

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS MATEMÁTICAS

UNIVERSIDAD POLITÉCNICA DE MADRID
ESCUELA TÉCNICA SUPERIOR DE INGENIEROS DE
TELECOMUNICACIÓN

**MÁSTER EN TRATAMIENTO ESTADÍSTICO COMPUTACIONAL DE
LA INFORMACIÓN**



TRABAJO DE FIN DE MÁSTER

**Análisis de redes cerebrales a través de nuevos paradigmas
computacionales. Teoría de grafos**

José Pedro Manzano Patrón

Director

Pedro José Zufiria Zatarain
Salvador Jiménez Burillo

Madrid, 2018

Resumen

El análisis multimodal de datos cerebrales (EEG, MEG, fMRI, DTI) está llamado a ser una de las grandes estrategias de búsqueda de nuevos neuro-biomarcadores en la medicina personalizada y en los futuros avances del sector. Sin embargo, hoy en día deben tratarse por separado al no existir un marco de análisis común. Por otro lado, las técnicas clásicas de análisis de redes cerebrales empiezan a presentar serios problemas para estudiar las variaciones en los patrones espacio-temporales de actividad y conectividad funcional.

Este trabajo revisa y estudia los principales métodos de teoría de grafos como una alternativa capaz de abordar todos estos retos bajo un mismo entorno computacional. Se abordarán aspectos básicos en la definición de la red, como el tipo de grafo (binario o ponderado) o métricas de conectividad. Tras dicha revisión, se realizará una comparación de las propiedades en redes funcionales y anatómicas de sujetos sanos y pacientes con distintas patologías neurológicas. Por último, se aplicarán técnicas de detección dinámica de comunidades para identificar el estímulo y la tarea llevada a cabo por un sujeto durante un experimento de datos simulados de fMRI.

Abstract

The multimodal analysis of brain data (EEG, MEG, fMRI, DTI) is one of the most promising strategies in the search of new neuro-biomarkers for the personalized medicine and in the future advances of the field. However, nowadays they are still processed separately due to the lack of a common analysis framework. On the other hand, classical analysis techniques of brain networks begin to present serious problems to study the variations in the spatio-temporal patterns of activity and functional connectivity.

This paper reviews and studies some of the main methods of Graph Theory as a potential alternative of addressing all these challenges under the same computational framework. Basic aspects in the definition of the network will be addressed, such as the type of graph (binary or weighted) or connectivity metrics. After this review, a comparison of the properties in functional and anatomical networks between healthy subjects and patients with different neurological pathologies will be carried out. Finally, dynamic community detection techniques will be applied to identify the stimulus and the task carried out by a subject during an experiment with simulated fMRI data.

Tabla de contenido

1. INTRODUCCIÓN	5
2. OBJETIVOS Y MOTIVACIÓN	6
3. ESTADO DEL ARTE	8
3.1. TEORIA DE GRAFOS.....	8
3.2. DETECCIÓN DE COMUNIDADES	11
4. METODOLOGIA	20
4.2. PARADIGMA EXPERIMENTAL	20
4.3. DECISIONES BASADAS EN LA BIBLIOGRAFÍA	22
4.4. HERRAMIENTAS UTILIZADAS.....	25
5. RESULTADOS	26
5.2. SIMULACIÓN DE DATOS	26
5.3. BÚSQUEDA PARÁMETROS DE LA RED.....	27
5.4. COMPARACIÓN SUJETOS SANOS vs. PACIENTES	34
5.4. DETECCIÓN DINÁMICA DE COMUNIDADES.....	41
6. CONCLUSIONES Y TRABAJO FUTURO.....	48
6.1. TRABAJO FUTURO.....	50
<i>Referencias</i>	52

1. INTRODUCCIÓN

Las redes son estructuras que pueden encontrarse a diario en la naturaleza y en la sociedad. Este tipo de relaciones son estudiadas por lo que se conoce como *network science* o ciencia de redes [1]. La teoría de grafos es quizá la herramienta matemática más conocida dentro de esta disciplina. En ella, las interacciones que suceden dentro de una red se formalizan a través de un conjunto de nodos y enlaces que representan cómo se relacionan sus elementos.

El cerebro puede ser interpretado como una de las redes más complejas que existen actualmente. La aplicación de esta disciplina al campo de la neurociencia computacional ha dado lugar a numerosos trabajos sobre conectividad cerebral en los últimos 10 años. Una de las ventajas de esta técnica es que permite realizar estudios a diferentes escalas temporales y espaciales (molecular o micro-escala, meso-escala, exa-escala o redes) [2,3,4,5] y sobre distintos tipos de asociaciones o conectividades: anatómica, funcional y efectiva [6,7,8,9] (aunque en este trabajo solo se considerarán la anatómica y la funcional). Estos análisis han permitido poner de manifiesto algunas características de las redes cerebrales de múltiples especies tanto a nivel global como local [10,11].

La actividad, su dinámica y la estructura cerebral están condicionadas por un sinfín de posibles factores: genética, edad, sexo, ciclo circadiano, capacidad y entrenamiento cognitivo, estado emocional, estado de salud, etc. Asociado a cada uno de estos factores existe un posible objeto de estudio. Una de las áreas que más atención atrae es caracterizar cómo afectan diferentes enfermedades a la estructura cognitiva de los sujetos [12,13,14,15], como el Alzheimer [16], esquizofrenia [17], autismo [18], etc. con el objetivo de desarrollar tratamientos o, simplemente, entender mejor la enfermedad [19,20].

Sin embargo, el campo donde más trabajo se está desarrollando actualmente es en el estudio de redes cerebrales y la dinámica inherente a ellas. El cerebro posee una organización modular y jerárquica que le permite realizar múltiples tareas de forma simultánea y sincronizada. La herramienta más adecuada para este propósito en teoría de redes es lo que se denomina como detección dinámica de comunidades, aunque aún afronta algunos problemas a la hora de definir las redes y la resolución espacio-temporal con la que resolver este tipo de problemas.

De este modo, la memoria se organiza como sigue: el capítulo 2. *Objetivos y Motivación* expone los argumentos y la motivación de este trabajo. El capítulo 3. *Estado del arte* constituye una amplia revisión de los principales retos en la aplicación de teoría de grafos a las redes cerebrales y sobre las estrategias más comunes para la detección dinámica de comunidades. El capítulo 4. *Metodología* explicará las herramientas y el procedimiento diseñado para evaluar y discutir los parámetros comentados en el capítulo 3. *Estado del Arte*. Los resultados de dichos experimentos se mostrarán en el capítulo 5. *Resultados*. La memoria finaliza con el capítulo 6. *Conclusiones y Trabajo Futuro* donde se discutirán los resultados obtenidos, su aplicación y posibles líneas de mejora. Adjunto a la memoria se proporcionan una serie de anexos que sirven de soporte al trabajo. Exceptuando el Anexo I. *Propiedades de un grafo*, los anexos están en formato notebook de Jupyter, accesibles en [21] tanto en ipynb, html o pdf.

2. OBJETIVOS Y MOTIVACIÓN

En octubre de 2017 el alumno de este trabajo presentó como Trabajo Final de Máster del máster en Inteligencia Artificial Avanzada de la UNED el diseño de una plataforma computacional de análisis dinámico de redes cerebrales [21]. La propuesta conceptual consiste en un marco común desde el que afrontar los dos principales retos en la neurociencia computacional actual: **el análisis dinámico**, donde se estudia la evolución de los patrones de actividad cerebral durante un experimento, y el **análisis multimodal**, cuyo objetivo es combinar diferentes tipos de datos (anatomía, consumo de glucosa, similitud en los patrones de actividad cerebral, etc.) con el objetivo de revelar nueva información que de forma unimodal no es posible capturar [22,23,24].

Una vez que se tiene esta nueva construcción sobre las relaciones anatómico-funcionales, se pueden realizar diferentes tipos de aproximaciones para localizar redes funcionales. Por ejemplo, en la memoria de dicho trabajo se propuso una técnica de *clustering* difuso que utiliza conocimiento experto para generar una serie de reglas que permitan identificar correctamente la situación de las posibles redes cerebrales en cada instante temporal. Gran parte del proceso es *data-driven* (esto es, no dependiente de parametrización), uno de los objetivos principales de la plataforma, y agnóstico en cuanto a paradigma experimental (es decir, el procedimiento analítico es el mismo tanto para casos de *resting-state* como *task-design*) y a tipo de sujeto (el procedimiento analítico es el mismo e independientes de si se trata de pacientes enfermos, sujetos sanos, especies animales, etc.).

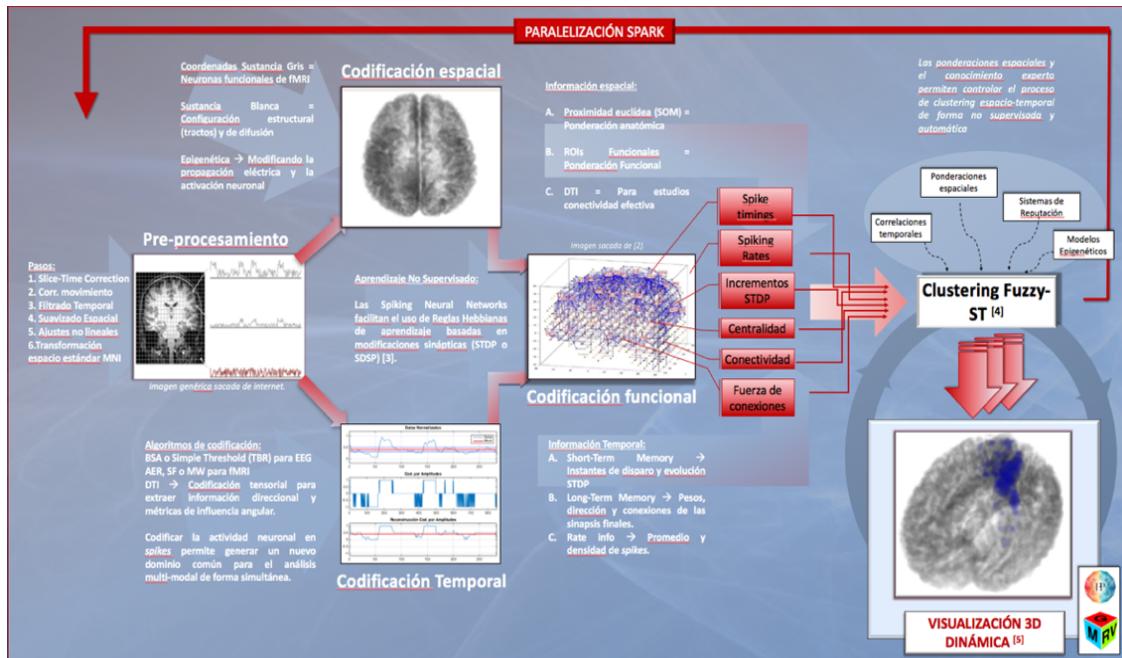


Fig. 1. Diseño computacional de la plataforma presentada en [21], actualmente en desarrollo. La imagen fMRI se somete a un proceso de codificación espacio-temporal.. Una vez construida la representación anatómico-funcional, se aplican reglas bio-inspiradas como Spike-Timing Dependant Plasticity (STDP) para extraer las relaciones funcionales de la red.

