are vector space (over real field):

$$\mathbf{V} \xrightarrow{f} \mathbf{U}$$

A linear function is sometimes referred to as linear map, linear morphism, vector space morphism. A linear map has to preserve vector space operations: vector addition and scalar multiplication. Formally,

$$\mathbf{f}(\mathbf{u} + \mathbf{v}) = \mathbf{f}(\mathbf{u}) + \mathbf{f}(\mathbf{v})$$

and

$$\mathbf{f}(c\mathbf{u}) = c\mathbf{f}(\mathbf{u})$$

where c is scalar and $\mathbf{u}, \mathbf{v} \in \mathrm{i}m^{-1}(f)$ and $\mathbf{f}(\mathbf{u}), \mathbf{f}(\mathbf{v}) \in \mathrm{i}m(f)$. A map is isomorphic if it is bijective (onto and one-to-one). A map is endomorphic if $\mathbf{V} = \mathbf{U}$. Furthermore, If $\mathbf{U} = \mathbf{V}$

and f is bijection, then f is isomorphic. Often, a linear map is represented by a matrix (in finite dimension vector spaces). Some of the useful cases are rotation, reflection matrices.

Basis and Linear Independence

Given any vector space with finite dimension n over the field of real numbers, a basis is a set with minimum number of vectors in R^n, necessary and sufficient, to *describe* any generic vector in such space. Formally, given a set of basis vectors $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, ...\mathbf{b}_m\}$

$$\forall \mathbf{x} \in R^m : \mathbf{x} = \sum_{\mathrm{i}=1}^{m} \mathbf{b}_{\mathrm{i}}$$

Note that $\mathcal{B}$ is not unique. The most desirable basis is a set of linear independent vectors that are also mutually orthogonal.

Four Fundamental Subspaces

Given any matrix $\mathbf{A} \in R^{m \times n}$, the four fundamental subspaces of such matrix are:

1. row space of $\mathbf{A}$, the set of all vectors $\mathbf{x}$ such that $\mathbf{A}\mathbf{x} \neq \mathbf{0}$;
2. null space of $\mathbf{A}$, the set of all vectors $\mathbf{x}$ such that $\mathbf{A}\mathbf{x} = \mathbf{0}$;
3. column space of $\mathbf{A}$, the set of all vectors $\mathbf{y}$ such that $\mathbf{A}^T\mathbf{y} \neq \mathbf{0}$;
4. Left-null space of $\mathbf{A}$, the set of all vectors $\mathbf{y}$ such that $\mathbf{A}^T\mathbf{x} = \mathbf{0}$;

The dimension of a subspace is the number of linearly independent vectors required to span that subspace. In other words, let f be a linear map from R^n to R^m, represented by an $m \times n$ matrix $\mathbf{A}$ with rank $r \leq \min(m, n)$ then the followings follows:

- r is the dimension of the column space, i.e. the number of linearly independent vectors in $\mathbf{R}^m$ needed to describe the image of the linear operator, i.e. $im(f)$;
- $n - r$ is the dimension of null space (also known as the kernel of the linear operator $ker(f)$;

- $m - r$ is the dimension of left-null space (also known as the co-kernel of the linear operator $coker(f)$;

Theorem .A.1. *Let $f : V \to W$ be a linear operator, then: $Rank(f) + Null(f) = dim(V)$ where $Rank(f) = dim(\text{im}(f))$ and $Null(f) = dim(ker(f))$.*

Orthogonality and Orthogonal Matrix

For any generic vectors $\mathbf{a}, \mathbf{b} \in R^m$, they are said to be orthogonal if and only if

$$\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{0}_m$$

Definition .A.1. *A matrix $\mathbf{A}$ is orthogonal if $\mathbf{A}^T = \mathbf{A}^{-1}$.*

It follows that $\forall \text{i}, \text{j} \in [n]$,

$$\mathbf{A}^T_{\text{i}*}\mathbf{A}_{*\text{j}} = \begin{cases} 1, \text{i} \neq \text{j} \\ 0, \text{i} = \text{j} \end{cases}$$

Note that the same logic applied to the rows of $\mathbf{A}$.

Singular/Eigen Value Decomposition

At the center stage of high-dimensionality data processing techniques are Singular/Eigen Value decomposition. These techniques supports the unravelling of independent/orthogonal modes of complex longitudinal data. We first define the Singular Value decomposition as follows:

Definition .A.2. *For each $\mathbf{A} \in R^{m \times n}$ of rank r, there are orthogonal matrices $\mathbf{U} \in R^{m \times m}$ and $\mathbf{V} \in R^{n \times n}$ and a diagonal matrix $\mathbf{D}_{r \times r} = diag(\sigma_1, \sigma_2, ..., \sigma_r)$ such that:*

$$\mathbf{A} = \mathbf{U} \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}^T = \mathbf{U}\Sigma\mathbf{V}^T$$

with non-increasing order of $\sigma_\text{i}, \forall \text{i} \in [r]$.

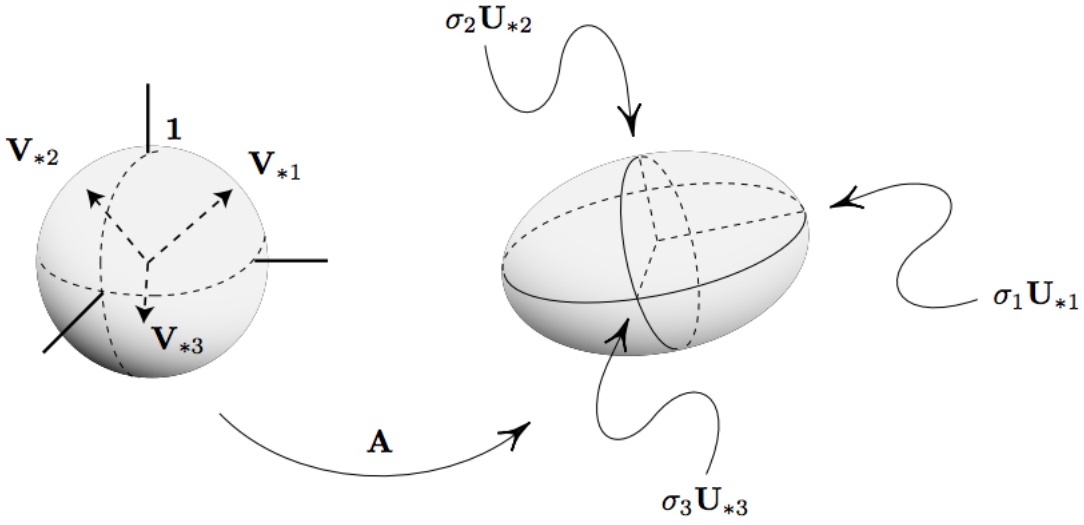


Figure S1. A geometrical illustration of SVD transformation in 3 dimensional geometry. Figure is reproduced at the courtesy of Carl Meyer's textbook on Linear Algebra [165].

In this case, $\mathbf{U}, \mathbf{V}$ are the column and row space of $\mathbf{A}$, respectively. The rank of the matrix is the minimum between the dimension of column and row space. By definition,

- the sum between the dimension of row space and null space equals m (total number of rows in $\mathbf{A}$);
- the sum between the dimension of column space and left-null space equals n (total number of columns in $\mathbf{A}$);

In a majority of applications, matrices are symmetric which can be written as a product of two matrices:

$$\begin{aligned}\mathbf{A} &= \mathbf{B}\mathbf{B}^T = \mathbf{U}\Sigma\mathbf{V}^T(\mathbf{U}\Sigma\mathbf{V}^T)^T \\ &= \mathbf{U}\Sigma\mathbf{V}^T\mathbf{V}\Sigma^T\mathbf{U}^T \\ &= \mathbf{U}\Sigma^2\mathbf{U}^T \\ &= \mathbf{U}\Lambda\mathbf{U}^T\end{aligned}$$

This is call the eigen decomposition of symmetric matrices (also known as semi-definite matrices). Because of the square nature of singular values of $\mathbf{B}$, symmetric matrices are guaranteed to have real, non-negative eigen values. Note that full-rank symmetric matrices have strictly positive eigen values by construction.

.B Polytope Theory

Polytope theory is a branch of mathematics that studies the geometry of shapes in a $d-$dimensional Euclidean space, R^d. Given a set of points $W = \left\{x_1, x_2, ...x_{|W|}\right\}$ for which $x_{\mathrm{j}} \in R^d, \forall \mathrm{j} \in [|W|]$, a convex hull formed by such set of points are mathematically represented by

$$\mathbf{Conv}(W) = \left\{ \sum_{\mathrm{j}=1}^{|W|} \alpha_{\mathrm{j}} x_{\mathrm{j}} \mid \sum_{\mathrm{j}=1}^{|W|} \alpha_{\mathrm{j}} = 1, \alpha_{\mathrm{j}} \geq 0, \forall \mathrm{j} \in [|W|] \right\}$$

where d is called the ambient space dimension. Moreover, if $|W| \geq d+1$, we recall that points in W are in general position if no hyperplane, i.e. flat of dimension $d-1$ contains more than d points, [166]. Otherwise, i.e. $|W| \leq d$, there exist(s) point(s) that are affinely dependent on other points in W.

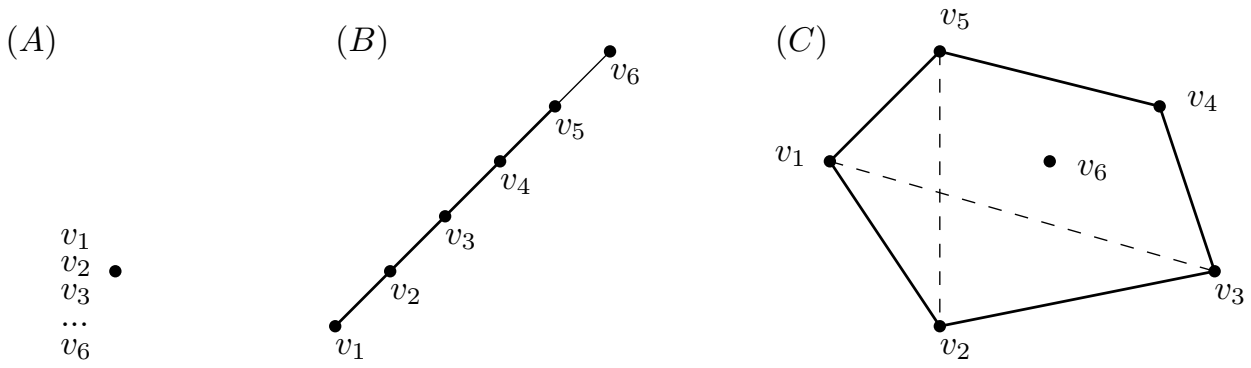


Figure S2. Convex Hull Demonstrations for a 2-dimensional space over the real field R^2.

In the above figure 3, given that $W = \{v_1, v_2, ..., v_6\}$, we demonstrate three possible scenarios of convex hull formed by W in morphospace Ω. Case *(A),(B),(C)* correspond to the polytope dimension of $h = 0, 1, 2$, respectively. Here we see that $\{v_1, v_6\}$ and $\{v_1, v_2, v_3, v_4, v_5\}$ forms the Pareto front in Case (B) Case (C), respectively. In case (C), v_6 belongs to the interior of the hull. Further, in case (B) and (C), we see that the hull vertices, i.e. points

belong to the Pareto front of the hull, are $\{v_1, v_5\}$ for case (B) and $\{v_1, v_2, v_3, v_4, v_5\}$ for case (C). Given the nature of this space, the first two scenarios are statistically rare. In the third scenario, we see that all 5 points constitute the boundary of $\mathbf{conv}(W)$. Further, we see that some type A pairs of points, graphically represented by solid lines, are $(v_1, v_5), (v_2, v_3)$ while some type B pairs, represented by dashed lines, are $(v_2, v_4), (v_3, v_5)$.

Providing that points in W in R^d, the approximated volume induced by the convex hull $\mathbf{Conv}(W)$ can be calculated through the formation of Delaunay Triangulation process [166]. The volume of the convex hull is denoted as $\mathbf{Vol}(\mathbf{Conv}(W))$. In R^d, the convex hull dimension can take on the values

1. $h = 0$ which constitutes a point in R^d, $\mathbf{Vol}(\mathbf{Conv}(W)) = 0$
2. $h = 1$ which constitutes a line segment, $\mathbf{Vol}(\mathbf{Conv}(W)) = \sup(d(x_{\mathrm{i}}, x_{\mathrm{j}})), \forall x_{\mathrm{i}}, x_{\mathrm{j}} \in W$ where $d(x_{\mathrm{i}}, x_{\mathrm{j}})$ denotes the pre-defined metric distance between two generic points.
3. $h = 2$ and $h \geq 3$ which constitutes the notion of area and volume, respectively.

For $h \geq 2$, convex hull volume is calculated using Qhull package implemented in Matlab, see [167]. In general, as pointed out also in [167], computing $\mathcal{V}-$ or $\mathcal{H}-$ polytope metric volume is NP-hard (see also [168], [169]) with the availability of efficient approximating algorithms.

.C Probability theory and stochastic processes

Some remarks on Probability Theory

Let $\mathcal{X}$ be a (finite) sample space and $x \in \mathcal{X}$ be elements in such space. We can X to be a random variable (truly, it can also be thought as function), if X assign probability measure in $[0, 1]$ to a specific value $x \in \mathcal{X}$. An event defined on $\mathcal{X}$ takes on the probabilistic measures as defined by the fundamental axioms in probability theory. From the sample space perspective, an event is any possible subset of $\mathcal{X}$. By definition,

$$P(A = \{\mathcal{X}\}) = 1$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Also, if the sample space can be partitioned into non-overlap, independent events A_{i} then the law of total probability is also obtained as follows:

$$P(\cup_{\mathrm{i}} A_{\mathrm{i}}) = P(\{X\}) = 1$$

Moreover, $P(A \cap B) = 0$ does not guarantee that the two events are mutually exclusive. Two events are mutually exclusive are such that the existence of one guarantees the non-existent of the other one. Coincidentally, $P(A \cap B) = 0$ in this case also.

Joint events and probability. Often, different events governed by different probability distributions can happen simultaneously. In such case, one can construct the contingency table which records the likelihood of any particular realized combinations and their corresponding joint probability $P(A_{\mathrm{i}} \cap B_{\mathrm{j}} \cap C_k ...)$. For instance, two random variables (representing two probability distribution) and their interactions can be recorded in a 2-dimensional contingency table.

Independent events Two events are independent if and only if the probability of knowing an event does not increase the odds of the other event to take place:

$$P(A \mid B) = P(A)$$

As a corollary, the joint probability of independent events can be computed as

$$P(A \cap B) = P(A \mid B) \times P(B) = P(A)P(B)$$

Bayesian statistics. One of the most significant corollary of conditional probability between, say 2 events, are the fact that the posterior conditional probability of one event (say A dependent on B) can be computed using **a priori** conditional probability (B dependent on A) as follows:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

In most applications of Bayes' approach, typically one can argue that $P(A)$ and $P(B)$ are constant (model independent). Hence,

$$P(A \mid B) \sim P(B \mid A)$$

This is of monumental importance in network modeling and inference as described in later sections.

Markov Chains

Assume we have a set of states $S = \{s_1, s_2, ..., s_r\}$ of r states and that at each step, the likelihood of moving from one state to another is expressed through the *transition probability*. In finite domain, this dynamic can be described in a matrix, called transition probability matrix $\mathbf{P}$. It is trivial to see that the n^{th} step probability has close-form formula $\mathbf{P}^n$. For example, a two-step probability of going from state i to state j can be computed as follows:

$$p_{\mathrm{ij}}^{(2)} = \sum_k p_{\mathrm{ik}} p_{k\mathrm{j}}$$

Some of the key characteristics of this matrix involves the row-sum adds up to 1 (given the current state of the process, the process can only visit states in S), e.g. $\sum_{\mathrm{j}} p_{\mathrm{ij}} = 1$.

Absorbing Markov Chain. One of the most important classes of Markov Chain with many real-world applications are Absorbing/Terminating Markov Chain. In this case, there are two types of states: transient and absorbing. Absorbing states has self-probability of 1 (guaranteed to revising itself in the next step once the process is currently in such state). The canonical form of transition probability matrix is given by:

$$\mathbf{P} = \begin{pmatrix} \mathbf{Q} & \mathbf{R} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$$

where $\mathbf{Q}$ captures transition probability between r_1 transient states; $\mathbf{I}$ is the identity matrix of size r_2; and $\mathbf{R}$ represents the probability going from a transient state to an absorbing state. By construction,

$$|S| = r = |S_{trans}| + |S_{abs}| = r_1 + r_2$$

where S_{trans} and S_{abs} are set containing transient and absorbing states, respectively.

Non-terminating Markov Chain is the opposite of absorbing chain for which no terminating state exists. In canonical form, $\mathbf{P} = \mathbf{Q}$.

Steady state distribution. A natural quantity to consider is in the long-term, what would be the percentage of time a random particle (walker) is found to be in any state in S. For non-terminating chain, the matrix rank is always one less than full; hence, there is only one single eigen pair (value/vector) exists:

$$\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}$$

For terminating chain, on the other hand, the process is guaranteed to be in absorbing states in the long run.

Time-to-absorption For terminating chain, $\mathbf{Q}$ is invertible (its inverse exists) and the time to absorption can be computed using the row sum of the fundamental matrix $\mathbf{Z} = \sum_{\mathrm{i}=1}^{\infty} \mathbf{Q}^{\mathrm{i}} = (\mathbf{I} - \mathbf{Q})^{-1}$ as follows:

$$\tau = \mathbf{Z}\vec{\mathbf{1}}$$

where $\vec{\mathbf{1}}$ is vector of all ones with the same dimension as the fundamental matrix. Moreover, $\mathbf{N} = [N_{\mathrm{ij}}]$ represents the number of times the random walker spend in state j, given that it starts in state i, for both i, j are transient states.

.D Generative models of Networks

Practically in most (if not all) scientific disciplines in the modern world, some generic form of system exists. Systems can be thought as a collection of elements and their interactions. From a mathematical standpoint, in order to gain meaningful insights to those systems, one

needs to create some scaffolding representation to represent the system elements and their corresponding interactions.

A graph is denoted as $G(V, E)$ where V and E are sets of vertices and edges in such network, respectively. $G(V, E)$ can be represented by $\mathbf{A}_G = \mathbf{A}(\mathrm{ij}) = [w_{\mathrm{ij}}]$, in which $w_{\mathrm{ij}} \in [0, 1]$ represents coupling strength between node i and j. The strength of node $\mathrm{i} \in V(G)$ is denoted as k_{i}, typically stored in the diagonal matrix $\mathbf{K}$ for $\mathbf{K}(\mathrm{ii}) = k_{\mathrm{i}}$.

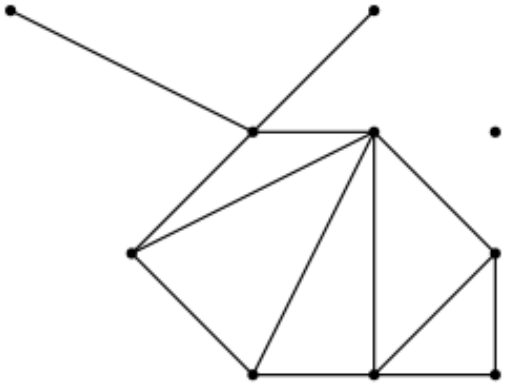


Figure S3. A toy example for a graph with 10 nodes and 13 edges.

Euler is the pioneer in thinking about solving a problem using a set of node and edges in the famous problem of Königsberg bridges in eighteenth century. In this problem, one is supposed to find a path that visits all island such that all bridges are only allowed to cross once, see figure S4.

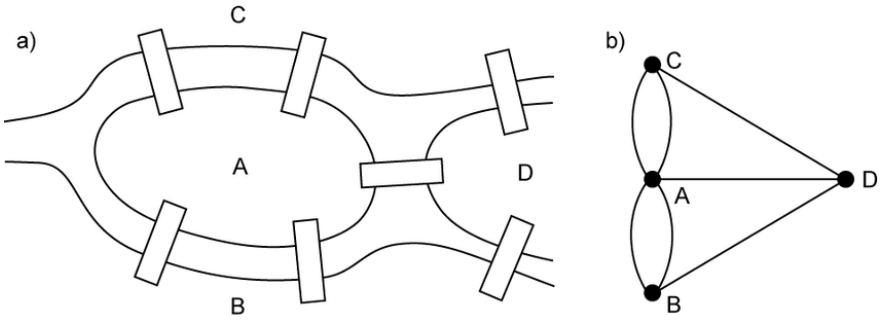


Figure S4. The first problem that was approached using graph theory is the Königsberg bridges. Panel (A) represents the pictorial demonstration of the bridges connecting neighboring islands; panel (B) shows the corresponding graph-theoretical representation of the problem.

The birth of social networks among other networked systems in biology and finances elucidated the necessity to understand these complex systems' behaviors. This gave rise to an research area called network science in which systems are referred to as networks. Network science can be thought as a derivative field emerging out of graph theory. Both fields bears obvious similarity and remarkably differences. Specifically, graph-theoretical research focuses on proving graph properties rigorously while network-science research contributions elucidate showing network empirical properties, using - for instance - simulations. By construction, the field of network science is much more interdisciplinary than graph-theory.

Network exists in every corner of lives, as we know it. In fact, network can be found from the smallest viable biological scale such as protein-protein interaction networks, to the largest scale such as the interaction of starts in the galaxy. If one can define two things: what are the elements of the system, and how those elements interact among each other, one has formalize a network. Here, we list out some of the networks that one typically exposes to, spanning different disciplines:

- **The Web**. One can also refer this as the Internet, which is a network in which a web-page (with *html* address) is a node and their reference to other web-pages form an edge between them;

- **The Social Networks**. The birth of social media has give rise to one of the most popular network: the social networks. These graphs can simply understood as a web of users (nodes) and their pairwise connection (whether they are connected as friends). Those networks can be either binary or weighted depending on the particular usage.

- **Human Brain network**. Simply put, these networks can be effectively divided into two types: structural and functional. Structural brain networks elucidates the hard-wires among all given pairs of brain regions of interest (ROIs). Functional brain networks represents ROIs' functional coupling strength. A more detailed treatment of functional brain networks is available in the subsequent section of this dissertation.

- **Financial network**. The emergent of time-series financial instruments such as stock prices has opened new window of opportunities for financial network modellers to

investigate this complex dynamics under network perspective. A simple example of weighted network would be constructing a network of stock picks where edges are computed by Pearson correlations between two stocks' time-series over a fixed window of period.

The aforementioned networks are just some primal examples of networked systems across multiple disciplines. Networks simply exist across every scale, every discipline as long as its elements and the corresponding interactions can be quantified.

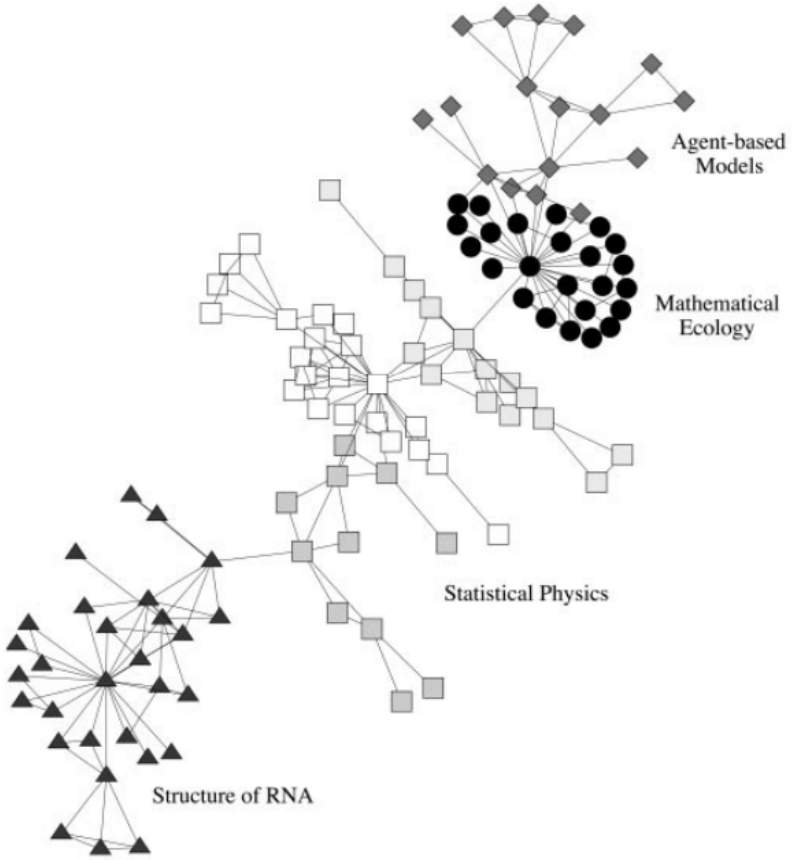


Figure S5. An example of a network of Santa Fe scientific collaboration network with identified communities from detection algorithm. Figure is adapted from [170].

