A morphospace framework to assess cognitive flexibility based on brain functional networks

Duy Duong-Tran, Enrico Amico, Bernat Corominas-Murtra, Mario Ventresca and Joaquín Goñi

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 for Science and Technology, Am Campus 1, A-3400 Klosterneuburg, Austria
4 Purdue Institute of Inflammation, Immunology and Infectious Disease, Purdue University, West Lafayette, IN, USA
5 Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA

* Corresponding email: jgonicor@purdue.edu

Abstract

Unfolding how the brain functionally shifts within the cognitive space remains an unresolved question. From a brain connectivity perspective, there exist two main concepts: cognitive shifts and cognitive flexibility. Although the former is the proxy of the latter, the biggest challenge, in terms of bridging these two concepts, lies in the fact that cognitive shifts are governed by topological rules whereas cognitive flexibility is purely numerical. In this paper, we bridge the aforementioned concepts while preserving the complexity of cognition by proposing a formalism based on a 2D network morphospace that quantifies trapping and exit characteristics of network subsystems, naturally interpreted as functional communities. We show that the constructed measurements reflect the emergent phenomenon of higher-order cognitive states in addition to being able to quantify cognitive flexibility, as a direct output. Leveraging this analytic framework, cognitive shifts among traversedly integrated/segregated states of cognition are shown to be projected from subject specific cognitive signatures. The evidence of individual fingerprint emerged from cognitive flexibility domain legitimizes the quest to explore the intrinsic relationship between flexibility of functional networks and behavioral measures, including fluid intelligence. The constructed multi-linear models using flexibility descriptors demonstrate an above chance level of specificity. Finally, through the associations between behavioral measures and flexibility theory, we found that frontoparietal (FP) activation level, expressed through FP preconfiguration, and default mode network (DMN) efficiency, expressed through DMN preconfiguration, are positively correlated with all behavioral measures.
1 Introduction

Recent advances in understanding levels of integration and segregation of human brain networks, during rest and task-evoked states [1,7,13,59,60,64] have highlighted the importance of network neuroscience in understanding the association between human cognition and brain function [10,41]. Despite such progress, modulations of brain functional networks, as well as their specificity to different cognitive tasks are key unresolved questions in Neuroscience. As the brain traverses through integrated and segregated states of cognition, it leads to what has been recently denominated as cognitive shifts [60]. Mechanically, those shifts are governed by topological reconfigurations within the functional connectome.

Deeply related to the concept of cognitive shifts, there is the concept of cognitive flexibility [21,58,60]. Cognitive flexibility, i.e. the capability of switching among integrated and segregated states of cognition, has been compartmentalized into (functional) reconfiguration [58,60] and preconfiguration [55]. Moreover, it has also been associated with a measure of brain adaptation efficiency, with the underlying assumption that efficient functional adaptations are those that require less reconfiguration [55]. Hence, we hereby utilize the concept of cognitive flexibility interchangeably with functional reconfiguration and preconfiguration.

The first challenge to establishing a bridge between the notions of cognitive shifts and cognitive flexibility is that, by construction, the former lies in a topological domain embedded in a specific geometry whereas the latter is purely numeric. Hence, when developing a framework that connects these concepts, it needs to be geometrical in nature, and with a well-defined metric in which measurements can be established. The second challenge lies in their common denominator: cognition. Assuming that such a framework is possible, how do we reflect cognition while bridging shifts and flexibility? From a complexity standpoint, the mind-brain continuum is a phenomenon that can be explained through the philosophical concept of emergence [9]. Operationally, cognitive
shifts emerge from microscopic elements in the system in response to a cognitive demand [9, 44]. The emergence phenomenon refers to the idea that human cognition cannot be simply reduced to assessing independent subsystems or to merely adding up their contributions [9]. Instead, human cognition can only emerge and hence, potentially be explained, based on complex interactions of subsystems [9]. To address the second question, the proposed measurements need to reflect the emergence phenomenon in which complex micro-scale neuronal interactions give rise to the observable phenomenon at the global scale, i.e. cognition.

The primary aim of this paper is to construct an analytical framework in which the notion of (cognitive) shifts and flexibility can be fully assessed while also capturing some of the characteristics of the brain complexity. This aim leads to the following question: What is the spatial scale that better reflects cognition and models shifts? As noted by Bassett et al., brain complexity is typically shown through modular structures and function of subsystems that are capable of facilitating higher levels of functionality such as cognition [9]. Specifically, mesoscopic structures exhibit modular characteristics that can adapt themselves configuratively in response to some cognitive demands without perturbing the global functionality [9]. Secondly, it is well-noted that integrated and segregated states existing within the rich repertoire of functional connectomes is a phenomenon that emerges in mesoscopic structures or sub-circuits [57, 60]. One possible way to interpret the term mesoscopic structures in the context of functional connectomes is that these are subsystems in the brain that collectively represent unique modes of cognition. As such, these subsystems may be directly referred to as independent functional traits [2, 4] that cooperate spontaneously to maintain the functionality of the cognitive space [4]. Furthermore, the integrative core of cognition could be represented as a network constructed by a low-dimensional manifold [57, 58, 60]. Alternatively, mesoscopic structures can also be defined as subsets of brain regions sustaining and/or modulating
different functions. In this case, mesoscopic structures are directly related to the set of elements that belong to the subsystem, rather than independent modes that constitute the cognitive space. In order to build a framework in which cognitive shifts and flexibility can be quantified and assessed, neither a macro- nor a micro-scopic approach would be sufficient.

There exist two common approaches when examining how mesoscopic reorganizations affect important characteristics of the mind-brain continuum under various cognitively demanding tasks. The first approach deciphers the notion of newly emergent subsystems as the subject engages in different tasks. Many interesting concepts have been proposed relying on this approach, such as individualized (atlas-free) parcellation \[43, 68\] or task-dependent atlases \[53\]. On the other hand, the alternative is to maintain a set of baseline subsystems under all scenarios. To construct a framework that builds on the concept of cognitive shifts, the first approach prevents the traceability of changes between cognitive states that are *comparable*, i.e. the lack of referenced architecture network, and *compatible*, i.e. changes among different equivalent subsystems under various tasks for the same or even different subjects. Moreover, from a brain complexity standpoint, fixing a subsystem to reflect cognitive shifts allows us to also monitor the constant evolution and adaptation of the different subsystems as the global system responds to inputs from the external environment \[9\].

Establishing a set of baseline subsystems, interpreted as functional communities, hence, plays a critical role in building a framework to assess cognitive shifts. Yet, which baseline subsystems could one use to accomplish this goal? To this end, one can immediately refer to the so called intrinsic network \[17\], that is prominently grounded in the resting state architecture. This view is also supported by Shine et al. in which the human brain is hypothesized to exhibit and maintain a core integrative network that dynamically reconfigures during cognitive demands \[59\]. In short, the baseline subsystems to study cognitive flexibility should rely on resting state networks. Having established the baseline
at rest, cognitive shifts can be viewed as topological perturbations departing from resting state as reflected by edge-wise changes in the connectomic domain. That is, the goal of this study is not about unravelling newly emerged communities that re-organize the cognitive space. Instead, here we pursue the definition and assessment of network theoretical properties of those \textit{a priori} functional modules as different tasks are performed. In such regards, we also share the viewpoint of Cole et al., from a whole-brain perspective, in which tasks are hypothesized to modify the intrinsic network to meet cognitive demands \cite{17}. Thanks to recent advancement in the emergent field of network neuroscience \cite{14, 15, 25}, subsystems possessing highly executive functions are consistently observed with a high level of reproducibility across many individuals under resting conditions \cite{27, 39, 60}. As a direct consequence, subsystems or functional communities detected at rest are usually referred to as functional networks (FNs).

To accomplish the primary aim of the paper, we see that functional communities are a useful mesoscopic view of the human brain to study the notion of cognitive shifts. Then, to depict the changes, one needs to map cognitive shifts into brain connectivity. Specifically, how do we disentangle cognitive shifts such that the complexity \cite{5, 9, 24} of emergent phenomenon, i.e. high-order cognition, is reflected? How many measurements are needed to sufficiently describe such a notion? Finally, how do we host different theoretical branches in a unified environment so that the notion of cognitive flexibility is captured? Similar types of questions were successfully addressed through the construction of phenotype spaces, also called \textit{morphospaces}, when assessing very heterogeneous systems and domains, see \cite{6, 18, 29, 40, 45, 54, 61, 65}. In this framework, quantitative traits of global or local network topology are conceptualized through the Cartesian coordinates that define an abstract space, the \textit{morphospace}. Any potential topological configuration is represented by a point in this multidimensional space, whose structure highlights the special properties under study. Interestingly, Bassett et al. \cite{9} also identified a fruitful opportunity in utilizing metric
geometry framework as a complementary research avenue to provide further insights in cognitive neuroscience. Through the development of a modularity morphospace, we assess, simultaneously, the level of segregation and integration of each functional network while higher-order cognition and cognitive shifts emerge from complex interaction among system elements.

Finally, studies on cognitive flexibility have shown associations between individual capacity of change within the cognitive space and behavioral scores [55]. This motivates the second aim, which is investigating individual fingerprints integrated in the domain of cognitive shifts. Do cognitive shifts intrinsically carry individual cognitive signatures? If so, are these traits direct consequence of functional community flexibility? This leads to a series of derivative questions that involve the understanding of dynamical behaviors of FNs, which ultimately shape cognitive shifts, or whether functional reconfiguration/preconfiguration (as suggested in [55] for the measure of efficiency) are related to behavioral measures such as fluid intelligence.

In this paper, we construct a 2D modularity morphospace to accomplish the aforementioned aims. The morphospace is constructed using two measurements: module trapping efficiency and module exit entropy. In the next section, the theoretical treatment of this framework is provided. It is followed by the result section where we apply such framework to the 100 unrelated subjects dataset in the Human connectome project. The details of the dataset and functional connectome construction are located in the supporting information (SI).

2 Theoretical construction of Modularity Morphospace and cognitive flexibility

2.1 The necessity of a modularity morphospace

As functional connectomes are combinatorial objects, any topological changes can be represented by nodes and edges. Yet, between rest and task-based
conditions, straightforward topological comparison such as numerical difference between functional edges is not sufficient to infer complex cognitive changes \[17\]. This is because simple numerical comparisons are not capable of disentangling the complexity of cognition. Miller and Wallis \[44\] have shown that executive function and higher order cognition emerge from the neuronal level. Applying similar logic into the brain connectivity domain, the finest scale elements of functional connectomes are represented by nodes, for a given spatial scale. Hence, in order to model cognitive shifts, we need to create measurements that integrate interactions among the finest scale elements of the subsystems that eventually constitute higher-order complexity. Simply considering functional edges, from the magnitude standpoint \[17\], does not drive such analogy for modelling purpose.

To accomplish this goal, two relevant theories – on top of the intrinsic network characterization of the system – that can satisfy the aforementioned requirements, are (i) stochastic processes modelling theory \[20, 34\] and (ii) information-theory \[19, 37, 56\]. Stochastic modelling provides 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 pair-wise or local interaction can fully describe its behavior. Information theory provides the quantification of uncertainty induced by topological structures. It allows a fine-grained approach that is complementary with the stochastic modelling approach, as seen in \[51, 52\], especially in the domain of community structures in complex networks.