Por lo tanto, este trabajo tiene un doble propósito:

- Servir como una **revisión general y actualizada** de la aplicación de teoría de grafos al estudio de redes cerebrales. Generalmente, las *reviews* existentes en la literatura solo se centran en uno de los campos de estudio (detección de comunidades, comparación con alguna patología concreta, etc.) y, dado el rápido avance de la neurociencia computacional, es difícil encontrar documentación actualizada y que englobe los diferentes aspectos de esta metodología. Por esta razón, la revisión del estado del arte ocupará una parte considerable de esta memoria y contará con múltiples referencias.
- **Implementación de las principales técnicas revisadas** para estudiar la teoría de grafos como alternativa a los métodos de *clustering* en la detección de redes cerebrales (a través de detección de comunidades) y su aplicación a datos tanto funcionales como estructurales para describir las principales propiedades de la red. Si los resultados son satisfactorios, las herramientas desarrolladas y las conclusiones obtenidas podrían servir para añadir adecuadamente esta metodología a la plataforma.

3. ESTADO DEL ARTE

3.1. TEORIA DE GRAFOS.

La teoría de grafos es una herramienta de las matemáticas y ciencias de la computación que estudia las propiedades de una red. Formalmente, un grafo $G = (V, E)$ es una pareja en la que V es un conjunto no vacío de vértices y E es un conjunto de aristas o enlaces. Este trabajo se centrará en grafos no dirigidos, lo que quiere decir que V consta de pares no ordenados de vértices, tales como $\{i, j\} \in E$, donde i, j son adyacentes y E se representa con una línea no orientada que une dichos vértices.

Una propiedad interesante de las redes es que pueden variar y evolucionar a lo largo del tiempo, lo que da origen a una red temporal [25]. Una red temporal es una red $\mathcal{M} = (V, w, l)$ con V vértices, $w: V \times V \rightarrow R$ como función de pesos en los enlaces, y l como función de etiquetado que permite identificar cada vértice v como una pareja de nodos por capa, de tal forma que $l(v) = (i, t) = i_t$ donde v es el nodo i en el instante t . Los enlaces entre cada par de nodos quedan definidos por w como $w(i_s, j_t)$. Este enlace será denominado *intra-layer* o intra-capa si $s=t$ (pertenecen a la misma ventana temporal) y entre-capas o *inter-layer* en caso contrario. Por otro lado, de manera general:

- Se considera que las redes son diagonales y ordinales, lo que significa que no existirá ningún enlace entre i_t, j_s a menos que $i = j$ y $s = t + 1$.
- Se define A_{ijt} como la matriz de adyacencia para la ventana temporal o capa t y B_{it} la matriz de enlaces entre el nodo i en las capas t y $t + 1$. También se puede considerar B_{ijt} como la matriz de enlaces entre los nodos i, j en las capas t y $t + 1$, aunque no se utilizará en los experimentos de este trabajo.

Hay 2 principales focos de estudio en redes temporales: cómo los patrones de actividad cambian dada una estructura de red fija (*dynamics on networks*) y cómo la red evoluciona, es decir, la reconfiguración de nodos y/o enlaces (*dynamics of networks* o *evolutionary networks*). Aunque actualmente los estudios se centran en un caso u otro, en el futuro serán algo inseparable [24].

Nótese que la definición de red temporal asume una topología multicapa, donde cada capa equivale a un instante o ventana temporal. Es decir, una red temporal es una red multicapa pero una red multicapa no siempre es una red temporal [26,27]. A lo largo de la memoria se hará referencia a capa, ventana temporal, *snapshot* o *layer* de manera indiferente.

Con estas definiciones, parece inmediato asociar la topología de una red a la estructura cerebral: las neuronas serían representadas por nodos y los enlaces serían los axones que unen cada par de neuronas, ponderados por la fuerza sináptica. Sin embargo, esto encierra una gran complejidad tanto a nivel técnico (no existen técnicas de adquisición de imagen que puedan conseguir tal resolución de todo el cerebro), a nivel computacional (son miles de millones de nodos, modelos de funcionamiento biológico diferentes en la mayoría de ellos, y un número mucho mayor de enlaces) y, por supuesto, a nivel analítico (enlaces direccionales, bidireccionales, dinamismo de la red, creación/eliminación de nodos/enlaces, factores genéticos/químicos/ambientales, etc.) [28,29,30,31].

Todo esto ha impulsado que los métodos de ciencia de redes se hayan expandido en múltiples direcciones, abandonando las simples métricas descriptivas hacia un estudio sobre las dinámicas de la red y la evolución de las comunidades a lo largo del tiempo [32] o simulaciones de

generative models [33] que tratan de enlazar comportamientos neurofisiológicos de gran escala con la arquitectura física del cerebro [34,35].

No obstante, la *network neuroscience* es un campo realmente joven (aproximadamente 10-15 años, aunque es en los últimos 5 donde mayor número de publicaciones y avances se han producido) y aún existen multitud de retos que afrontar. Obviando la variabilidad que puede existir en este tipo de estudios y/o las implicaciones que tiene el pre-procesamiento de datos cerebrales sobre la información resultante, la teoría de grafos no está exenta de críticas: desde la sensibilidad que existe respecto de la definición de los nodos y los enlaces de la red hasta otros aspectos inherentes en casi todos los métodos de neurociencia computacional, como la resolución espacio/temporal (a veces limitada por la propia técnica de adquisición) o los problemas de fiabilidad y reproducibilidad de los resultados. En este sentido, unas de las ramas donde más se está trabajando actualmente es en la formalización de test estadísticos que permitan expresar la robustez y significancia de los resultados así como para la comparación entre redes [36,37,38]

En este trabajo se estudiarán dos tipos de asociaciones: la conectividad anatómica y la funcional. La construcción de las redes y, por lo tanto, la interpretación de los resultados para cada una de ellas puede ser algo diferente. El Anexo I. *Propiedades de un grafo* es una guía de referencia útil para consultar algunas de las características y propiedades de las redes a las que se harán alusión a partir de este capítulo.

3.1.1. TIPOS DE REDES CEREBRALES

a) Análisis funcional

La conectividad funcional se define como la similitud de actividad BOLD (*Blood Oxygen Level-Dependant*) que tienen dos regiones del cerebro en una ventana de tiempo dada, es decir, que se comporten igual o realicen el mismo proceso [39,40]. Con este análisis pueden encontrarse diversas redes funcionales que son (des)activadas en función de la tarea, situación o enfermedad. Esta asociación (**enlaces**) entre regiones cerebrales (**nodos**) puede definirse como una simple correlación lineal entre las series temporales o con métricas más complejas como Información Mutua, Entropía, *Phase-coherence* o *Cross-correlations* (**métricas**) [28,29]. El resultado de esta asociación entre nodos formaría lo que se define como **matriz de conectividad o adyacencia** sobre la que construye el grafo.

Para obtener datos funcionales es necesaria su adquisición a través de resonancia magnética funcional (fMRI). De esta forma, se adquiere una imagen 3D de todo el cerebro cada TR segundos (donde TR=2 seg. aproximadamente), de donde se extrae la actividad asociada a cada *voxel* en ese instante. Una vez que la imagen se somete a un complejo procesamiento (filtros, *data cleaning*, registro a *templates* anatómicos estándares que permitan comparaciones entre sujetos, etc.), se pueden extraer las series temporales de cada región, que representan el contraste BOLD en cada momento de adquisición [41].

En este punto surgen varias cuestiones importantes sobre la formación del grafo:

- ¿Cómo se definen los nodos? Una adquisición común de 3T proporciona más de 100.000 voxels. Por definición, las señales son muy ruidosas e inestables, por lo que es común delimitar Regiones de Interés (ROIs) sobre las que trabajar [42]. Dentro de las múltiples opciones, las principales se basan en utilizar atlas funcionales predefinidos [43,44], basarse en regiones anatómicas como las áreas de Brodmann [45], o hacer un Análisis

de Componentes Independientes (ICA) [46] y utilizar estas componentes como nodos de la red. En cualquiera de los tres casos, a cada uno de estos nodos se le asigna la serie temporal promedio de todos los voxels que haya en su interior, lo que implica una tremenda pérdida de posible información relevante (aunque también se elimina ruido y se consigue una mayor estabilidad).

- ¿Qué métrica de asociación utilizar para construir la matriz de adyacencia? Cuanto más compleja sea la métrica, más complicada será la interpretación de la red. Cuanto más simple, menos información puede que estemos capturando entre las series temporales de cada región. Aparece así un *trade-off* a la hora de definir el tipo y magnitud de los enlaces. Existen propuestas alternativas a la correlación de Pearson para capturar no linealidades, como la información mutua, *phase-synchronization* o correlaciones a partir de las *wavelets* [47].

Por otro lado, aunque las métricas basadas en correlaciones suelen ser sencillas y rápidas de computar, son propensas a “generar” enlaces entre nodos acoplados de forma indirecta (*transitive closure*). Esta tendencia suele resultar en redes excesivamente densas, clusterizadas y modulares, con múltiples dependencias que surgen de las interacciones indirectas de la estructura anatómica subyacente (pueden no ser enlaces funcionales reales, por tanto). Aunque las correlaciones parciales [48] pueden resolver este problema y proporcionar redes más dispersas, requieren de más muestras para conseguir estabilidad, los resultados son más difíciles de interpretar y pueden estar eliminando algunas interacciones reales. Por todo esto y por su fácil interpretación, la correlación lineal sigue siendo la opción más común.