Random graphs and random network models

In this section, some common network models that are relevant to this dissertation are reviewed from the two books: Generating random networks and graphs by Coolen and colleague [171] and ii) Random graphs by Alan Frieze [172]. As mentioned in the community section, network models need to be more versatile to incorporate the wider array of community classes to be more applicable.

Why network models? In order to study the collective behavior of the network population that inherits some topological features of the at-hand ensemble, one needs to make inferences on the latent generative mechanism that such ensemble is created from. This is called network inference. The opposite direction is defined to be network synthesis. Together, network inference and synthesis is collectively referred to as network modelling. The field of network modelling is important to many research endeavors because of three reasons:

1. In some particular applications, the limitation of data availability rises the needs of synthesizing artificial data with desired properties.

2. To test the hypothesis of whether newly acquired data belongs to the current population of networks.

3. To construct null models. In many applications, one needs to compare the pre-hypothesized structured network with some arbitrary structure-less network through modelling. Perhaps the most used model, at least in comparing to the network at hand is the structured-less counterpart, is the random graph.

Classical random graph models. The original random graph model is dated back to the celebratory work by Erdos and Rényi (ER) with the proposal on $G(n, p)$ with edge

existence of probability p and non-existence of $1-p$. Hence, for binary graphs, the probability that a particular ensemble is observed is:

$$\begin{aligned} P(A) &= \prod_{\mathrm{i<j}} p\delta(A_{\mathrm{ij}}, 1) + (1-p)\delta(A_{\mathrm{ij}}, 0) \\ &= \prod_{\mathrm{i<j}} p^{A_{\mathrm{ij}}}(1-p)^{1-A_{\mathrm{ij}}} \\ &= p^{\sum_{\mathrm{i<j}} A_{\mathrm{ij}}}(1-p)^{\binom{n}{2}-\sum_{\mathrm{i<j}} A_{\mathrm{ij}}} \end{aligned}$$

where δ is the typical Kronecker delta function, assuming all edges sampled independently from an identical distribution (Binomial in this case). It is trivial to observe that ER random graph assign equal probability to all ensemble with the same number of existing edges m whose expectation $\langle m \rangle = p\binom{n}{2}$.

Random graphs with topological constraints. Often, using trivial random models such as ER graphs does not suffice for practical applications where networks are shown to have heavy-tail, scale-free distribution. In such case, constructing null models with desirable topological constraints is necessary. Assume that $\Omega_\mu \mid \mu \in [k]$ represents k desired features with associated measure μ. There are two ways of incorporating topological constraints:

- Hard constraint: all generated ensembles must have the pre-defined features Ω_μ;
- Soft constraint: collection of ensembles must have this feature, in expectation. In other words,

  $$\Omega_\mu(A) = \sum_{\mathrm{i}} P(\mathcal{A}_{\mathrm{i}})\Omega_\mu(\mathcal{A}_{\mathrm{i}}), \forall \mathcal{A}_{\mathrm{i}} \in \mathcal{G}$$

  where $\mathcal{A}$ is generated ensemble from graph population $\mathcal{G}$ and A is the at-hand network.

An example of constraint-based graph generation is the generation of ER random graph $G(n,p)$ in which:

- Hard constraint synthesis would restrict that all realized ensemble from $G(n,p)$ must have exactly $\Omega_\mu(\mathcal{A}_{\mathrm{i}}) = p\binom{n}{2}$ edges;
- Soft constraint synthesis would relax the above condition to enforce $\sum_{\mathrm{i}} p(\mathcal{A}_{\mathrm{i}})m(\mathcal{A}_{\mathrm{i}}) = p\binom{n}{2}$, where $\Omega_\mu(\mathcal{A}_{\mathrm{i}}) = m(\mathcal{A}_{\mathrm{i}})$ is the number of edges in ensemble $\mathcal{A}_{\mathrm{i}}$.

In the next section, a well-studied network model with community structures is reviewed independently.

Critical phases in ER Random models. In [173], [174], among others, it was proved that the random graph explicitly experience six distinguishable ranges. Namely,

1. Range 1: $p = o(1/n)$ i.e. $lim_{n\to\infty} pn = 0$ or p does not scale with n. In this range, it is proven with high probability that all connected component of G is, at most, a tree. Furthermore, a tree of size k only possible when $p = \Theta(n^{-\frac{k}{k-1}})$ i.e. p is bounded between two constant c_1, c_2 times $n^{-\frac{k}{k-1}}$. Other results related to this range relates to the number of connected components can be inferred by Poisson distribution with mean $\lambda = \frac{(2c)^{k-1}k^{k-2}}{k!}$.

2. Range 2: $p = \frac{c}{n}$ for $c \in [0, 1]$. In this range, all connected components are either trees or uni-cyclic components (a tree with an additional edge). Most of the vertices in G ($n - o(n)$ are now belongs to an connected component). More importantly, the "largest" component in G has order

$$\frac{1}{\alpha}(log(n) - \frac{5}{2}loglog(n)) = O(log(n))$$

3. Range 3: $p = \frac{1+\mu}{n}$ This is the most exciting range in ER random graph because of the so-called *double-jump*. Specifically, they proved that before this range, as we observed, most of the connected component has size $O(1)$ where as $p = \frac{1}{n}+$, the size of the largest component is now $O(n)$. This is the range that I will base most of the analysis and simulation construction on.

4. Range 4, 5, and 6: These are ranges where other phenomena such as what happens to other nodes that are not in the giant component or how likely is the graph connected are rigorously proved. However, we are not going to investigate into these ranges due to the limitation of project scope.

Relationship between $\mathcal{G}_{n,p}$ and $\mathcal{G}_{n,m}$. As an realized ensemble G can be either drawn from $\mathcal{G}_{n,p}$ or its counterpart $\mathcal{G}_{n,m}$, one can differentiate the philosophical difference between

two "methods". While $\mathcal{G}_{n,p}$ draws the number of edges from the Binomial distribution ($\mu = np$,$\sigma^2 = np(1-p)$), $\mathcal{G}_{n,m}$ fixed the number of edges. Hence, drawing a graph in $\mathcal{G}_{n,m}$ uniformly at random will resulted in a labelled realization G with probability $\binom{\binom{n}{2}}{m}^{-1}$.

The asymptotic behavior of those two methods are the same as long as, roughly,

$$\binom{n}{2}p = m$$

Equivalently,

$$p = \frac{2m}{n^2}$$

In other words, $\mathcal{G}_{n,m}$ can be thought as the hard-constraint counterpart versus the soft-constraint one $\mathcal{G}_{n,p}$.

The random graph evolution. A graph process is defined as follows:

Definition .D.1. *A graph process ([175], [176]) $G_0 = ([n], |E_{n,p_0}| = 0), G_1, G_2 ....., G_{\left\{N=\binom{n}{2}\right\}} = K_n$ such that G_t is formed by G_{t-1} by adding an edge at random to a pair of vertices.*

If one considers $\mathcal{G}_{n,p}$ then $p = p(n) = 0 \to 1$ while, analogously, $\mathcal{G}_{n,m}$ then $m = m(n) = 0 \to \binom{n}{2}$. One nice thing about the graph process is that it only adds 1 edge per step. Hence the number of edges equals to the number of steps (ticks in Netlogo).

Critical Phase with respect to m:

Putting things together, we would like to understand in what step during the evolution of graph process does the double jump phenomenon takes place.

$$m = \frac{n^2 p}{2} = \frac{n^2 \frac{1+\mu}{n}}{2} = \frac{(1+\mu)n}{2} = \frac{n}{2} \tag{1}$$

Simulation Description. Since the critical phase transition is, now, known $\mathrm{E}(d_\mathrm{i}) = 1$, the investigated statistics are b_1 *ticks* before and b_2 *ticks* after the critical phase transition, given any graph size (n). Note that the critical phase transition is when the average degree of the ensemble $\mathrm{E}(d_\mathrm{i}) = 1 \forall \mathrm{i} \in [n]$. We will set up the experiment as follow:

1. Collect the statistics of the giant component size t_1 *ticks* before the critical phase for a given graph i.e. given n.

2. Collect the statistics of the giant component size t_2 *ticks* after the critical phase for a given ensemble.

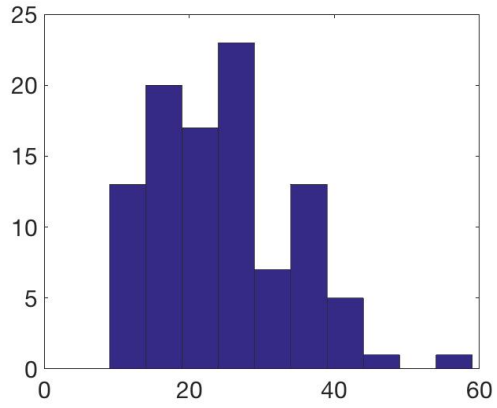
In this dissertation, $t_1 = 5$ and $t_2 = 10$ (ticks in Netlogo Software) is chosen to capture the behavior of the critical phases. It is important to note that the ER graph experience 6 different phase changes but only the change between Range II and Range III are sharp. To test the phase transition of ER random graphs, two questions need to be addressed:

1. How much does the giant component size grow between t_1 and t_2 ticks i.e. 'just' before and after the double-jump period?

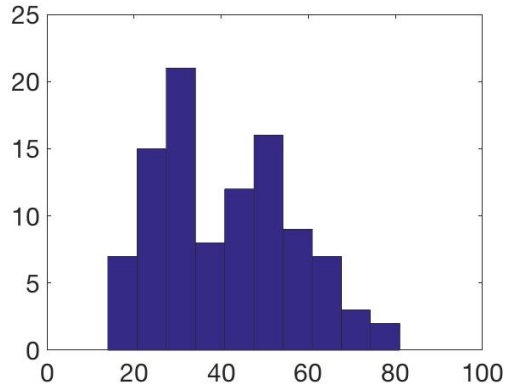
2. Do the giant component needs the entire edge set $\binom{n}{2}$ to acquire all possible nodes? Or is it way earlier in the process?

Simulation Results. In this section, the results (mostly through frequency of giant component sizes in a given graph at three different ticks: namely, t_1, t_2, t_3) are presented. The two graph sizes we will use are $n = 200, 400$.

•Case $G \sim \mathcal{G}_{(n=200,p_t)}$ Out of 100 simulations, the average giant component size difference



(a) Histogram of the random graph with $m = \frac{n}{2} - 5$ to examine the giant component size before critical phase (5 *ticks* before). The histogram has $(\mu = 25.2, \sigma^2 = 91.36)$



(b) Histogram of the random graph with $m = \frac{n}{2} + 10$ to examine the giant component size after critical phase (10 *ticks* after). The histogram has $(\mu = 41.32, \sigma^2 = 241.73)$.

Figure S6. The component size before and after (sharp) phase transition of $G \sim \mathcal{G}_{n,p_t}$ with size $n = 200$.

before/after the critical phase is approximately 15 nodes, on average, apart. If we put this

into perspective, 15 *ticks* addition to graph G which has at most $\binom{200}{2}$ edges. The expansion of giant component is very significant.

To examine the behavior of the giant component size "long" after the critical phase (but still way far from saturated clique size), we choose $w = 40$ for $n = 200$. To be able to choose the saturation coefficient w, one need to run the simulation a couples of times to find the spot where almost all nodes belong to the giant component. The number of ticks

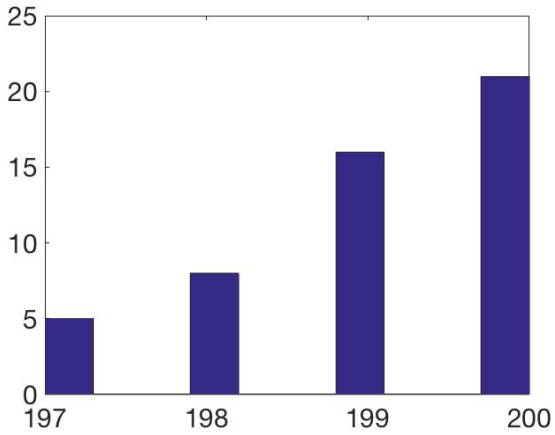
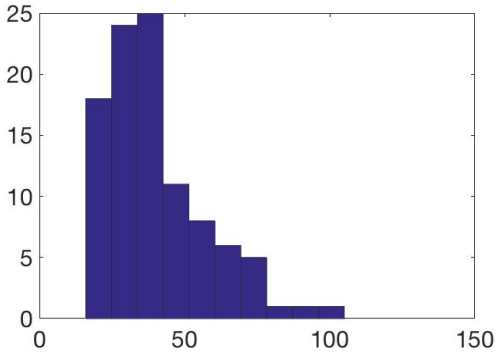


Figure S7. The component size after 500 ticks. The giant component size mean and variance are $(199.06, 0.99)$.

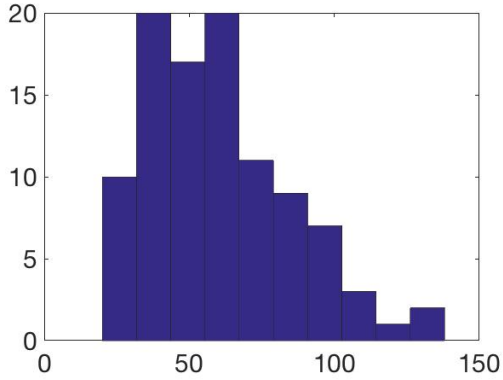
is computed as $\frac{n(n-1)}{80} = 500$ ticks; in addition, the giant component size mean and variance are recorded using the software as follows: $(\mu, \sigma^2) = (199.06, 0.99)$, see figure S7.

•Case $G \sim \mathcal{G}_{(n=400,p_t)}$ The next figure represents the size of the giant component where there are $\binom{400}{2} - \frac{\binom{400}{2}}{40}$ edges left to be realized from the model. The number of ticks is computed as $\frac{n(n-1)}{80} = 1000$ ticks; moreover, the giant component size mean and variance are recorded using the software as follows: $(\mu, \sigma^2) = (397, 3.75)$, see figure S9. One can observe the similar phenomenon where there are still lots of edges left to be added but the size of the giant component is now practically $n = 400$.

In both cases ($n = 200$ and $n = 400$), simple empirical exploration confirms the giant component sharp phase transition elegantly. This is one of the (if not the) first phase transition, discovered and rigorously proven in a random graph models. This discovery of



(a) Histogram of the random graph with $m = \frac{n}{2} - 5$ to examine the giant component size before critical phase (5 *ticks* before). The histogram has $(\mu = 40.7, \sigma^2 = 314.77)$



(b) Histogram of the random graph with $m = \frac{n}{2} + 10$ to examine the giant component size after critical phase (10 *ticks* after). The histogram has $(\mu = 60.5, \sigma^2 = 634.77)$.

Figure S8. The component size before and after (sharp) phase transition of $G \sim \mathcal{G}_{n,p_t}$ with size $n = 400$.

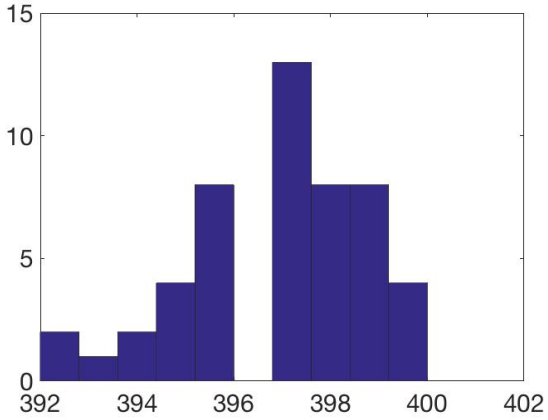


Figure S9. The component size after 1000 ticks. The giant component size mean and variance are $(397, 3.75)$

phase transition has also established a standard quest for other random graph models created after the famous ER model.

Structured graphs and Stochastic Block Models

Inevitably, Stochastic block models (SBMs) is one of the most studied graph model with hypothesized communities. The model is dated back to sociology in the early work

of Holland [177]. The model also appeared in different discipline with different names such as heterogeneous random graphs in graph theory. In figure S10, the plotted ensemble is

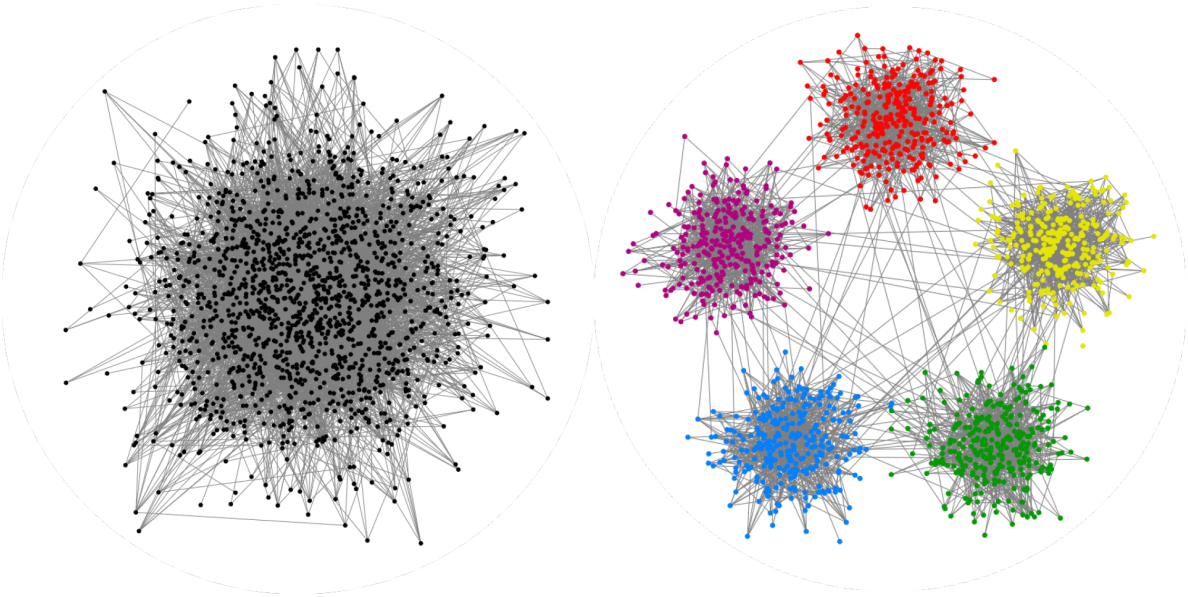


Figure S10. An example of a graph with 1000 nodes equality divided into 5 communities with drastic difference in connectivity probability between and within communities. Figure is adapted from [22].

generated from SBM with within-community and between-community probability are $\frac{1}{50}$ and $\frac{1}{1000}$, respectively.

Model Description. Some of the key components of basic SBM is defined here. Other fundamental mathematical notations are referred to Linear Algebra section.

- $G = [a_{uv}] = \begin{cases} FC, & weighted - graphs \\ M, & binarized - graphs \end{cases}$: network/graph (e.g. functional connectomes (FCs) in the context of this work);
- $V(G) = \{u\}$, and $E(G) = \{uv \mid u, v \in V(G)\}$ be set of vertices and edges, respectively;
- $|V(G)| = n$ and $|E(G)|$ are the size and order of network, respectively;

- $\{G_n\}\, \forall n \in \mathcal{N}$ is the graph sequence; in empirical domain, the number of graphs in the sequence is defined as $|\,\{G_n\}\,| = N$;
- k: number of communities/clusters;
- $\sigma = [\sigma_u] \in [k]^n$ is the pre-defined, well-understood community assignment in vector form of length n. It is the mathematical map $\sigma \coloneqq \{u \mapsto \mathrm{i}, \forall u \in [n], \mathrm{i} \in [k]\}$. In general, σ is also referred to as a graph partition;
- $\Omega = [|\Omega_{\mathrm{i}}|]$ is the vector containing cardinality of community where

$$|\Omega_{\mathrm{i}}| = |\,\{u \mid \sigma_u = \mathrm{i}\}\,|, \forall \mathrm{i} \in [k], u \in [n]$$

- C is the statistical summary of edge properties within and between communities in matrix form. Mathematically,

$$C = \begin{cases} C_{bin} \in \mathcal{N}_+^{k \times k} \\ C_{\mathrm{wei}} \in \mathcal{R}^{k \times k} \end{cases}$$

  where $C_{bin} \in N_+^{k \times k}$ denoted the simple edge count matrix within/between communities and C_{wei} denoted the weighted edge sum (also within/between communities);
- $C_{max} \in N_+^{k \times k}$ is the maximum number of edges within/between communities;
- $p = [p_{\mathrm{i}}]$: the probability that a node u belongs to community $\mathrm{i} \in [k]$; $P = diag(p)$ is a $k \times k$ matrix filled with p_{i} in the diagonal;
- $Q = [Q_{\mathrm{ij}}] \in R^{k \times k}$ is the expected node degree matrix, i.e. the expected number of connections a node in community i has with community j;
- s_n: scalable factor of degree regime in a graph sequence G_n where $n \in N$;
- $W = [w_{\mathrm{ij}}]$ is the edge probability between 2 nodes in community i and j, respectively.[1];

[1]↑It is worth-noting that if w_{ij} is the same for all $\mathrm{i}, \mathrm{j} \in [k]$, then SBM collapses to classical ER random graph model

- $PQ = nP\left[\frac{W}{s_n}\right] = nPW$ is the community profile matrix where i column is the expected number of edges that community i has with all communities. Note that for weak-recovery (detection), scaling factor $s_n = 1$.

Phase transitions in SBM. Stochastic Block Models (SBM) has recently gaining traction due to exciting developments in both theoretical and practical perspective. Specifically, in theoretical domain, phase transitions in **detecting** communities (or more generally, mesoscopic structures) were discovered through the measure Signal-to-noise Ratio (SNR) [22]. In brain connectivity domain, SBM has demonstrated its advantages in exploring and uncover diverse types of brain functional sub-circuits (e.g. dis-assortative or core-periphery) beyond the traditional assortative mesoscopic structures [151], [152]. Specifically, Sandon and Abbe, in [22], laid out a comprehensive treatment of mesoscopic recovery criteria for any pair of networked systems and *a priori* set of communities (or functional networks in brain connectomic domain) as follows:

1. Weak Recovery (also known as detection):
2. Almost Exact Recovery;
3. Exact Recovery:

Definition .D.2. *Weak recovery (of a ground-truth partition) can be rigorously thought as the existence of an algorithm that infer a partition that agrees with the ground-truth one up to* $\max_{\mathrm{i}} p_{\mathrm{i}}+, \forall \mathrm{i} \in [k]$. *This level of accuracy is the minimal requirement for most community detection methods.*

Theorem .D.1. *(Sandon and Abbe [22]) Let* $(G, \sigma) \sim SBMn, p, \frac{s_n Q}{n}$ *for* p, Q *arbitrary and* $s_n = 1$. *If* $SNR > 1$, *then weak recovery is efficiently solvable; where*

$$SNR = \frac{\lambda_2^2}{\lambda_1}$$

and λ_{i} *is the* i^{th} *eigen value of the community profile matrix* PQ.