Given a functional network (abbreviated as FN, and denoted as \(\mathcal{C}\)) in a functional connectome \(G\), denoted as \(\mathcal{C} \subset G\), a modularity morphospace is constructed to assess functional network behaviors through the lenses of topologically-induced dynamics \((x - axis)\), and information theory \((y - axis)\). Details regarding the construction of \(\mathcal{C}\) are available in the SI. The rest of this section describes the proposed measurements that ultimately shape the topology of this modularity morphospace.
In subsequent sections, we show the reflection of emergent phenomenon in each proposed measurement along with its philosophical meaning and computation. Lastly, we address how these two measurements express unique features of cognition in sub-section D where we introduce the notion of feature orthogonality.

2.2 Module Trapping Efficiency

2.2.1 Definition and Formulation

Module Trapping Efficiency, denoted as \( TE \) (unit: \( \frac{\text{steps}}{\text{weight}} \)), quantifies the capacity of a module to prevent a random walker from leaving the FN \( \mathcal{C} \).

Specifically, it is the topologically-driven measure assessing the internal flow sustainability of module \( \mathcal{C} \subseteq G \) (unit: \( \text{steps} \)) relative to its exiting strength (unit: \( \text{weight} \)). The flow-preserving characteristic of a community, denoted as \( \tau \), assesses how well the subgraph performs as a functional community through two conceptual modularity landscapes: density-based, \([26,46]\), and flow/pattern-based, \([38,51,52]\) simultaneously.

Mathematically, given an induced subgraph \( \mathcal{C}(V_C, E_C) \subseteq G(V, E) \), the proposed formula to quantify internal-flow preservation on a subgraph (or functional community) is as follows:

\[
TE = \frac{||\tau||_2}{L_C} \tag{1}
\]

where \( ||\tau||_2 \) is the 2-norm of the mean time to absorption vector from nodes in \( \mathcal{C}(V_C, E_C) \) to exiting nodes that are defined to have at least one connection, i.e. strictly positive functional edge, to any member of module \( \mathcal{C} \). The denominator \( L_C \) is the 1-norm of the vectorized functional connectome submatrix induced by nodes in \( \mathcal{C} \) and its corresponding exiting nodes.

Note that normalizer \( L_C \) has two distinct functions: (i) as \( \tau \) naturally scales with community’s size, normalizer \( L_C \) acts as a damping factor to the measure; (ii) the total ”leakage” of \( \mathcal{C} \) to \( G \setminus C \), it represents the level of integration between...
subsystem \( C \) and the rest of the network \( G \setminus C \).

### 2.2.2 Interpretation

We see that the analogy between TE and emergent phenomenon of cognition lies in the fact that \( \tau \) is quantified through edge-induced dynamics in the stochastic process. In the SI, we also provide a toy example that shows the sensitivity of \( \tau \). Additionally, \( \mathcal{L}_C \) is the tallied effect from weighted edges that depict the interactions among the finest scale elements of the subsystem \( C \) with \( G \setminus C \).

Neuroscientifically, as FNs are established as \textit{a priori} within the rich repertoire of cognitive states, module trapping efficiency assesses the underlying dynamics induced from the topology of a specific FN \( C \subset G \) under different cognitive states.

![Diagram showing morphospace measurements examples](image.png)

**Figure 1. Morphospace Measurements - examples.** In the above demonstration, all induced sub-graphs have the same cardinality (\(|C| = 8\)) with different number of exits (connections to \( G \setminus C \)). Nonetheless, depending on their topological structures, the corresponding morphospace measurements have distinctive values. Note that nodes within the boundary belong to functional modules whereas the remaining nodes are exiting nodes, as defined in SI.
2.3 Module Exit Entropy

2.3.1 Definition and Formulation

Module Exit Entropy (denoted as \( EE \) being in the range \( EE \in (0, 1] \) and unitless) assesses the degree of uncertainty in selecting a specific exit for a random walker within FN \( C \).

Specifically, it is an information-theoretically driven measure addressing the module communication preference, at a node level. As module trapping efficiency assesses how a FN topology drives its underlying local/global dynamics, module exit entropy evaluates the overall preference evaluated at a module exit channels.

Given an induced subgraph \( C(V_C, E_C) \) with \( \bar{m} \) exits (absorbing states with characteristic absorption rate per Terminating Markov process construction - see SI for details), the exiting node entropy, denoted as \( H_e \), measures the level of uncertainty of which exiting node is preferred. Module exit entropy is mathematically formalized as:

\[
EE = \frac{H_e}{\bar{N}_C} = -\frac{\sum_{i=1}^{\bar{m}} \psi_i \log(\psi_i)}{\log(\bar{m})} \tag{2}
\]

where preferential exit probability can be calculated as \( \psi^T = 1^T_{|V_C|} \Psi \left[ 1^T_{|V_C|} \Psi 1_{\bar{m}} \right]^{-1} \), see SI for details. The normalizer, \( \bar{N}_C = \log(\bar{m}) \), is the maximum entropy obtained from a module in which all exit nodes has the same absorption rate.

2.3.2 Interpretation

As far as the relationship between \( EE \) and emergent phenomenon of higher-order cognition is concerned, we see that the exit entropy is really driven from the node’s perspective. In fact, exiting nodes determine the maximum amount of information needed to escape the module if there is no particular preference in escaping strategy imposed by the random walker. In terms of functional brain networks, instead of evaluating the module characteristics of FN purely
through topologically induced dynamics, module exit entropy facilitates the understanding of collective behavior from \( C \) to other FNs through its outreach channels (edges formed by nodes in \( C \) and exiting nodes in \( G \setminus C \)).

![Figure 2. Tasks-Sensitivity Study](image)

**Figure 2. Tasks-Sensitivity Study** Given the design of individual axes, the first test for each proposed measurement is its capacity to identify a task, i.e. task-sensitivity (see details in SI). Here, the notion of task-sensitivity is the degree, at which a task is recognized by its cognitive requirement imposed on the subject. It is measured by intra-class correlation. Each dot in this plot represents one subject. The analysis uses all seven fMRI tasks and rest as available in HCP data.

### 2.4 Modularity Morphospace formalism

According to our formalism of the morphospace, two distinct features of each FN in brain graphs are addressed by a point \( u(C) \) in the Euclidean space \( \Omega \subset (0, M) \times [0, 1] \) where \( M < \infty \).

\[
u(C) = (u(TE(C)), u(EE(C))) \in \Omega
\]  

(3)

Given a functional brain network \( G \) with highly-putative parcellation that results in \( l \) induced subgraphs \( C \subset G \), we can obtain \( l \) points \( u(C) \) corresponding to \( l \) FNs in network \( G \).

In general, module trapping efficiency, \( u(TE(C)) \) must be finitely bounded (see more details in SI). However, in the context of the investigated dataset, a better bound is possible. This is due to two driving factors: connectome sparsity
and edge weights [6], and SI for further details. In the context of investigated data, we address the upper bound for $\text{TE}$ as: $\max(\text{TE}) = M = 1$. In terms of $u(\text{EE}(C))$, its numerical range $u(\text{EE}(C)) \in (0, 1]$ is well-defined as it is driven purely from information-theory. Hence, $\Omega \subset (0, 1) \times [0, 1]$.

Ideally, in morphospace design, each Cartesian axis should be explicitly "orthogonal", see [6], in the sense that reflects one unique feature of understudied objects, in this case functional modules $C$. In the case of brain networks, a given module can possess a high value of $\text{TE}$ and $\text{EE}$, simultaneously. Hence, numerically, orthogonality is not always guaranteed, although this is definitely not a necessary and sufficient condition for the morphospace to be well-defined or provide useful insights. Network-theoretically, the morphospace assesses non-overlapping traits of modular structures for a given network, i.e. topologically-driven dynamics and communication.
Figure 3. Cognitive flexibility: Geometrical Presentations: functional reconfiguration and preconfiguration for all FNs are graphed using group average result. Specifically, for each subject/task/FN combination, we have two connectomes (Test and Retest); we average two visits to obtain mean functional connectome. We then acquire 100 pairs of morphospace Cartesian products for such combination. Finally, we average these coordinates across all subjects to obtain one Cartesian product for one specific task and FN configuration. One immediate observation drawn from such presentation is that the morphospace framework reconfirms, quantitatively, that default mode network is acting more as an assortative subsystem at rest - as seen in the lower right regime - as opposed to under task-evoked conditions - as seen in the top left corner of the same space, [32]. Another observation is that in terms of sub-system associativity measured by TE, the lower bound of subcortical convex hull is, approximately, the upper bound of other FNs, with the exception of visual network, in the cortical region. Note that $x-$ and $y-$ axis are purposely not scaled in the same range so that all tasks, task-centroid, and rest can be visualized easier.
2.5 Defining functional reconfiguration and preconfiguration through modularity morphospace

2.5.1 Cognitive flexibility

As mentioned in the Introduction, the notion of cognitive flexibility, for the $i^{th}$ subject, is equivalently compartmentalized into two components: (i) FN (task) reconfiguration and (ii) FN rest-to-[task-general] preconfiguration. We then propose a mathematical relation between them as follows:

$$F_i = f(R_{i}^{FN}, P_{i}^{FN})$$

where $F_i$ represents cognitive flexibility for subject $i$ which is comprehensively formalized through all considered FN reconfiguration and preconfiguration. Here, we do not infer the function $f$ that maps subsystem functional reconfiguration and preconfiguration into subject (cognitive) flexibility. Rather, we provide directly the measures that quantify (functional) reconfiguration and preconfiguration of subsystems that constitute $i^{th}$ subject’s cognitive flexibility.
2.5.2 Functional Reconfiguration

**Definition 1** Functional reconfiguration is defined to be the cognitive capacity in exploring its potential when switching from the current (task) cognitive state to another, and is quantified as

\[ R_{i}^{FN} = \text{Vol}(\text{Conv}(W_{i}^{FN})) \] (5)

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

It is important to note that, given the concept of functional reconfiguration, first order measurements such as distance and its variants, thereof, are not sufficient. We provide further evidence in SI for the relevance of second order measurement such as area in 2D modularity morphospace.

**Remark 1** Notice that the ambient space dimension in this case is \( d = 2 \); hence, at most, we can constitute the notion of area (see SI for more details). Furthermore, for each functional community, (task) reconfiguration for the \( i^{th} \) subject is the capacity to collectively modify pairwise interactions, i.e. changes in functional coupling strengths between pairs of nodes, that eventually constitute higher-order cognition in response to task demands [11, 12, 58, 60]. Further, without any specific assumptions, all tasks are given the same level of importance and hence, no task is weighted more than others.

2.5.3 Functional Preconfiguration

**Definition 2** Functional preconfiguration is defined as the brain capability, arranging itself topologically, to switch from resting state configuration (cognition without the presence of task) to a task-general position, and is quantified as
follows:

$$P_{FN}^{i} = || \text{Rest}_{FN}^{i} - \eta_{W_{FN}^{i}} ||_2$$  \hspace{1cm} (6)

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

It is important to note that given the definition of functional preconfiguration, first order measurement suffices, see SI for further details.