- ¿Qué ventana temporal? El *sliding-window* es el método más común para el análisis dinámico. Dado que los tiempos de adquisición en fMRI rondan los 2 segundos, aparece un *trade-off* entre la potencia estadística (es decir, el número de puntos temporales en cada ventana) y la capacidad de capturar las fluctuaciones temporales. Por ejemplo, utilizar intervalos de 30 TRs produce ventanas temporales mínimas de alrededor de un minuto. Se ha demostrado que en función de la ventana temporal, las dinámicas capturadas (y, por lo tanto, los resultados) pueden ser completamente diferentes [49].

En experimentos donde se desconoce la periodicidad del patrón que se quiere capturar se acaba optando por utilizar toda la serie temporal, sin enventanado, lo que estaría suponiendo que el comportamiento del cerebro es estático (red estática), cuando se ha demostrado que existen fluctuaciones y patrones de actividad a lo largo del tiempo [24,50].

- Redes funcionales: Las redes funcionales son regiones co-activadas de la misma forma a lo largo del tiempo con el fin de llevar a cabo alguna función. Un ejemplo de red funcional es la que se forma por aquellas áreas en el cerebro (regiones dorsolaterales y prefrontales, la corteza insular y el SMA/pre-SMA) cuya actividad se incrementa (y con patrones de activación similares) cuando es necesario realizar una tarea de forma consciente; se le denomina *Task-positive Network* o *Dorsal Attention Network* (DAN) [43,44]. En análisis de grafos, esta búsqueda puede realizarse a través de detección de comunidades. Sin embargo, las comunidades se suelen suponer sin solapamiento espacial (sobre todo si se utilizan regiones anatómicas para definir los nodos), lo que se desvía del funcionamiento real, ya que una misma región cerebral puede estar procesando información de dos redes funcionales distintas pero simultáneas (véase la amígdala o el hipotálamo). Esto conduce de nuevo a pérdida de información, ya que el

nodo acabará perteneciendo (en el mejor de los casos) a la comunidad a la que pertenezca con más fuerza.

Es importante recordar que la conectividad funcional no está sujeta a la conectividad anatómica (aunque sí que puede verse influida [51,52]), por lo que dos regiones separadas pueden estar conectadas funcionalmente.

b) Análisis estructural o anatómico

Conceptualmente, se pueden tomar consideraciones muy parecidas para el análisis de grafos sobre datos anatómicos. En este caso, los enlaces sí son físicos y representan los tractos de sustancia blanca que unen las distintas regiones cerebrales. Generalmente, las regiones unidas corresponden a la corteza y a la sustancia gris del cerebro, donde se produce el procesamiento neuronal que se analiza con los datos funcionales. Existen múltiples formas de adquisición, siendo DTI (*Difussion Tensor Imaging*) una de las más precisas y conocidas [53,54].

La construcción de los nodos aquí sí se que corresponde con regiones anatómicas. La fuerza de los enlaces viene definida por el volumen o densidad de axones que unen cada una de las regiones, extrayendo así estimaciones sobre la probabilidad y la fuerza de la conectividad entre un nodo u otro. Consideraciones explicadas en el apartado anterior como el tamaño de las regiones (ROIs), solapamiento entre regiones o las métricas utilizadas siguen presentes en este análisis.

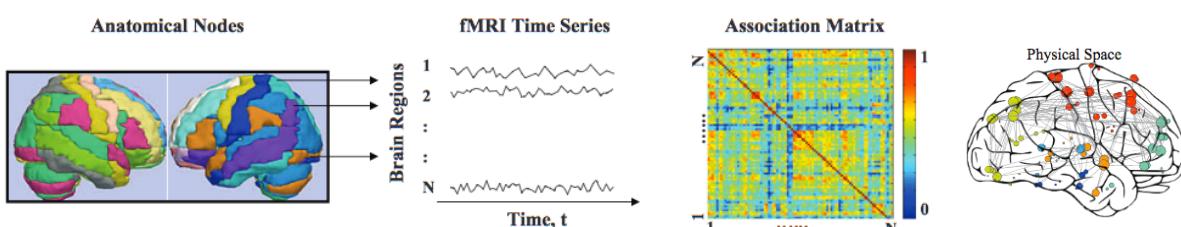


Fig.2 Extraído de [55]. Análisis funcional de datos cerebrales a partir de resonancia magnética funcional. Tras procesarse la imagen, se seleccionan las regiones de interés (ROIs), de las que se extraen las series temporales BOLD promedio. A través de métricas de asociación, como la correlación, se genera la matriz de conectividad, que dará lugar al grafo a analizar.

3.2. DETECCIÓN DE COMUNIDADES

Similar a un proceso de *clustering*, una comunidad es un subgrafo de la red que está densamente conectado de forma interna y dispersamente conectado con el resto de la red. Es uno de los temas más importantes en ciencia de redes desde que Girvan and Newman publicasen su trabajo en 2002 [56]. Desde entonces, cientos de publicaciones han abordado distintos aspectos sobre estas particiones del grafo como el solapamiento entre comunidades [57], las descomposiciones jerárquicas [58] o el estudio dinámico sobre cómo varían las comunidades y sus relaciones con el resto de elementos de la red a lo largo del tiempo [59,60].

Realmente, la detección de comunidades engloba un conjunto de algoritmos y métricas que sirven para caracterizar una red. El hecho de que figuren como un apartado diferente al *Anexo I. Propiedades de un grafo* es su relevancia para este trabajo. La forma más popular de encontrar esta descomposición de la red en particiones es optimizando una función de calidad (*quality*

function) como, por ejemplo, maximizar la diferencia entre las densidades relativas de conexiones intracomunitarias e intercomunitarias.

Desde que Broadmann publicara su trabajo sobre la organización modular de las funciones del cerebro [45], esta propiedad ha sido una de las más estudiadas (y validadas) sobre las redes cerebrales. Así, la optimización de la modularidad permite particionar los nodos de la red en comunidades de manera que la fuerza total de las conexiones dentro de la comunidad sea mayor que los enlaces entre comunidades o que la obtenida por un modelo nulo. Las referencias [61,62] ofrecen recientes *reviews* sobre estructuras de comunidades en redes. Las referencias [63,64] hablan sobre algunas de las consideraciones y problemas que se deben tener en cuenta a la hora de optimizar una función de este tipo para la detección de comunidades.

Como se puede ver en el *Anexo I. Propiedades de un grafo*, la modularidad de un grafo ponderado no-dirigido viene dado por el índice Q de modularidad (el subíndice S hace referencia a *static*):

$$Q_S(c_i, c_j) = \frac{1}{2m} \sum_{ij} [A_{ij} - \gamma R_{ij}] \delta(c_i, c_j) \quad (1)$$

donde $2m$ es la suma de todos los enlaces de la red (se usa para normalizar), A_{ij} es la matriz de adyacencia, R_{ij} es la matriz de adyacencia del modelo nulo correspondiente, γ es el parámetro de resolución espacial o estructural (hablaremos de él más adelante), c_i y c_j son las comunidades asignadas a los nodos j e i , y $\delta(c_i, c_j)$ es la delta de Kronecker, donde $\delta(c_i, c_j) = 1$ si $c_i = c_j$ y 0 en cualquier otro caso. Es decir, Q_S es una función vectorial que se optimiza variando, principalmente, los valores del vector c que contiene la asignación de comunidad para cada nodo. Q_S está acotado entre [-1,1] y proporciona una medida comparativa entre la densidad de los enlaces dentro de las comunidades respecto de la de los enlaces fuera de la comunidad.

En este caso, optimizar Q se corresponde con un proceso de maximización; teóricamente, cuanto mayor Q_S , mayor densidad intra-grupos, menor densidad entre-grupos y mejor será la partición encontrada. Sin embargo, este proceso de maximización implica iterar sobre todas las posibles combinaciones, algo poco factible computacionalmente en redes grandes (es un problema NP-completo [65] y para resolverlo se acude a métodos heurísticos como el algoritmo de Louvain [66,67,68]. Este método está basado en una búsqueda en *grid* que resuelve el problema en un tiempo $O(n \cdot \log(n))$. El método consta de dos fases diferenciadas:

- 1) Se realiza una búsqueda de comunidades a nivel local: El algoritmo se inicializa considerando cada nodo i como una comunidad.

Para cada nodo i de la red, calcula la variación obtenida en Q_S que se produce entre eliminar i y asignarlo a una comunidad vecina c_{in} . Por ejemplo, si se utiliza el modelo nulo de Girvan-Newman ($R_{ij} = \frac{k_i k_j}{2m}$), este incremento vendría dado por:

$$\Delta Q_S = \left[\frac{\sum_{in} + 2k_{i,in}}{2m} - \left(\frac{\sum_{tot} + k_i}{2m} \right)^2 \right] - \left[\frac{\sum_{in}}{2m} - \left(\frac{\sum_{tot}}{2m} \right)^2 - \frac{k_i}{2m} \right] \quad (2)$$

donde \sum_{in} es la suma de todos los pesos de los enlaces dentro de la comunidad adonde i ha sido movida, \sum_{tot} es la suma total de todos los pesos y $k_{i,in}$ es la suma de los pesos en los enlaces entre i y el resto de nodos en la comunidad a la que ha sido asignada.

Así, tras calcular este incremento para todas las comunidades, el nodo i será asignado a aquella donde ΔQ_S sea mayor. Si no hay incremento, el nodo se mantiene en la misma comunidad. Este proceso es repetido iterativamente para todos los nodos de la red hasta que ningún incremento en Q_S es posible.

- 2) En la segunda fase del algoritmo, se agrupa a todos los nodos en la misma comunidad y se crea una nueva red donde los nodos son las comunidades de la fase anterior. Todos los enlaces entre nodos de la misma comunidad ahora estarán representados por autoenlaces y los enlaces que había de múltiples nodos en la misma comunidad a un nodo en una comunidad diferente estarán representados por enlaces ponderados entre comunidades. Una vez que se crea la nueva red, se puede volver al paso 1) y repetir el proceso con esta nueva red.

El algoritmo de Louvain sigue un procedimiento estocástico que padece un problema de degeneración, o lo que es lo mismo, existe un número elevado de posibles estructuras no óptimas en las que el algoritmo puede converger. La mejor manera de combatir este problema hasta el momento es repetir el proceso de optimización un número suficientemente alto de veces para que pueda llegarse a una partición consenso [69,70], bajo la idea básica de que la verdadera estructura de la red sea detectada en mayor proporción que el resto de posibles estructuras.