Weak recovery (of a given ground-truth communities) means that the recovered partition (from some algorithm) beats the random guess, i.e. $\max_{\mathrm{i}} p_{\mathrm{i}}$ by a small factor . The criteria

for weak-recovery is driven by a hard threshold approach presented in the below theorem. As we can see here, to satisfy weak recovery criteria, we do not need the graph to be connected, asymptotically. Loosely speaking, we only need the graph sequence to have a large connected component. In other words, we only need $\{G_{n\in N}\}$ to be in the constant degree regime, i.e. $s_n = 1$. An example of weak recoverability is that suppose we have a network with n nodes and two ground-truth communities of equal size (i.e. $\frac{n}{2}$ nodes for each community), if the graph is in weak recovery regime,we can recover the true community membership of each node with the probability, somewhat, larger than 50% by a small amount, say 5%. We see that if an ensemble is generated in constant degree regime, one can arbitrarily assign any community membership to isolated nodes, i.e. leafs, and hence, exact recovery is not possible in this regime. Finally, for exact recovery, since W scales with n through the factor s_n, the community profile matrix M is consequently grows with the factor s_n, as well.

Definition .D.3. *Exact recovery (of a ground-truth partition) can be rigorously thought as the existence that has the probability of inferring the correct node memberships of* $O_n(1)$ *nodes to be* $1 - o_n(1)$ *as* $n \to \infty$.

Theorem .D.2. *(Abbe et al. [22]) Exact recovery in* $SBM(n, p, \frac{s_n Q}{n})$ *is solvable and efficiently so if*

$$I_+(p, Q) = min_{\mathrm{i,j}\in[k]} D_+((PQ)_\mathrm{i}, (PQ)_\mathrm{j}) > 1$$

where

$$D_+((PQ)_\mathrm{i}, (PQ)_\mathrm{j}) = max_{t\in[0,1]} \sum_x (PQ)_\mathrm{j}(x) f_t \left[\frac{(PQ)_\mathrm{i}(x)}{(PQ)_\mathrm{j}(x)}\right]$$

and

$$f_t = 1 - t + ty - y^t$$

Recall that $(PQ)_\mathrm{i}$ is the i^{th} column of the community profile matrix. Essentially, the Chernoff-Hellinger distance/divergence (CHD) measure "how difference" any two column of such matrix can be distinguished and that if the minimum CHD passes the hard threshold of 1 then the graph is now in a exact-recoverability regime. In other words, latent node membership recoverability is almost guaranteed. For example, given an unknown *a priori* set of FNs and a given graph sequence $\{G_{n_l}\}$ that satisfied the Theorem 2, one can recovery

such set of FNs almost exactly (say 95%) with some neglect-able errors, i.e. some small fraction of nodes might be classified incorrectly (say 5%). To be in this regime, the graph sequence needs to be asymptotically connected, e.g. scaling factor must exists $s_n > 1$.

Graph Sequence and required topological features for recovery

Definition .D.4. *A graph sequence is a mathematical series of graph ensembles generated by some fixed rules. Mathematically, it is denoted by* $\{G_l\}$ *where* l *is the sequence index.*

For a example, one can generate an ER random graph sequence, denoted as $G_n(n, p)$ with fixed p with limiting graph $G_{n=\infty}(n, p)$ (also known as graphon). With respect to the recent developments in $SBM(k, p, W)$ theory, it is important to note the theoretical scaling characteristics of model parameters such as k, p, W and make necessary assumptions, i.e. which scales with n, which stays constant as graph size grows to ∞. Firstly, it is common to assume that p, k does not scale with n; hence, the number of communities and their respective sizes do not grow with n [22]. In other words, communities are assumed to have linear sizes [22]. We also see that $Q = [Q_{\text{ij}}]$ is constant but W scales with n through scaling factor s_n.

Moreover, matrix W has theoretical ties with an important topological characteristic of a graph sequence, e.g. degree regime in the graph sequence $\{G_l\}$. The importance of degree regime lies on its relations with graph connectivity. Specifically, there are two important degree regimes that are relevant for graph partition recoverability:

- **Constant degree regime**: In this regime, connectivity pattern is fixed (independent of Schaefer granularity levels). Asymptotically, node degrees do not scale with graph size, i.e. $W = O(n^{-1})$. In random graph theory, this is the degree where ER graph is expected to have a giant component. This is the minimum requirement for the weak-recovery criteria which will be formally defined later.

- **Diverging degree regime**: In this regime, connectivity pattern varies with graph sequence sizes. Asymptotically, node degrees do scale with graph size at a scalable factor s_n, i.e. $W = O(log(n)n^{-1})$. In random graph theory, this degree regime generates connected ER ensemble, in expectation. This is a minimum requirement for exact recovery (defined in later section).

.E Community Detection Methods of Networks

Communities on networks

Originated from sociology context [178], communities, also known as meso-scale structures, typically refers to a proper subset of nodes possessing certain degree of similarity. Originally, "similarity" is thought in terms of internal connectivity, e.g. the number of connections that nodes within the same community exerts, with respect to other nodes in other communities. This leads to the most popular definition of a community is **assortative** clusters: the notion of particularly dense subgraphs in sparse graph. The intuition of such definition emerges from a social context, where group of members exerts higher internal density relatively to external ones. This definition of assortativity is, later, referred to as $density - based$ clusters. Philosophically, the concepts of communities can be thought from a $flow - based$ (also known as pattern-based) perspective where communities are proper subsets of nodes that sustain underlying internal flow of information within itself. Different measurements were also proposed and analysed (see [107], [109] for comprehensive overviews). Beyond density and flow based communities, there are other classes of communities such as dis-assortative or core-periphery. These classes of communities are not conventional, at least compared to the early notion of sociological communities. As the field of community detection grows, they are logically, more often, referred to as mesoscopic structures of networks.

Rather than an exhaustive argument for which particular view one should always adapt, it is heavily driven by application. For instance, in networks where dynamical interactions among entities are emphasized in edge weights, $flow - based$ communities tends to represent meaningful meso-scale structures. The notion of similarity shared by an induced subnetwork also legitimizes disassortative communities in which external edge connectivities outweigh internal ones. This allows the foundation to extend the traditional notion of communities above assortativity. Collectively, these are called community classes; there exist three primary ones: assortative, dis-assortative, and core-periphery. Despite the classical and intuitive notion of assortative communities, real-world networks are often hypothesized to possess a diverse classes of meso-scale structures. For instance, in the context of brain

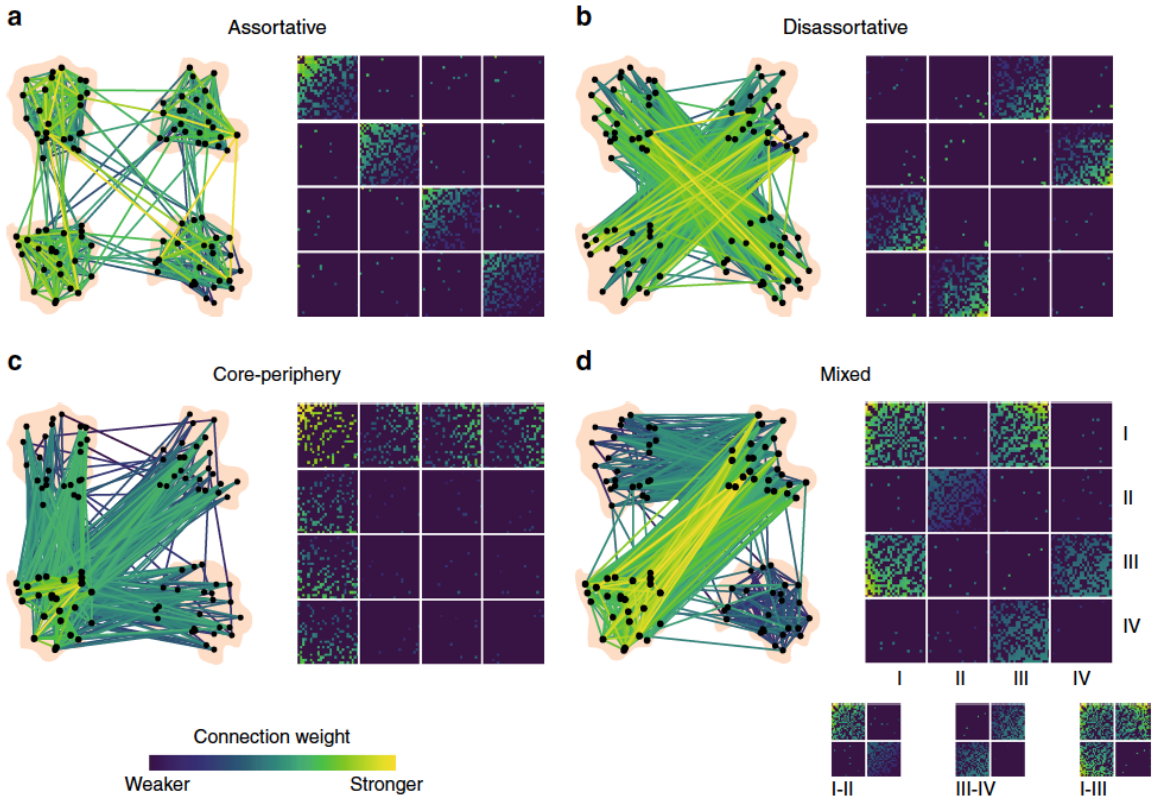


Figure S11. An example of four main community classes: assortative (**a**), dis-assortative (**b**), core-periphery (**c**), and mixed (**d**). Figure is adapted from [151]

functional connectomic, the co-existence of multiple community classes is proven particularly important to explain higher-order cognition in human, see [151]. In this and the next section, approaches to finding communities for a given network are reviewed. There are two most common approaches:

- objective function;
- statistical inference.

Objective-function approach

Each of the two approaches has its advantages and shortcomings which are to be reviewed in subsequent sections. For objective based methods, one typically starts with the philosophy

Table 2. Common Objective-based Community Detection Methods

Category	Method	Objective Function	Underlying Mechanics	Notations	Refs.
Density-based	**Modularity**	$Q = \frac{1}{2m}\sum_{i,j}(A_{ij} - p_{ij})\delta_{c_i,c_j}$	Quantify the difference between at-hand graph and a random counterpart that preserves the degree distribution.	• A_{ij}: binary edge (non-)existence between node i and j; • p_{ij} is the probabilistic counterpart representing the likelihood of edge existence between node pair i and j; • δ_{c_i,c_j}: Kronecker Delta which equals to 1 if node i, j are in the same community, i.e. $c_i = c_j$ and 0, otherwise, i.e. $c_i \neq c_j$.	[155], [179], [153], [86], [180] among many others.
Flow-based	**info-map**	$L(M) = qH(\mathcal{Q}) + \sum_{i=1}^{m} p^i H(\mathcal{P}^i)$	Total amount of information (measured by entropy) required by the random walker to fully exploit meaningful communities.	q module-switching probability; $H(\mathcal{Q})$: between-module entropy; p^i within-module probability; $H(\mathcal{P}^i)$: within-module entropy	[181], [88], [89].
	Stability	$r(t) = \max_H \{\min_{0\leq s\leq t} \sum_{i=1}^{c}(R_s)_{ii}\}$	Sustainability of random walker to stay inside communities after discrete time step t.	c: number of clusters, given a partition; H community assignment (0-1 matrix) for nodes in graph G; R_s is a $s-$step dependence of transfer probabilities between clusters.	[182], [183].

(which class of communities is hypothesized in the given network - this is domain-specific). The two most common views are

- density-based communities - which are more suitable in undirected networks;
- flow based communities - which are suitable for directed networks.

The below table summarizes some of the most known community detection method(s) for assortative communities in a given network driven by two aforementioned views. Per figure S12, both Q score and informap approaches yields different partitions. However, these partitions are technically assortative with the only difference lies in the philosophical perspective in which Q-score driven partitions do not consider the underlying flow of information among the elements of the same communities as infomap.

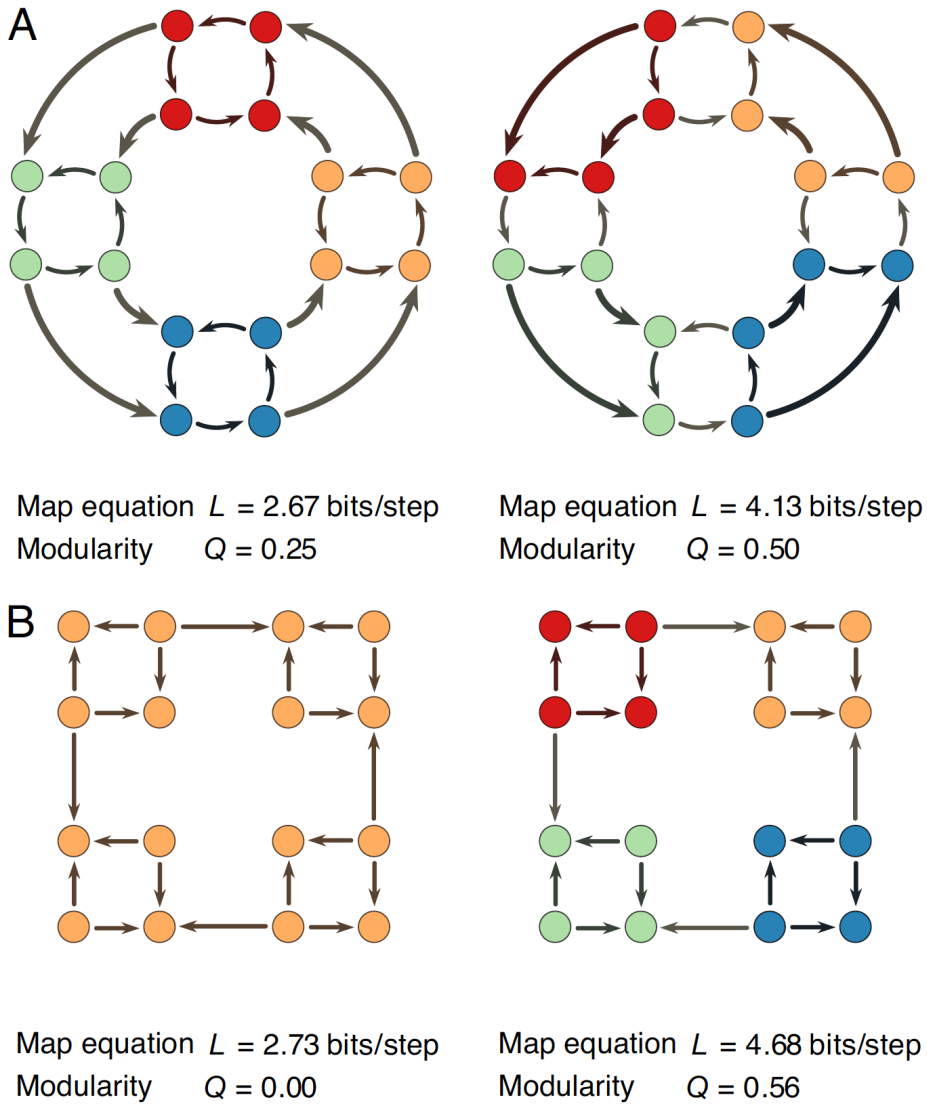


Figure S12. An example of different partition results from maximizing modularity score versus infomap. In both cases, infomap rewards communities with meaningful underlying flow of information whereas modularity rewards simply on internal vs. external edges. Note that in case B, infomap result shows that no meaningful mesoscopic structures exists while modularity shows 4 distinct communities. Figure is reproduced with permission from [88].

In this dissertation, the Newman's Q-score is extensively focused because of its significance in pioneering community detection methods. Specifically, for a given partition can be computed as follows:

$$
\begin{aligned}
Q(\sigma, \alpha = 1) = Q(\sigma) &= \sum_{u,v} (A_{uv} - \alpha P_{uv}) \delta(\sigma_u, \sigma_v) \\
&= \frac{1}{2m} \sum_{uv}^{184} \left\{ A_{uv} - \frac{d_u d_v}{2m} \right\} \delta(\sigma_u, \sigma_v)
\end{aligned}
$$

where $2m$ and d_u is graph and node degree, respectively:

$$\forall u \in V(G) : d_u = \sum_v A_{uv} \quad \& \quad 2m = \sum_u d_u$$

and the default scaling factor $\alpha = 1$; this scaling factor is typically used to scan the hierarchical structure of community in a network.

There is more than one way to model the null model P_{uv}. The Newman's approach is $P_{uv} = \frac{d_u d_v}{2m}$ which represents the random graph (no particular community structures) with the same empirical degree sequence. In theory, one can assume Poisson Distribution for node degree (like in the WSBM case). In the case of Newman's Q, this null model is built with respect to the empirical network degree. Further, tuning parameter is, by default, set at $\alpha = 1$ and the delta function is defined as

$$\delta(\sigma_u, \sigma_v) = \begin{cases} 1, \sigma_u = \sigma_v \\ 0, \sigma_u \neq \sigma_v \end{cases}$$

If *a priori* partition σ is known, then the modularity score can be written as a blockage format (only survived terms in within communities: $\forall u, v \in V(G) \mid \sigma_u = \sigma_v = \mathrm{i} \in [k]$) as follows:

$$\begin{aligned} Q(\sigma) &= \frac{1}{2m} \sum_{uv} \left\{ A_{uv} - \frac{d_u d_v}{2m} \right\} \delta(\sigma_u, \sigma_v) \\ &= \sum_{\mathrm{i}=1}^{\mathrm{i}=k} \left[\frac{\sum_{u,v \in \mathrm{i}} A_{uv}}{2m} - \frac{\sum_{u,v \in \mathrm{i}} d_u d_v}{(2m)^2} \right] \\ &= \sum_{\mathrm{i}=1}^{\mathrm{i}=k} \left[\frac{C_{\mathrm{ii}}}{2m} - \sum_{u,v \in \mathrm{i}} \frac{d_u}{2m} \frac{d_v}{2m} \right] \\ &= \sum_{\mathrm{i}=1}^{k} \left[\frac{C_{\mathrm{ii}}}{2m} - \left[\frac{s_{\mathrm{i}}}{2m} \right]^2 \right] \\ &= \sum_{\mathrm{i}=1}^{k} (P_\sigma(\mathrm{ii}) - P^Q_{null}(\mathrm{ii})) \end{aligned}$$

Because:

$$\sum_{u,v\in \mathrm{i}} A_{uv} = C_{\mathrm{ii}}$$

and

$$\sum_{u,v\in \mathrm{i}} d_u d_v = \sum_{\sigma_u=\mathrm{i}} d_u^2 + 2\sum_{u\neq v} d_u d_v = (\sum_{\sigma_u=\mathrm{i}} d_u)^2 = (s_1)^2$$

where s_1 is the total number of half-edges (stubs) that originates from nodes in community i.

We also look at the null model from another perspective that the event of a stub (half-edge) exist at node u with probability $P_u(stub) = \frac{d_u}{2m}$, likewise, at node v with $P_v(stub) = \frac{d_u}{2m}$. These two independent event needs to happen sequentially to form a edge between node u and v with probability

$$P^Q_{null} = P_{uv}(\mathrm{e}dge) = \frac{d_u}{2m}\frac{d_v}{2m}, \forall \sigma_u = \sigma_v = \mathrm{i} \in [k]$$

Note that here, no community indication is available for either node u or v which implies a null model (i.e. random partition) of σ (ground-truth).

Q-score can be applied to both binarized or weighted graph. In this case, for each threshold value, Q-score is computed for the weighted group-average FCs across Schaefer granularity levels. Maximizing modularity has been shown to unravel assortative communities while SBM has been shown to uncover different types of community, beyond assortative ones [151].

It is worthy to mentioned that the majority of community detection methods using objective functions revealing different perspectives on *assortative* communities (not necessarily different classes of communities such as assortative versus dis-assortative and core-periphery).

Network Inference using Stochastic Block Models

(Binary) SBM

. Since SBM is a generative model, one ought to talk about how to synthesize ensembles using such models, e.g. network synthesis and how to infer SBM parameters, using the

observable ensembles, e.g. network inference. We note that in the context of our problem, we have a slightly different starting point as the partition is not latent but in general, partitions are often inferred. Networks with existent ground-truth partition are very rare; furthermore, those ground-truth cannot be defined in an absolute sense. The majority of SBM is defined as follows:

$$G \sim SBM(k, p, W, \sigma)$$

In some particular applications, the partition is not latent (i.e. known as a prior). Specifically,

$$(G, \sigma, k) \sim SBM(p, W)$$

.

In the case of $G \sim SBM(k, p, W, \sigma)$, SBM seeks an partition that divide network G into k communities. The probability that two nodes are connected to each other is governed by probability W_{σ_u,σ_v}. To fit SBM onto a network, one needs to estimate $W = [w_{\mathrm{ij}}], \forall \mathrm{i}, \mathrm{j} \in [k]$ (this means that k is *a priori* condition for fitting) along with community label $\sigma_u, \forall u \in [n]$. Assuming that each edge is drawn independently from identical distributions, then, the probability that a network $G = A = [a_{uv}]$ is generated (synthesized) from *a priori* W and σ (prior beliefs) is as follows:

$$P(A \mid W, \sigma) = \prod_{u>v} W_{\sigma_u,\sigma_v}^{a_{uv}} (1 - W_{\sigma_u,\sigma_v})^{1-a_{uv}}$$

for symmetric networks. From the inference standpoint, the Bayesian posterior probability can be computed as follows:

$$P(\sigma \mid A) = \frac{\sum_W P(A \mid W, \sigma) P(W, \sigma)}{P(A)}$$

where $P(W, \sigma)$ is Bayesian prior beliefs. If there is only one W (*hard constraint*, Piexoto) that is comparable to network A and partition σ then we can drop the summation notion which results in:

$$\begin{aligned} P(\sigma \mid A) &= \frac{P(A \mid W, \sigma)P(W, \sigma)}{P(A)} \\ &= \frac{exp\{-ln(P(A \mid W, \sigma)) - ln(P(W, \sigma))\}}{P(A)} \end{aligned}$$

The hard constraint assumption is very standard technique to isolate eventually partition σ for inference purpose. Note that adjacency structure A is of course "hard" (there is only one ensemble A).

Since $P(A)$ is also fixed, maximization of posterior probability $P(\sigma \mid A)$ is equivalent to maximizing

$$-ln(P(A \mid W, \sigma)) - ln(P(W, \sigma))$$

which is also understood as minimization of the description length (DL, measured *in bits*) of ensemble A using partition σ. Once again, hard constraint assumption yields the description length ultimately only depends on:

$$DL = -ln(P(A \mid W, \sigma))$$

In binary SBM case, it follows that:

$$\begin{aligned} DL &= -ln(\prod_{u>v} W_{\sigma_u, \sigma_v}^{a_{uv}} (1 - W_{\sigma_u, \sigma_v})^{1-a_{uv}}) \\ &= -\sum_{u>v} a_{uv} ln(W_{\sigma_u, \sigma_v}) + (1 - a_{uv}) ln(1 - W_{\sigma_u, \sigma_v}) \end{aligned}$$

Hence, minimization of DL is equivalent to maximizing the log likelihood function.

Weighted SBMs

. The assumption of binary edges could be unfitting to some applications, including the functional brain networks where there is a need to express different level of functional

coupling strength numerically. In such case, we need to introduce the structure of covariates (denoted as $x = [x_{\sigma_u,\sigma_v}]$) to model the weights. In this case, the prior is written as follows:

$$P(x, A \mid \sigma) = P(x \mid A, \sigma)P(A, \sigma)$$

It follows that the posterior probability becomes:

$$P(\sigma \mid A, x) = \frac{P(A \mid x, \sigma)P(x, \sigma)}{P(A)}$$

Using similar technique in binary case, one can estimate covariate structure first so that joint probabilities with x, e.g. $P(x, \sigma)$ and $P(A, x)$, does not alter the posterior distribution behavior. Hence, the posterior belief is proportional to the priors which can be written as follows:

$$P(\sigma \mid A, x) \sim P(A \mid x, \sigma)$$

Ultimately, the task of finding "ground-truth" partition σ depends on the likelihood of prior beliefs. This is equivalent to maximizing $P(A \mid x, \sigma)$. The first task is of course to estimate x as A is already available. We notate the covariate structure x to be more integrated with probability distribution parameter notations $P(X = x)$. Specifically, we assume that the realized FC edge weights are drawn from some distributions with specific parameter(s).