**Remark 2** For each FN, functional preconfiguration for the $i^{th}$ subject, represents the collective changes, as reflected through pairwise coupling interactions among nodes in functional connectomes, from resting state to a generic task-engaged configuration, represented by $\eta_{W_{FN}^{i}}$ in modularity morphospace. In other words, it is the brain capability to initialize its functional readiness between resting and task-evoked conditions, [55].
**Figure 5. Cognitive flexibility Analysis** Panels 1 and 2 analyze functional reconfiguration and preconfiguration, respectively, from both magnitude and subject-sensitivity viewpoints. Panels A’s report quantified individual FN’s p-/reconfiguration. Panel B’s demonstrate R and P’s subject sensitivity and its corresponding null model (Description in S.I.). Firstly, for majority of FNs, preconfiguration carry more subject functional fingerprints than reconfiguration of the same functional module. Secondly, null model subject sensitivity maintain stable magnitude of approximately 0.2, independently of FN and cognitive flexibility components. Lastly, panel C merges all 16 terms through descending order of subject sensitivity; merged subject identifiability is color-coded, consistently with individual cognitive flexibility ingredient (Reconfiguration: Blue; Preconfiguration: Red). Color available online.

### 3 Results

The modularity morphospace formalized in section 2 is used to assess the 100 unrelated subjects (HCP, Q3 release). It includes (test and retest) resting state and seven fMRI tasks: gambling (GAM), relational (REL), social (SOC), working memory (WM), language processing (LANG), and emotion (EMOT). Functional connectomes have 360 cortical brain regions and 14 subcortical regions. The functional communities evaluated in the morphospace were the seven functional networks (FNs) as reported by Yeo and colleagues, namely: visual (VIS),
somatomotor (SM), dorsal attention (DA), ventral attention (VA), frontoparietal (FP), limbic (LIM), default mode (DMN); and a last one including subcortical regions (SUBC). Additional details about the dataset are available at SI.

3.1 Task-Sensitivity analysis

The first analysis consists of evaluating the task sensitivity of module trapping efficiency and module exit entropy across subjects. In other words, we evaluate if the task-dependent flexibility of the FNs as the subject performs different tasks are captured by these two measurements. To do so, we obtain the task-based intra-class correlation (ICC) for each FN for each subject. Results including ICC estimations for all subjects and all FNs are shown in Fig. 2A for module trapping efficiency and in Fig. 2B for module exit entropy. Overall, most of the subjects for each and all FNs have high positive ICC values. Furthermore, module exit entropy displays a higher task sensitivity than module trapping efficiency. VIS, VA and DMN have the highest task-sensitivity when evaluating module trapping efficiency whereas VA and FP have the highest when evaluating module exit entropy. LIM is the FN with the lowest ICC values for both coordinates. Moreover, with the exception of LIM, for all other FNs high intra-class correlation (ICC) values in one coordinate do not necessarily imply similar tendency in the other coordinate. For instance, VA has the third lowest task-sensitivity mean value in TE whereas, in EE, it has the highest mean score. Similarly, FP has second lowest average score in TE while in EE, third highest.

3.2 Quantifying cognitive flexibility of functional networks

The modularity morphospace as defined in section 2 allows the quantification of cognitive flexibility. We can now compute and subsequently assign a number, utilizing formulas 5 and 6 to measure functional reconfiguration and preconfiguration respectively, given a functional community.

The group average geometrical presentation of functional communities is
shown in Fig. 3, functional reconfiguration of FNs are shown as filled convex hull whereas preconfiguration of FNs are shown as dashed lines from rest to the corresponding task hull geometric centroid. In Fig. 4, the cognitive flexibility for all functional networks is presented within the same axis window to ease comparison. Based on Fig. 4, VIS network polytope, representing group-average behavior, locates in the lower regime, relative to the rest of the FNs in a fixed window. Moreover, from a group-level perspective (as reported in figure Fig. 3 and Fig. 4), with the exception of VIS and SUBC, all other FNs seems to be packed in a similar spot in this morphospace.

Figures 5.1A and 5.2A summarize functional reconfiguration and preconfiguration values respectively, for test and retest fMRI sessions for all subjects and FNs. It can be observed that the VIS system displays the largest functional reconfiguration (see figure 5.1A). From Fig 5.2A, functional preconfigurations display a more comparable magnitude among all FNs.

3.3 Subject specificity of cognitive flexibility

The formulation of functional flexibility (in terms of preconfiguration and reconfiguration) enables to assess these properties at the subject level. As shown in section 3A, such formulation is built upon two task sensitive measures, namely module trapping efficiency and exit entropy. To what extent is cognitive flexibility a fingerprint of subjects and to what extent we may detect so with this framework?

This question is supported by recent findings [55], pointing out the intrinsic relationship between flexibility and behavioral measures. In Fig 5.1B, we, firstly, analyze subject sensitivity of functional reconfiguration for each FN. For the considered parcellation, [73], all functional communities result in consistently higher individual sensitivity compared to their corresponding null model (see SI for details on the null model). Analogously, subject sensitivity of functional preconfigurability is analyzed in Fig 5.2B. FNs with high preconfiguration
mean values, Fig 5.2A, tends to have significant capacity to differentiate among
subjects, i.e. between-subject variability, similarly defined in [72]. Specifically,
FP, DMN and VA display a high degree of subject sensitivity. As seen in
reconfiguration analysis, all eight FN subject sensitivity (as measured by ICC
scores) are higher than their corresponding null models. In addition, we also see
that functional preconfigurations dominated the subject sensitivity ranking, as
shown in Fig. 5C.

Subject sensitivity results uncovered that functional preconfiguration of FNs
is systematically higher than its reconfiguration counterparts (Fig. 5.1.B and
5.2.B). Furthermore, FP and DMN are among the FNs with highest subject
fingerprints in preconfiguration. Subject fingerprints of subsystem cognitive
flexibility are comprehensively arranged in descending order in figure Fig 5.C.

3.4 Cognitive flexibility and behavior

As shown in section C, we found evidence of subject fingerprints through our
proposed framework to assess cognitive flexibility. In this section we address
the question of whether such framework has predictive capacity for a number
of widely studied behavioral measures. To accomplish this, we assess four
behavioral measures: episodic memory, verbal episodic memory, fluid intelligence
and general intelligence. The reason we chose these specific behavioral measures
is based on results shown in section C. In particular, we identified FN and DMN
functional networks as of having high subject sensitivity. Furthermore, previous
findings have also associated the neurological behavior of these two functional
networks with behavioral measures related to memory and intelligence [31,55,67].

As fluid intelligence reflects subject capacity to solve novel problems, general-
ized intelligence \( g \) reflects not only fluid intelligence traits but also crystallized
intelligence, i.e. subject’s cognitive capacity through acquired knowledge, as
mentioned in [55] with respect to the original work of [16]. The early notion
of general intelligence is established through Spearman’s [63] so-called *positive*
manifold for which no single task-performance can fully describe. Quantification of \( g \) can be accomplished using subspace extraction techniques such as explanatory factor analysis \cite{22}, or principle component analysis (PCA) \cite{55}. Here we used the PCA approach described in \cite{55} to quantify general intelligence \( g \).

To propose the intrinsic relationship between cognitive flexibility and behavioral measures, we use two fundamental concepts: (i) cognitive flexibility is associative with behavioral measures at the subject level \cite{55}; (ii) cognitive flexibility is equivalently compartmentalized into two components: FN reconfiguration and preconfiguration, as proposed in sections 2 and 3. Consequently, we propose the following composite relationship:

\[
\triangleright_i = \Gamma(\mathcal{F}_i) \\
= \Gamma(f(\mathcal{R}^{FN}_i, \mathcal{P}^{FN}_i)) \\
= \Upsilon(\mathcal{R}^{FN}_i, \mathcal{P}^{FN}_i)
\]

where \( \Upsilon = \Gamma \cdot f \) is the composite function between \( f \) from equation (4) (corresponding to concept (ii) - stated above) and function \( \Gamma \) which corresponds to concept (i). Additionally, \( \triangleright_i \) represents some behavioral measure. Neuroscientifically, equation 7 is motivated from findings in \cite{36}; whereas equation 8 is due to hypothesis proposed in \cite{55}.

Having established a plausible connection between behavioral measures and \( \mathcal{P}^{FN}_i \) and \( \mathcal{R}^{FN}_i \), equation (9) can be viewed as a multi-linear model (MLM) using FN reconfiguration and preconfigurations as independent variables (or predictors). The MLM is constructed iteratively, starting with the descriptor with the highest individual sensitivity. In each iteration, the next ranked descriptor according to Fig. 5C is appended to the existing ones. The best MLM, which determines the number of linear descriptors included the model, is selected based on the model p-value.

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 cognitive flexibility and behavioral measures. If there exists such robustness, 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 values. Specifically, cognitive flexibility, 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.

To test the level of specificity in the model, we perform 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, for each simulation, 5 correlation values are obtained. It can be shown that the mean of these five values per simulation follows a normal distribution (details shown in SI). Lastly, to provide the level of specificity of linear descriptors, we provide, on top of this, a corresponding null model where the same descriptors are evaluated to predict random vectors of appropriate size.

To truly scrutinize our model and its ability to predict the behavioral measures, we rely completely on the cognitive flexibility predictors ranked in descending order of subject specificity.

| MLM terms /coefficients | Constant | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_4$ |
|-------------------------|----------|-----------|-----------|-----------|-----------|-----------|
| Episodic Mem.           | 0.6      | 2.9       | -9.3      |           |           |           |
| Verb. Epi. Mem          | 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      |           |

Table 1. Multi-linear regression models with corresponding standardized $\beta$ coefficients. Dependent variables for each model are: episodic memory, verbal episodic memory, fluid intelligence ($gF$) and general intelligence ($g$).

One important observation drawn from the top panels of Fig.6 is that
| MLM terms /coefficients | Constant | $p^{FP}$ | $p^{DMN}$ | $p^{DA}$ | $p^{SUBC}$ | Entire Model |
|-------------------------|----------|----------|----------|----------|-----------|-------------|
| Episodic Mem.           | 0        | 0.57     | 0.01     |          |           | 0.03        |
| Verb. Epi. Mem          | 0        | 0.02     | 0.77     | 0.17     | 0.03      | 0.04        |
| $gF$                    | 0        | 0.3      | $9 \times 10^{-4}$ |          |           | 0.004       |
| $g$                     | 0.03     | 0.44     | 0.16     | 0.57     | 0.05      | 0.05        |

Table 2. 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.

the predictive power, as more linear descriptors are added to iterative MLMs, decreases as the absence of subject fingerprint in cognition traits increases, within the cognitive flexibility domain. This result highlights the importance of the subject sensitivity descending order of the predictors. Specifically, as individual specificity reduces from left to right, the differential correlations, i.e. the different between two consecutive correlation values, decrease.
Figure 6. Cognitive flexibility correlation analysis to behavioral measures: Top panels are iterative multi-linear regression model (MLM); bottom panels are model specificity (MS) for corresponding behavioral measure. Highly subject sensitive predictors, as described in figure Fig 5.C, are prioritized in the iterative process. For each behavioral measure, the optimal MLM is indicated by the [*] notation. Model specificity for each behavioral measure are then tested using procedure described in section 3D. Interestingly, the null model’s empirical distribution behave almost similar with first moment situated close to zero. Final step in model specificity involves the pair t-test between the null model and the actual distribution. The null hypothesis for this test is that their corresponding mean values is the same, i.e. $H_0: \mu_0 = \mu_1$. All four behavioral measures pair t-tests reject the null hypothesis.