3.2.1. DETECCIÓN DINÁMICA DE COMUNIDADES

Anteriormente se introdujeron las redes temporales, las cuales experimentan una evolución a lo largo del tiempo [25]. Esta evolución podría resumirse en que nuevos nodos y enlaces aparecen/desaparecen y/o se modifican a lo largo del tiempo (y con ellos, el resto de propiedades de la red). El estudio de las dinámicas en la red se centra en las variaciones que suceden dentro de la red, lo que tiene una repercusión directa en los pesos de los enlaces (*dynamics of network*).

En Palla et. al [71] se da una primera aproximación para el estudio de la dinámica de las comunidades, distinguiendo 6 posibles operaciones entre comunidades (*dynamics of network*): **Crecimiento, Contracción, Unión, División, Nacimiento y Muerte**. En el trabajo de Cazabet et al. [72] se identifica una séptima: **resurgimiento**.

Sin embargo, estas operaciones no pueden ser calculadas simultáneamente por ninguno de los algoritmos existentes. Algunos algoritmos se centran en proporcionar resultados robustos y estables, por lo que las operaciones ofrecen información sobre nacimientos y muertes de comunidades pero no sobre crecimiento y contracción o unión/división; la semántica de estas operaciones no está del todo clara (depende mucho del contexto de aplicación) y no suelen ser consideradas frecuentemente.

Por ejemplo, consideremos el caso de unión entre dos comunidades A y B, que se vuelven lo suficientemente similares (primer problema: ¿cómo se define *similares* y cuánto es *suficiente*?) para unirse. Si se unen, una de las comunidades ya no existe más, es decir, una de las comunidades absorbe a la otra. *¿Cómo decidimos cuál absorbe a cuál?* Entre algunas de las posibles opciones:

- a. Mantener la más antigua/nueva.
- b. Mantener la más grande/pequeña.

- c. Eliminar la que mayor porcentaje de enlaces tenga en la otra.
- d. Que mueran las dos y se crea una nueva comunidad C. Esta opción abre el debate sobre la continuidad en la evolución de una red (cuando está evolucionando o cuando pasa a ser otra distinta).

Por otro lado, se unen o se genera una nueva comunidad C, *¿qué hacer con los nodos que no pertenecen a la intersección de A y B?* Algunas ideas pueden ser:

- o Integrar o no todos esos nodos en la comunidad resultante.
- o Evaluar cada nodo para ver si deberían pertenecer a la comunidad con algún tipo de métrica.
- o Mantenerlos si pertenecían a la comunidad que absorbe a la otra; eliminarlos de la red en caso contrario.

Una vez definidas estos conceptos, se puede entender mejor la complejidad del problema.

En cuanto al estudio de las variaciones en los patrones de actividad que suceden dentro de la red, con el tiempo han ido surgiendo distintas propuestas, siendo las dos más conocidas *Instant Optimal* [73,74] y *Cross-time Community Detection* [75,76].

- **Instant Optimal**

La idea es generar una secuencia de redes estáticas para simular la evolución. Las comunidades son identificadas en cada instante t (*snapshot*), independientemente de los estados anteriores. Tras ello, busca hacer un match correcto entre las comunidades detectadas en el instante t con las detectadas en $t+1$. Es decir, aunque el reparto sea el mismo, la etiqueta o comunidad asignada a cada grupo puede cambiar de un instante a otro, lo que se conoce como problema de *matching*. Se hace necesario por tanto realizar un análisis de consenso para unificar la asignación de comunidades a cada nodo.

Es el primer método que se propuso y tiene la ventaja de que permite reutilizar todas las técnicas tradicionales de detección de comunidades. El problema del *matching* también es algo clásico para el que existen soluciones. Por otro lado, dado que la búsqueda de comunidades es independiente de cada estado, es un proceso fácilmente paralelizable.

Sin embargo, hoy en día está perdiendo popularidad debido a los problemas de inestabilidad que sufre. Está demostrado que ofrece resultados muy diferentes al hacer pequeñas modificaciones sobre una red. De hecho, dado que los algoritmos de optimización de Q son estocásticos, repetir el proceso dos veces en la misma red no tiene que proporcionar las mismas particiones. Por lo tanto, cuando se observan variaciones entre dos *snapshots*, no se puede deducir si es debido al algoritmo o a variaciones reales de la red. De nuevo, se justifica así el procedimiento común de repetir múltiples veces el proceso de optimización y la aplicación de métodos de consenso que permitan extraer una partición verosímil o frecuente.

- **Cross-time Community Detection**

La detección de comunidades solo se estima una vez, pero teniendo en cuenta simultáneamente todos los pasos de la evolución de la red. Es decir, solo existe una descomposición en comunidades de manera global. Existen dos alternativas principales:

- Crear una sola red a partir de todos los instantes temporales y realizar la búsqueda de comunidades sobre esa gran red. Así, los nodos de la red serían todos los nodos de todos los *snapshots* y los enlaces pueden unir tanto nodos de un mismo *snapshots* como enlazar nodos de diferentes instantes temporales.
- Seleccionar o diseñar una función que pueda ser optimizada directamente en cada instante temporal. Son comunes las funciones a partir de la modularidad o de Modelos estocásticos por bloques (SBM, por sus siglas en inglés Stochastic Block Models [77]).

Dado que se estudia toda la evolución de la red como un mismo objeto, la coherencia global de la red y las condiciones de estabilidad se verifican. Sin embargo, al tratarlo como “una gran red estática”, no se asegura que se capturen los comportamientos dinámicos. De hecho, si se actualiza la red con un nuevo instante temporal, es necesario recalcular todo de nuevo ya que no es posible mantener el histórico de resultados de la red. Por otro lado, esto también conlleva que sea estructuralmente imposible manejar operaciones como la unión o separación de redes a lo largo del tiempo.

En cualquiera de estos dos métodos, la decisión sobre qué forma una comunidad o no, sigue estando basada en la optimización de una métrica. Como se vio anteriormente, dada la naturaleza de las redes cerebrales, la métrica más utilizada es la modularidad de las redes funcionales. En [26] se propuso una extensión de Q_S para casos de redes temporales multicapa:

$$\begin{aligned} Q_{ML}(P) &= \frac{1}{2\mu} \sum_{ijlr} \{(A_{ijl} - \gamma_l R_{ijl}) \delta_{lr} + \delta_{ij} \omega_{jlr}\} \delta(c_{il}, c_{jr}) \\ &= \sum_t^L \sum_{ij} (A_{ijl} - \gamma_l R_{ijl}) \delta(c_{it}, c_{jt}) + \sum_t^{L-1} \sum_i 2\omega \delta(c_{it}, c_{i_{t+1}}) \end{aligned} \quad (3)$$

donde A_{ijl} es la matriz de adyacencia para las l capas, R_{ijl} es su análogo para el modelo nulo correspondiente, γ_l es la resolución estructural de la capa l , c_{il} y c_{jr} son la asignación de los nodos i, j en las capas l y r , respectivamente. El parámetro introducido, ω_{jlr} , se denomina acoplamiento entre-capas (*interlayer coupling parameter*) o resolución temporal, y proporciona la fuerza de la conexión de un nodo entre dos capas consecutivas. μ es el peso total de los enlaces, dado por $\mu = \frac{1}{2} \sum_{jr} \kappa_{jr}$; la fuerza $\kappa_{jl} = k_{jl} + c_{jl}$ del nodo j en la capa l viene dado por la suma de su fuerza intra-capa $k_{jl} = \sum_i A_{ijl}$, y su fuerza entre-capas $c_{jl} = \sum_r \omega_{jlr}$.

Una vez que la función objetivo está definida es importante tener en cuenta una serie de consideraciones [75], como la validación estadística de los resultados con modelos nulos. El siguiente apartado está dedicado a la validación estadística y a la búsqueda óptima de parámetros, aspectos de los que dependerán fuertemente los resultados.

3.2.2. COMPARACIÓN DE REDES CEREBRALES. MODELOS NULOS

Las redes de datos cerebrales se construyen sobre datos adquiridos y procesados previamente. Aunque se hacen esfuerzos por estandarizar este proceso lo máximo posible, las fuentes de ruido y la variabilidad de cada experimento pueden influir enormemente en los resultados. Esto

conlleva que, si la generación de la red no es trivial, la comparación entre grafos tampoco lo es. Comparaciones entre sujetos sanos y enfermos, entre sujetos con la misma patología o incluso sobre el mismo sujeto en distintas condiciones han llegado a mostrar resultados incluso contradictorios.

En este sentido, el objetivo de utilizar modelos nulos es doble: calibrar y optimizar los resultados y validarlos estadísticamente comprobando si las variaciones son aleatorias o no.

a) Aplicación a redes estáticas

Para el caso de redes cerebrales estáticas suelen utilizarse como modelos nulos la **red de Erdos-Renyi**, **una red en malla** y la **red de tipo estrella**, pues está demostrado que la red cerebral va a estar comprendida entre el espacio que dejan estos tres tipos de redes.

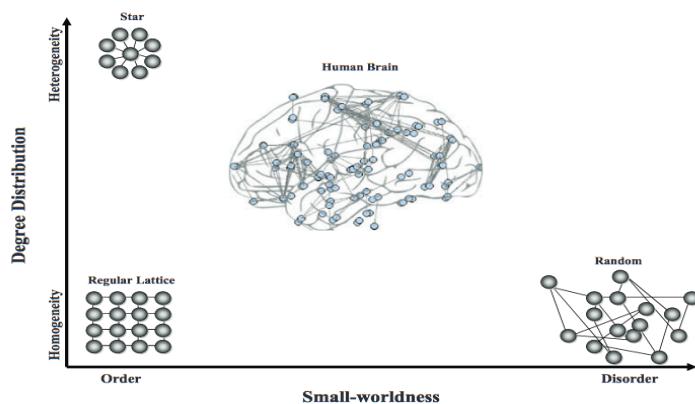


Fig. 3. Extraído de [55]. Las redes cerebrales son complejas y en ocasiones no se pueden reproducir los mismos resultados incluso bajo las mismas condiciones o experimentos. Sin embargo, algunas características como su distribución de grado o propiedades de mundo pequeño se encuentran de manera robusta.