Model Selection

There is different approaches to model selection (e.g. which edge weight distribution one should use, given the empirical data). In the context of this paper, given the empirical distribution of FC edge weight and the usage of absolute functional connectivity (non-negative pair-wise edges), we short-list two candidate distribution: exponential (continuous), and Poisson (discrete counterpart). Each choice has its pros and cons. For instance, choosing exponential distribution allows us to stick with continuous ensembles of functional edge weights which is consistent with how FC edges are computed using Pearson correlations. However, in continuous distribution, the probability of FC edge takes on a particular value is zero by

definition. Yet, the functional connectome is sparse [151], e.g. majority of pairwise interactions between two brain regions is non-existent. Hence, using exponential distribution will not suffice. A common approach is to use a different distribution such as Binomial to model edge (non-)existent and connect the two distribution using weighted average (see method section in [151] for further details). This will force modeler to make an precursor assumption on weight value which is not ideal.

On the other hand, using Poisson distribution (a discrete counterpart of exponential distribution) offer us a distinct advantage of modeling a non-zero probability of getting zero-valued functional edges. Recall that the Poisson probability density function is as follows:

$$f(k, \lambda) = P(X = k) = \frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}$$

Clearly, $P(X = 0) = \mathrm{e}^{-\lambda} > 0$ which ultimately depends on λ inference based on empirical observations of functional edges. Nonetheless, the shortcoming of using discrete distribution is precisely the advantage of using exponential one: being able to model edge in a continuous manner. To overcome this shortcoming of discrete distribution usage, we convert functional couplings (computed by Pearson correlations which are nicely bounded between $[-1, 1]$) to percentage point and rounding to nearest integer. For instance, if a functional edge has value of $a_{uv} = 0.588$, the weighted graph will take $a_{uv} = 59$. The reason for rounding to nearest integer is because Poisson distribution takes on non-negative integer values N^+. Note that using Poisson Distribution, no essential topological changes is made to the original FC other than rounding functional couplings to nearest integer.

In this paper, we use Poisson Distribution for degree sequence as proposed by Karrer and Newman [154]. In the next sections, we review the inference procedure (as proposed in [154]) for both assumptions:

- Non-degree-corrected WSBM;
- Degree-Corrected WSBM.

WSBM Inference Procedure: In this paper, we use method as described in https://graph-tool.skewed.de/ by Tiago Piexoto. Further treatments on weighted SBM can be found

at [184]. After we review the inference approaches, we compare the philosophical similarity and difference between WSBM inference and Q score modularity.

Standard (Non degree-corrected) WSBM

The review of non-degree-corrected (NDC) WSBM is provided in a well-cited paper by Karrer and Newman [154], the author assumed such distribution for multi-graph ensembles, where edges can take on integer values larger than 1. In this case, prior probability can be written as follows:

$$P(A \mid x, \sigma) = \prod_{u<v} \frac{(x_{\sigma_u,\sigma_v})^{A_{uv}} \mathrm{e}^{-x_{\sigma_u,\sigma_v}}}{A_{uv}!} \times \prod_{u} \frac{(x_{\sigma_u,\sigma_u})^{A_{uu}/2} \mathrm{e}^{-x_{\sigma_u,\sigma_u}}}{(A_{uu}/2)!}$$

It is important to note that the expected adjacency structure in this case is

$$\mathrm{E}(A_{NDC}) = YxY^T$$

where $Y \in [0,1]^{n\times k}$ be the node community membership matrix, i.e. $y_{ul} = 1$ if and only if node u is in community $l \in [k]$. Note that self-loop edge weight cannot be counted twice. For symmetric networks where $A_{uv} = A_{vu}$ and $x_{\mathrm{ij}} = x_{\mathrm{ji}}$, the above prior probability can be written as follows:

$$P(A \mid x, \sigma) = \frac{\prod_{\mathrm{ij}} x_{\mathrm{ij}}^{C_{\mathrm{ij}}/2} exp(-\frac{1}{2}|\Omega_{\mathrm{i}}||\Omega_{\mathrm{j}}|x_{\mathrm{ij}})}{\prod_{u<v}(A_{uv}! \prod_u 2^{A_{uu}/2}(A_{uu}/2)!)}$$

where $|\Omega_{\mathrm{i}}|$ is the cardinality of community i, C_{ij} is the counted number of edges between community i and j which can be simply computed by:

$$C_{\mathrm{ij}} = \sum_{u,v} A_{uv} \delta_{\sigma_u,\mathrm{i}} \delta_{\sigma_v,\mathrm{j}}$$

where δ is the Kronecker delta function as defined in the main text. Similar to the binary case, the log-function is then be:

$$logP(A \mid x, \sigma) = \sum_{\mathrm{ij}} (C_{\mathrm{ij}} log(x_{\mathrm{ij}}) - |\Omega_{\mathrm{i}}||\Omega_{\mathrm{j}}|x_{\mathrm{ij}}) + \Theta(G)$$

where $\Theta(G)$ is the quantity dependent on ensemble G (such as $|\Omega_{\mathrm{i}}|$ or A_{uv}) which has no impact onto the logarithmic function behavior (i.e. not impacting the optimal value of this function). The inference process reduces to maximizing:

$$L(x, \sigma) = \sum_{\mathrm{ij}} (C_{\mathrm{ij}} log(x_{\mathrm{ij}}) - |\Omega_{\mathrm{i}}||\Omega_{\mathrm{j}}|x_{\mathrm{ij}})$$

Note that here, we drop A (ensemble adjacency structure) just to ease notation usage and emphasize which variable(s) the likelihood function depends on. To do optimize the above function, one can just use the differential calculus as follows:

$$\frac{dL}{dx_{\mathrm{ij}}} = L_{\mathrm{ij}} = \frac{d}{dx_{\mathrm{ij}}} \left[\frac{C_{\mathrm{ij}}}{x_{\mathrm{ij}}} - |\Omega_{\mathrm{i}}||\Omega_{\mathrm{j}}| \right]$$

Setting the first derivative to zero, e.g. $L = 0$, we obtain:

$$\hat{x}_{\mathrm{ij}} = \frac{C_{\mathrm{ij}}}{|\Omega_{\mathrm{i}}||\Omega_{\mathrm{j}}|}$$

Note that now we have estimate x, i.e. the covariate structure, the likelihood function can be written as follows:

$$L(\hat{x}, \sigma) = \sum_{\mathrm{ij}} (C_{\mathrm{ij}} log(x_{\mathrm{ij}})) - 2m$$

Dropping constant $2m$ (ensemble node's degree sum) and substitute the estimated covariate $\hat{x}$, the log-likelihood function can now be written as:

$$L(\sigma) = \sum_{\mathrm{ij}} C_{\mathrm{ij}} log \left\{ \frac{C_{\mathrm{ij}}}{|\Omega_{\mathrm{i}}||\Omega_{\mathrm{j}}|} \right\}$$

Using simple algebra, the log-likelihood function can be rewritten to:

$$\begin{aligned} L(\sigma) &= 2m \sum_{\mathrm{ij}} \frac{C_{\mathrm{ij}}}{2m} \left[log \left\{ \frac{C_{\mathrm{ij}}/2m}{|\Omega_{\mathrm{i}}||\Omega_{\mathrm{j}}|/n^2} \right\} - log \left\{ \frac{n^2}{2m} \right\} \right] \\ &= \sum_{\mathrm{ij}} \frac{C_{\mathrm{ij}}}{2m} log \left\{ \frac{C_{\mathrm{ij}}/2m}{|\Omega_{\mathrm{i}}||\Omega_{\mathrm{j}}|/n^2} \right\} + \Theta(G) \end{aligned}$$

where, again, Θ is a constant function based on ensemble G.

Let Y and Z be the random variables representing community assignment on one end of a stub (half-edge). Then we can build a joint probability distribution between Y and Z as follows:

$$P_\sigma = P_\sigma(Y, Z) = \frac{C_{\mathrm{ij}}}{2m}, \forall \mathrm{i}, \mathrm{j} \in [k]$$

On the other hand, the randomized counterpart distribution of these random variables (with the same *a priori* partition σ) is

$$P_{null}^{WSBM} = \frac{|\Omega_{\mathrm{i}}||\Omega_{\mathrm{j}}|}{n^2}$$

In this case, edge formation (from two stubs) are completely at random with probability $\frac{|\Omega_{\mathrm{i}}|}{n}$ and $\frac{|\Omega_{\mathrm{j}}|}{n}$ for each stub. Comprehensive, the likelihood function becomes:

$$\begin{aligned} L(\sigma) &= \sum_{\mathrm{ij}} P_\sigma(\mathrm{ij}) log \left\{ \frac{P_\sigma(\mathrm{ij})}{P_{null}^{WSBM}(\mathrm{ij})} \right\} \\ &= \sum_{\mathrm{ij}} P_\sigma(\mathrm{ij}) \left[log(P_\sigma(\mathrm{ij})) - log(P_{null}^{WSBM}(\mathrm{ij})) \right] \end{aligned}$$

On the other hand, the Kullback-Leibler Divergence between two probability distributions $P(x)$ and $Q(x)$ is defined to be:

$$D_{KL}(P \mid\mid Q) = \sum_{x \in \mathcal{X}} P(x) log \left\{ \frac{P(x)}{Q(x)} \right\}$$

where $x \in \mathcal{X}$ is the random variable takes on values in the sample space $\mathcal{X}$. Then the log-likelihood function above can be thought as an information theoretic measurement between

the "ground-truth" probability distribution $x(\sigma)$ and the corresponding null distribution $x(null)$. If we only look at what happen within communities, the quality function becomes:

$$L_{within}(\sigma) = \sum_{\mathrm{i}} P_{\sigma}(\mathrm{ii}) \left[log(P_{\sigma}(\mathrm{ii})) - log(P_{null}^{WSBM}(\mathrm{ii})) \right]$$

If we substitute the estimated P_{σ} and P_{null} above to this equation, we obtain:

$$L_{within}(\sigma) = \sum_{\mathrm{i}} \frac{C_{\mathrm{ii}}}{2m} \left\{ log \left[\frac{C_{\mathrm{ii}}}{2m} \right] - log \left[\frac{|\Omega_{\mathrm{i}}|^2}{n^2} \right] \right\}$$

Of course, we also have the log-function described the differential information description requirement between communities:

$$L_{between}(\sigma) = \sum_{\mathrm{i} \neq \mathrm{j}} \frac{C_{\mathrm{ij}}}{2m} \left\{ log \left[\frac{C_{\mathrm{ii}}}{2m} \right] - log \left[\frac{|\Omega_{\mathrm{i}}||\Omega_{\mathrm{j}}|}{n^2} \right] \right\}$$

We will compare the within-community L_{within} with the modularity function Q score in the subsequent sections.

Degree-corrected WSBM

Fot the degree corrected (DC) WSBM case, a new hyper-parameter is introduced into the model θ_r (arbitrary constant terms that are $o(x_{\sigma_r,\sigma_s})$, i.e. constant terms that get absorbed into x_{ij}). The prior probability can now be written as follows:

$$\begin{aligned} P(A \mid \theta, x, \sigma) = & \prod_{u<v} \frac{(\theta_u \theta_v x_{\sigma_u,\sigma_v})^{A_{uv}} \mathrm{e}xp(-\theta_u \theta_v x_{\sigma_u,\sigma_v})}{A_{uv}!} \\ & \times \prod_{u} \frac{(\theta_u^2 x_{\sigma_u,\sigma_u})^{A_{uu}/2} \mathrm{e}xp(-\theta_u^2 x_{\sigma_u,\sigma_u})}{(A_{uu}/2)!} \end{aligned}$$

where $\sum_u \theta_u \delta_{\sigma_u,\mathrm{i}} = 1$ (with δ is the Kronecker delta function as usual). Basically, θ_u represents the probability that an half-edge (stub) in community i originated from u itself in which

$\sigma_u = \mathrm{i}$. It is noteworthy that the expected value of adjacency structure in this case is no longer just x_{σ_u,σ_v} but instead:

$$\begin{aligned}\mathrm{E}(A_{DC}) &= [\,\mathrm{E}(a_{uv})] = \theta_u x_{\sigma_u,\sigma_v}\theta_v \\ &= diag(\theta) Y x Y^T diag(\theta)\end{aligned}$$

where $diag(\theta) = diag([\theta_u])$ is the diagonal matrix contains the θ_u weights of node u. The priors can then be condensed as follows:

$$P(A \mid \theta, x, \sigma) = \frac{\prod_u \theta_u^{d_u} \prod_{\mathrm{ij}} x_{\mathrm{ij}}^{C_{\mathrm{ij}}/2} \mathrm{e}xp(-\frac{1}{2}x_{\mathrm{ij}})}{\prod_{u<v} A_{uv}! \prod_u 2^{A_{uu}/2}(A_{uu}/2)!}$$

with d_u being node u degree. The log-likelihood function is then

$$\begin{aligned}L = logP(A \mid \theta, x, \sigma) &= 2\sum_u d_u log\theta_u + \sum_{\mathrm{ij}} \{C_{\mathrm{ij}} log x_{\mathrm{ij}} - x_{\mathrm{ij}}\} \\ &= L_1 + L_2\end{aligned}$$

where d_u is node u degree and, again, ignoring constant terms $\Theta(G)$ which are terms containing A_{uv}. The goal is to maximize this log-function, compartment-ally, with respect to the normalization condition $\sum_u \theta_u \delta_{\sigma_u,\mathrm{i}} = 1$. We look at them separately (again, ignoring constant if any). Maximizing $L_2 = \sum_{\mathrm{ij}} \{C_{\mathrm{ij}} log x_{\mathrm{ij}} - x_{\mathrm{ij}}\}$ is straight-forward by taking derivative with respect to x_{ij}. Specifically,

$$L_2 = \frac{dL_2}{dx_{\mathrm{ij}}} = \frac{C_{\mathrm{ij}}}{x_{\mathrm{ij}}} - 1 = 0 \rightarrow \hat{x}_{\mathrm{ij}} = C_{\mathrm{ij}}$$

$$
\begin{aligned}
L_1 &= \sum_u d_u log\theta_u = \sum_{\mathrm{i}} d_u \delta_{\sigma_u,\mathrm{i}} log\theta_u \\
&= \sum_{\mathrm{i}} \left\{ \sum_{u|\sigma_u=\mathrm{i}} d_u log\theta_u \right\} \quad s.t. \quad \sum_{u|\sigma_u=\mathrm{i}} \theta_u = 1 \\
&= \sum_{\mathrm{i}} \left\{ s_{\mathrm{i}} \sum_{u|\sigma_u=\mathrm{i}} \frac{d_u}{s_{\mathrm{i}}} log\theta_u \right\} \quad s.t. \quad \sum_{u|\sigma_u=\mathrm{i}} \theta_u = 1
\end{aligned}
$$

where $s_{\mathrm{i}} = \sum_{u|\sigma_u=\mathrm{i}} d_u$ is the number of half-edges in community i. Note that there are $|\Omega_{\mathrm{i}}|$ terms of θ_u for each community. We see that is the entropy of the probability distribution representing the random variable θ, e.g. the probability that an edge in community i lands on u for which $\sigma_u = \mathrm{i}, \forall \mathrm{i}$. This entropy is minimized when

$$
\hat{\theta}_u = \frac{d_u}{\sum_u d_u}
$$

Here, it is important to note that if we choose random uniform distribution for random variable θ (e.g $\hat{\theta}_u = \frac{1}{|\Omega_{\mathrm{i}}|}$), we obtain minimized L_1 which reduces L.

Difference between Non-degree-corrected and Degree-corrected model

Plugging in the estimated parameters for both cases of WSBM, we obtain:

$$
L_{NDC} = \sum_{\mathrm{ij}} C_{\mathrm{ij}} log \left[\frac{C_{\mathrm{ij}}}{|\Omega_{\mathrm{i}}||\Omega_{\mathrm{j}}|} \right]
$$

and

$$
\begin{aligned}
L_{DC} &= \sum_{\mathrm{ij}} C_{\mathrm{ij}} log \left[\frac{C_{\mathrm{ij}}}{s_{\mathrm{i}} s_{\mathrm{j}}} \right] \\
&= \sum_{\mathrm{ij}} \frac{C_{\mathrm{ij}}}{2m} log \left[\frac{C_{\mathrm{ij}}/2m}{(s_{\mathrm{i}}/2m)(s_{\mathrm{j}}/2m)} \right]
\end{aligned}
$$

which is the Kullback-Leibler divergence between $P(x)$ (same as in the NDC case) and $Q_{DC}(x)$. In other words,

$$P_{null}^{WSBM} = \frac{s_{\mathrm{i}} s_{\mathrm{j}}}{(2m)^2}$$

for the DC case. Recall that for the NDC case, the null model is:

$$P_{null}^{WSBM} = \frac{|\Omega_{\mathrm{i}}||\Omega_{\mathrm{j}}|}{n^2}$$

Thus, the best fit to the NDC WSBM is the partition that most surprises the Erdos-Reyni random counterpart while for DC WSBM case, it is the group assignment that is most surprising to the random model with the same empirical degree sequence.

A conceptual comparison between Newman's modularity and SBM inference likelihood function

In this section, we compare the NDC, DC WSBM and Q score approach by first revisiting their formulas:

1. the NDC log-likelihood function (within communities):

$$L_{within} = \sum_{\mathrm{i}} \frac{C_{\mathrm{ii}}}{2m} \left\{ log \left[\frac{C_{\mathrm{ii}}}{2m} \right] - log \left[\frac{|\Omega_{\mathrm{i}}|^2}{n^2} \right] \right\}$$

2. the DC log-likelihood function (within communities):

$$L_{within} = \sum_{\mathrm{i}=1}^{k} \frac{C_{\mathrm{ii}}}{2m} \left\{ log(C_{\mathrm{ii}}/2m) - log \left[\frac{s_{\mathrm{i}}^2}{(2m)^2} \right] \right\}$$

3. Q-score modularity function, using community block format:

$$Q = \sum_{\mathrm{i}=1}^{k} \left[\frac{C_{\mathrm{ii}}}{2m} - \left[\frac{s_1}{2m} \right]^2 \right]$$

It is very interesting (yet, not surprising) that the two most known method for community detection is based on a similar principled of comparing a structure-less counterpart (that has

some similar topological characteristic) with the network at hand (which is hypothesized to have some latent structure of communities). In both approach, the hypothesized distribution of random variable Y and Z for within-community is actually the same:

$$P^{NDC-WSBM} = P^{DC-WSBM} = P^{Q} = \frac{C_{\mathrm{ii}}}{2m}$$

Obviously, there is a difference between the null model choice between Q score and NDC WSBM approach as the later does not feature the observed network degree sequence while Newman's Q method actually does. This shortcoming is resolved with the DC WSBM approach as mentioned in the previous section. In fact, the DC WSBM and Q score null model is actually the same. Specifically, for DC WSBM, the null model is

$$P_{null}^{DC-WSBM} = \frac{s_{\mathrm{i}}^2}{(2m)^2}$$

while for modularity approach, it is also:

$$P_{null}^{Q} = \frac{s_{\mathrm{i}}^2}{(2m)^2}$$

What, then, is the shortcoming of Q score approach? It does not emphasize what happen with the "between" community dynamics. To be clear, one can rewrite Q score so that it reflects between- community edges as proposed by Fortunato [109] as follows:

$$\begin{aligned} Q &= \sum_{\mathrm{i}=1}^{k} \left[\frac{C_{\mathrm{ii}}}{2m} - \left[\frac{s_1}{2m} \right]^2 \right] \\ &= \frac{-1}{m} \left[\left\{ m - \frac{1}{2} \sum_{\mathrm{i}} C_{\mathrm{ii}} \right\} - \left\{ m - \sum_{\mathrm{i}} (\frac{s_{\mathrm{i}}^2}{4m}) \right\} \right] \\ &= \frac{-1}{m} \left[Cut - \mathrm{E}(Cut) \right] \end{aligned}$$

where $Cut = \left\{ m - \frac{1}{2} \sum_{\mathrm{i}} C_{\mathrm{ii}} \right\}$ being the number of inter-community edges and $\mathrm{E}(Cut)$ is its corresponding expected counterpart. Basically, modularity would like to maximize the within community edges (equivalently, minimize the between community edges). Hence, it biases towards assortative community assignments.

On the other hand, the log function of WSBM has also incorporate what happens between communities through $L_{between}$. This is why SBM inference method shines over traditional Q maximization techniques for its ability to uncover a more diverse classes of community structures beyond assortative one.

An information-theoretical comparison between Newman's modularity and SBM-inference partitions

Given the node set of N elements $S = \{s_{\mathrm{i}} \mid \mathrm{i} \in [N]\}$, To quantify the amount of shared information between two partitions (e.g. $\sigma^1 = \{\sigma_r^1 \mid r \in [R]\}$ and $\sigma^2 = \{\sigma_c^2 \mid c \in [C]\}$), a normal approach would be using mutual information score, which can be quantified as follows:

$$MI(\sigma^1, \sigma^2) = \sum_{r=1}^{r=R} \sum_{c=1}^{c=C} P_{\sigma^1\sigma^2}(r, c) log \frac{P_{\sigma^1\sigma^2}(r, c)}{P_{\sigma^1}(r) P_{\sigma^2}(c)}$$

where R, C is the number of clusters in partition vector σ^1 and σ^2, respectively; $P_{\sigma^1\sigma^2}$ and $P_{\sigma^1}(\mathrm{i})$ are the joint and marginal probability distribution, respectively between two discrete random variables representing two realized partitions. A typical initialization step is to build a contingency table which indicates the number of common nodes has in common between cluster σ_r^1 and σ_c^2:

$$\begin{pmatrix} n_{11} & n_{12} & \cdots & n_{1C} \\ \ddots & \cdots & \ddots & \cdots \\ n_{R1} & n_{R2} & \cdots & n_{RC} \end{pmatrix}$$

where n_{rc} represents the number of common entities between cluster σ_r^1 and σ_c^2; the row and column marginal sums are denoted as $\vec{a} = a_r$ and $\vec{b} = b_c$, respectively. By construction $\sum_r a_r = \sum_c b_c = N$. The expected mutual information for a random partition with the same contingency table has closed-form formula as proposed in [185] as follows:

$$\begin{aligned} \mathrm{E}(MI(\sigma^1, \sigma^2)) = \sum_r \sum_c \sum_{max(1, a_r + b_c - N)}^{min(a_r, b_c)} & \frac{n_{rc}}{N} log \left\{ \frac{N \times n_{rc}}{a_r b_c} \right\} \\ & \times \frac{a_r! b_c! (N - a_r)! (N - b_c)!}{N! n_{rc}! (a_r - n_{rc})! (b_c - n_{rc})! (N - a_r - b_c + n_{rc})!} \end{aligned}$$

The entropy associated with the two partitions are:

$$H(\sigma^1) = \sum_{r=1}^{R} P_{\sigma^1}(r) log(P_{\sigma^1}(r))$$

where $P_{\sigma^1}(r) = \frac{|\sigma_r^1|}{N}$ is the probability that a element picked randomly from set S belongs to σ_r^1. Analogously,

$$H(\sigma^2) = \sum_{\mathrm{j}=1}^{C} P_{\sigma^2}(\mathrm{j}) log(P_{\sigma^2}(\mathrm{j}))$$

Finally, putting all components together, adjusted (for chance) normalized mutual information (AMI) can be computed as follows:

$$AMI = \frac{MI - \mathrm{E}(MI)}{max(H(\sigma^1), H(\sigma^2)) - \mathrm{E}(MI)}$$

Note that there are other ways to average the independent entropy of the two partitions such as arithmetic. In this paper, we use the maximum between the two entropy quantities.

APPENDIX: NETWORK NEUROSCIENCE PUBLICATION

Citation: **Duong-Tran, D.**, Abbas, K., Amico, E., Corominas-Murtra, B., Dzemidzic, M., Kareken, D., Ventresca, M., Goñi, J. (2021). A morphospace of functional configuration to assess configural breadth based on brain functional networks. Network Neuroscience, 5(3), 666–688.