4 Discussion

The neuroscientific community has progressed towards large publicly available datasets such as the Human Connectome Project [70], the 1,000 Functional Connectomes Project, [42], as well as the development of guidelines for data analysis and sharing in neuroimaging, [47]. These initiatives have provided a tremendous opportunity to explore different aspects of human cognition, in which cognitive shifts emerge among the most critical concept to be fully understood. In this context, the concept of cognitive flexibility is a direct proxy of cognitive shifts or traversely integrated and segregated states of functional brain networks. However, up to date, there was a lack of a mathematical framework that bridges
and quantifies those two concepts.

One of the plausible explanation for this gap is the fact that typically, a framework is either topological (suitable to study integrated and segregated states of cognition) or numeric (which is suitable for quantifying the notion of flexibility as a real positive value). The primary contribution of this framework is that being an abstract space with a well-defined metric, integrated and segregated states of cognition are now well-defined through points in the Euclidean space $\Omega$; moreover, cognitive flexibility (or equivalently subsystem functional reconfiguration and preconfiguration) is now well-defined using conceptualized measures in such space. In this regard, the framework is the bridge that links the two aforementioned concepts. Another important aspect of the modularity morphospace is that it allows for exhaustive and continuous exploration of cognitive shifts, which are truly continuous $^{21,59}$. Interestingly, Bassett et al. had also pointed out the necessity to utilize metric geometry in modelling and understanding cognition and human brain complexity $^9$.

In terms of morphospace design, based on section 3A, we observe further evidence of orthogonality $^6$ between the two morphospace dimensions (trapping efficiency and exit entropy). In other words, high task-sensitivity value in one measurement does not imply similar value in the other axis (Fig. 2). Specifically, we see this through two functional communities: VA and FP. Interestingly, the limbic network is the one that exerts the lowest task sensitivity, as measured in intra-class correlation (see Fig. 2A, 2B and results section 3A). Notice that similar findings were observed in $^1$ when using Jensen-Shannon divergence as a distance metric of functional connectivity. It could be that, maintaining minimal cognitive loads, allows this functional network to act as an intermediate station that sustains the neuromodulatory system under various evoked conditions. One could interpret such result as follows: through the lenses of cognitive flexibility theory domain, limbic network is, to some extent, absent of task fingerprints. Indeed, all other functional communities exert very high levels of
task-identifiability, in both morphospace measurements. This establishes the foundation to comprehensively investigate such morphospace combining both measurements.

Among many important concepts laid in the foundation of brain cognitive flexibility theory, functional reconfiguration \cite{55} has been considered a crucial one. Recently, functional preconfiguration has been also introduced. Both play an integrating part in describing the general notion of cognitive flexibility (denoted as $\mathcal{F}$ in this paper). Schultz et al. \cite{55} have established an association between cognitive flexibility and cognitive reconfiguration. To do so, they compare the whole-brain functional connectome of one task with respect to rest, as measured by Pearson’s Correlation coefficient. Although correlation appears to be appealing to measure one-task-to-rest similarity, the challenge to quantify, at a comprehensive level, total cognitive flexibility when all tasks are considered simultaneously remains. Specifically, note that in \cite{55}, the proposed notion of cognitive flexibility performance, suggested through general efficiency, is quantified using most explained-variance eigen mode after measuring Pearson’s correlation between resting FC and three distinct task FCs. Such notion is, however, challenging to generalize into cognitive flexibility when more and more tasks are considered. This is due to the difficulty in comprehensively choosing one (or more) eigen vectors in describing the notion of cognitive flexibility. In this context, the second contribution of our modularity morphospace is its independence of any dimensionality-reduction technique. Moreover, this modularity morphospace also legitimizes a rigorous framework to quantify two concepts of cognitive flexibility theory: reconfiguration and preconfiguration across traversedly integrated and segregated stages of the brain functional networks.

The formalism of cognitive flexibility has, first of all, allowed us to look at these components in terms of magnitude, (see Fig. 3, Fig. 4, and Fig. 5, 1A,2A and results section 3B). In terms of functional reconfiguration magnitude
(Fig. 3, Fig. 4, Fig 5.1A), interestingly, Shine et. al. [58] suggests anatomical evidence through pupil dilation, administered through neuromodulatory systems shift from segregated to integrated cognitive states across different task-evoked conditions. Essentially, our results could show quantifiable evidences between visual network reconfiguration, through the lenses of cognitive flexibility theory, and demonstrate that our measures capture shifts among states of cognition.

Additionally, Cole et al. [17] have shown that resting architecture network modifies itself to fit task-engaged cognitive requirement, only through a subtle number of changes in functional edges. Numerically, small changes constitute by functional edges between rest and task-based connectivity might not be significant when looking at them from a discrete standpoint. However, as we incorporate complexity theory [9] to understand such changes, we see that combinatorially, these functional edge changes constitute a higher order of cognition. Moreover, we observe that while such changes might be negligible on a global scale, they are however more noticeable when looking at subsystems or functional brain networks, as clearly observed in the VIS network. As far as functional preconfiguration is considered (figure Fig. 3, Fig. 4, Fig. 5.2A), a more homogeneous tendency is observed among the functional communities. Interestingly, for any functional network, preconfiguration values tend to be higher than reconfiguration values. One can think of preconfiguration as a convex hull volume of two points in which $d = 1$, which simply constitutes the notion of distance.

The construction of cognitive flexibility for subsystems in brain networks has also allowed us to unravel and compare individual cognitive traits existing in each functional network (figure Fig. 5.1B,2B,C and section 3C). In this paper, we also assess individual sensitivity through cognitive flexibility. Specifically, after quantifying reconfiguration and preconfiguration for all functional communities, we ask the question of whether these quantities incorporate information about individual cognitive traits, as presented in figure Fig 5.C. We notice that reconfigurations and preconfigurations of the functional communities display het-
erogeneous levels of subject fingerprints. Moreover, functional preconfigurations tend to exert higher levels of subject sensitivity than functional reconfigurations. Overall, this suggests that functional preconfigurations are more specific to individuals than functional reconfigurations for any given FN.

Leveraging this result, subject traits allow us to make inferences on the relationship between certain functional networks with behavioral measures such as general ($g$) or fluid intelligence ($gF$). Schultz et al. [55] provide evidence, using whole-brain connectomes, towards implicit association between functional reconfiguration and the subject $gF$ and $g$. Specifically, it is conjectured that smaller (more efficient) whole-brain reconfigurations under tasks are indicative of better performance. Moreover, authors also suggest that individuals with more efficient global preconfigurations between rest and task-general FC might be associated to having higher intelligence. Nonetheless, Kruschwitz et al. [36] show that a whole-brain efficiency approach to establish association with intelligence might not be a suitable choice. In particular, it was shown that general ($g$), crystallized and fluid intelligence ($gF$) are not associated with global network efficiency of functional connectomes. Considering that the concept of general efficiency can be understood as a global network efficiency measure, the necessity to look deeper into subsystem cognitive flexibility emerges. Thus, this goal motivates the inquiry to find intrinsic relationships between subsystem-level cognitive flexibility with behavioral measures. Quantifying the functional network flexibility through subject sensitive cognitive traits allows us not only detour from a whole-brain approach [36] but also, maintain the utilization of function flexibility theory to explore possible correlations with $g$, $gF$ and other behavioral measures.

Another significant contribution of this paper is that it confirms the plausible relationship between fluid intelligence with frontoparietal and default mode network. Specifically, as pointed out in [67], high levels of fluid intelligence correspond with the high levels of activation of frontoparietal network; such result
is also confirmed by [31] when performing the three-back working memory task. In cognitive flexibility domain, we see that this is also confirmed through positive coefficient $\beta_1$ representing frontoparietal preconfiguration. One can see that the level of activation of the FP network is an indicator of its corresponding distance from FP at rest with respect to task-general position (geometric centroid) in the morphospace. On the other hand, default mode network efficiency (based on normalized path length) has been shown to be negatively correlated to IQ measures [69]. Specifically, brain areas that constitute the default mode are found to have significant (negative) correlations with IQ. Analogously, when viewed under cognitive flexibility theory, we observe the negative signs on DMN preconfiguration in predicting $g$ and $gF$ (see Table 1). Comprehensively, a critical observation drawn from the selected multi-linear model is that on the global scale, it is shown that brain reconfiguration correlates negatively to intelligence measures, [55]. However, when investigating the meso-scale of brain functional networks, we see that this observation can not simply be extrapolated and generalized to all functional communities.

Along with significance contributions, this study has several shortcomings. The first one being the choice of parcellations. Firstly, going from voxel to region of interest (ROI) level, we used the parcellation proposed by Glasser et al. [27]. Subsequently, going from ROIs to mesoscopic structures, the parcellation proposed by Yeo et al. [73], with 7 FNs and sub-cortical network was used in this study. One could evaluate similar questions to other highly putative parcellations of the human brain. The second drawback is the quality of input data to compute points in the morphospace. The third drawback is that negative functional couplings are not considered due to the mechanical limitation of the morphospace. In the context of trapping efficiency measurement, having negative functional edges does not support subsequent construction, especially after it is ranged in the non-negative domain. The fourth one is that sensitivity of points in this space are not thoroughly studied. This could impact the magnitude of
cognitive flexibility components.

With these acknowledged drawbacks, future studies should incorporate a sensitivity study on the behavior of those points $\mathbf{u}(\mathcal{C})$ with respect to the input connectomes. Specifically, future studies can focus on this space behavior with the introduction of, possibly, Gaussian noise into functional connectomes and re-analyze the results. Additionally, future studies could also study the trajectory of resting periods using dynamical functional connectivity, [30]. In this quest, the morphospace is a well-suited tool to accomplish such aim. Additionally, one can also test the continuity of cognitive shifts between integrated and segregated brain states using the *modularity* morphospace.
Methodology

We provide the detailed information on materials and methods in SI. In short, all necessary mechanics collected from multiple disciplines and general set-up for matrix computations are described SI Preliminaries and Methods. Data set is consisted of high-resolution functional connectivity matrices describing human cerebral cortex and sub-cortex (see SI Data Description). The construction of morphospace and the formalized notion of cognitive flexibility are described in SI Morphospace analysis and cognitive flexibility section. Multi-linear model and model specificity are described in SI Behavioral Measure Analysis.

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 acknowledges financial support from Purdue Industrial Engineering Frontier Teams Network Morphospace Award and from Purdue Discovery Park Integrative Data Science Initiative. We thank Olaf Sporns and Kausar Abbas for valuable comments.

Author Contributions

D-T.D., J.G., E.A. designed research; D-T.D. and J.G. performed research; D-T.D., J.G., E.A., C.-M.B., M.V. contributed new reagents/analytic tools; D.-T.D., J.G., E.A. analyzed data; and D.-T.D., J.G. wrote the paper with E.A., C.-M.B., and M.V. providing comments and edits. The authors declare no conflict of interest.
Supporting Information

This document purpose is to collaborate on the machinery of the morphospace and other aspects such as the data set and brain atlas used to analyze the data. The aim is to provide further analytic results in conjunction with the ones that are already presented in the main paper.

A Preliminaries

In this section, we provide the necessary theory to construct modularity morphospace and analyze the results obtained from such construction.