Con este tipo de aproximaciones se ha demostrado en trabajos previos que existen características anatómicas y funcionales que se mantienen de forma cualitativa en redes cerebrales [38.78]:

- Small-worldness (alto *clustering* y camino característico corto): un buen diseño anatómico debería ser capaz de encontrar un equilibrio óptimo entre regiones especializadas en determinadas funciones (segregación) y una cantidad de enlaces entre módulos lo suficientemente alta como para dotar al sistema de robustez y rapidez en la transmisión de información (integración). Estas son propiedades de redes de mundo pequeño, que pueden encontrarse en redes cerebrales. [79]
- Modularidad: indicando la organización del cerebro en distintas subredes que, a su vez, tienen una organización jerárquica interna (tanto a nivel funcional como estructural). [80,81] y con nodos altamente conectados dentro de los módulos (*hubs*) [82], formando lo que se conoce como un *integrative core* o *rich club* [83,84].
- Distribución de grado heterogénea: asociado con la propiedad anterior, tiene un *broad-scale distribution*, indicando la presencia en la red de numerosos conjuntos densamente

conectados con nodos con un número bajo de conexiones (no sigue una distribución en ley de potencias).

Nuevas líneas de investigación se centran en el rol de la estructura geométrica y la efectividad costo-espacial (*wiring cost* [80]) en la conectividad de redes anatómicas. Todas estas características físicas son las que condicionan en parte importantes aspectos funcionales, como una comunicación eficiente entre distintos nodos de la red que faciliten el procesamiento de la información o la robustez ante lesiones.

Por último, para poder comparar redes es apropiado disponer de a) el mismo número de nodos o b) el mismo número de enlaces, ya que mucha de las métricas dependen tanto del tamaño como de la densidad de conexiones. En caso contrario, **se podrían hacer inferencias cualitativas pero de ningún modo cuantitativas**.

Por ejemplo, conforme la densidad de conexiones crece, existe un incremento en la eficiencia global (más enlaces hacen más fácil llegar de un nodo a otro) pero también un aumento en el coste topológico. En redes cerebrales, esta eficiencia crece de forma supralineal hasta alrededor del 20% de densidad de conexiones, donde el coste-efectividad de la red sería positivo. A partir de una densidad del 30%, esta relación llega a su máximo y comienza a decrecer, por lo que no tiene sentido poner densidades mayores.

b) Modelos nulos en detección de comunidades

Dado que las métricas de modularidad suelen presentar un alto grado de degeneración en datos reales, los modelos nulos también pueden ser aplicados para corregir el ruido a la hora de realizar las particiones. Además, los algoritmos de detección de comunidades llevan intrínsecos procesos estocásticos que pueden producir particiones óptimas (desde el punto de vista que Q es el máximo encontrado) pero, en cierto modo, aleatorias (sin significado para el campo de estudio en cuestión). Por lo tanto, el análisis de los resultados puede carecer de sentido sin una comparación adecuada con un *benchmark* o modelo estadístico que los valide [75].

Así, tras construir la red A es necesario seleccionar un modelo nulo R adecuado para Q_{ML} . El modelo nulo más común y generalizable en redes no dirigidas es el de Newman-Girvan [32,61]:

$$R_{ij} = \frac{k_i k_j}{2m} \quad (4)$$

donde $k_i = \sum_j A_{ij}$ (la fuerza del nodo i) y $m = \frac{1}{2} \sum_{ij} A_{ij}$. Esta definición puede ser extendida para el caso temporal o multicapa como:

$$R_{ijl} = \frac{k_{il} k_{jl}}{2m_l} \quad (5)$$

Usando este modelo, Q es optimizado de forma que las particiones identificadas tienen más conexiones (en el caso de redes binarias) o una mayor densidad total de conexiones (para grafos ponderados) que la esperada para el modelo nulo. El modelo de Newman-Girvan es particularmente útil para el caso de nodos categóricos (no ordenados), como en las redes

cerebrales, donde puede surgir una conexión entre cualquier par de nodos de la red (de hecho, suelen ser redes totalmente conectadas).

- **Modelos nulos optimizados para redes temporales**

Aunque el modelo de Newman-Girvan es aplicable a redes temporales, no aprovecha las propiedades de que las redes sean construidas a partir de similitudes en las series temporales (en este caso, las series temporales que reflejan la actividad BOLD). Por ello, también puede ser interesante generar modelos nulos ad-hoc para redes derivadas de series temporales. En [75] se proponen algunas alternativas:

a) **Temporal null-model:**

Un modelo simple que consiste en randomizar los puntos temporales de la serie temporal BOLD de cada nodo por separado antes de calcular la matriz de adyacencia. Sin embargo, en este primer modelo no se preserva la media ni la varianza de la distribución original, lo que suele producir una red con pesos muy pequeños en los enlaces. Utilizando propiedades del espacio de Fourier, es posible generar series temporales aleatorizando la fase de las series originales, de forma que se preserva la media, varianza y la función de autocorrelación [85]. Con este segundo método se consiguen distribuciones y redes más similares a la original. De este modo, las comparaciones con el modelo nulo serán estadísticamente más rigurosas.

b) **Intra-layer null models:**

Aleatorizando la estructura de conexiones internas de la comunidad de forma que se mantenga una determinada propiedad de interés. Por ejemplo, si se quiere mantener la distribución original de la fuerza de los enlaces [86], se puede utilizar un método basado en permutaciones que reubique los enlaces en esa capa de forma uniformemente aleatoria. Dado que no tiene en cuenta ninguna propiedad en especial de la estructura de la red, puede ser utilizado tanto para redes binarias como ponderadas.

c) **Inter-layer null models:**

En este caso se pueden realizar múltiples aleatorizaciones para validar tanto el resultado temporal como los enlaces entre capas. Alterando el orden de las capas o modificando la identidad de los nodos (es decir, a qué comunidad pertenecen) se puede generar un *temporal null model*. Otra posibilidad es permutar los enlaces que unen nodos entre una capa y otra, generando un *nodal null model*, para comprobar cómo la identidad de los nodos pueden influir en la organización y estructura de la red.

La utilización de cada uno de estos modelos dependerá de qué propiedad de la red quiera validarse. Una vez definido el modelo nulo, tan solo queda comparar los resultados con los de la red real. Por otro lado, dados los problemas de degeneración (gran número de posibles particiones óptimas de Q) y el procedimiento heurístico del algoritmo de Louvain, los resultados pueden variar considerablemente de una iteración a otra. Es lo que se conoce como *optimization variance*. Por ello, será necesario repetir el proceso de optimización de Q un número elevado de veces C y luego computar la media (o mediana) y los estadísticos necesarios para obtener el resultado final. Una vez obtenidos los valores medios (o medianos) de las métricas que se quieran comparar bastará con realizar un t-test entre las métricas de la red real y las del modelo nulo.

Es importante resaltar que en el caso del modelo nulo aparecen más fuentes de incertidumbre denominadas *randomization variance*, fruto de cómo se ha construido el modelo nulo que, por definición, utiliza distribuciones aleatorias y puede originar miles de posibles combinaciones diferentes.

c) Modelos nulos para la búsqueda de parámetros óptimos en *Cross-time Community Detection*

En este tipo de algoritmo, los resultados obtenidos van a ser tremadamente dependientes tanto de la elección de los parámetros como del modelo nulo generado. En el caso de optimizar la modularidad en redes multicapa es necesario especificar los parámetros de resolución estructural o espacial γ y la resolución temporal o acoplamiento entre capas ω , que definirán la escala a la que realizamos nuestro análisis. Para ello, se pueden comparar la estructura obtenida con esos parámetros con la obtenido usando un modelos nulo. Cabe recordar que:

- Variando γ se ajusta el tamaño mínimo de las comunidades detectadas en una misma capa. Es un parámetro que se especifica por capa. Cuanta mayor resolución, menor puede ser el tamaño mínimo de las comunidades y mayor número de comunidades podrán ser detectadas.
- A través de ω se puede controlar el número de comunidades descubiertas a través de las capas; cuanto mayor sea ω , menos comunidades. Es un parámetro que se fija por nodo. Sin embargo, poco se ha discutido respecto a su optimización. De hecho, el caso común es mantenerlo constante a lo largo de la evolución temporal para todos los $\omega_{ij} = 1$.

En sistemas reales es difícil encontrar un punto óptimo a través de estas búsquedas, por lo que se hace necesario la aplicación de conocimiento a priori para poder elegir unos valores adecuados de los parámetros de resolución. En caso de no tenerlo, se puede seleccionar un rango óptimo de valores para los parámetros y elegir aquella combinación que proporcione mayores diferencias estadísticamente significativas respecto del modelo nulo. Algunos trabajos en esta dirección se centran en realizar amplias búsquedas en *grid* sobre el plano (γ, ω) generado [75].

Para comparar la estructura de comunidades en función de los parámetros de resolución es útil definir una función que cuantifique el valor medio de la Q optimizada. Para ello, se puede utilizar el valor medio sobre las C optimizaciones de la métrica de similitud que se haya escogido y compararlo con el valor obtenido del modelo nulo a través del z-score (*Rand coeff.* [87]), también denominada como *mean partition similarity*.

$$z_{\alpha\beta} = \frac{2}{\sigma_{\omega_{\alpha\beta}}} (\omega_{\alpha\beta} - \frac{M_\alpha - M_\beta}{M_{\alpha\beta}}) \quad (6)$$

donde $\sigma_{\omega_{\alpha\beta}}$ es la desviación estándar de $\omega_{\alpha\beta}$ (acoplamiento entre las particiones α y β), M_α es el número de pares de nodos en la partición α , M_β el análogo en β y $M_{\alpha\beta}$ es el número de pares asignados a la misma comunidad tanto en α como en β .

4. METODOLOGIA

En el capítulo 3 se han mostrado los conceptos básicos de cómo generar un grafo a partir de datos funcionales o estructurales en neuroimagen. Lo primero que cabe destacar, es la cantidad de simplificaciones que es necesario asumir con el objetivo de poder llevar a cabo el análisis y que los resultados sean, en cierto modo, robustos y reproducibles, sobre todo en análisis funcionales. Algunas de las decisiones que han de tomarse a la hora de construir el grafo son: el criterio utilizado para la definición de nodos y enlaces, el empleo de redes binarias o ponderadas, dirigidas o no dirigidas, estudiar las propiedades a nivel de sujeto o a nivel de grupo, la selección de las métricas de conectividad, etc.