Link: https://doi.org/10.1162/netn_a_00193

NETWORK NEUROSCIENCE

an open access journal

Citation: Duong-Tran, D., Abbas, K., Amico, E., Corominas-Murtra, B., Dzemidzic, M., Kareken, D., Ventresca, M., & Goñi, J. (2021). A morphospace of functional configuration to assess configural breadth based on brain functional networks. *Network Neuroscience*, *5*(3), 666–688. https://doi.org/10.1162/netn_a_00193

DOI:
https://doi.org/10.1162/netn_a_00193

Supporting Information:
https://doi.org/10.1162/netn_a_00193

Received: 2 January 2020
Accepted: 17 March 2021

Competing Interests: The authors have declared that no competing interests exist.

Corresponding Author:
Joaquín Goñi
jgonicor@purdue.edu

Handling Editor:
Claus C. Hilgetag

The MIT Press

METHODS

A morphospace of functional configuration to assess configural breadth based on brain functional networks

Duy Duong-Tran[1,2], Kausar Abbas[1,2], Enrico Amico[1,2,3,4], Bernat Corominas-Murtra[5], Mario Dzemidzic[6], David Kareken[6], Mario Ventresca[1,7], and Joaquín Goñi[1,2,8]

[1]School of Industrial Engineering, Purdue University, West Lafayette, IN, USA
[2]Purdue Institute for Integrative Neuroscience, Purdue University, West Lafayette, IN, USA
[3]Institute of Bioengineering/Center for Neuroprosthetics, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
[4]Department of Radiology and Medical Informatics, University of Geneva, Switzerland
[5]Department of Zoology, Institute of Biology, Karl-Franzens University Graz, Graz, Austria
[6]Department of Neurology, Indiana University School of Medicine, Indianapolis, IN, USA
[7]Purdue Institute of Inflammation, Immunology, and Infectious Disease, Purdue University, West Lafayette, IN, USA
[8]Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA

Keywords: Functional reconfiguration, Functional configural breadth, Resting-state networks, Functional connectomes

ABSTRACT

The quantification of human brain functional (re)configurations across varying cognitive demands remains an unresolved topic. We propose that such functional configurations may be categorized into three different types: (a) network configural breadth, (b) task-to task transitional reconfiguration, and (c) within-task reconfiguration. Such functional reconfigurations are rather subtle at the whole-brain level. Hence, we propose a mesoscopic framework focused on functional networks (FNs) or communities to quantify functional (re)configurations. To do so, we introduce a 2D network morphospace that relies on two novel mesoscopic metrics, trapping efficiency (TE) and exit entropy (EE), which capture topology and integration of information within and between a reference set of FNs. We use this framework to quantify the network configural breadth across different tasks. We show that the metrics defining this morphospace can differentiate FNs, cognitive tasks, and subjects. We also show that network configural breadth significantly predicts behavioral measures, such as episodic memory, verbal episodic memory, fluid intelligence, and general intelligence. In essence, we put forth a framework to explore the cognitive space in a comprehensive manner, for each individual separately, and at different levels of granularity. This tool that can also quantify the FN reconfigurations that result from the brain switching between mental states.

AUTHOR SUMMARY

Understanding and measuring the ways in which human brain connectivity changes to accommodate a broad range of cognitive and behavioral goals is an important undertaking. We put forth a *mesoscopic* framework that captures such changes by tracking the topology and integration of information within and between functional networks (FNs) of the brain. Canonically, when FNs are characterized, they are separated from the rest of the brain network. The two metrics proposed in this work, trapping efficiency and exit entropy, quantify the topological and information integration characteristics of FNs while they are still

embedded in the overall brain network. Trapping efficiency measures the module's ability to preserve an incoming signal from escaping its local topology, relative to its total exiting weights. Exit entropy measures the module's communication preferences with other modules/networks using information theory. When these two metrics are plotted in a 2D graph as a function of different brain states (i.e., cognitive/behavioral tasks), the resulting morphospace characterizes the extent of network reconfiguration between tasks (functional reconfiguration), and the change when moving from rest to an externally engaged "task-positive" state (functional preconfiguration), to collectively define network configural breadth. We also show that these metrics are sensitive to subject, task, and functional network identities. Overall, this method is a promising approach to quantify how human brains adapt to a range of tasks, and potentially to help improve precision clinical neuroscience.

Network configural breadth:
Represents, for an FN, a given individual's repertoire of cognitive and emotional states through functional configurations while performing different tasks. In practice, how well the entire "cognitive space" is sampled depends on the number and nature of the tasks. The functional network configural breadth, for a given subject and a given FN, is compartmentalized into two components: (a) FN (task) reconfiguration and (b) FN rest-to-[task-positive] preconfiguration.

Task-to-task transitional reconfiguration:
Represents the specific shift in the network functional configuration of an FN when a subject switches between distinct cognitive/mental tasks. For instance, task transitions and accompanying reconfigurations will occur when a subject transitions from quiet reflection to engage in a spatial problem-solving task, or from a lexical retrieval to a decision-making paradigm.

Within-task reconfiguration:
Represents specific network functional configuration changes of an FN that may occur within a single task. This phenomenon has been assessed at the whole-brain level, showing the presence of distinct brain states within a task. For instance, within-task reconfiguration can be tracked by using dynamic (sliding-window) functional connectivity.

INTRODUCTION

Human behavior arises out of a complex interplay of functional dynamics between different brain networks (Bassett & Gazzaniga, 2011). These interactions are reflected in functional network (FN) reconfigurations as subjects perform different tasks or are at rest (Amico, Abbas, et al., 2019; Amico et al., 2020; Cole, Bassett, Power, Braver, & Petersen, 2014). One of the network neuroscience challenges is to develop a comprehensive framework to quantify the brain network (re)configurations across different mental states and cognitive tasks. To that end, configurations across a collection of cognitive tasks can be conceptualized at three distinct levels of granularity:

- **Network configural breadth** represents, for an FN, a given individual's repertoire of cognitive and emotional states through functional configurations while performing different tasks. In practice, how well the entire "cognitive space" (Varona & Rabinovich, 2016; Varoquaux et al., 2018) is sampled depends on the number and choice of the tasks. This concept is inspired by Schultz and Cole (2016).
- **Task-to-task transitional reconfiguration** represents the specific shift in network functional configuration when a subject switches between cognitive/mental tasks (Douw, Wakeman, Tanaka, Liu, & Stufflebeam, 2016; Gonzalez-Castillo et al., 2015). For instance, task transitions and accompanying reconfigurations will occur when a subject transitions from quiet reflection to engage in a spatial problem-solving task, or from a lexical retrieval to a decision-making paradigm.
- **Within-task reconfiguration** represents specific network functional configuration changes that may occur within a single task. This phenomenon has been assessed at the whole-brain level, showing the presence of distinct brain states within a task (Bassett et al., 2011; Betzel, Satterthwaite, Gold, & Bassett, 2017; J. M. Shine et al., 2016; J. M. Shine et al., 2019; J. M. Shine & Poldrack, 2018).

While brain network configural properties are task and subject dependent (Schultz & Cole, 2016), task-induced functional (re)configurations are rather subtle in whole-brain functional connectomes, even when comparing task with rest (Cole et al., 2014). In addition, mesoscopic structures (e.g., functional networks of the brain) exhibit modular characteristics that adapt to cognitive demands without significantly affecting the rest of the system where higher levels of cognition emerge through the changing interactions of subsystems, instead of pairwise edge-

Module trapping efficiency (TE): Quantifies the capacity of an FN to act as a segregated module and hence contain (or trap) a signal within its local topology.

Module exit entropy (EE): Quantifies the uncertainty of a signal in taking a specific exiting node while escaping the local topology of an FN.

Functional magnetic resonance imaging (fMRI): A noninvasive imaging modality that estimates brain activity by detecting changes associated with levels of blood oxygenation. The rationale of this technique relies on the fact that there is an association between blood oxygenation and neuronal activation.

Functional reconfiguration: Quantifies the flexibility of an FN as a subject adapts to different cognitive tasks (excluding rest). In this work, it is represented by a two-dimensional spatial volume derived from a given FN's **EE** and **TE** coordinate values across different cognitive tasks.

Resting-state networks: Spontaneous brain activity is organized into a robust and reproducible (across subjects) set of localized and distributed networks, denoted resting-state networks (RSNs). One of the most common sets of RSNs divides the cortex into seven RSNs: visual (VIS), somatomotor (SM), dorsal attention (DA), ventral attention (VA), limbic (LIM), frontoparietal (FP), and default mode network (DMN). RSNs can be characterized by their functional connectivity in terms of within-network cohesion and between-network integration. RSNs can also be referred to as functional networks (FNs).

level interactions (Bassett et al., 2011). Hence, a mesoscopic scale (as the one provided by functional networks or communities/modules) may uncover differential patterns of (re)configuration (Mohr et al., 2016), across functional subcircuits, which might otherwise not be detectable at other scales. Traditionally, a mesoscopic assessment of functional brain networks would involve the *detection* of functional communities (Sporns & Betzel, 2016) either based on topology (density-based; Newman, 2006a, 2006b) or based on the information flow (flow-based; Rosvall, Axelsson, & Bergstrom, 2009; Rosvall & Bergstrom, 2008). These approaches, however, are not designed to *track* the dynamic behavior of a priori set of communities across time, tasks, and/or subjects. The primary aim of this work is to clearly define and quantify different configurations that FNs can assume, as well as measure their nature of reconfigurations switching between a seemingly infinite number of cognitive states. From a graph-theoretical perspective, FNs and their corresponding reconfigurations are described by two attributes: topology and communication. From a system dynamic perspective, FNs can be characterized by segregation and integration (Sporns, 2013) properties across which the human brain reconfigures across varied cognitive demands (J. Shine et al., 2018; J. M. Shine et al., 2016; J. M. Shine et al., 2019; J. M. Shine & Poldrack, 2018). To formally capture these diverse characteristics of FNs, we constructed a mathematically well-defined and well-behaved 2D "mesoscopic morphospace" based on two novel measures defined for nonnegative, undirected, weighted functional connectomes: trapping efficiency (**TE**) and exit entropy (**EE**). Trapping efficiency captures the level of segregation/integration of a functional network embedded in the rest of the functional connectome and quantifies the extent to which a particular FN "traps" an incoming signal. Exit entropy captures the specificity of integration of an FN with the rest of the functional connectome, and quantifies the uncertainty as to where (in terms of exit nodes) that same signal would exit the FN. In summary, this mesoscopic morphospace is a representation of the cognitive space as explored within and between cognitive states, as reflected by brain activity in fMRI. Such representation relies on FN reconfigurations that can be tracked, at an individual level, and at different granularity levels in network (re)configurations.

By using this 2D **TE**, **EE**-based morphospace, we formally study network configural breadth (Figure 1A), the most global and coarse grain exploration of the cognitive space, and its subsequent functional configuration components. To that end, we formally define measures of (a) functional reconfiguration (capacity of an individual to reconfigure across widely differing cognitive operations) and (b) functional preconfiguration (efficiency of transition from resting state to task-positive state (Schultz & Cole, 2016)), for potentially any community or FN. These measures are quantified for resting-state networks (Yeo et al., 2011) on the 100 unrelated subjects from the Human Connectome Project (HCP) dataset. We then study how such quantification is related to measures of cognitive abilities, such as fluid intelligence.

A MESOSCOPIC MORPHOSPACE OF FUNCTIONAL CONFIGURATIONS

The *mesoscopic morphospace* proposed here is a two-dimensional space built upon trapping efficiency and exit entropy measures for assessing functional networks or communities of functional connectomes. In this framework, functional connectomes must be undirected (symmetrical) weighted graphs, with *nonnegative* functional couplings. This framework allows for any a priori partition into functional communities. In this work, we assess the resting-state functional networks as proposed by Yeo et al. (2011) as the a priori FNs. Also, we use functional connectivity (without incorporating structural connectivity information), which is a quantification of statistical dependencies between BOLD time series of brain regions, and it can be used as a proxy of communication dynamics in the brain (Fornito, Zalesky, & Bullmore, 2016). Under this section, further technical details that are not mentioned in the main text will be directed to different subsections in the Supporting Information.

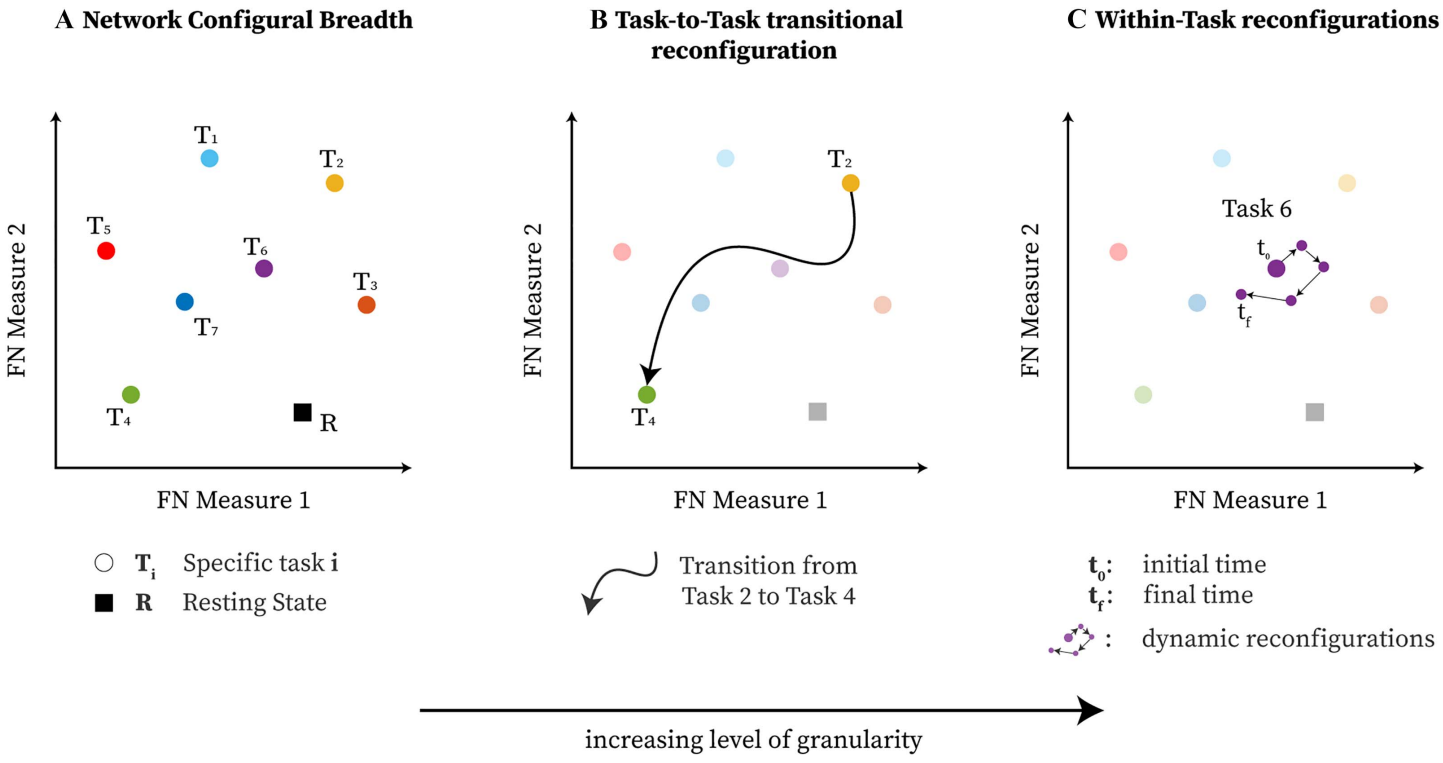


Figure 1. The three types of brain (re)configurations that can be represented by a mathematical space parameterized by, in this case, two generic phenotypic measures of functional communities of the brain: (**A**) network configural breadth, which represents changes across a number of cognitive demands; (**B**) task-to-task transitional reconfiguration; and (**C**) within-task reconfiguration.

Functional connectome/connectivity (FC) matrix:
A network representation of the functional coupling between brain regions. Such coupling is usually measured by quantifying the statistical dependencies between time series of brain regions (e.g., pairwise Pearson's correlation, mutual information) as obtained by functional magnetic resonance imaging (fMRI).

Computing Mechanistic Components for Morphospace Measures

A mesoscopic morphospace is constructed to assess functional network behaviors through two focal lenses: level of segregation/integration (using graph topology), and specificity of integration (using information theory). We first define all necessary components to compute **TE** and **EE** as follows:

(a) The whole-brain functional connectome (FC) is graph-theoretically denoted by $G(V, E)$, where V is the set of vertices (represented by the regions of interest, ROIs) and E is the set of edges (quantified by functional couplings between pairs of ROIs). The whole-brain FC is mathematically represented by an adjacency structure denoted as $\mathbf{A} = [w_{ij}]$, where i, j are indexed over vertex set V and $w_{ij} \in [0, 1]$ are functional couplings.

(b) Using a predefined set of FNs, a functional community (graph-theoretically denoted as $G_{\mathcal{C}}(V_{\mathcal{C}}, E_{\mathcal{C}})$ or for short) is defined to have the corresponding node set $V_{\mathcal{C}} \subset V$ and edge set $E_{\mathcal{C}} \subset E$ for which the union over all FNs exhaust the vertex and edge set of G such that

$$\cup V_{\mathcal{C}} = V \quad \text{and} \quad \cup E_{\mathcal{C}} = E.$$

(c) For a given functional community $\mathcal{C} \subset G$, define the set of states (or equivalently, vertices) S that contains the set of transient states (denoted as $S_{trans} = V_{\mathcal{C}}$), and absorbing states (denoted as $S_{abs} = \{j \mid w_{ij} > 0;\ j \notin V_{\mathcal{C}},\ \forall\ i \in V_{\mathcal{C}}\}$) such that

$$S = S_{trans} \cup S_{abs}.$$

(d) We mathematically denote a whole-brain FC as $\mathbf{A} = [w_{ij}]$ (see the Constructing Functional Connectomes section of the Supporting Information for more details), where

i and j are brain regions (from now on denoted as vertices or states) of the specified parcellation or atlas. Each matrix $\mathbf{A}$ represents a single subject, single session, single task whole-brain FC. We assess the whole-brain FC with respect to organizations into FNs, here denoted by $\mathcal{C}$. For a specific $\mathbf{A}$ and a specific $\mathcal{C}$, we obtain an induced submatrix $\mathbf{A}_{\mathcal{C}}$ by extracting the corresponding rows and columns of matrix $\mathbf{A}$ using only the vertices that belong to S, which results in the following matrix:

$$\mathbf{A}_{\mathcal{C}} \in (0,1)^{|S|\times|S|}.$$

We note that the row and column order of the states (or vertices) of $\mathbf{A}_{\mathcal{C}}$ respects the order of $S = S_{trans} \cup S_{abs}$ with transient states followed by absorbing ones, which results in a blockage structure:

$$\mathbf{A}_{\mathcal{C}} = \begin{array}{c} \\ \textbf{Transient} \\ \textbf{Absorbing} \end{array} \begin{array}{c} \begin{array}{cc} \textbf{Transient} & \textbf{Absorbing} \end{array} \\ \begin{pmatrix} \mathbf{A}(S_{trans}, S_{trans}) & \mathbf{A}(S_{trans}, S_{abs}) \\ \mathbf{A}(S_{abs}, S_{trans}) & \mathbf{A}(S_{abs}, S_{abs}) \end{pmatrix} \end{array},$$

where $\mathbf{A}(S_{trans}, S_{trans})$ means that we extract the submatrix of $\mathbf{A}$ that corresponds to states in S_{trans} for the rows (first argument) and S_{trans} for the columns (second argument).

(e) For any functional network $\mathcal{C}$, using the induced adjacency structure $\mathbf{A}_{\mathcal{C}}$ in the previous step, define each vertex in S to be a state in the stochastic process and construct the corresponding terminating Markov chain by computing the following:

- the normalization of $\mathbf{A}_{\mathcal{C}}$ by the nodal connectivity strength:

$$\mathbb{Q} = \mathbf{D}_{\mathcal{C}}^{-1}\mathbf{A}_{\mathcal{C}} \in (0,1)^{|S|\times|S|},$$

where $\mathbf{D}_{\mathcal{C}}$ is the weighted degree sequence matrix filled with the node strength (defined by the row [or equivalently, column] sum of $\mathbf{A}_{\mathcal{C}}$) in the diagonal entries and zeros for the off-diagonal elements:

$$\mathbf{D}_{\mathcal{C}} = \left[d_{ij}\right] = \begin{cases} \sum_{j=1}^{j=|V_{\mathcal{C}}|} w_{ij}, \forall i = j \\ 0, \forall i \neq j \end{cases},$$

where i, j are indexed over S. Note that the order of rows and columns of $\mathbb{Q}$ and $\mathbf{D}_{\mathcal{C}}$ also respect the order of S.

- the transition probability matrix of the terminating Markov chain:

$$\mathbf{P} = \begin{array}{c} \\ \textbf{Transient} \\ \textbf{Absorbing} \end{array} \begin{array}{c} \begin{array}{cc} \textbf{Transient} & \textbf{Absorbing} \end{array} \\ \begin{pmatrix} \mathbb{Q}(S_{trans}, S_{trans}) & \mathbb{Q}(S_{trans}, S_{abs}) \\ \mathbf{0}_{|S_{abs}|\times|S_{trans}|} & \mathbf{I}_{|S_{abs}|} \end{pmatrix} \end{array},$$

where $\mathbf{0}_{|S_{abs}|\times|S_{trans}|}$ is the matrix of all zeros (size $|S_{abs}|$ rows by $|S_{trans}|$ columns); $\mathbf{I}_{|S_{abs}|}$ is identity matrix of size $|S_{abs}|$; the index $\mathcal{C}$ for $\mathbb{Q}$ and $\mathbf{P}$ is dropped for simplicity.

(f) Using matrix $\mathbf{P}$, we extract the submatrix induced by states in S_{trans} (denoted by $\mathbf{P}|_{S_{trans}}$). Note that $\mathbf{P}|_{S_{trans}} = \mathbb{Q}(S_{trans}, S_{trans})$ because rows and columns of $\mathbf{P}$ respect the order of S. We then compute the fundamental matrix (denoted as $\mathbf{Z}$; Kemeny & Snell, 1960), which contains the mean number of steps a specific transient state in S_{trans} is visited, for any pair of transient states in S_{trans}, before the random walker is absorbed by one of the states in S_{abs}:

$$\mathbf{Z} = (\mathbf{I}_{|S_{trans}} - \mathbf{P}|_{S_{trans}})^{-1} \in \mathbb{R}_{+}^{|S_{trans}|\times|S_{trans}|}.$$

(g) Compute the mean time to absorption (denoted as τ), which contains the mean number of steps that the random particle needs to be absorbed by one of the states in S_{abs}, given that it starts in some state in S_{trans}:

$$\tau = \mathbf{Z}\mathbf{1}_{|S_{trans}|} \in \mathbb{R}_{+}^{|S_{trans}|\times 1},$$

where $\mathbf{1}_{|S_{trans}|}$ is the all one vector of size $|S_{trans}|$.

(h) Compute the absorption probability matrix (denoted as Ψ), which contains the likelihood of being absorbed by one of the absorbing states, given that the stochastic process starts in some transient state:

$$\Psi = \mathbf{Z}\left[\mathbf{P}\big|_{S_{trans},S_{abs}}\right] \in \mathbb{R}_{+}^{|S_{trans}|\times|S_{abs}|},$$

where $\mathbf{P}|_{S_{trans},S_{abs}}$ is the subtransition probability matrix induced from (row) state S_{trans} and (column) state S_{abs}. Hence, $\mathbf{P}|_{S_{trans},S_{abs}} = \mathbb{Q}(S_{trans}, S_{abs})$.