A.1 Notations

In this section, we are to establish some of the key mathematical notations used throughout the paper. Specifically, scalar is italicized, $a$. A vector is denoted as bold letter, $a$. Matrix is notated as capitalized bold letter, $A$. If $r \in [q]$ where $q \in \mathbb{N}^+$, it means that $r$ accepts integer values from 1 up to $q$. For any given vector $a$, the average of its entries is denoted as $\langle a \rangle$. Given any set $S$, its cardinality is denoted as $|S|$.

A.2 Graph Theory & Linear Algebra

In terms of graph theory, a finite dimensional undirected/loopless network is denoted as $G(V, E)$ where $V$ and $E$ are sets of vertices and edges in such network, respectively; the size and order of such network are denoted as $|V(G)| = n$, $|E(G)| = m$.

Graph-theoretically, $G(V, E)$ can be represented by $A_G = A(ij) = [w_{ij}]$, in which $w_{ij} \in [0, 1]$ represents coupling strength between node $i$ and $j$ in matrix $A$. The strength of node $i \in V(G)$ is denoted as $k_i$, in which stored in either a vector format or the diagonal matrix $K$ for $K(ii) = k_i$. 
In terms of linear algebra, given any two vector \( \mathbf{a}, \mathbf{b} \in \mathbb{R}^n \), \( \mathbf{a}^T \mathbf{b} \) and \( \mathbf{a} \mathbf{b}^T \) denote inner and outer product, respectively. Default vector is organized in column fashion; otherwise, specified.

A generic matrix \( \mathbf{F} \) with entries valued in a continuous interval \([x, y]\), \( z_1 \) rows and \( z_2 \) columns is denoted as \( \mathbf{F} \in [x, y]^{z_1 \times z_2} \). Further, if we want to induce a sub-matrix from the original matrix \( \mathbf{F} \) based on a specific set of rows, denoted as set \( S_{rows} \), and columns, denoted as set \( S_{columns} \), we use the notation: \( \mathbf{F}|_{S_{rows}, S_{columns}} \). If the set of rows and columns are matched (both denoted as \( S_w \)), then we will ease notation by using \( \mathbf{F}|_{S_w} \).

### A.3 Markov Chain Theory

Network diffusion is often modeled by a random walk, i.e. the behavior of (random) particles based on some pre-defined probabilistic rules induced from the at-hand network. Given a set of states, \( S \), and probabilistic rule, \( P \) (defined on \( S \)), the resulting Markov chain is notated as \( M = (S, P) \). Furthermore, the transition probability matrix \( P = [p_{ij}] \), is constructed from the adjacency structure of \( G(V, E) \) utilizing local information, i.e. the probability of transitioning from state \( i \) to \( j \) is \( \frac{w_{ij}}{\sum_j w_{ij}}, \forall i, j \in [S] \).

Given the general settings, we define an absorbing stochastic process as follows: Terminating (Absorbing) Markov chain for a given state set \( S \), is used to model stochastic behavior of systems in which the random particle will eventually be absorbed in one (of the absorbing states), regardless of which transient state it starts at. Specifically, let us define \( S_{trans} \) and \( S_{abs} \) to be the set containing all transient and absorbing states, respectively, such that

\[
S = S_{trans} \cup S_{abs}
\]

Given the transition probability matrix \( P \), one can compute the mean number of time a specific transient state, say \( s_j \), is visited, given that the random walker
starts at state $s_i$. Since such quantity is available for all transient state, the outcome can be presented in a matrix format, denoted as $Z \in \mathbb{R}^{|S_{\text{trans}}| \times |S_{\text{trans}}|}$. Analytically, the closed-form formula is given by

$$Z = (I_{|S_{\text{trans}}|} - P_{|S_{\text{trans}}|})^{-1}$$

where $P_{|S_{\text{trans}}|}$ represents the induced (sub)-stochastic process based only on the set of transient states $S_{\text{trans}}$ and $I_{|S_{\text{trans}}|}$ is the identity matrix of dimension $|S_{\text{trans}}|$. In Markov chain theory, this matrix is often referred to as the fundamental matrix \[34\].

The mean time to absorption, denoted as $\tau \in \mathbb{R}^{|S_{\text{trans}}|}$, is defined to be the average number of steps that a random walker needs to visit other states, given that it starts in some transient state - say state $s_i$, before getting absorbed by one of the absorbing states. Numerically, it can be computed using the fundamental matrix $Z$ as follows:

$$\tau = Z1_{|S_{\text{trans}}|} \in \mathbb{R}^{|S_{\text{trans}}| \times 1}$$

The absorption probabilities is defined to be the likelihood of being absorbed by one of the absorbing state, given that the stochastic process starts in some transient state. This quantity is available for each transient-absorbing state pair; hence, it can be represented using a matrix format as follow:

$$\Psi = Z \left[ P_{|S_{\text{trans}}, S_{\text{abs}}|} \right]$$

where $\Psi \in \mathbb{R}^{S_{\text{trans}} \times S_{\text{abs}}}$ and $P_{|S_{\text{trans}}, S_{\text{abs}}|} \in \mathbb{R}^{|S_{\text{trans}}| \times |S_{\text{abs}}|}$ is the sub-stochastic matrix induced from row state $S_{\text{trans}}$ and column state $S_{\text{abs}}$.

### A.4 Polytope Theory

Given a set of points,

$$W = \{ x_1, x_2, ..., x_{|W|} \}$$
for which \( x_j \in \mathbb{R}^d, \forall j \in |W| \), a convex hull formed by such set of points are mathematically represented by

\[
\text{Conv}(W) = \left\{ \sum_{j=1}^{|W|} \alpha_j x_j \mid \sum_{j=1}^{|W|} \alpha_j = 1, \alpha_j \geq 0, \forall j \in |W| \right\}
\]

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. Otherwise, i.e. \( |W| \leq d \), there exist(s) point(s) that are affinely dependent on other points in \( W \).

Given the convex hull induced by points \( x_i \) in \( W \), the convex hull dimension is defined to be the geometrical dimension of the polytope formed by points the hull. The convex hull dimension, denoted as \( h \), is governed largely by the number of points participating, i.e. \( |W| \). Given points \( x_i \in W \), these points belong either to

- the boundary (Pareto front) - sometimes, these points are referred to as vertices of the hull [74];
- the interior of the hull

For each pair of points that form the Pareto front (i.e. vertices of the hull), let us define two type of point pairs as follows:

- type A pairs are such that their convex combination belongs to the boundary of convex hull, denoted as \( \delta(\text{conv}(W)) \);
- type B pairs are such that their convex combination belongs to the interior of hull \( W \), denoted as \( \text{int}(\text{conv}(W)) \)

Providing that points in \( W \) are in general positions in \( \mathbb{R}^d \), the approximated volume induced by the convex hull \( \text{Conv}(W) \) can be calculated through the formation of Delaunay Triangulation process [74]. Note that the volume of a convex hull depends on its dimension which is upper-bounded by the dimension of
the ambient space $d$. The volume of the convex hull is denoted as $\text{Vol}(\text{Conv}(W))$.

In $\mathbb{R}^d$, the convex hull dimension can take on the values

1. $h = 0$ which constitutes a point in $\mathbb{R}^d$,

$$\text{Vol}(\text{Conv}(W)) = 0$$

2. $h = 1$ which constitutes a line segment,

$$\text{Vol}(\text{Conv}(W)) = \text{sup}(d(x_i, x_j)), \forall x_i, x_j \in W$$

where $d(x_i, x_j)$ denotes the pre-defined metric distance between two generic points.

3. $h = 2$ which constitutes the notion of area.

4. $h \geq 3$ which constitutes the notion of volume.

**Figure S1.** Given that $W = \{v_1, v_2, \ldots, v_6\}$, we demonstrate three possible scenarios of convex hull formed by $W$ in morphospace $\Omega$. 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 $\text{conv}(W)$. 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 $(v_2, v_4), (v_3, v_5)$.

Convex hull volume is calculated using Qhull package implemented in Matlab,
see [48]. In general, as pointed out also in [48], computing \( V^- \) or \( H^- \) polytope metric volume is NP-hard (see also [23], [35]) with the availability of efficient approximating algorithms.

In the context of modularity morphospace which has a 2-D format, convex hull dimension is, at most, \( h_{max} = 2 \), see Fig. S1 for further details. Only the first three cases can occur (since \( \max(h) \) equals to the ambient space dimension) although, in general, higher-dimensional morphospace can accommodate the notion of volume.

B Methods

In this section, we apply the theoretical frameworks introduced in the preliminaries to functional connectomes of rest and various task-based conditions. We show how the adjacency and stochastic structure of functional community is built based on the chosen parcellation (of which nodes belong to which functional network). We then provide further explanation on how to compute different measures that will eventually construct the morphospace.

B.1 Adjacency Structure

Given any proper subset of the network \( G \), we obtain a functional community \( C \subseteq G \) by sampling the vertex set \( V_C = V(G) \) and \( E_C = E(G|_C) \). Let us define \( c \in C \) and \( j \in J \) (\( J \) has cardinality \( |J| = \bar{m} \)) be nodes in community \( C \) and nodes that makes at least one connection to \( c \in C \). It’s trivial that \( J \subseteq G \). The adjacency structure, which is formed by \( C \) with itself and the rest of the network, i.e. \( G \setminus C \), denoted as \( \tilde{A}_C \), has four components:

- Upper Left: \( A_C \in \mathbb{R}^{[C \times C]} \) represents the induced adjacency structure from nodes in community \( C \in G \);

---

1. Given any proper subset \( C \), the complement of \( C \) is defined to be \( \bar{C} \).
2. The set of edges of community \( C \) is the set of edges formed by nodes in \( C \) restricted to network \( G \).
• Upper Right: Let \( j \in \mathcal{J} \) be the set of nodes that has at least one connection(s) with nodes in \( \mathcal{C} \). Matrix \( \mathbf{A}_C \) is appended by \( \bar{m} \) columns (to the right of \( \mathbf{A}_C \)) to encode the number of exits, denoted as \( e_{i,j} \), that node \( i \in \mathcal{C} \) makes with \( j \in \mathcal{J} \), i.e. vector \( \tilde{\mathbf{A}}_C(i, |\mathcal{C}| + j) = e_{i,j} \). We denote this matrix to be \( \mathbf{A}|_{\mathcal{C},\mathcal{J}} \in \mathbb{R}^{|\mathcal{C}| \times |\mathcal{J}|} \) in which \( \bar{m} = |\mathcal{J}| \);

• Lower left: matrix are all zeros entries, i.e. \( 0^{\bar{m} \times |\mathcal{C}|} \);

• Lower right: Identity matrix of size \( \bar{m} \), i.e. \( \mathbf{I}_{\bar{m}} \). Note that self-loops are, by default, part of the construction for these exit nodes to build absorbing states.

\[
\tilde{\mathbf{A}}_C = \begin{bmatrix}
\mathbf{A}_C & \mathbf{A}|_{\mathcal{C},\mathcal{J}} \\
0^{\bar{m} \times |\mathcal{C}|} & \mathbf{I}_{\bar{m}}
\end{bmatrix}
\]

B.2 Stochastic Structure

Let \( \mathbf{P}_C \) be an absorbing Markov Chain induced from nodes in community \( \mathcal{C} \), i.e. \( \mathbf{P}_C = \left\{ 1_{|\mathcal{C}|+\bar{m}} \tilde{\mathbf{A}}_C 1_{|\mathcal{C}|+\bar{m}} \right\}^{-1} \tilde{\mathbf{A}}_C \) and \( \tilde{\mathbf{A}}_C(|\mathcal{C}| + j, i) = 0 \). Structurally, \( \mathbf{P}_C \) is made of 4 components:

• Upper left: \( \mathbf{Q}_C \in [0, 1]^{|\mathcal{C}| \times |\mathcal{C}|} \) be the matrix containing entries with values between 0 to 1. Such matrix can be extracted directly from the first \( |\mathcal{C}| \) rows and columns of \( \mathbf{P}|_{\mathcal{C}} \).