En ocasiones se tiene conocimiento a priori sobre qué regiones son activadas durante un experimento o se pueden extraer los parámetros y configuraciones adecuadas a partir de la literatura. Sin embargo, la mayoría de veces hay que definirlas ad-hoc para el problema al que se hace frente o directamente experimentar con las posibles opciones. Abordar todas estas cuestiones supone un elevado número de posibles combinaciones. En este trabajo se propone una metodología basada en la variación de un único parámetro en cada prueba, siempre que no sea necesario extraerlo de la bibliografía. Es indudable que las interacciones que surjan de la variación simultánea de dos (o más) parámetros no serán observadas, pero este es el enfoque que permite extraer conclusiones de forma más robusta y sencilla. Con todo esto, se propone:

1. Proponer un paradigma adecuado para el análisis.
2. Especificar todos los parámetros posibles basándose en la literatura para reducir la combinatoria de experimentos a realizar.
3. En aquellos casos donde no sea posible, realizar una serie de pruebas experimentales donde solo se varíe un parámetro en cuestión. Tras estas pruebas, se decidirá cuál es la configuración óptima de parámetros para el análisis.

4.1. PARADIGMA EXPERIMENTAL

Como se comentaba al comienzo de la memoria, uno de los objetivos de este trabajo era inspeccionar herramientas para el análisis multimodal y dinámico a través de la teoría de grafos. Por ello, se proponen dos experimentos diferenciados:

1. **Comparación entre redes estáticas funcionales y anatómicas.** Una vez definida la configuración de parámetros adecuada, se procederá a realizar una comparación entre sujetos sanos y pacientes con diversas patologías, tanto a nivel anatómico como funcional. En concreto, los análisis serán:
 - Conectividad funcional en sujetos sanos, con ADHD y Autismo, a partir de matrices de conectividad binarias de datos de fMRI (functional Magnetic Resonance Imaging).
 - La conectividad anatómica en sujetos sanos, en sujetos de avanzada edad con el gen APOE4 indicador de Alzheimer, y en sujetos con autismo. Los datos serán extraídos de experimentos de DTI (*Difussion Tensor Imaging*).
 - Los modelos nulos utilizados serán el de estrella, el de Erdos-Renyi [92] y una red en malla [1].

CASO	FUNCIONAL (fMRI)	ANATÓMICO (DTI)
Sujetos sanos	1000 Functional Connectomes (Baltimore)	NKI_Rockland
Alzheimer	-	UCLA_CCN_APOE_DTI
ADHD	ADHD200_CC200 (KKI)	-
Autismo	UCLA_Autism 2	UCLA_Autism 2

Tabla 1. Estudios del que se han extraído las redes anatómicas y funcionales que se compararán (resultados en 5.3. Comparación sujetos sanos vs. pacientes). De todos ellos se extrae la misma información: matriz de asociación o conectividad (de la que obtendremos la matriz de adyacencia), nombre de cada nodo (en el mismo orden que la matriz) y coordenadas espaciales del centroide de cada región en el cerebro.

Por último, procede destacar que los datos funcionales han sido adquiridos bajo condiciones de reposo (*resting-state*), donde el sujeto no realiza ningún tipo de tarea. Este es uno de los estándares actuales que han logrado poner de manifiesto diferencias en las redes cerebrales entre sujetos sanos y enfermos [93,94]. Todos los datos están disponibles en [95].

2. Detección dinámica de comunidades. Este será un problema de *dynamics of networks*, ya que se quieren estudiar las fluctuaciones en los patrones de actividad cerebral que repercuten de forma directa sobre el peso de los enlaces. Por otro lado, la estrategia implementada será de *Instant-optimal* dado que se ajusta mejor a la plataforma propuesta en [21]; estos son que el algoritmo tenga el menor número de parámetros a seleccionar y que sea lo más paralelizable posible.

Para verificar los resultados de la detección dinámica de comunidades y decidir qué configuración de parámetros es óptima se utilizarán datos simulados, ya que nos permiten conocer y controlar el *ground truth* [96]. En cuanto a los modelos nulos, Girvan-Newman se utilizará para extraer las propiedades de la red y durante la búsqueda de comunidades (mediante el algoritmo de Louvain), mientras que se utilizará un modelo nulo temporal y un modelo nulo nodal para comparar los resultados con la red simulada.

Los datos simulados seguirán el siguiente diseño:

- Cada componente independiente (IC) está compuesta por una serie temporal y una región cerebral. Así, se simulará la actividad temporal (serie temporal, TC) de 30 ICs.
- El paradigma propuesto está compuesto de un estímulo auditivo periódico en bloques de 20 TRs de duración (ON) y 30 TRs de reposo (OFF), durante un total de 250 TRs. Comenzando en reposo, habrá un total de 5 bloques o períodos de estimulación completos (ON), aunque la serie temporal acaba durante el último bloque de activación (véase fig.5).
- Este estímulo auditivo emitirá la orden de realizar una tarea motora o no. La tarea motora solo se producirá en los bloques 2º, 4º y 5º.
- Las componentes auditivas, frontales y la red de atención (*Dorsal Attention Network*, DAN) se activarán con cada estímulo auditivo. De forma simultánea, durante la tarea de los bloques 2, 4 y 5, también se co-activarán las regiones motoras y el precúneo (aunque la serie temporal de este último tendrá más ruido).

- Durante los períodos de reposo, se activará la *Default Mode Network* (DMN), lo que implica una correlación negativa con las regiones que se activan durante la estimulación.
- La generación de las series temporales de aquellos nodos activados durante la estimulación sigue el modelo convolución del diseño por bloques con la *double-gamma Hemodynamic Response Function* (HRF) [97].
- El resto de componentes serán series temporales aleatorias con ruido blanco (distribución gaussiana).

Para obtener más información sobre el resto de parámetros de la simulación se puede consultar el *Anexo II A. Parámetros de la simulación de datos fMRI*.

4.2. DECISIONES BASADAS EN LA BIBLIOGRAFÍA

- ¿Redes estáticas o dinámicas? ¿Anatómicas o funcionales?

Existe una clara dualidad e interacción entre la estructura anatómica que rige la funcionalidad de cada región y viceversa, cómo la actividad cerebral (a nivel químico y eléctrico) puede modificar la estructura física del cerebro. Una de las primeras aproximaciones con técnicas de neuroimagen que demuestran esta evidencia a nivel de redes es la predicción de patrones de activación en *resting-state* a partir de la estructura de los tractos de sustancia blanca [78]. Así, aplicaciones que están actualmente en estudio se dirigen a cómo la conectividad estructural entre regiones cambia a lo largo de la vida [51] o tras un proceso de neuro-rehabilitación o aprendizaje cognitivo [19,20].

Dado que los cambios estructurales suceden en períodos largos de tiempo, habría que tener acceso a conjuntos de datos de este tipo. Por lo tanto, la **detección dinámica de comunidades se realizará solo en redes funcionales**, mientras que para comparar la **estructura anatómica y las redes funcionales** entre pacientes y sujetos sanos es más adecuado utilizar redes **estáticas** que pongan de manifiesto las diferencias globales.

- ¿Dirigido o no dirigido?

Definir direcciónalidad con la resolución temporal de la RM funcional no es tarea sencilla [98]. Esto derivaría en otro tipo de estudio mucho más complejo denominado conectividad efectiva [8], basado en analizar cómo unas regiones influyen en la actividad de otras a partir de los flujos de información. En cuanto al estudio anatómico, la cantidad de tractos resumidos en cada enlace es tan alto que se puede asumir bidireccionalidad de la información con certeza. Así, nuestras **redes serán no dirigidas**.

- ¿Binario o ponderado?

Los enlaces binarios denotan la presencia o ausencia de conexiones, mientras que los ponderados ofrecen información también sobre la fuerza de esas conexiones. En redes anatómicas, esta fuerza puede representar el tamaño o la densidad de los tractos, mientras que en análisis funcionales y causales suelen representar la magnitud de la asociación (correlación en este caso).

La mayoría de los estudios utilizan grafos binarios, ya que facilita la caracterización de las redes y la creación de un modelo nulo para comparaciones estadísticas. Sin embargo, aunque la ponderación añada variabilidad a los resultados, proporciona información complementaria sobre la organización de la red. Dado que la mayoría de métodos son aplicables tanto a redes binarias como ponderadas, **estudiaremos experimentalmente las diferencias que se encuentran entre ambos tipos de modelos.**

- ¿Cómo definimos los nodos?

Se ha demostrado que tanto asumir homogeneidad espacial dentro de la ROI [99] como consistencia espacio-temporal [100] es completamente impreciso y diferente para cada ROI, por lo que queda mucho trabajo por hacer en este sentido.

Por otro lado, los atlas funcionales se definen de una forma limitada, basados en experimentos con una serie de tareas específicas. Es posible que utilizar esta definición de nodos conlleve el riesgo de obviar la organización funcional en cualquier otro tipo de tarea. Así, **los nodos de la red serán definidos a partir de regiones anatómicas** con el objetivo de no “sesgar” el análisis funcional. Para conocer el atlas anatómico utilizado en la comparación de sujetos sanos y pacientes se acudir a la descripción de cada dataset en [95], mientras que la descripción sobre las regiones simuladas en la detección de comunidades se encuentra en [101]. En realidad, consecuencias similares se podrían considerar al elegir atlas anatómicos: al promediar la información de todos los vóxels de su interior es posible que estemos abarcando más de una región funcional y se esté eliminando información relevante. Este tipo de paradojas lleva a la conclusión de que es necesario desarrollar métodos alternativos para definir los nodos de la red.

- ¿Cómo definimos los enlaces?

Como se ha comentado anteriormente, la correlación lineal de Pearson es la métrica *standard*. Sin embargo, carece de la capacidad de encontrar relaciones no lineales, por lo que la información mutua es otra de las opciones más comunes en la literatura. En este trabajo introduciremos la métrica de *distance correlation* [102] como alternativa a la información mutua para encontrar dependencias no lineales entre nodos. También se estudiará el efecto de la correlación parcial para definir experimentalmente qué métrica utilizar en los experimentos.

- ¿Medidas locales o globales?

Generalmente, las propiedades de una red (grado, intermediación, *clustering*, etc.) suelen estar evaluadas sobre los nodos. De este modo, las métricas globales de la red representan un valor promedio de las métricas locales obtenidas en cada nodo. Dada la gran cantidad de métricas existentes y el alto número de nodos en este tipo de redes, lo normal es proporcionar únicamente las métricas globales de la red en su conjunto, dado que comparar dos redes a partir de métricas locales puede ser una tarea de sintetización muy compleja. Este problema se hace más notable en el caso de redes temporales.