Module Trapping Efficiency

Module trapping efficiency, denoted as **TE** (unit: $\frac{steps}{weight}$), quantifies a module's capacity to contain a random particle from leaving its local topology, that is, $\mathcal{C}$. Specifically, through FN topology, we want to assess its level of *segregation/integration*, measured by the L_2 norm of τ (unit: *steps*), that is, the mean time to absorption of nodes in $\mathcal{C}$, normalized by its total exiting strength (unit: *weight*), measured by

$$\mathcal{L}_{\mathcal{C}} = \sum_{i\in S_{trans}, j\in S_{abs}} A_{ij} = \mathbf{A}(S_{trans}, S_{abs}).$$

Mathematically, trapping efficiency is quantified as follows:

$$\mathbf{TE} = \frac{\|\tau\|_2}{\mathcal{L}_{\mathcal{C}}}. \tag{1}$$

We see that the mean time to absorption vector, τ, is dependent on both **density-based** (Fortunato, 2010; Newman, 2006b) and **flow-based** (Malliaros & Vazirgiannis, 2013; Rosvall et al., 2009; Rosvall & Bergstrom, 2008) modularity. The mean-time-to-absorption vector τ for which τ_i contains the average number of steps a random walker needs to escape the FN topology, given that it starts from node i. This means that the numerical values in τ are always greater than or equal to 1. We chose to use L_2 norms because it squares the input values of the vector and thus enhances our capacity to quantify FN (re)configuration. On the other hand, the denominator $\mathcal{L}_{\mathcal{C}}$ is a simple statistical summary of the module "leakages" to the rest of the cortex. Since all the values in $\mathcal{L}_{\mathcal{C}}$ are between (0, 1), L_2 norm would have diminished the differences across FNs. Hence, we chose L_1 norm for the denominator. The role of $\mathcal{L}_{\mathcal{C}}$ is to account for potential differences in trapping efficiency due to community size. Numerically, higher **TE** indicates that a module is more segregated (or equivalently, less integrated). This is because the FN topology traps the incoming signal efficiently, relative to its exiting edges when embedded in the cortex. **TE** value ranges are given in Figure 2.

Module Exit Entropy

Module exit entropy (denoted as **EE**, and in the range **EE** $\in$ (0, 1] and unitless) assesses the normalized level of uncertainty in selecting an exiting node in S_{abs} of a random particle that starts in $\mathcal{C}$. The exit entropy, denoted as $\mathcal{H}_e$, measures the level of uncertainty exiting node

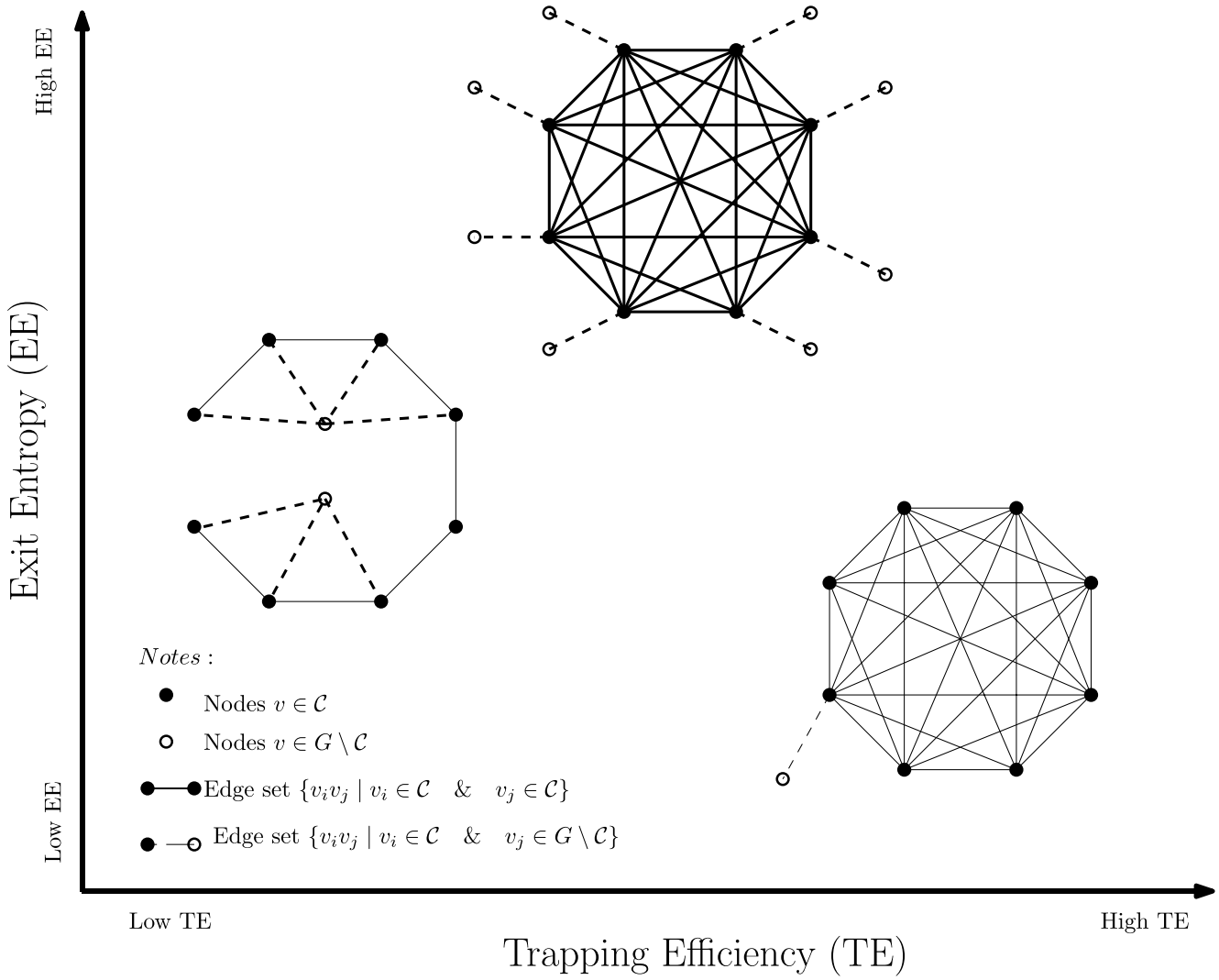


Figure 2. Morphospace measurements, examples. All three induced subgraphs have the same cardinality ($|\mathcal{C}| = 8$) with a different number of exits (connections to $G \setminus \mathcal{C}$). Nonetheless, depending on their topological structures, the corresponding morphospace measurements (**TE** and **EE**) have rather distinct values.

$j \in S_{abs}$ (outside of the module) is preferred. Module exit entropy is mathematically formalized as

$$\mathbf{EE} = \frac{\mathcal{H}_e}{\mathcal{N}_\mathcal{C}} = \frac{-\sum_{i=1}^{|S_{abs}|} \psi_i \log(\psi_i)}{\log(|S_{abs}|)}, \tag{2}$$

where preferential exit probability is the probability vector that contains $|S_{abs}|$ entries that represents the likelihood that exit signal selects a specific exiting state $j \in S_{abs}$ such that $\Sigma_{j \in S_{abs}} \psi_j = 1$.

The numerator of **EE**($\mathcal{C}$), that is, $-\sum_{i=1}^{|S_{abs}|} \psi_i \log(\psi_i)$, measures the degree to which channels of communication between nodes in S_{trans} and S_{abs} are preferred for a fixed task/subject. It is noteworthy that **EE** is not influenced by the (cumulative) magnitudes (of functional connectivity values) that connect nodes from within the FN to outside (exiting) nodes. It is only affected by the distribution of such values. In particular, homogeneous distributions display high entropy levels, and uneven distributions favoring certain exiting node(s) display low entropy. To

demonstrate this point, an example is provided in the Supporting Information under the Module Exit Entropy section. The normalizer, $\mathcal{N}_{\mathcal{C}} = \log(|S_{abs}|)$, is the maximum entropy obtained from a module in which all exit nodes have the same absorption rate. Numerically, a high **EE** would denote the homogeneous integration within the rest of the system, whereas a low **EE** would indicate a preferential communication or integration of the module with the rest of the system. In terms of functional brain networks, module exit entropy facilitates the understanding of collective behavior from $\mathcal{C}$ to other FNs through its outreach channels (edges formed by nodes in $\mathcal{C}$ and exiting nodes in $G \setminus \mathcal{C}$. This is because entropy measures the level of uncertainty in communication; hence, lower entropy means higher specificity in communication between the FN with the rest of the cortex. **EE** value ranges are given in Figure 2.

The Definition of the Mesoscopic Morphospace Ω

The two distinct features of each FN in brain graphs are addressed by a point $\mathbf{u}(\mathcal{C})$ in $\Omega \subset (0, M) \times [0, 1] \subset \mathbb{R}^2$ as follows:

$$\mathbf{u}(\mathcal{C}) = (\mathbf{TE}(\mathcal{C}), \mathbf{EE}(\mathcal{C})) \in \Omega, \tag{3}$$

where $M < \infty$. For a given subject and task, a functional brain network G is obtained with a predefined parcellation that results in I induced subgraph $\mathcal{C} \subset G$. We can then obtain I points $\mathbf{u}(\mathcal{C})$ corresponding to I FNs in network G.

In general, trapping efficiency $\mathbf{TE}(\mathcal{C})$ is finitely bounded by construction (see more details in the Module Trapping Efficiency section in the Supporting Information). However, a better bound is possible for the HCP dataset used for this study. This is due to two driving factors: connectome sparsity and edge weights (Avena-Koenigsberger, Goñi, Solé, & Sporns, 2015). We address the upper bound for **TE** as $\max(\mathbf{TE}(\mathcal{C})) = M = 1$. In terms of $\mathbf{EE}(\mathcal{C})$, its numerical range $\mathbf{EE}(\mathcal{C}) \in (0, 1]$. Hence, $\Omega \subset (0, 1) \times [0, 1]$ for this dataset.

THE NETWORK CONFIGURAL BREADTH FORMALISM

Studying the manifold topology defined in this 2D mesoscopic morphospace theoretically requires an infinite amount of points. In finite domain with discrete sampling of the morphospace, polytope theory, a mathematical branch that studies object geometry, allows us to create a reasonable scaffold presentation with well-defined properties to formally define and quantify configural components of the functional networks.

Polytope theory is a branch of mathematics that studies the geometry of shapes in a d-dimensional Euclidean space, $\mathbb{R}^d$. Given a set of points in this space, $W = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{|W|}\}$, a convex hull formed by W is represented by

$$\mathbf{Conv}(W) = \left\{ \sum_{j=1}^{|W|} \alpha_j \mathbf{x}_j \Big| \sum_{j=1}^{|W|} \alpha_j = 1, \alpha_j \geq 0 \right\}.$$

One can compute the notion of volume of the convex hull enclosed by $\mathbf{Conv}(W)$, denoted as $Vol(\mathbf{Conv}(W))$. Given that the morphospace is 2D, the manifold dimension can be from 0 up to 2. In the Supporting Information under the Polytope Theory section, further details on volume computation are defined.

The functional network configural breadth, for the ith subject, is compartmentalized into two components:

- FN (task) reconfiguration and
- FN rest-to-[task-positive] preconfiguration.

We then propose a mathematical relation between network configural breadth with FN reconfiguration and preconfiguration as follows:

$$\mathcal{F}_i = f\left(\mathcal{R}_i^{FN}, \mathcal{P}_i^{FN}\right), \tag{4}$$

where $\mathcal{F}_i$ represents configural breadth for subject *i*th. Here, we provide directly the measures that quantify (functional) reconfiguration and preconfiguration of FNs for *i*th subject's configural breadth. Tasks are assigned the same level of importance, and hence, no task is weighted more than others.

Functional Reconfiguration

Definition 1. *Functional reconfiguration in this work is represented by a two-dimensional spatial volume derived from given FN's **EE** and **TE** coordinate values across different cognitive tasks. As such, it represents an example of "cognitive space" (Varona & Rabinovich, 2016; Varoquaux et al., 2018) within a functional domain that spans a variety of network states under various task-evoked conditions. We quantify this as*

$$\mathcal{R}_i^{FN} = Vol\left(\mathbf{Conv}\left(W_i^{FN}\right)\right), \tag{5}$$

where W_i^{FN} represents the set containing all investigated task coordinates of subject i's FN; $\mathbf{Vol}(Conv(W_i^{FN}))$ is the convex hull volume induced by points in W_i^{FN}.

For a given subject *i*th's FN, note that $\mathbf{Conv}(W_i^{FN}$ represents the broad span (breadth) of task configurations for a given functional community. Subsequently, $\mathcal{R}_i^{FN}$ represents the amount of breadth as measured by the volume of $\mathbf{Conv}(W)$. Functional reconfiguration for a given subject's FN, denoted as $\mathcal{R}_i^{FN}$, is geometrically depicted in Figure 3.

Functional Preconfiguration

Functional preconfiguration: Reflects, for an FN, the ease of functional transition from a resting-state configuration to a task-positive state. In this work, it is represented using Euclidean distance between **TE** and **EE** coordinates of resting state and geometric centroid of the cognitive tasks.

Definition 2. *Functional preconfiguration reflects the topologically distributed equipotentiality that is theoretically designed to enable an efficient switch from a resting-state configuration to a task-positive state (Schultz & Cole, 2016), and is quantified as follows:*

$$\mathcal{P}_i^{FN} = \left\| Rest_i^{FN} - \eta_{W_i^{FN}} \right\|_2, \tag{6}$$

where $\eta_{W_i^{FN}}$ is the geometrical centroid of W_i^{FN}; $\mathcal{P}_i^{FN}$ measures the distance between rest to task general position (represented by $\eta_{W_i^{FN}}$). It is defined with the selected metric space, in this case it is the 2 norm in Euclidean space.

Note that functional preconfiguration can be viewed as $Vol(\mathbf{Conv}(W))$ where the convex hull is defined solely by two points: FN's rest and FN's geometrical centroid of task convex hull, that is, $W = \{R_i^{FN}, \eta_{W_i^{FN}}\}$. In such regards, the notion of $Vol(\mathbf{Conv}(W))$ is also suitable to describe the configural breadth between rest and task-positive location. Functional preconfiguration is geometrically depicted in Figure 3.

RESULTS

The mesoscopic morphospace formalized in the Mesoscopic Morphospace of Functional Configurations section is used to assess network configural breadth in terms of functional preconfiguration and reconfiguration for the 100 unrelated subjects of the HCP 900-subject data release (Van Essen et al., 2013; Van Essen et al., 2012). This dataset includes (test and retest)

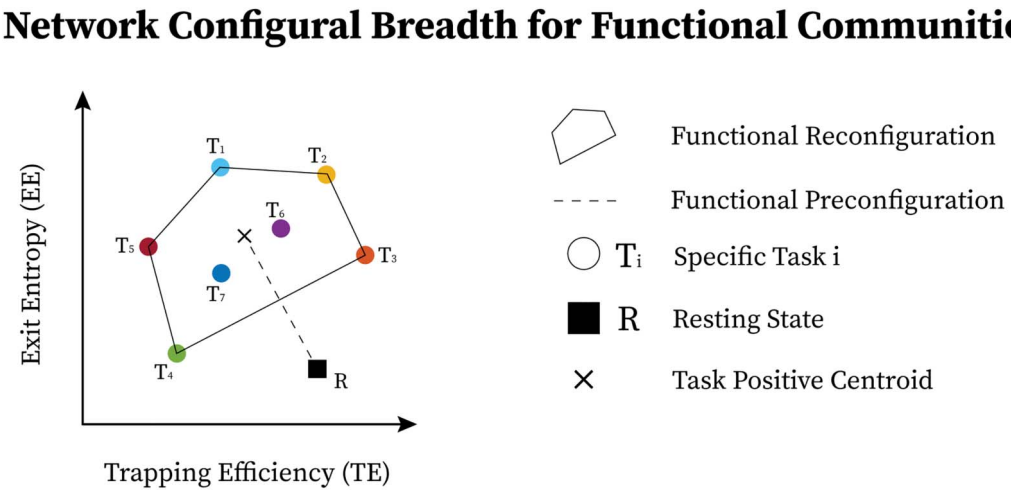


Figure 3. Functional network configural breadth is geometrically represented using two predefined morphospace measures. Specifically, for mesoscopic structures such as communities in functional brain networks, the first measure is trapping efficiency (**TE**) while the second is exit entropy (**EE**). In this case, tasks T1 to T5 belong to the convex hull (e.g., Pareto front; further details are available in the Supporting Information under the Polytope Theory section), while T6 and T7 are in the interior enclosed by the convex hull.

sessions for resting state and seven fMRI tasks: gambling (GAM), relational (REL), social (SOC), working memory (WM), language processing (LANG), emotion (EMOT), and motor (MOT). Whole-brain functional connectomes estimated from this fMRI dataset include 360 cortical brain regions (Glasser et al., 2016) and 14 subcortical regions. The functional communities evaluated in the morphospace include seven cortical resting-state FNs from Yeo et al. (2011); visual (VIS), somatomotor (SM), dorsal attention (DA), ventral attention (VA), frontoparietal (FP), limbic (LIM), default mode (DMN), and one composed of subcortical regions (SUBC). Additional details about the dataset are available in the Supporting Information, HCP Dataset and HCP Functional Data sections.

Task and Subject Sensitivity

Within- and between-subject task sensitivity. We first evaluate the capacity of module trapping efficiency and exit entropy to differentiate between tasks **within** subject (Figure 4A). For both test and retest sessions of each subject, we compute the **TE** and **EE** metrics for each FN. We compute these values for all eight fMRI conditions. We compute the intraclass correlation coefficient (ICC), with test and retest (per subject) being the repeated measurements and task being the class variable (**TE** in Figure 4A, top and **EE** in Figure 4A, bottom, respectively, where each ICC is computed using a 2 [test, retest] by 7 [tasks] design, and the ICC reflects task within-subject sensitivity). For most subjects, ICC values in all FNs are high and positive values. **EE** displays a higher within-subject task sensitivity than **TE**. Specifically, **TE** in VIS, DA, and DMN most distinguished between the cognitive tasks, whereas **EE** in VA and FP was best at distinguishing the within-subject task-based configural changes. The ICC values for both coordinates were the lowest for LIM.

We then evaluate the degree to which morphospace metrics capture cohort-level configural changes. To test this, for each morphospace metric (**TE** or **EE**), we compute ICC of each FN with subjects as the repeated measures and task as the class variable (Figure 4B). We performed the evaluation separately for test and retest sessions as denoted by gray and dark bars, respectively, for **TE** (Figure 4B, top) and **EE** (Figure 4B, bottom). **EE** captures cohort-level task-

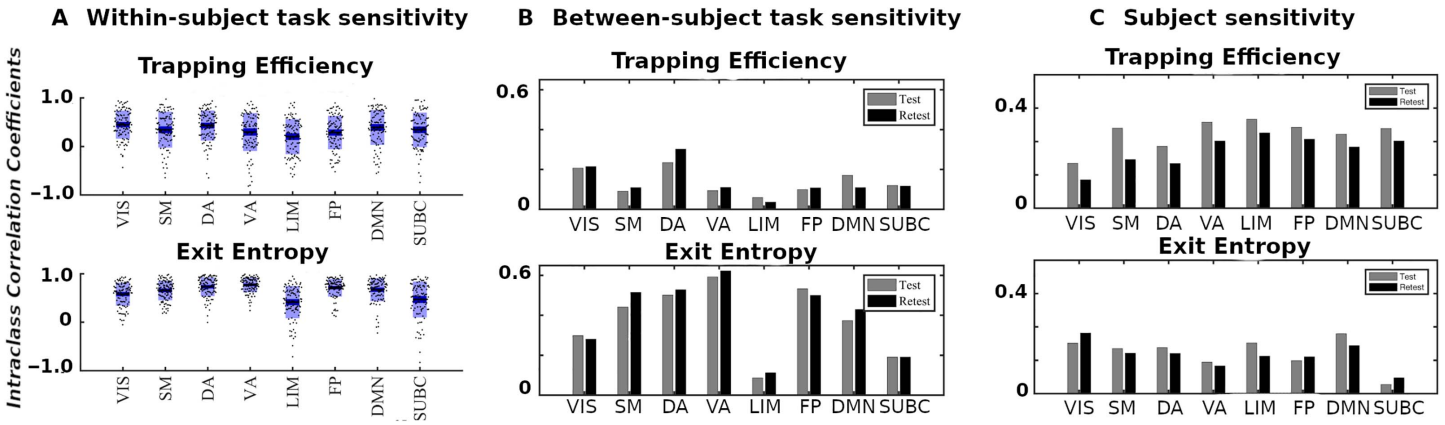


Figure 4. Morphospace measures and their task and subject sensitivity measured by intraclass correlation coefficients for each functional network. (**A**) Within-subject task sensitivity of module trapping efficiency (**TE**) and exit entropy (**EE**) for each FN per subject. (**B**) Between-subject task sensitivity of **TE** (top) and **EE** (bottom). (**C**) Subject-sensitivity ICC of **TE** (top) and **EE** (bottom).

based signatures as ICC values are consistently higher than those of **TE**. Interestingly, LIM has the lowest cohort-level task-based sensitivity for both morphospace metrics.

Subject sensitivity across tasks. Here, we compute ICC considering the tasks (fMRI conditions) the repeated measurements and considering subjects the class variable (Figure 4C). It is noteworthy that **TE** is superior in uncovering subject fingerprints, compared with **EE**, for the majority of FNs. This is complementary to **EE** being more task-sensitive.

TE and **EE** are disjoint features. Results in the Task and Subject Sensitivity section suggest that **TE** and **EE** have the differentiating capacity to highlight nonoverlapping characteristics of objects under consideration, that is, task- and subject-based FNs. First of all, for within-subject task differentiation (Figure 4A), FNs with high ICC values in one measure do not necessarily show a similar tendency in the other. For instance, VA has the third lowest mean **TE** value in characterizing within-subject task differentiation but it has the highest mean **EE** score. Similarly, FP has the second lowest average **TE** score and the third highest **EE** score, indicating that each of the two measures captures unique aspects of a given FN. Second, evidence of disjoint features is shown through the ICC results in cohort-level task-sensitivity (Figure 4B) and subject-sensitivity (Figure 4C) configural changes. Indeed, **TE** is superior in detecting subject fingerprints, while **EE** is better in unraveling task fingerprints. The idea is that, for a given studied object (i.e., task-based FNs), configurations are shown to "stretch" in exclusive/disjoint directions (subject-sensitive trapping efficiency and task-sensitive exit entropy).

Quantifying Network Configural Breadth on Functional Networks

The mesoscopic morphospace allows the quantification of network configural breadth. For a given functional community, we compute functional reconfiguration (degree of configurations across tasks) and preconfiguration (distance from rest to task-positive state), using Formulas 5 and 6, respectively.