• Upper right: \( \mathbf{U} \in [0, 1]^{|\mathcal{C}| \times \bar{m}} \) contains the probability of begin absorbed by a specific exit.

• Lower left contains all zeros entries, i.e. \( 0^{\bar{m} \times |\mathcal{C}|} \);

• Lower right: \( \mathbf{I}_{\bar{m}} \).

\(^3\)This sub-matrix will have real entries of size \( |\mathcal{C}| \times \bar{m} \).
\[ \mathbf{P}_C = \begin{bmatrix} \mathbf{Q}_C & \mathbf{U} \\ \mathbf{0}^{\bar{m} \times |\mathcal{C}|} & \mathbf{I}_{\bar{m}} \end{bmatrix} \]

B.3 Assumption - (Global) topological Connectedness

It is important to be aware that the network at hand might be fragmented (disconnected). In such case, one can simply find proceed with the largest connected component. Moreover, there is not guarantee that functional community induced from global adjacency structure is connected. In this case, we will present the procedure to define and calculate the morphospace in the later section without impacting the induced topological structure of the functional networks.

B.4 Further computations

Based on the construction of stochastic process, we apply the mean time to absorption computation to states in functional community \( \mathcal{C} \). Specifically, it can be computed using the fundamental matrix \( \mathbf{Z} \) as follows:

\[
\tau = \mathbf{Z}_{|\mathcal{C}|} \mathbf{1}_{|\mathcal{C}|} \in \mathbb{R}^{|\mathcal{C}| \times 1}
\]

where \( \mathbf{Z}_\mathcal{C} = (\mathbf{I}_{|\mathcal{C}|} - \mathbf{Q}_\mathcal{C})^{-1} \) as defined in the Markov chain theory.

With the construction of \( \mathbf{P}_\mathcal{C} \), the preferential exit matrix represents the likelihood of being absorbed by a specific exit, starting from any state in \( \mathcal{C} \), which can be quantified as:

\[
\Psi = \mathbf{ZU} \in \mathbb{R}^{|\mathcal{C}| \times \bar{m}}
\]

From there, we normalize \( \Psi \) row-wise by summing over the rows of such matrix which results in a row vector \( \mathbf{1}_{|\mathcal{V}_\mathcal{C}|}^T \Psi \in \mathbb{R}^{1 \times \bar{m}} \). Finally, the last step is to
normalize this vector to make it a probability vector, as discussed in the main text.

C Data

In this section, we provide the details related to the dataset we used to analyse the notion of cognitive flexibility. We also provide information related to the brain atlas.

C.1 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. [27]. Similarly to reference [2][3], 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](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’ flexibility by utilizing Resting State Networks (FNs), see [73], which includes seven functional networks (FNs): Visual (VIS), Somato-Motor (SM), Dorsal Attention (DA), Ventral Attention (VA), Limbic (LIM), FrontalParieto (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.

C.2 HCP Dataset

The fMRI dataset used in this paper is available in the Human Connectome Project (HCP) depository [http://www.humanconnectome.org/](http://www.humanconnectome.org/), with Released
Q3. The processed functional connectomes obtained from this 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.

C.3  HCP Functional Data

The fMRI data from the 100 unrelated subjects in the HCP Q3 release were employed in this study \cite{71, 70}. The two resting-state functional MRI acquisitions (HCP filenames: \texttt{rfMRI\_REST}\textsubscript{1} and \texttt{rfMRI\_REST}\textsubscript{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, \cite{28, 71, 70} for details. This release includes also data from seven different fMRI tasks: gambling (\texttt{tfMRI\_GAMBLING}), relational or reasoning (\texttt{tfMRI\_RELATIONAL}), social (\texttt{tfMRI\_SOCIAL}), working memory (\texttt{tfMRI\_WM}), motor (\texttt{tfMRI\_MOTOR}), language (\texttt{tfMRI\_LANGUAGE}, including both a story-listening and arithmetic task), and emotion (\texttt{tfMRI\_EMOTION}). Per \cite{28, 8}, 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 \cite{8, 28, 62} for further details on the HCP dataset. All tasks and resting functional MRIs are equally weighted importance. In other words, no particular weight is assigned to any specific tasks.

C.3.1  Pre-processing

We used the standard HCP functional pre-processing pipeline, which includes artifact removal, motion correction and registration to standard space, as described in \cite{28, 62} 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 $[39,50]$. For task fMRI data, aforementioned steps are applied, with a relaxation for bandpass filter $[0.001 \text{ Hz}, 0.25 \text{ Hz}]$. 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 sections.

C.3.2 Post-Processing

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 average the resting functional connectome in the first day (which contains morning/afternoon scans) and call it Test. By the same token, we obtain Retest FC for resting condition.

FC’s matrix entries: For all considered fMRI images in this paper, we first threshold negative correlations. This is purely technical because one of the morphospace axis is built upon stochastic ground; hence, numerically it is not possible to utilize negative entries. The remaining matrix are, then, squared.

C.4 Improve Individual fingerprint

To improve identifiability in human functional connectome, we utilize data dimensional reduction technique described in Amico et al. $[3]$. Specifically, the framework is casted into an optimization problem where individual connectomes should look more similar to themselves, compared to others.

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

The identifiability matrix, denoted as $I \in \mathbb{R}_{n \times n}^+$, is task-based for which $I_{ij} \mid i, j = [n]$ represents the similarity - measured by Pearson Correlation between individual $i$-Test and $j$-Retest vectorized upper-triangular (functional) connectome matrices.
(under original and reconstructed conditions). Collaboratively, for rest or any given task, $I_{ii}$ $i \in [n]$ 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.

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

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 $PC$.

$$FC_{Recon}^k = \mu^k + \sum_{i=1}^{k} w_i^k PC_i$$

where $w_i^k$’s are weights corresponding to $PC_i$’s. In each step, reconstructed FCs are mapped, surjectively, from connectome space to identifiability score space. Hence, $k$ is found by computing $\text{argmax}_k [I_{diff}]$.

We apply this framework to rest and all available tasks for 100 unrelated subjects in HCP dataset, see Fig. S2 for further details. After the optimization procedure, the surviving PC components are then used to reconstruct the functional connectomes. For most tasks and rest, the optimal number of components surviving is approximated the number of subjects. We see that the result is plausible as if, theoretically, test and retest acquisitions of the same subject should not add dimension to the dataset.
In order to optimize the individual fingerprints of the subjects used in this study, we used the framework proposed by Amico et al. [3]. The method aims to reconstruct functional connectivity of each individual such that the connectome profile between scans are most "aligned" and, at the same time, maximize the connectome differences with other subjects in the study set. Such a method is applied for both resting and 7 task-evoked conditions. The input data has dimension of, at most, 200 as we utilized 100 unrelated subjects with two scanning patterns per subject. In each subgraph above, the optimal reconstructed number of orthogonal components is indicated by a black dot.

D Morphospace Analysis

In this section, we analyze both morpho-subspace in depth. We first introduce the formulation of each axis and then provide further characteristics and/or requirements/assumptions, if any.

D.1 The Coordinates of modularity morphospace

In this section, we describe the modularity morphospace Cartesian axes in greater details. The two coordinates are Module Trapping Efficiency (TE) and Exit Entropy (EE). Thus, for any functional module, $C \subset G$, we define a point in morphospace $\Omega$ to be

$$u(C) = (TE(C), EE(C))$$
D.2 Module Exit Entropy

This is the $y$-coordinate of $u(\mathcal{C})$.

Module exit entropy represents communicating preferences of $\mathcal{C}$ with respect to the rest of network $G$ from information theoretical viewpoint. The magnitude of this measure demonstrates the informational-theoretically integrative duties of a given community in an evolving network setting.

$$EE(\mathcal{C}) = -\sum_{i=1}^{\bar{m}} \psi_i \log(\psi_i) \log(\bar{m})$$

D.2.1 Numerator

The numerator of $EE(\mathcal{C})$, i.e. $-\sum_{i=1}^{\bar{m}} \psi_i \log(\psi_i)$, 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 belongs to other functional communities in $G$. Therefore, $-\sum_{i=1}^{\bar{m}} \psi_i \log(\psi_i)$ does not concern or incorporate the connectivity strength, represented by $w_{ij} \forall i \in \mathcal{C}, j \in \mathcal{J}$.

D.2.2 Numerical Range

The maximum value it can take on is 1 which represents no particular preference from $\mathcal{C}$ to $G\setminus\mathcal{C}$. Note that, conceptually, this is not the same as $L_S$ incorporated in $\mathcal{TE}(\mathcal{C})$ (as discussed later) as $EE(\mathcal{C})$ is driven solely from information theoretical viewpoint of the module with respect to its external topology. On the other hand, numerically, the value towards 0 demonstrates extreme preference in terms of integrative duties. It means that nodes in $\mathcal{C}$ have very specific nodes outside of $\mathcal{C}$ such that established channels of communication take place.

D.2.3 EE’s Normalization

This is the coordinate where normalization is possible. Note that since entropy is normalized by its maximum value (i.e. $\log(\bar{m})$), the number of exits $\bar{m}$ 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.

D.3 Module Trapping Efficiency

This is the $x$-coordinate of $u(C)$.

Module trapping efficiency assesses the characteristic of a functional commu-
nity based on how well it sustains its executive function under rich repertoire of
task-evoked conditions, relatively to its integration role, simultaneously. Recall
that module Trapping efficiency is formalized as followed:

$$\text{TE}(C) = \frac{||\tau||_2}{\mathcal{L}_C}$$

D.3.1 $\text{TE}(C)$ and $\mathcal{C}$-connectedness

Even though we have assumed that $G$ is connected, i.e. every nodes can
reach every other nodes in finite steps, there is no guarantee that the induced
topological structure, $\mathcal{C}$, as defined in the general setting section is connected. As
pointed out in [38] among others, a meaningful cluster should, at minimum, be
connected. To overcome this technicality, while preserving the original topology
of the induced subgraph, we have added a small perturbation to edges with zero
weight, i.e.

$$a_{ij} = \epsilon \iff a_{ij} = 0 \ | \forall i', j' \in \mathcal{C}$$

D.3.2 $\tau$ and topological sensitivity

In terms of sensitivity (to topological perturbations), since $\tau = (I_{|\mathcal{C}|} - Q_{\mathcal{C}})^{-1} 1_{|\mathcal{C}|}$,
it is trivial to see that $\tau$ is unique and specific to $Q_{\mathcal{C}}$. 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 $\tau$ would be altered.
For realistic network where topological symmetries are rare, the $L_2$-norm of $\tau$ 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, [46] but under TE, these two configurations are much different, see figure Fig. S3 for details.