De nuevo, la decisión de extraer solamente métricas globales puede implicar una pérdida considerable de información. Por ello, es posible estudiar métricas locales de aquellas regiones que se conozcan a priori que pueden tener implicación en el problema (de forma que se busca el nodo o los nodos de interés directamente) o se dispone de alguna herramienta de visualización adecuada que permita conocer la información de manera sencilla e intuitiva.

En este trabajo se extraerán generalmente las métricas globales, aunque es posible ofrecer información local en casos específicos (p.e., regiones involucradas en una tarea conocida).

- ¿Redes individuales o por grupo de sujetos?

De nuevo, se presenta uno de los dilemas por antonomasia en neurociencia computacional: ¿promediar o no promediar? Las consideraciones tomadas anteriormente se repiten aquí. Promediar un grupo de sujetos de mismas características conseguirá extraer aquellos rasgos comunes, pero eliminará la variabilidad personal. Este procedimiento es el óptimo cuando se trata de comparar en general “sujetos sanos” con “pacientes enfermos” o grupo experimental y grupo control. Sin embargo, es probable que algunos de estos resultados no sean reproducibles (y aparezcan otras características nuevas) cuando intenten extraerse a partir de un sujeto particular al que se le trata de dar un diagnóstico.

En este trabajo se mostrarán algunas comparaciones, tanto de red estática como de análisis dinámico, entre grupos e individuos. No obstante, en el resto de los casos **se trabajará con el promedio de un grupo de sujetos** de características similares.

- Otros aspectos

En las matrices de conectividad funcional pueden encontrarse valores negativos de correlación (o de la métrica que se utilice). El significado biológico de estas **correlaciones negativas** en redes cerebrales aún no está bien definido, por lo que suelen tomarse los valores absolutos o mantener solo aquellos enlaces con valores positivos.

Otro aspecto importante es la construcción de un grafo binario a partir de uno ponderado. Para ello, es necesario seleccionar el valor de un umbral o **threshold**. Todos los enlaces con un peso por encima de este umbral se consideran como significativos y se mantienen en el grafo con peso 1; en caso contrario, se elimina el enlace (peso 0). Dependiendo de este umbral, se modifica directamente el número de enlaces y la fuerza con la que se consideran equivalentes o significativos, lo que equivale a que las propiedades de la red resultante puedan variar.

Dado que no existen guías claras respecto a estos dos parámetros, **se estudiarán** diferentes valores para el **threshold** y se compararán los resultados obtenidos a trabajar solo con **correlaciones positivas** o con el **valor absoluto**.

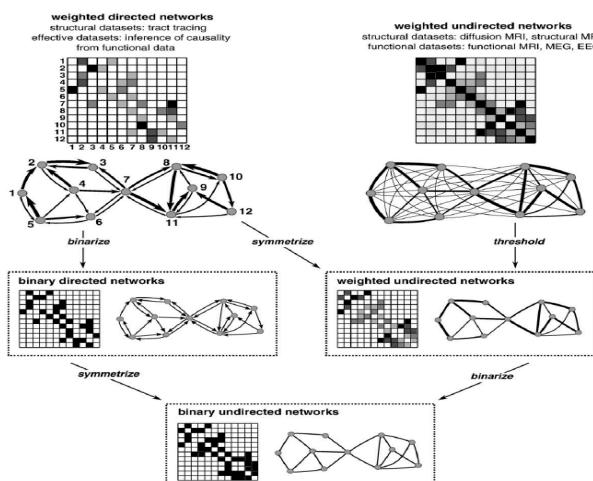


Fig.3. Extraído de [28]. Posibles análisis a partir de la matriz de conectividad: análisis anatómico, funcional o causal; a su vez, pueden ser dirigidos o no y ponderados o binarios.

4.3. HERRAMIENTAS UTILIZADAS

Dado que en este trabajo se realizan múltiples pruebas, hay múltiples resultados y figuras en cada apartado. Con el objetivo de no exceder en el tamaño de la memoria, se ha decidido mostrar y comentar solo aquellos resultados más relevantes y utilizar anexos a modo de material suplementario con el código y/o procedimiento detallado para cada experimento. En concreto:

- Para aquellos lectores no familiarizados con la teoría de grafos, puede encontrarse una descripción de las métricas y propiedades básicas de la red en el *Anexo I. Propiedades de un grafo*.
- Para la simulación de los datos, se utilizó la *Toolbox SimTB* de Matlab [101]. Todo el proceso y configuración se puede seguir en los *Anexo II A. Parámetros de la simulación de datos fMRI* y *Anexo II B. Simulación de datos fMRI*.
- El estudio de parámetros óptimos se realizó en Python. Para seguir todo el proceso acuda al *Anexo III. Estudio de parámetros óptimos*.
- Para el análisis anatómico-funcional de sujetos sanos y pacientes se utilizó la librería NetworkX de Python [103]. El código y pasos del mismo en *Anexo IV. Análisis anatómico-funcional*. Para ayudar a la interpretabilidad, también se han utilizado herramientas de visualización de [95].
- Finalmente, se utilizó parte de *Brain Connectivity Toolbox* (BCT) de Matlab para el análisis dinámico y la detección de comunidades [104] por estar diseñada especialmente para este tipo de problemas. No obstante, ha sido necesario el desarrollo de código adicional o la utilización de otras Toolboxes para este trabajo. El código de la función *main* que llama a todas estas subrutinas y realiza el análisis dinámico se puede encontrar en el *Anexo V. Detección dinámica de comunidades*, mientras que las funciones y subrutinas se encuentran en la carpeta *Code_DynamicCommunityDetection* en [21]. Dentro de cada subrutina hay una pequeña explicación del código y la referencia de dónde se ha extraído el algoritmo.

Todos los anexos mencionados en este apartado están disponibles en [21]. El formato original es interactivo a través de *notebooks* (*en Jupyter o Matlab LiveEditor*), pero también se proporcionan copias en ficheros estáticos como PDF y HTML (formato en el que se recomienda visualizar para mantener el formato). Si no se dispone de las herramientas para trabajar con el formato *notebook*, existen herramientas online para poder hacerlo (<http://jupyter.org/try> para Jupyter y <https://es.mathworks.com/products/matlab/live-editor.html> para Matlab). Para cualquier problema con la ejecución del código o instalación de las librerías, contacte con josepman@ucm.es

5. RESULTADOS

5.1. SIMULACIÓN DE DATOS

En la figura 4 se muestra la serie temporal (TC) que corresponde a cada componente independiente (IC) simulada. Se puede ver cómo las componentes S4, S5 (lóbulos frontales), S18 (*Dorsal Attention Network*, DAN), S24 (Precúneo), S27 y S28 (componentes auditivas) están perfectamente correlacionadas y se activan cada vez que reciben el estímulo. Las componentes S23 y S24 (componentes motoras) solo se co-activan con las anteriores en las 3 ocasiones que tienen que pulsar el botón. S7 (lóbulo bífrontal) también aparecerá como correlacionada, aunque es una señal más ruidosa. La componente S8 (*Default Mode Network*, DMN) está correlacionada negativamente con todas las anteriores, pues se activa cuando el sujeto entra en estado de reposo. Las series temporales del resto de componentes son ruido blanco.

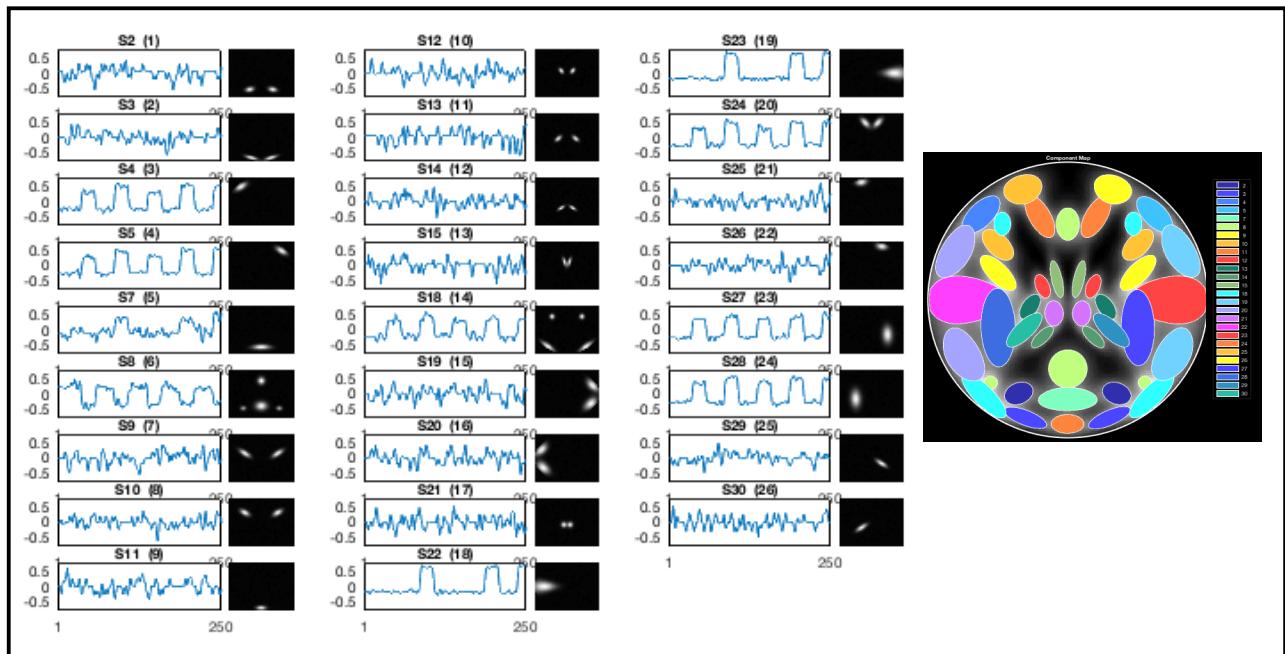


Fig 4. A la izquierda: Series temporales y mapas espaciales (en blanco sobre el cuadro negro) de cada componente simulada. A la derecha, el mapa espacial de todas las componentes, sin las componentes de líquido cefalorraquídeo (CSF), sustancia blanca (WM) y el sinus droupout signal que han sido eliminadas.

En la figura 5 se puede ver el *delay* modelado de la respuesta hemodinámica a cualquier estímulo.