Group-average results. The group-average behavior of functional communities is shown in Figure 5. Functional reconfiguration of FNs are shown as filled convex hulls, whereas preconfiguration of FNs are shown as dashed lines from rest to the corresponding task hull geometric centroid. To facilitate comparing network configural breadth across all functional networks,

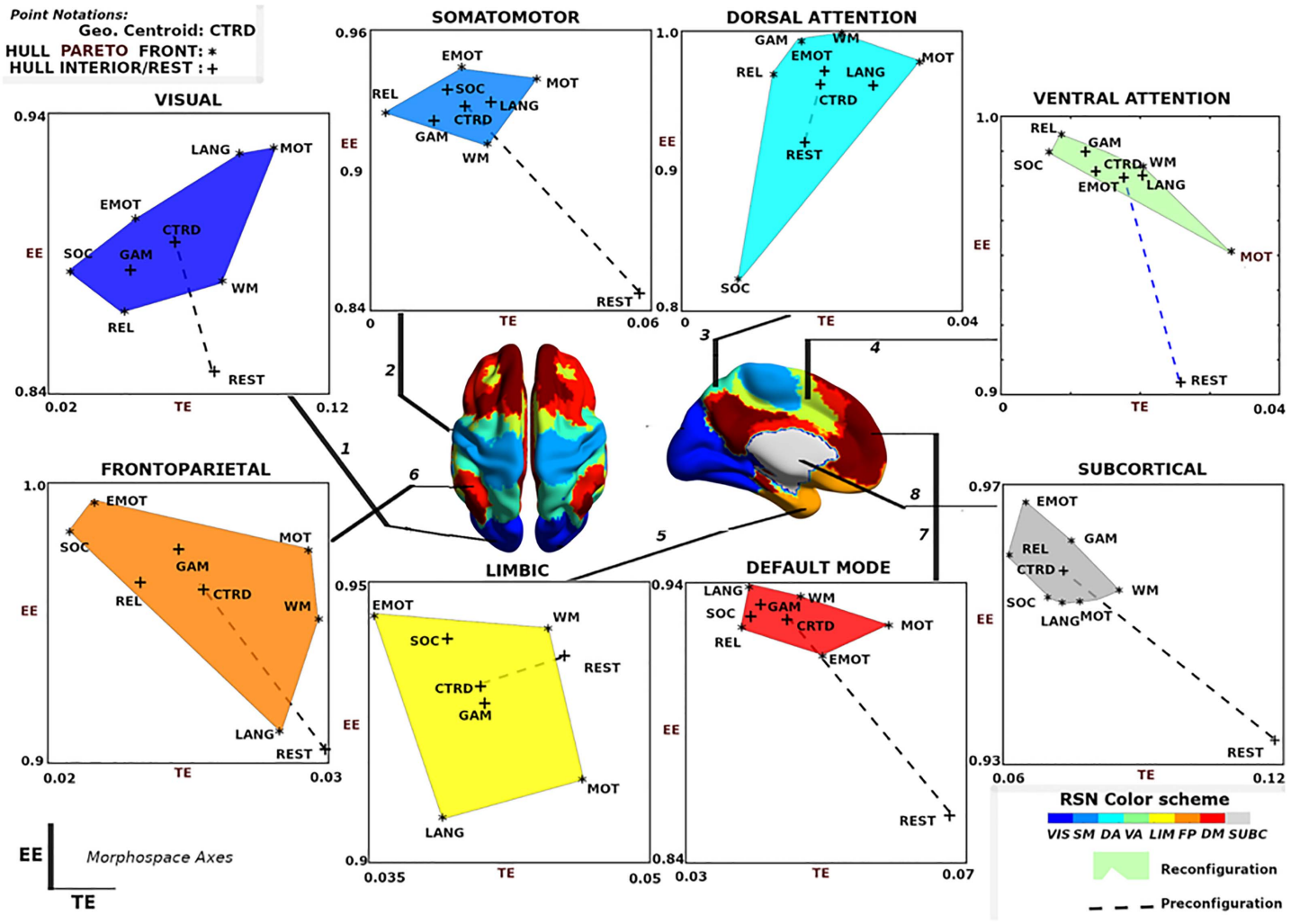


Figure 5. Visualization of network configural breadth. Functional reconfiguration and preconfiguration for all FNs are represented using group average of individual subjects' coordinates. Task coordinates in this space are represented by either an asterisk (*) or a plus (+) symbol. The asterisk symbol is used for those tasks that are part of the Pareto front of the convex hull; the plus symbol represents either the resting state or task that belongs to the interior of the convex hull. Note that x- and y-axis are purposely not scaled in the same range so that the full range of values for all tasks, task-centroid, and rest can be more easily visualized.

these same convex hulls are shown in Figure 6A with the same x- and y-axis values. VIS network polytope, representing group-average behavior, is lower in **EE** relative to other FNs.

With the exception of VIS and SUBC, all other FNs cluster in a similar, high **EE** / low **TE** area of the morphospace (Figure 6A). It should be noted that different tasks and subject populations (e.g., older or clinical groups) might cluster FNs differently. We also note that the subcortical polytope is relatively high in exit entropy. However, the subcortical parcellation might not optimally reflect the functional and/or structural makeup of various subcortical regions (e.g., role of the basal ganglia in the motor system), so these results should be interpreted cautiously.

One observation drawn from such a presentation is that the morphospace framework reconfirms, quantitatively, that functional dichotomy of the brain between task-positive and rest

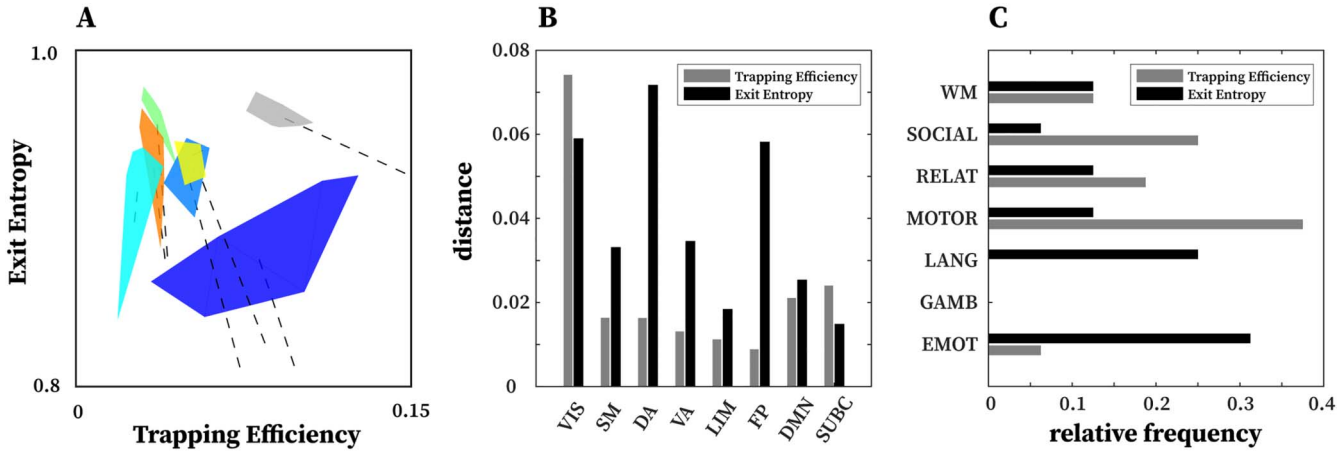


Figure 6. Network configural breadth insights on functional networks and tasks. (**A**) An illustration of network configural breadth for all functional communities. Polytope colors are analogous to the scheme shown in Figure 5. For each functional community, the dashed line represents the amount of functional preconfiguration, whereas the polytope volume represents the amount of functional reconfiguration. (**B**) Maximal distance is computed using the maximum pairwise distance between two tasks for a given functional network. (**C**) Relative frequency with which a task appears in the maximal distance normalized by 16 (8 FNs and 2 tasks per FN).

state (Fox et al., 2005). Specifically, the default mode network acts more as a segregated module with high level of integration specificity at rest - as seen in the lower right regime with high **TE**, low **EE** values - as opposed to under task-evoked conditions - as seen in the top left corner with low **TE**, high **EE** values (Figure 5, default mode; Fox et al., 2005; Greicius, Krasnow, Reiss, & Menon, 2003).

Another observation is that in terms of segregation level measured by **TE**, the lower bound of subcortical convex hull is, approximately, the upper bound of other FNs, with the exception of the visual network. Figures 7.1A and 7.2A also summarize functional reconfiguration and preconfiguration, respectively, for test and retest fMRI sessions in all subjects and FNs. Here, the VIS system displays the largest functional reconfiguration (see Figure 7.1A). From Figure 7.2A, functional preconfigurations display a more comparable magnitude among all FNs.

Further evidence of disjoint feature is also displayed in Figure 6B and 6C. In Figure 6B, maximal distance is computed using pairwise distances for two given tasks for a specific FN. The result shows that for a given FN, the two measures complement each other and in many cases, stretch the cognitive space in one direction or the other. For instance, in the case of DA and FP, the maximal distance in **EE** is very high but low for **TE**, whereas in VIS and SUBC, **TE** maximal distance is higher than that of **EE**. Furthermore, in Figure 6C, only specific tasks (e.g., motor and emotion) push the cognitive space in a particular direction (which is captured by maximal distance computation). Evidence of disjoint features is also illustrated by the relative frequency of motor and emotion tasks for which **TE** and **EE** are complementary.

Subject specificity of pre- and reconfiguration of functional networks. The formulation of network configural breadth (in terms of preconfiguration and reconfiguration) enables us to assess these properties at the subject level.

In Figure 7.1B and 7.2B, we use ICC to analyze the ability of morphospace measures (in the form of reconfiguration, panels Figure 7.1, and preconfiguration, panels Figure 7.2) to reflect subject identity within each FN. For all FNs from Yeo et al. (2011), the ICCs suggest that

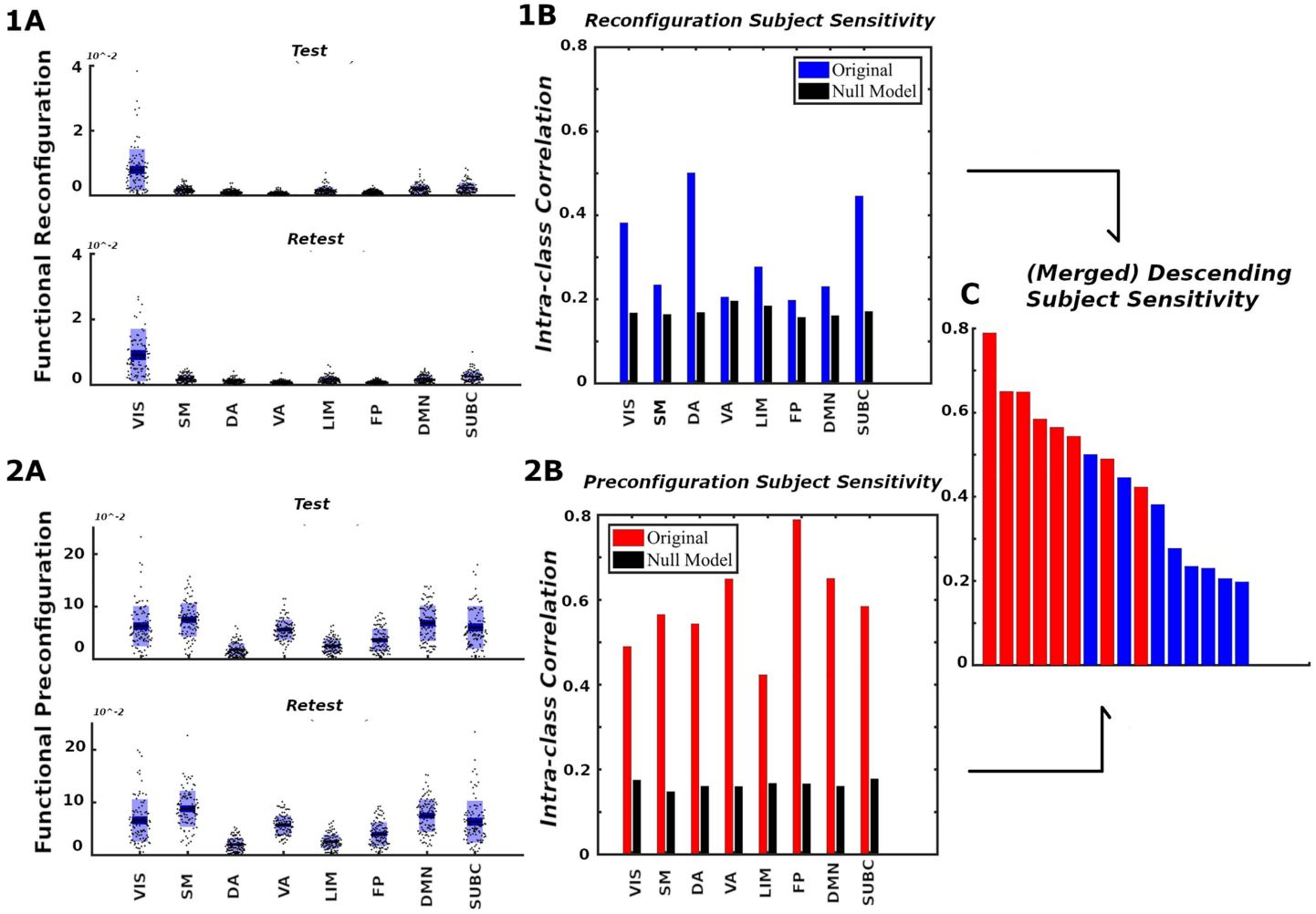


Figure 7. Network configural breadth, subject specificity analysis. Panels **1** and **2** show functional reconfiguration and preconfiguration, respectively, from both magnitude and subject-sensitivity viewpoints. For each functional network, the (**A**) panels report subject's preconfiguration and reconfiguration values whereas the (**B**) panels quantify subject sensitivity. Reconfiguration and preconfiguration measures are displayed in blue and red, respectively. Panel (**C**) merges all 16 configural breadth terms in descending order of subject sensitivity.

subjects can be differentiated from each other when contrasted against a corresponding null model (for details, see the Supporting Information, Subject Sensitivity section). We see that subject-sensitivity scores of all eight FNs for both pre- and reconfigurations are higher than their corresponding null models. Finally, for a fixed FN, functional preconfigurations dominated the subject sensitivity ranking, as illustrated by Figure 7C. Furthermore, FP, DMN, and VA preconfigurations are among the FNs with the highest subject fingerprints in overall subject-sensitivity ranking.

Network Configural Breadth and Behavior

Network configural breadth, compartmentalized into FN reconfiguration $\mathcal{R}^{FN}$ and preconfiguration $\mathcal{P}^{FN}$, shows a high level of subject sensitivity. This allows us to assume that $\mathcal{F}_i$ is associated with an individual's behavioral measures (denoted as $\succcurlyeq_i$ for subject ith). Several

studies reported that FP and DMN networks are associated with memory and intelligence (Gray, Chabris, & Braver, 2003; Schultz & Cole, 2016; Tschentscher, Mitchell, & Duncan, 2017). Therefore, we evaluated whether the outlined framework reflects four widely studied cognitive/behavioral measures, related to memory and intelligence: episodic memory, verbal episodic memory (verb. epi. mem.), fluid intelligence *gF*, and general intelligence *g*. While fluid intelligence reflects subject capacity to solve novel problems, general intelligence, *g*, reflects not only fluid intelligence, *gF*, traits but also crystallized (i.e., acquired) knowledge (Cattell, 1963, and typically denoted as *gC*). The early notion of general intelligence is conceptualized by Spearman's positive manifold (Spearman, 1904) that cannot be fully described using a single task. Quantification of *g* can be accomplished using subspace extraction techniques such as explanatory factor analysis (Dubois, Galdi, Paul, & Adolphs, 2018) or principal component analysis (PCA; Schultz & Cole, 2016). In this work, we quantified *g* using the PCA approach described in Schultz and Cole (2016). Mathematically, we propose the following composite relationship:

$$\succcurlyeq_i = \Upsilon\left(\mathcal{R}_i^{FN}, \mathcal{P}_i^{FN}\right). \tag{7}$$

Having established a plausible connection between behavioral measures and $\mathcal{P}^{FN}$, $\mathcal{R}^{FN}$, Equation 7 can be viewed as a multilinear model (MLM) using FN preconfiguration and reconfiguration as independent variables (or predictors). The MLM is constructed iteratively, starting

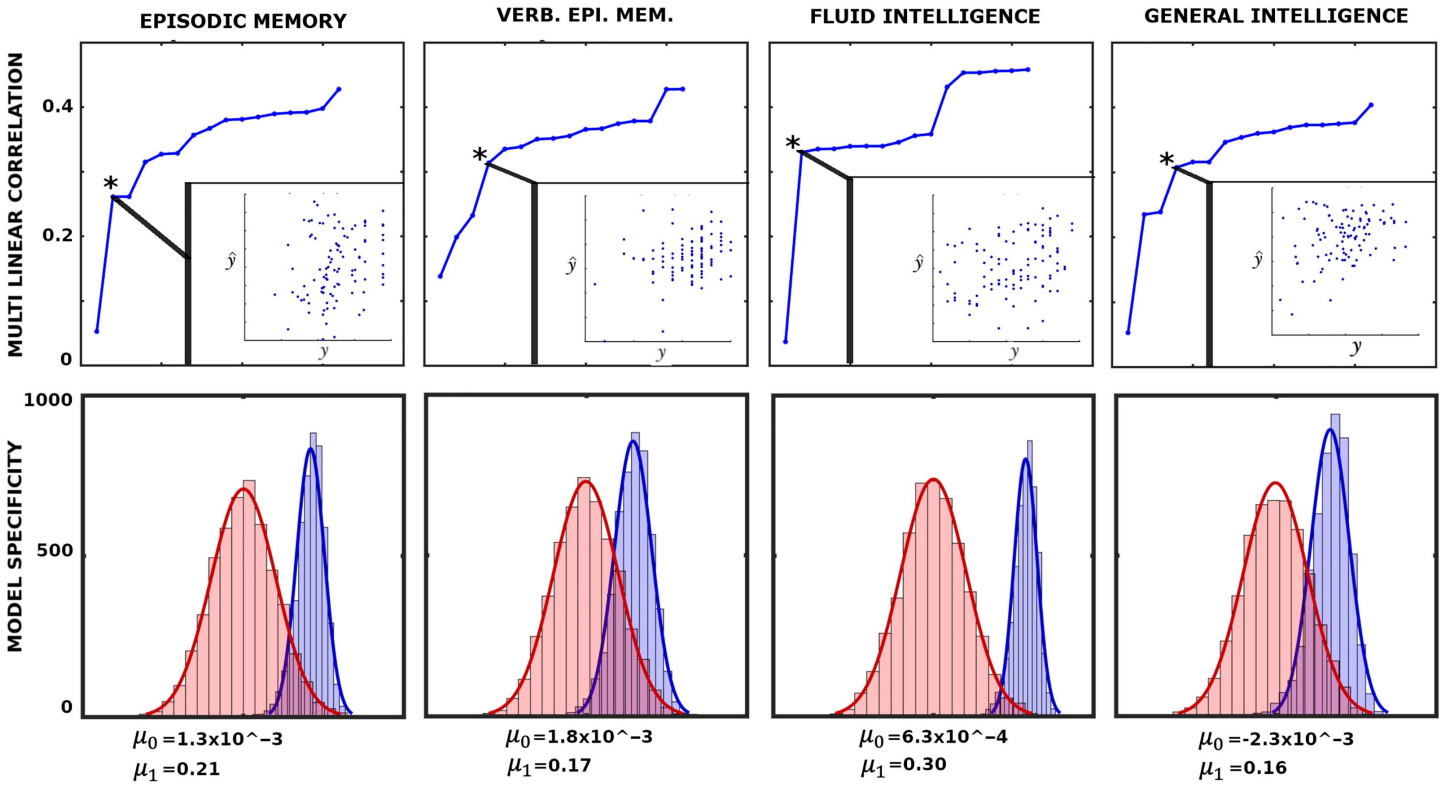


Figure 8. Associations between network configural breadth and behavior. The x-axis represents functional network preconfiguration and reconfiguration terms, that is, $\mathcal{P}_i^{FN}$ and $\mathcal{R}_i^{FN}$, ordered in decreasing subject fingerprints (as shown in Figure 7C). The top panels illustrate iterative multilinear regression model (MLM), while the bottom panels show model specificity (MS) for corresponding behavioral measures. Asterisk represents the optimal MLM with lowest *p* value. Further details are available in the Supporting Information, Behavioral Measure Analysis section.

with the descriptor with the highest individual fingerprints in Figure 7C. In each iteration, the subsequently ranked descriptor (according to Figure 7C) is appended to the existing ones. The best MLM (denoted with an asterisk in Figure 8), which determines the number of linear descriptors included the model, is selected based on the model p value.

To test the level of specificity in the model, we performed 2,000 simulations of k-fold cross validation where $k = 5$ between the selected MLM and the corresponding behavioral measure. Specifically, for each cross validation (per simulation), we obtain a correlation between the 20 left-out values (y) with the predicted values ($\hat{y}$). Hence, in each simulation we obtained five correlations and their mean value. It can be shown that those means follow a normal distribution (details shown in the Supporting Information). Lastly, to provide the level of specificity of linear descriptors, we present a corresponding null model where the same descriptors are evaluated to predict random vectors of appropriate size. To test our model and its ability to predict the behavioral measures, we rely completely on network configural breadth predictors ranked in descending order of subject specificity.

The top panels in Figure 8 show that as more linear descriptors (FN's functional pre- and reconfigurations) are added to iterative MLMs, variance associating with behavioral/cognitive performance measures decreases with linear descriptors that bear less subject sensitivity. This result highlights the importance of appending linear predictors in descending order with respect to the subject sensitivity. Specifically, as individual specificity reduces from left to right (Figure 7C), the differential correlations, that is, the difference between two consecutive correlation values, decreases.

DISCUSSION

In this work, we fill an existing gap in the field of network neuroscience by proposing a mathematical framework that captures the extent to which subject-level functional networks, as estimated by fMRI, reconfigure across diverse mental/emotional states. We first propose that brain networks can undergo three different types of (re)configurations: (a) network configural breadth, (b) task-to-task transitional reconfiguration, and (c) within-task reconfiguration. Unlike other existing frameworks (Schultz & Cole, 2016; J. M. Shine et al., 2019; J. M. Shine & Poldrack, 2018), the framework presented here can be applied to all three reconfiguration types. As a first step, we focus on assessing the broadest aspect of reconfiguration, that is, network configural breadth. We postulate, based on previous literature (Cole et al., 2014), that macroscale (whole-brain) and microscale (edge-level) reconfigurations of brain networks are subtle, and hence difficult to disentangle. At the same time, mesoscopic structures in the brain (e.g., functional networks, FNs) reconfigure substantially across different mental/emotional states as elicited by different tasks (Mohr et al., 2016). The framework presented here constitutes the first attempt to formalize such (re)configurations of mesoscopic structures of the brain, and quantify the behavior of a reference set of FNs with changing mental states. We set forth a mathematically well-defined and well-behaved 2D network morphospace using novel mesoscopic metrics of trapping efficiency (**TE**) and exit entropy (**EE**). This morphospace characterizes not only the topology of FNs but also the flow of information within and between FNs. We show that this morphospace is sensitive to FNs, tasks, subjects, and the levels of cognitive performance. We show that both of these measures are highly subject-sensitive for some FNs, while preconfiguration is highly subject-sensitive for all of them. Lastly, we also formalize and quantify the concepts of functional reconfiguration (the extent to which an FN has the capacity to reconfigure across different tasks) and functional preconfiguration (amount of transition from resting-state to a task-positive centroid). We thus construct a

formalism that can explore FN changes across different cognitive states in a comprehensive manner and at different levels of granularity.

Ideally, a morphospace framework (Avena-Koenigsberger et al., 2015; Avena-Koenigsberger, Misic, & Sporns, 2018; Corominas-Murtra, Goñi, Solé, & Rodríguez-Caso, 2013; Goñi et al., 2013; McGhee, 1999; Morgan, Achard, Termenon, Bullmore, & Vértes, 2018; Schuetz, Zamboni, Zampieri, Heinemann, & Sauer, 2012; Shoval et al., 2012; Thomas, Shearman, & Stewart, 2000) would have a minimal complexity and, in this particular case, capture distinct features of functional network changes. As discussed in Avena-Koenigsberger et al. (2015), metrics parametrizing a given morphospace should be disjoint. We see that, for any specific FN, high within-subject task sensitivity of **TE** does not necessarily imply a high value in **EE** and vice versa (e.g., VA and FP in Figure 4A). In addition, we see that both **TE** and **EE** offer their unique insights in capturing nonoverlapping features, with **TE** being more subject-sensitive and **EE** more task-sensitive at the cohort level (Figure 4B, 4C). Figure 6B highlights the disjoint nature of the two metrics as well, where we compute maximal distance per FN polytope in the TE and the EE axes separately. Results show that corresponding **TE** and **EE** maximal distances are disjoint and FN dependent. In other words, for a specific FN, the polytope is "stretched" in a particular task direction, where each morphospace measurement (**TE** or **EE**) unravels distinct properties. In Figure 6C, we further see that a subset of tasks dominantly contribute to the maximal distance computation, such as motion, language, and social tasks. Interestingly, we see that motion and language tasks can be considered "orthogonal" tasks with respect to **TE** and **EE**.

Interestingly, the limbic network possesses the lowest ability to distinguish between tasks (Figure 4). Similar behavior has been observed in Amico, Arenas, and Goñi (2019) when using Jensen-Shannon divergence as a distance metric of functional connectivity. In addition, the limbic network seems to work as a "relay" in brain communication (Amico, Abbas, et al., 2019). One potential explanation for this unique behavior is that the limbic network maintains a minimal cognitive load across various tasks, most of which comprises relaying information from one part of the brain to the others; it thus does not reconfigure as much across different mental states.