Figure S3. Schematic presentation between two binary graphs with the same number of internal and external edges. Internally, both configurations are formed by a clique (of size 8), i.e. $K_8$. 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 different structural dynamics which is effectively measured through $\tau$. Externally, the left configuration, denoted as $G_l$, has evenly-distributed exits while the right one, denoted as $G_r$, has congested/bottle-necked exit. Additionally, $G_l$ bears much more symmetrical structure, compared to $G_r$. Another important result is that having the maximal internal subgraph like the clique structure does not necessarily help the notion of a "trapped" random walker. As a 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. Note that $L_{C}$ is the same for both configurations; hence, we only compare $\tau$ behavior.

D.3.3 Numerical Range

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 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

\[ M = \max_k [\text{TE}(S_k)] \mid \forall k \]

Such value is well-defined to be finite. An alternative way to see the trivial bound of the measures is as follows: Let us consider the entire network \( G \), we have:

\[ \text{TE}(C \equiv G) = \frac{||\tau||_2}{L_C} = \frac{\infty}{0} = \infty \]

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.

In the context of the data set at hand, we can, however provide a better bound then finiteness. We proceed by obtaining the maximum value of \( \text{TE} \) when all subjects and all tasks are under consideration which yields the result

\[ \max_{\text{subjects}, \text{tasks}} (\text{TE}) = 0.5064 \]

One can relate this numerical value with two factors: Functional connectome density and edge strengths, see Fig. S4 for further details.

**D.3.4 TE’s Normalization**

Realistically, since larger communities carry more exits which is driven purely from a topological viewpoint, \( L_C \) is a logical choice to normalize the magnitude of \( \tau \).

Additionally, \( L_C \) 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 \( \text{TE} \) as it converges to \( \text{TE}(C \equiv G) \).

\(^4\)Note that \( \frac{\infty}{0} \) is undefined. However, in such case, we define this quantity to be unbounded which is the notion of infinity.
(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 S4. 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 $\text{TE}$.

D.4 Task sensitivity on individual subspace

D.4.1 Task-sensitivity: A definition

We define task sensitivity for each FN and subject is the degree at which a task is recognizable from others. Intra-class correlation (ICC) is calculated for rest and 7 tasks with two observations (Test/Retest) per category.

D.4.2 Null Model

In this section, we describe the null model that test task identifiability for each morpho-subspace: $\text{TE}$ and $\text{EE}$. As mentioned in the main text, robust task sensitivities suggested from all functional communities in both Cartesian coordinate lays solid foundation to foster the cooperation among two axes.

D.5 Morphospace Trajectory

D.5.1 Randomization Process

Given a weighted network $A = [a_{ij}] = [w_{ij}]$, we convert $A$ into a distance matrix, $D = [d_{ij}]$ as follows:

$$\mathcal{D} = [d_{ij}] = [w_{ij}]^{-1}$$
Since matrix $A$ with thresholded entries now have the value of zero which result in infinity in matrix $D$. We set those entries to zero as default.

Next, we apply randomized algorithm $X_{\text{swap}}$, see [33] for more details, 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.

The algorithm preserves network size, density and degree sequence (hence, degree distribution of original network). Ideally, we would like that the randomized counterpart, denoted as $A_{\text{rand}}$ to be as different as possible compared to the original matrix, denoted as $A_{\text{orig}}$. We quantify this objective by looking at the difference between two graphs

$$Diss = \frac{\sum_{i,j=1}^{n} |A_{\text{rand}}(ij) - A_{\text{orig}}(ij)|}{\sum_{i,j=1}^{n} A_{\text{rand}}(ij)}$$

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, see Fig. S5 for details. In our case, we pick subject 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 then used the same number of changes for the investigated tasks state in the subsequent section E.2.

### D.5.2 Why do we study trajectory?

The main drive for studying trajectory is because it could provide further evidence in the design of individual subspace measurements. Specifically, if done correctly, measurements should highlight unique characteristics of functional communities. Hence, any destruction of such topology, at global scale, would
Figure S5. Subject 100307 resting state FC are randomized with the aforementioned step number of changes and the corresponding dissimilarity indices between the original graph and the randomized one. Due to computational demand of the produce, we only run Xswap for resting FC and choose the acceptable number of changes to be $2^{15}$. We apply the same number of changes to four considered tasks to demonstrate the trajectory of points.

also be identified by the morphospace itself. Furthermore, the randomized graph (with topological preserved features) get assigned the parcellation as the original one. This assignment allows us to test the robustness in morphospace architecture. We demonstrate this through applying randomization procedure to the first subject in HCP data set for rest and the first four tasks in HCP dataset. The result is shown in Figure Fig. S6.
Figure S6. Morphospace Trajectory Analysis: Subject 100307 with resting and task-evoked FCs are randomized using \texttt{Xswap} procedure. 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 the top left corner. This regime of the morphospace represent random exiting strategy from module $C$ (high value of EE) and high degree of dis-assortativity (low TE). This is an important results to see that functional networks’ topology is truly well-defined and highly reproducible across subject domain. Additionally, it also shows the robustness of morphospace design as it is capable of pointing out the destruction of meaningful topology and underlying communication within the brain connectome. Note that black square dot denote functional community TE and EE with no randomization. Color available online.

E Cognitive Flexibility

E.1 On individual subspace

In this section, we analyze the notion of flexibility induced from (individual) subspace perspective. As defined in the main text, functional reconfiguration and preconfiguration viewed under morphospace framework really boils down to behaviors of FN coordinates under several task-evoked and rest condition. Thus, we propose the notion of subspace cognitive flexibility by calculating standard deviation of individual axis’ coordinates from rest and all available tasks. Specifically, for a fixed subject, cognitive flexibility on FN $j$, over all
available tasks and rest, can be quantified as follow:

\[ F_{TE}^j = \sigma(TE(C_j)) \]

\[ F_{EE}^j = \sigma(EE(C_j)) \]

where \( \sigma([\bullet](C_j)) \) represents standard deviation of [\bullet] subspace if FN \( C_j \) under all aforementioned conditions. One of the reason we propose this notion is because it helps to lay the understanding of convex hull volume and why it is chosen to represent reconfiguration on the entire morphospace. The notion of standard deviation provides a straightforward understanding of which FN degree of movement under a specific morpho-subspace.

**Figure S7. Cognitive flexibility defined by individual morphospace axis:** In terms of module trapping efficiency, within the cortical region, Visual network has dominant flexibility compared to other functional networks. This quantitatively suggests evidence towards possible pupils’ diameter dilation during task-engaging periods. Other networks, such as DA and FP, appear to have relatively small flexibility, compared to VIS. In terms of module exit entropy, although VIS still has the highest cognitive flexibility, we see a much more balanced, well-leveled tendency, among all FNs. Comprehensively, cognitive flexibility of VIS on both subspaces implies that this network seems to be very active topologically and informational-theoretically.

Based on Figure Fig. S7, we observe that high cognitive flexibility in one coordinate does not imply similar trend in the remaining one. For instance, SUBC has second highest \( F_{TE}^{SUBC} \) but smallest \( F_{EE}^{SUBC} \). Comprehensively, results
in Fig. S7 suggest exclusive features for a given FN. Although not perfect, this shows further evidence of the morphospace to cover non-overlapping features of a given sub-systems. Moreover, we also observe reproducibility of results for test and retest for both axes.

E.2 On entire morphospace

E.2.1 Cognitive flexibility - A definition

Recall that, in the main text, we define the equivalent notion of cognitive flexibility using functional reconfiguration and preconfiguration for a given FN.

\[ \mathcal{F}_i = f(P_{FN}^{i}, R_{FN}^{i}) \]

where \( P_{FN}^{i} \) and \( R_{FN}^{i} \) represent functional preconfiguration and reconfiguration, respectively.

E.2.2 Functional Reconfiguration

In the main text, we address that once the points are well-defined to represent tasks per each functional community, we need now the notion that highlights the capacity to shift within this space. This measure needs to integrate both the notion of capability (potential); and desire to change. In this section, we provide a deeper analysis of the drive behind the usage of volume of the convex hull.

\[ R_{FN}^{i} = \text{Vol}(\text{Conv}(W_{FN}^{i})); \]

First of all, 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 extent of spread among the Pareto points. To this end, if one only uses the notion of distance to measure reconfiguration, three candidates for reconfigurations
emerge:

1. Exhaustive distance among all pairs of points in the hull $W$, i.e. $\sum_{i,j} d(v_i, v_j) | i \neq j, i, j \in \text{Conv}(W)$;

2. Exhaustive distance among all pairs of points type A, i.e. $\sum_{i,j} d(v_i, v_j) | i \neq j, i, j \in \delta(\text{Conv}(W))$;

3. Exhaustive distance among all pairs of points type B, i.e. $\sum_{i,j} d(v_i, v_j) | i \neq j, i, j \in int(\text{Conv}(W))$;

However, we see that all three fall short in the following regards:

- extremely hollow (empty) interior space formed by hull vertices, i.e. vertices belong to Pareto front induced by task points per FN;

- linearity between task points

We see that this implies such interior space is non-accessible from the combinatorial viewpoint of connectomes. More specifically, the notion of distance does not cover the space of possibility imposed by the object under study, in this case the trapping and exiting characteristics of functional communities under various tasks and rest. Further, we see that option 1 above also uses points that belong to the interior of the hull. This does not support the concept of cognitive flexibility which highlight the capacity of pushing multiple forefronts in modularity morphospace.

We see that extreme cavity does not highlight the potential of cognitive shift within the hull identified by $\text{conv}(W)$. Moreover, based on figure S6, we see that morphospace trajectories are much more complicated than linearity; hence, suppose that the subject is asked to switch tasks, functional reconfiguration, as highlighted through modularity morphospace, measured by distance only covers linear trajectory (between the two chosen tasks). This is not a realistic assumption. Lastly, functional reconfiguration measured by aforementioned
methods does not effectively highlight the space of possibility as identified by $\text{Conv}(W)$. This is very important to model cognitive shifts within this space.

In order to strive beyond those shortcomings, we propose the notion of volume (or area in this morphospace) enclosed by the convex hull, i.e. $\text{Vol}((\text{Conv}(W))$.

Notice that the notion of area is also limited in the regard that if there is area of impossibility, i.e. cavity within the convex hull, this measure would also estimates the true shifting capacity. Nonetheless, compared to the notion of exhaustive distance sum which extremely underestimate shifting capacity, we see that the notion of area are much more robust to reflect both complexity of trajectory and cognitive adaptation capacity within this space. Under the construction of modularity morphospace in Euclidean space, the notion of area is well-defined. We discuss these limitations and strengths of modelling cognitive shifts using morphology in greater details in Discussion section - main paper.

### E.2.3 Functional Preconfiguration

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 maintext for functional preconfiguration:

$$P_{FN}^i = ||\text{Rest}_{FN}^i - \eta_{W_{FN}}||_2$$

where $\text{Rest}_{FN}^i$ and $\eta_{W_{FN}}$ 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.

E.2.4 Subject Sensitivity

- To quantify subject sensitivity (through cognitive flexibility), 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.