Brain network configuration is typically studied considering a specific task at multiple spatial and temporal scales (see Bassett et al., 2011; Betzel et al., 2017; Mohr et al., 2016; J. Shine et al., 2018; J. M. Shine et al., 2016; J. M. Shine et al., 2019; J. M. Shine & Poldrack, 2018). Previous investigations have mainly focused on the mechanism of how the brain traverses between high/low cognitive demands (Amico, Arenas, & Goñi, 2019; Avena-Koenigsberger et al., 2018; Bertolero, Yeo, & D'Esposito, 2015; J. M. Shine et al., 2019; Sporns, 2013), or on periods of integration and segregation at rest (J. Shine et al., 2018; J. M. Shine et al., 2019; J. M. Shine & Poldrack, 2018), defined in this paper as within-task reconfigurations. On the other hand, whole-brain configurations have also been investigated across different tasks (one configuration per task) with respect to rest, which led to the concept of general efficiency (Schultz & Cole, 2016). This approach would belong to a wider category that we formally generalize as the network configural breadth. The idea of general efficiency in Schultz and Cole (2016) relied on whole-brain FC correlations between task(s) and rest. While intuitive in quantifying similarity/distance between a single task and rest, quantification across multiple tasks becomes a challenge. Specifically, note that in Schultz and Cole (2016), general efficiency is quantified using the first eigenmode, which explains most of the variance, after measuring the correlation between resting FC and three distinct task FCs. As more and more tasks are included, using the first eigenmode would become less and less representative of the task-related variations present in the data (in this paper summarized as the network configural

breadth). The proposed network morphospace overcomes these limitations and can be used to study brain network (re)configurations across any number of tasks. It allows us to study different types of brain network (re)configurations, as mentioned above, using one comprehensive mathematical framework, which also facilitates a meaningful comparison between these seemingly disparate kinds of (re)configurations. Schultz and Cole (2016) proposed that configurations can be compartmentalized into two differentiated concepts: functional reconfiguration and preconfiguration. Note that although the term **reconfiguration** is also used in Schultz and Cole (2016), it is not referring to the action of switching among multiple mental/emotional states, that is, as represented by task-to-task transitional reconfiguration or within-task reconfiguration (as shown in Figure 1B and 1C). Rather, it refers to the overall competence in exploring the total repertoire of task space of each subject given its resting configuration. That is why when we translate the corresponding idea into the mesoscopic morphospace, we call it the network configural breadth. We have also incorporated the two concepts of functional pre- and reconfigurations into a well-defined mathematical space, which solves some of the technical difficulties (as discussed in the Mesoscopic Morphospace of Functional Configurations section) and generalizes these concepts to mesoscopic structures.

Brain network within-task reconfigurations have been almost exclusively qualitatively assessed. For instance, J. M. Shine et al. (2016) show that the whole-brain functional connectome traverses segregated and integrated states as it reconfigures while performing a task. They also found that integrated states are associated with faster, more effective performance. Our formalism of within-task reconfigurations permits assessing such reconfigurations in a quantitative manner. Potentially, such within-task reconfigurations could also be used to assess cognitive fatigue, effort, or learning across time.

Cole et al. (2014) have shown that the resting architecture network modifies itself to fit task requirements through subtle changes in functional edges. Numerically, small changes constituted by functional edges between rest and task-based connectivity might not be statistically significant when looking at edge level. Moreover, we also observe that while such changes might be negligible on a whole-brain global scale, they are more evident when looking at subsystems or functional brain networks, as clearly observed in the VIS network, relative to others. For functional preconfiguration (Figure 5, Figure 6, Figure 7.2A), this effect is observable in all the FNs. In essence, we are postulating that a mesoscopic exploration of changes in brain network configurations with changing mental states is more informative than a macroscopic or microscopic exploration.

A key feature of this morphospace is that, in order to study brain network (re)configuration, an FN is not removed from the overall network for exploration. On the contrary, both metrics that define the morphospace, namely **TE** and **EE**, account for a particular FN's place embedded within the overall functional brain network, in terms of both topological structure and flow of information. That is why it is important to begin with a reference set of FNs (e.g., RSNs), so as to study how these FNs adapt to changing mental states within the context of the overall network.

Another benefit of a mesoscopic framework is that we can compare individual cognitive traits in each FN, instead of the whole brain (Figure 7.1B, 7.2B). Specifically, after quantifying reconfiguration and preconfiguration for all FNs, we determine whether these quantities incorporate information about individual traits (Figure 7C). We observe different levels of subject fingerprint in different FNs for both re- and preconfiguration measures. This subject fingerprint heterogeneity across different FNs is consistent with previous literature on functional connectome fingerprinting (Amico & Goñi, 2018; Finn et al., 2015). Interestingly, functional preconfiguration

Table 1. Multilinear regression models with corresponding standardized β coefficients. Dependent variables for each model are episodic memory, verbal episodic memory, fluid intelligence (gF), and general intelligence (g).

	Constant	$\mathcal{P}^{FP}$	$\mathcal{P}^{DMN}$	$\mathcal{P}^{VA}$	$\mathcal{P}^{SUBC}$
MLM terms/coefficients	β_0	β_1	β_2	β_3	β_4
Episodic memory	0.6	2.9	−9.3		
Verbal episodic memory	0.5	11.8	−1.1	−8.8	−6.1
gF	0.7	5.1	−12		
g	0.8	3.9	−5.5	−3.6	−5.7

(amount of transition from a resting state to a task-positive state) displayed greater subject fingerprint than functional reconfiguration for all FNs. Based on this observation, we argue that to have better subject differentiability, we need to design tasks where the subject transitions from a stable resting state to a task-positive state and/or vice versa (Amico et al., 2020). This could be a significant step forward in precision psychiatry (Fraguas, Díaz-Caneja, Pina-Camacho, Janssen, & Arango, 2016), where we can identify regional brain dysfunction more precisely as a function of the type and degree of cognitive or emotional load.

Subject sensitivity of the proposed network morphospace framework is also supported by significant associations of the frontoparietal and default mode networks with fluid intelligence; see Tables 1 and 2. Specifically, as pointed out by Tschentscher et al. (2017), high fluid intelligence is associated with a greater frontoparietal network activation, which is also consistent with findings from a three-back working memory task (Gray et al., 2003). In the domain of network configural breadth, we observe a higher reconfiguration as represented by a positive frontoparietal functional preconfiguration coefficient (Table 1).

This study has several limitations. The framework was tested specifically on the Human Connectome Project dataset and using a single whole-brain parcellation. Alternative parcellations (Schaefer et al., 2018; Tian, Margulies, Breakspear, & Zalesky, 2020), additional fMRI tasks to better sample the cognitive space, and other datasets might offer further insights about the mesoscopic network morphospace (see Avena-Koenigsberger et al., 2015; Corominas-Murtra et al., 2013). In addition, we did not perform a sensitivity analysis on how small fluctuations in functional connectomes affect mapping into the network morphospace. Because of the nature of module trapping efficiency and exit entropy metrics, negative functional couplings were not considered, and hence were set to zero. In future work, other combinations of L_1 and L_2 norms, or even other norm choices, should be evaluated when defining trapping

Table 2. Multilinear models with corresponding p values. Note that we do not use stepwise linear model which discards descriptors that are not statistically significant. Column entire model shows the significance of the entire model.

	Constant	$\mathcal{P}^{FP}$	$\mathcal{P}^{DMN}$	$\mathcal{P}^{VA}$	$\mathcal{P}^{SUBC}$	
MLM terms/p values	p_0	p_1	p_2	p_3	p_4	Entire model
Episodic memory	0	0.57	0.01			0.03
Verbal episodic memory	0	0.02	0.77	0.17	0.03	0.04
gF	0	0.30	9×10^{-4}			0.004
g	0.03	0.44	0.16	0.57	0.05	0.05

efficiency. This would impact not only the magnitude of the morphospace measure but also the differentiating capacity of configuration across different functional networks.

Future studies should incorporate a sensitivity study of the behavior of this network morphospace with respect to small fluctuations in the input functional connectomes. Further studies could also incorporate structural connectivity information to inform both **TE** and **EE** measures when assessing the morphospace coordinates of functional reconfiguration. Additional exploration of different aspects of this morphospace could provide further insights. For example, location of the polytopes in the morphospace might improve individual fingerprint. An important aspect of the proposed mesoscopic network morphospace is that it allows for an exhaustive and continuous exploration of network reconfigurations, including those that are continuous in time (Douw et al., 2016; J. M. Shine et al., 2019), for example, if the subject performs several tasks within the same scanning session, including extended resting-state periods (such as the fMRI experiment done at Barnes, Bullmore, & Suckling, 2009). This would allow us to fully explore the cognitive space and gain a valuable insight into how different subjects adapt to different levels of cognitive demands. One can also study the trajectory of changing mental states using dynamic functional connectivity (Gonzalez-Castillo et al., 2015), which can easily be mapped to this morphospace for additional insights. Another potential avenue could be the application of this framework to characterize and understand different brain disorders.

In summary, this mesoscopic network morphospace is our first attempt to create a mathematically well-defined framework to explore an individual's cognitive space at different levels of granularity. It allows us to characterize the structure and dynamics of specific subsystems in the brain. This type of framework can be extremely helpful in characterizing brain dynamics at the individual level, in healthy and pathological populations, which in turn would pave the way for the development of personalized medicine for brain disorders.

METHODOLOGY

We provide detailed information on materials and methods in the Supporting Information. In short, all necessary mechanics collected from multiple disciplines and general setup for matrix computations are described in main text under the Mesoscopic Morphospace of Functional Configurations section and Supporting Information Preliminaries and Data sections. The dataset consists of high-resolution functional connectivity matrices describing human cerebral cortex and subcortex (see Supporting Information, Data). The construction of morphospace and the formalized notion of configural breadth are described in the Supporting Information, Morphospace Analysis section. Multilinear model and model specificity are described in Supporting Information, Behavioral Measure analysis section.

ACKNOWLEDGMENTS

Data were provided (in part) by the Human Connectome Project, WU-Minn Consortium (principal investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. JG acknowledges financial support from NIH R01EB022574 and NIH R01MH108467 and the Indiana Clinical and Translational Sciences Institute (Grant Number UL1TR001108) from the National Institutes of Health, National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award. MV and JG acknowledge financial support from Purdue Industrial Engineering Frontier Teams Network Morphospace Award and from Purdue

Discovery Park Data Science Award "Fingerprints of the Human Brain: A Data Science Perspective." We thank Dr. Olaf Sporns and Meenusree Rajapandian for valuable comments.

SUPPORTING INFORMATION

Supporting information for this article is available at https://doi.org/10.1162/netn_a_00193.

AUTHOR CONTRIBUTIONS

Duy Anh Duong-Tran: Conceptualization; Formal analysis; Investigation; Methodology; Writing – original draft. Kausar Abbas: Investigation; Writing – original draft. Enrico Amico: Conceptualization; Formal analysis; Methodology; Visualization. Bernat Corominas-Murtra: Conceptualization; Formal analysis; Investigation; Methodology. Mario Dzemidzic: Data curation; Methodology; Writing – original draft. David Kareken: Conceptualization; Supervision; Writing – original draft. Mario Ventresca: Conceptualization; Supervision. Joaquin Goñi: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Supervision; Writing – original draft.

FUNDING INFORMATION

Joaquin Goñi, National Institutes of Health (https://dx.doi.org/10.13039/100000002), Award ID: NIH R01EB022574. Joaquin Goñi, National Institutes of Health (https://dx.doi.org/10.13039/100000002), Award ID: R01MH108467.

REFERENCES

Amico, E., Abbas, K., Duong-Tran, D. A., Tipnis, U., Rajapandian, M., Chumin, E., … Goñi, J. (2019). Towards a mathematical theory of communication for the human connectome. *arXiv:1911.02601*

Amico, E., Arenas, A., & Goñi, J. (2019). Centralized and distributed cognitive task processing in the human connectome. *Network Neuroscience*, *3*(2), 455–474. https://doi.org/10.1162/netn_a_00072, PubMed: 30793091

Amico, E., Dzemidzic, M., Oberlin, B. G., Carron, C. R., Harezlak, J., Goñi, J., & Kareken, D. A. (2020). The disengaging brain: Dynamic transitions from cognitive engagement and alcoholism risk. *NeuroImage*, *209*, 116515. https://doi.org/10.1016/j.neuroimage.2020.116515, PubMed: 31904492

Amico, E., & Goñi, J. (2018). The quest for identifiability in human functional connectomes. *Scientific Reports*, *8*(1), 8254. https://doi.org/10.1038/s41598-018-25089-1, PubMed: 29844466

Avena-Koenigsberger, A., Goñi, J., Solé, R., & Sporns, O. (2015). Network morphospace. *Journal of the Royal Society Interface*, *12*(103), 20140881. https://doi.org/10.1098/rsif.2014.0881, PubMed: 25540237

Avena-Koenigsberger, A., Misic, B., & Sporns, O. (2018). Communication dynamics in complex brain networks. *Nature Reviews Neuroscience*, *19*(1), 17. https://doi.org/10.1038/nrn.2017.149, PubMed: 29238085

Barnes, A., Bullmore, E. T., & Suckling, J. (2009). Endogenous human brain dynamics recover slowly following cognitive effort. *PLoS ONE*, *4*(8), e6626. https://doi.org/10.1371/journal.pone.0006626, PubMed: 19680553

Bassett, D. S., & Gazzaniga, M. S. (2011). Understanding complexity in the human brain. *Trends in Cognitive Sciences*, *15*(5), 200–209. https://doi.org/10.1016/j.tics.2011.03.006, PubMed: 21497128

Bassett, D. S., Wymbs, N. F., Porter, M. A., Mucha, P. J., Carlson, J. M., & Grafton, S. T. (2011). Dynamic reconfiguration of human brain networks during learning. *Proceedings of the National Academy of Sciences*, *108*(18), 7641–7646. https://doi.org/10.1073/pnas.1018985108, PubMed: 21502525

Bertolero, M. A., Yeo, B. T., & D'Esposito, M. (2015). The modular and integrative functional architecture of the human brain. *Proceedings of the National Academy of Sciences*, *112*(49), E6798–E6807. https://doi.org/10.1073/pnas.1510619112, PubMed: 26598686

Betzel, R. F., Satterthwaite, T. D., Gold, J. I., & Bassett, D. S. (2017). Positive affect, surprise, and fatigue are correlates of network flexibility. *Scientific Reports*, *7*(1), 520. https://doi.org/10.1038/s41598-017-00425-z, PubMed: 28364117

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*(1), 1. https://doi.org/10.1037/h0046743

Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S., & Petersen, S. E. (2014). Intrinsic and task-evoked network architectures of the human brain. *Neuron*, *83*(1), 238–251. https://doi.org/10.1016/j.neuron.2014.05.014, PubMed: 24991964

Corominas-Murtra, B., Goñi, J., Solé, R. V., & Rodríguez-Caso, C. (2013). On the origins of hierarchy in complex networks. *Proceedings of the National Academy of Sciences*, *110*(33),

13316–13321. https://doi.org/10.1073/pnas.1300832110, PubMed: 23898177

Douw, L., Wakeman, D. G., Tanaka, N., Liu, H., & Stufflebeam, S. M. (2016). State-dependent variability of dynamic functional connectivity between frontoparietal and default networks relates to cognitive flexibility. *Neuroscience*, *339*, 12–21. https://doi.org/10.1016/j.neuroscience.2016.09.034, PubMed: 27687802

Dubois, J., Galdi, P., Paul, L. K., & Adolphs, R. (2018). A distributed brain network predicts general intelligence from resting-state human neuroimaging data. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *373*(1756), 20170284. https://doi.org/10.1098/rstb.2017.0284, PubMed: 30104429

Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., … Constable, R. T. (2015). Functional connectome fingerprinting: Identifying individuals using patterns of brain connectivity. *Nature Neuroscience*, *18*(11), 1664. https://doi.org/10.1038/nn.4135, PubMed: 26457551

Fornito, A., Zalesky, A., & Bullmore, E. (2016). *Fundamentals of brain network analysis*. Academic Press.

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, *486*(3), 75–174. https://doi.org/10.1016/j.physrep.2009.11.002

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, *102*(27), 9673–9678. https://doi.org/10.1073/pnas.0504136102, PubMed: 15976020

Fraguas, D., Díaz-Caneja, C. M., Pina-Camacho, L., Janssen, J., & Arango, C. (2016). Progressive brain changes in children and adolescents with early-onset psychosis: A meta-analysis of longitudinal MRI studies. *Schizophrenia Research*, *173*(3), 132–139. https://doi.org/10.1016/j.schres.2014.12.022, PubMed: 25556081

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., … Van Essen, D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, *536*(7615), 171–178. https://doi.org/10.1038/nature18933, PubMed: 27437579

Goñi, J., Avena-Koenigsberger, A., de Mendizabal, N. V., van den Heuvel, M. P., Betzel, R. F., & Sporns, O. (2013). Exploring the morphospace of communication efficiency in complex networks. *PLoS ONE*, *8*(3), e58070. https://doi.org/10.1371/journal.pone.0058070, PubMed: 23505455

Gonzalez-Castillo, J., Hoy, C. W., Handwerker, D. A., Robinson, M. E., Buchanan, L. C., Saad, Z. S., & Bandettini, P. A. (2015). Tracking ongoing cognition in individuals using brief, whole-brain functional connectivity patterns. *Proceedings of the National Academy of Sciences*, *112*(28), 8762–8767. https://doi.org/10.1073/pnas.1501242112, PubMed: 26124112

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, *6*(3), 316. https://doi.org/10.1038/nn1014, PubMed: 12592404

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences*, *100*(1), 253–258. https://doi.org/10.1073/pnas.0135058100, PubMed: 12506194

Kemeny, J. G., & Snell, J. L. (1960). *Finite Markov chains* (Vol. 356). Princeton, NJ: van Nostrand.

Malliaros, F. D., & Vazirgiannis, M. (2013). Clustering and community detection in directed networks: A survey. *Physics Reports*, *533*(4), 95–142. https://doi.org/10.1016/j.physrep.2013.08.002

McGhee, G. R. (1999). *Theoretical morphology: The concept and its applications*. Columbia University Press.

Mohr, H., Wolfensteller, U., Betzel, R. F., Mišić, B., Sporns, O., Richiardi, J., & Ruge, H. (2016). Integration and segregation of large-scale brain networks during short-term task automatization. *Nature Communications*, *7*(1), 1–12. https://doi.org/10.1038/ncomms13217, PubMed: 27808095

Morgan, S. E., Achard, S., Termenon, M., Bullmore, E. T., & Vértes, P. E. (2018). Low-dimensional morphospace of topological motifs in human fMRI brain networks. *Network Neuroscience*, 2(*2*), 285–302. https://doi.org/10.1162/netn_a_00038, PubMed: 30215036

Newman, M. E. (2006a). Finding community structure in networks using the eigenvectors of matrices. *Physical Review E*, *74*(3), 036104. https://doi.org/10.1103/PhysRevE.74.036104, PubMed: 17025705

Newman, M. E. (2006b). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*(23), 8577–8582. https://doi.org/10.1073/pnas.0601602103, PubMed: 16723398

Rosvall, M., Axelsson, D., & Bergstrom, C. T. (2009). The map equation. *The European Physical Journal: Special Topics*, *178*(1), 13–23. https://doi.org/10.1140/epjst/e2010-01179-1

Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, *105*(4), 1118–1123. https://doi.org/10.1073/pnas.0706851105, PubMed: 18216267

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., … Yeo, B. T. (2018). Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cerebral Cortex*, *28*(9), 3095–3114. https://doi.org/10.1093/cercor/bhx179, PubMed: 28981612

Schuetz, R., Zamboni, N., Zampieri, M., Heinemann, M., & Sauer, U. (2012). Multidimensional optimality of microbial metabolism. *Science*, *336*(6081), 601–604. https://doi.org/10.1126/science.1216882, PubMed: 22556256

Schultz, D. H., & Cole, M. W. (2016). Higher intelligence is associated with less task-related brain network reconfiguration. *Journal of Neuroscience*, *36*(33), 8551–8561. https://doi.org/10.1523/JNEUROSCI.0358-16.2016, PubMed: 27535904

Shine, J., Breakspear, M., Bell, P., Ehgoetz, K. M., Shine, R., Koyejo, O., … Poldrack, R. (2018). The dynamic basis of cognition: An integrative core under the control of the ascending neuromodulatory system. https://doi.org/10.1101/266635

Shine, J. M., Bissett, P. G., Bell, P. T., Koyejo, O., Balsters, J. H., Gorgolewski, K. J., … Poldrack, R. A. (2016). The dynamics of functional brain networks: Integrated network states during cognitive task performance. *Neuron*, *92*(2), 544–554. https://doi.org/10.1016/j.neuron.2016.09.018, PubMed: 27693256

Shine, J. M., Breakspear, M., Bell, P. T., Martens, K. A. E., Shine, R., Koyejo, O., … Poldrack, R. A. (2019). Human cognition involves the dynamic integration of neural activity and neuromodulatory systems. *Nature Neuroscience*, *22*(2), 289. https://doi.org/10.1038/s41593-018-0312-0, PubMed: 30664771

Shine, J. M., & Poldrack, R. A. (2018). Principles of dynamic network reconfiguration across diverse brain states. *NeuroImage*, *180*, 396–405. https://doi.org/10.1016/j.neuroimage.2017.08.010, PubMed: 28782684

Shoval, O., Sheftel, H., Shinar, G., Hart, Y., Ramote, O., Mayo, A., … Alon, U. (2012). Evolutionary trade-offs, pareto optimality, and the geometry of phenotype space. *Science*, 1217405. https://doi.org/10.1126/science.1217405, PubMed: 22539553

Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology*, *15*(2), 201–292. https://doi.org/10.2307/1412107

Sporns, O. (2013). Network attributes for segregation and integration in the human brain. *Current Opinion in Neurobiology*, *23*(2), 162–171. https://doi.org/10.1016/j.conb.2012.11.015, PubMed: 23294553

Sporns, O., & Betzel, R. F. (2016). Modular brain networks. *Annual Review of Psychology*, *67*, 613–640. https://doi.org/10.1146/annurev-psych-122414-033634, PubMed: 26393868

Thomas, R., Shearman, R. M., & Stewart, G. W. (2000). Evolutionary exploitation of design options by the first animals with hard skeletons. *Science*, *288*(5469), 1239–1242. https://doi.org/10.1126/science.288.5469.1239, PubMed: 10817998

Tian, Y., Margulies, D. S., Breakspear, M., & Zalesky, A. (2020). Hierarchical organization of the human subcortex unveiled with functional connectivity gradients. *bioRxiv*. https://doi.org/10.1101/2020.01.13.903542

Tschentscher, N., Mitchell, D., & Duncan, J. (2017). Fluid intelligence predicts novel rule implementation in a distributed frontoparietal control network. *Journal of Neuroscience*, *37*(18), 4841–4847. https://doi.org/10.1523/JNEUROSCI.2478-16.2017, PubMed: 28408412

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., & Ugurbil, K. (2013). The WU-Minn Human Connectome Project: An overview. *NeuroImage*, *80*, 62–79. https://doi.org/10.1016/j.neuroimage.2013.05.041, PubMed: 23684880

Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T., Bucholz, R., … WU-Minn HCP Consortium. (2012). The Human Connectome Project: A data acquisition perspective. *NeuroImage*, *62*(4), 2222–2231. https://doi.org/10.1016/j.neuroimage.2012.02.018, PubMed: 22366334

Varona, P., & Rabinovich, M. I. (2016). Hierarchical dynamics of informational patterns and decision-making. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1832), 20160475. https://doi.org/10.1098/rspb.2016.0475, PubMed: 27252020

Varoquaux, G., Schwartz, Y., Poldrack, R. A., Gauthier, B., Bzdok, D., Poline, J.-B., & Thirion, B. (2018). Atlases of cognition with large-scale human brain mapping. *PLoS Computational Biology*, *14*(11), e1006565. https://doi.org/10.1371/journal.pcbi.1006565, PubMed: 30496171

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., … Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(3), 1125–1165. https://doi.org/10.1152/jn.00338.2011, PubMed: 21653723