F Behavioral Measure Analysis

F.1 Iterative Multi-Linear Regression Model (MLM)

F.1.1 Model Description

We apply iteratively multi-linear correlation models (MLM) to correlate $F_i = f(R_{FN}^i, P_{FN}^i)$ with various behavioral measures, $\triangleright_i$. We hypothesize that highly subject sensitive predictor, as described in Fig. 5C (main text) should be prioritized in MLM model.

Iteratively, we start by using only 1 predictor ($P_{FP}^i$); in every subsequent step, we append one extra predictor to the existing one(s), again, accordingly per panel 2C. At the end of iterative process, we consequently obtain 16 MLMs.
F.1.2 Optimal MLM - A selection process

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.

F.2 Model Specificity (MS)

F.2.1 Model Description

We further test the strength of our hypothesis by splitting available data into two subsets: inquiry and validation set. Specifically, we first extract the optimal number of predictors by applying the procedure described in Section F.1.2.

We then proceed with the model specificity by creating 2000 simulations; for each simulation - indexed by \( j = \{1, 2, 3, ..., J = 2000\} \) - we first find a randomized order of indices from 1 to 100, denoted as \( \vec{d} \), and divide them into five batches (indexed by \( i = \{1, 2, ..., I = 5\} \)) of 20 subjects. In other words, each batch of 20 randomly picked subjects, indexed by the set \( W_i \), 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: \( \vec{d} = W = \bigcup W_i \). 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 \( \vec{\beta} \in \mathbb{R}^{|*|} \) where \(|*|\) denotes the optimal MLM driven by procedure described above (Notice that we use the same notation in the main text under Fig. 6 as well). These corresponding coefficients are then used to predict the remaining 20 unused data points, indexed by \( w \in W_i \), denoted as \( \hat{y} \).

\[
\hat{y}_w = \hat{\beta}_0 + \left[ \left\{ P_{w}^{FN}, R_{w}^{FN} \right\}|*| \right] \hat{\vec{\beta}}
\]
where \( \{P^N_w, R^N_w\}^{[*]} \in \mathbb{R}^{+,*} \) is the \([*]\)-tupled vector representing functional preconfiguration, reconfiguration, obeying the descending order of concatenated subject sensitivity in Fig 5.C located in the main text.

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 \( R_i, \forall i = 1, 2, ..., I = 5 \). 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

\[
R_j = \sum_{i=1}^{I} R_{ij} = \langle R_{.,j} \rangle
\]

\[
\sigma_j = \sqrt{\frac{\sum_{i=1}^{I} (R_{ij} - R_j)^2}{I}}
\]

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

\[
\mu_0 = \frac{\sum_j \sum_i R_{ij}}{I \times J}
\]

\[
\sigma_0 = \sqrt{\frac{\sum_{j=1}^{J} \sigma_j^2}{J}}
\]

F.2.2 MS’s null model

Similarly to the MLMs, we want to test the authenticity of VM’s 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 \( \mu_1 \))

\[
R_{j}^{rand} \sim N(\mu_1, \sigma_1)
\]
F.2.3 Model Specificity

Pair t-tests are applied between two aforementioned distributions to test the capacity of cognitive flexibility predictors towards behavioral measures. Interestingly, given the investigated behavioral measures, all null model empirical distributions have very similar first and second moments.

References

1. E. Amico, A. Arenas, and J. Goñi. Centralized and distributed cognitive task processing in the human connectome. arXiv preprint arXiv:1802.00334, 2018.

2. E. Amico and J. Goñi. Mapping hybrid functional-structural connectivity traits in the human connectome. Network Neuroscience, pages 1–17, 2018.

3. E. Amico and J. Goñi. The quest for identifiability in human functional connectomes. Scientific reports, 8(1):8254, 2018.

4. 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, 148:201–211, 2017.

5. P. Anderson. Perspective: Complexity theory and organization science. Organization science, 10(3):216–232, 1999.

6. A. Avena-Koenigsberger, J. Goñi, R. Solé, and O. Sporns. Network morphospace. Journal of the Royal Society Interface, 12(103):20140881, 2015.

7. A. Avena-Koenigsberger, B. Misic, and O. Sporns. Communication dynamics in complex brain networks. Nature Reviews Neuroscience, 19(1):17, 2018.
8. 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*, 80:169–189, 2013.

9. D. S. Bassett and M. S. Gazzaniga. Understanding complexity in the human brain. *Trends in cognitive sciences*, 15(5):200–209, 2011.

10. D. S. Bassett and O. Sporns. Network neuroscience. *Nature neuroscience*, 20(3):353, 2017.

11. 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*, 108(18):7641–7646, 2011.

12. D. Bavelier, D. M. Levi, R. W. Li, Y. Dan, and T. K. Hensch. Removing brakes on adult brain plasticity: from molecular to behavioral interventions. *Journal of Neuroscience*, 30(45):14964–14971, 2010.

13. M. A. Bertolero, B. T. Yeo, and M. D’Esposito. The modular and integrative functional architecture of the human brain. *Proceedings of the National Academy of Sciences*, 112(49):E6798–E6807, 2015.

14. E. Bullmore and O. Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186, 2009.

15. E. T. Bullmore and D. S. Bassett. Brain graphs: graphical models of the human brain connectome. *Annual review of clinical psychology*, 7:113–140, 2011.

16. R. B. Cattell. Theory of fluid and crystallized intelligence: A critical experiment. *Journal of educational psychology*, 54(1):1, 1963.
17. 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*, 83(1):238–251, 2014.

18. 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*, 110(33):13316–13321, 2013.

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

20. J. L. Doob and J. L. Doob. *Stochastic processes*, volume 7. Wiley New York, 1953.

21. 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*, 339:12–21, 2016.

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

23. M. E. Dyer and A. M. Frieze. On the complexity of computing the volume of a polyhedron. *SIAM Journal on Computing*, 17(5):967–974, 1988.

24. J. Flum and M. Grohe. *Parameterized complexity theory*. Springer Science & Business Media, 2006.

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

26. S. Fortunato. Community detection in graphs. *Physics reports*, 486(3):75–174, 2010.
27. 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*, 536(7615):171–178, 2016.

28. 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*, 80:105–124, 2013.

29. 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*, 8(3):e58070, 2013.

30. 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*, 112(28):8762–8767, 2015.

31. J. R. Gray, C. F. Chabris, and T. S. Braver. Neural mechanisms of general fluid intelligence. *Nature neuroscience*, 6(3):316, 2003.

32. 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*, 100(1):253–258, 2003.

33. S. Hanhijärvi, G. C. Garriga, and K. Puolamäki. Randomization techniques for graphs. In *Proceedings of the 2009 SIAM International Conference on Data Mining*, pages 780–791. SIAM, 2009.

34. J. G. Kemeny, J. L. Snell, et al. *Finite markov chains*, volume 356. van Nostrand Princeton, NJ, 1960.
35. L. Khachiyan. Chapter iv. complexity of polytope volume computation. *New trends in discrete and computational geometry*, 10:91, 1993.

36. J. Kruschwitz, L. Waller, L. Daedelow, H. Walter, and I. Veer. General, crystallized and fluid intelligence are not associated with functional global network efficiency: A replication study with the human connectome project 1200 data set. *NeuroImage*, 171:323–331, 2018.

37. S. Kullback. *Information theory and statistics*. Courier Corporation, 1997.

38. F. D. Malliaros and M. Vazirgiannis. Clustering and community detection in directed networks: A survey. *Physics Reports*, 533(4):95–142, 2013.

39. 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*, 5:4, 2011.

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

41. J. D. Medaglia, M.-E. Lynall, and D. S. Bassett. Cognitive network neuroscience. *Journal of cognitive neuroscience*, 27(8):1471–1491, 2015.

42. M. Mennes, B. B. Biswal, F. X. Castellanos, and M. P. Milham. Making data sharing work: the fcp/indi experience. *Neuroimage*, 82:683–691, 2013.

43. M. Milano, P. H. Guzzi, and M. Cannataro. Using multi network alignment for analysis of connectomes. *Procedia Computer Science*, 108:1155–1164, 2017.

44. E. Miller and J. Wallis. Executive function and higher-order cognition: definition and neural substrates. *Encyclopedia of neuroscience*, 4(99-104), 2009.
45. 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*, 2(02):285–302, 2018.

46. M. E. Newman. Modularity and community structure in networks. *Proceedings of the national academy of sciences*, 103(23):8577–8582, 2006.

47. T. E. Nichols, S. Das, S. B. Eickhoff, A. C. Evans, T. Glatard, M. Hanke, N. Kriegeskorte, M. P. Milham, R. A. Poldrack, J.-B. Poline, et al. Best practices in data analysis and sharing in neuroimaging using mri. *Nature neuroscience*, 20(3):299, 2017.

48. T. G. C. H. Page. Qhull the geometry center home page, 1995.

49. 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*, 72(4):665–678, 2011.

50. 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*, 84:320–341, 2014.

51. M. Rosvall, D. Axelsson, and C. T. Bergstrom. The map equation. *The European Physical Journal-Special Topics*, 178(1):13–23, 2009.

52. M. Rosvall and C. T. Bergstrom. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105(4):1118–1123, 2008.

53. 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*, page 431833, 2018.
54. R. Schuetz, N. Zamboni, M. Zampieri, M. Heinemann, and U. Sauer. Multidimensional optimality of microbial metabolism. *Science*, 336(6081):601–604, 2012.

55. D. H. Schultz and M. W. Cole. Higher intelligence is associated with less task-related brain network reconfiguration. *Journal of neuroscience*, 36(33):8551–8561, 2016.

56. C. E. Shannon. A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423, 1948.

57. 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.

58. 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*, 92(2):544–554, 2016.

59. 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*, pages –, 2019.

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

61. 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*, page 1217405, 2012.

62. 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*, 80:144–168, 2013.

63. C. Spearman. "general intelligence," objectively determined and measured. *The American Journal of Psychology*, 15(2):201–292, 1904.

64. O. Sporns. Network attributes for segregation and integration in the human brain. *Current opinion in neurobiology*, 23(2):162–171, 2013.

65. R. Thomas, R. M. Shearman, and G. W. Stewart. Evolutionary exploitation of design options by the first animals with hard skeletons. *Science*, 288(5469):1239–1242, 2000.

66. B. Thomas 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*, 106(3):1125–1165, 2011.

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

68. 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*, 9(5):e96196, 2014.

69. M. P. van den Heuvel, C. J. Stam, R. S. Kahn, and H. E. H. Pol. Efficiency of functional brain networks and intellectual performance. *Journal of Neuroscience*, 29(23):7619–7624, 2009.

70. 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*, 80:62–79, 2013.
71. D. C. Van Essen, K. Ugurbil, E. Auerbach, D. Barch, T. Behrens, R. Bucholz, A. Chang, L. Chen, M. Corbetta, S. W. Curtiss, et al. The human connectome project: a data acquisition perspective. *Neuroimage*, 62(4):2222–2231, 2012.

72. H. Xie, V. D. Calhoun, J. Gonzalez-Castillo, E. Damaraju, R. Miller, P. A. Bandettini, and S. Mitra. Whole-brain connectivity dynamics reflect both task-specific and individual-specific modulation: A multitask study. *Neuroimage*, 2017.

73. 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*, 106(3):1125–1165, 2011.

74. G. M. Ziegler. *Lectures on polytopes*, volume 152. Springer Science & Business Media, 2012.