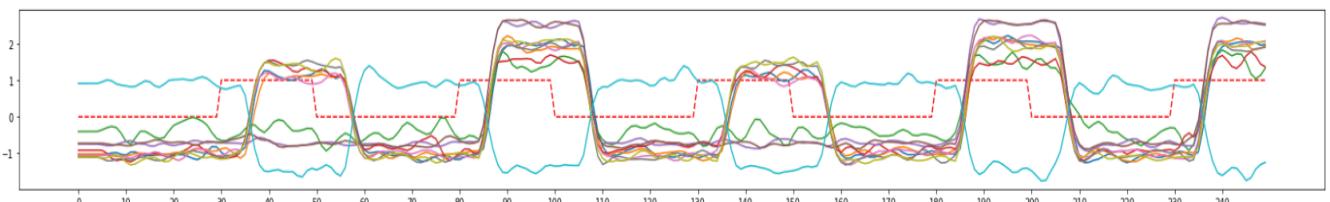


Fig 5. En línea roja discontinua, el diseño por bloque de la tarea-estímulo. La línea azul representa la TC de la DMN. En verde y magenta se representan las TCs de las componentes motoras (solo se activan en el 3º, 4º y 5º estímulo, que se corresponden con los bloques de la tarea de pulsar el botón). El resto de TCs corresponden a las componentes estimuladas (auditivas, DAN, Precúneo y lobos frontales).

En el Anexo II. B Simulación de datos fMRI se ofrece más información y características de los datos simulados.

5.2. BÚSQUEDA PARÁMETROS DE LA RED

Una vez simulados los datos y conocido el *ground truth*, podremos realizar los análisis y extraer conclusiones sobre qué configuración de parámetros podría ser la más adecuada.

- **Efecto de la binarización, valor del umbral y el signo de la asociación.**

Observando las figuras 6 y 7, se puede concluir que manteniendo el umbral alrededor de 0.4-0.5 se puede mantener parte de la información de la red, tanto en redes estructurales como funcionales. Disminuyendo más el umbral produce una explosión de enlaces no significativos; una correlación menor de 0.4 no refleja gran similitud funcional, por lo que biológicamente no sería aceptable. Como se comentaba en la revisión de la literatura, se puede observar que el aumento de enlaces no es lineal con el decremento del *threshold*; las correlaciones de actividad cerebral suelen ser realmente bajas, por lo que llega un punto (*thresholds* bajos) en el que el número de enlaces se dispara, la red se vuelve muy densa, sobre todo para redes funcionales.

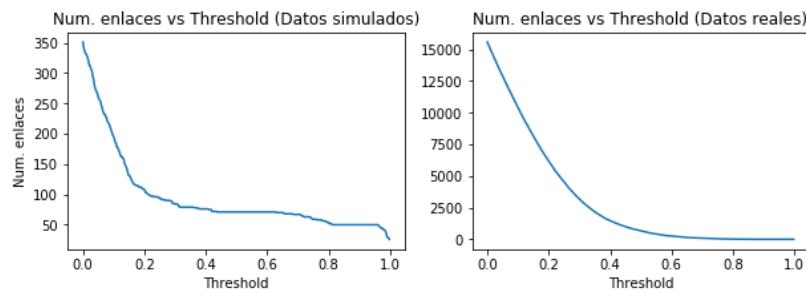


Fig. 6. Redes binarias: Disminución del número de enlaces en función del umbral para la red simulada (izquierda) y una red real (derecha).

Por otro lado, se puede comparar el comportamiento de la red en función del signo ya que existen correlaciones negativas. En redes anatómicas esto puede carecer de sentido y es apropiado utilizar valores absolutos que indiquen si existe conexión estructural entre dos nodos o regiones y cómo de fuerte es dicha conexión. Además, las redes anatómicas suelen ser moderadamente dispersas, por lo que no se corre el riesgo *a priori* de generar un excesivo número de enlaces utilizando el valor absoluto.

En cuanto a las redes funcionales, sí que tiene sentido mantener la información de correlaciones negativas ya que ofrece información sobre qué tipo de interacción tienen dos nodos. Por ejemplo, la simulación se diseña para que la red por defecto (DMN, *Default Mode Network*) se desactive cuando la tarea aparecía y se activase en estado de reposo. Si eliminamos las correlaciones negativas o trabajamos con el valor absoluto esta información se pierde o es más difícil de localizar e interpretar. No obstante, dado el ruido existente, puede dificultar aún más el proceso de análisis, razón por la que suelen tomarse solo los valores positivos en trabajos de exploración de datos (sin hipótesis *a priori*).

- **Individual vs. grupal y anatómico vs. funcional**

Realizando un primer análisis exploratorio (fig. 8), se puede deducir que las matrices de conectividad funcional están mucho más densamente conectadas que las anatómicas, ya que no necesitan proximidad ni conectividad anatómica para tener relación entre regiones. Además, al basarse en una métrica de similitud y dado lo ruidoso que son los datos cerebrales, es posible

que surjan muchos más enlaces que en el anatómico, que es una cuantificación directa de la imagen estructural de RM.

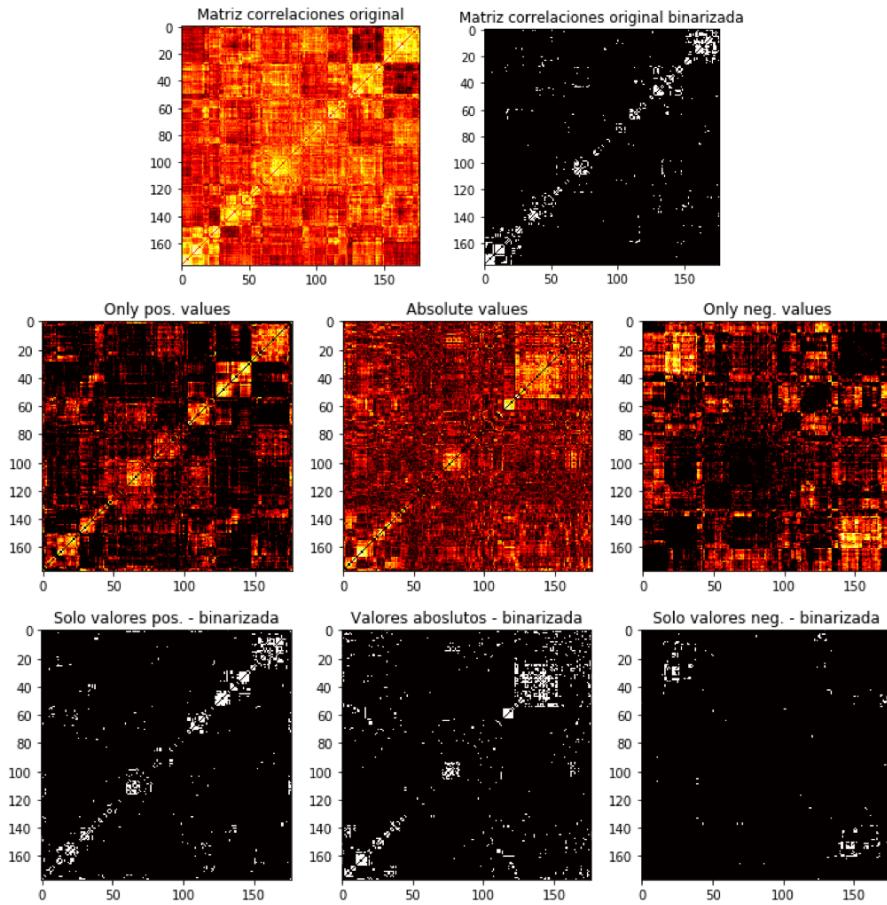


Fig. 7. Matrices de conectividad fMRI de sujeto sano en distintas condiciones: Arriba la matriz original ponderada (izquierda) y binarizada con $Thr = 0.5$ (derecha). En la segunda fila, solo los valores mayores que 0 de la matriz ponderada (izquierda), matriz con valor absoluto (centro) y solo los valores negativos. En la fila de abajo, mismo orden, esta vez con la matriz binarizada con $Thr = 0.5$

La otra característica que se aprecia en la figura 8 es la disminución de conectividad en las matrices funcionales promediadas con 10 sujetos, mientras que en el caso anatómico se mantienen bastantes similares. Es decir, gran parte de estas conexiones son fruto de la variabilidad interpersonal y no reflejan ninguna característica global del sistema. Por ello, si el objetivo es ofrecer unos resultados robustos, replicables y que reflejen un comportamiento general es recomendable trabajar con la información grupal en lugar de a nivel individual, aunque es evidente que habrá información interesante que pueda perderse.

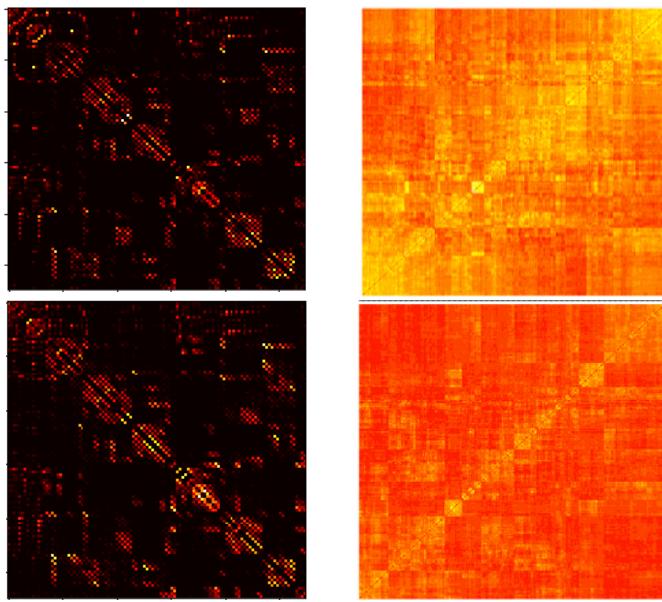


Fig. 8. Matrices de conectividad de: Arriba izquierda: DTI paciente con autismo. Arriba derecha: fMRI del mismo paciente. Abajo izquierda: DTI promedio de 10 sujetos con Autismo. Abajo derecha: fMRI promedio de los mismos 10 sujetos.

▪ Métricas de conectividad

Trabajando ahora sobre los datos simulados, como era de esperar, la correlación de Pearson detecta todas las fuentes que están altamente correlacionadas entre ellas y con la tarea (ICs 3,4,5,6,14,18,19,20,23 y 24 en las matrices mostradas). De hecho, muestra dos propiedades interesantes: 1) la correlación altamente negativa de la IC 6 (DMN) con las demás y una correlación un poco más débil de las ICs 18 y 19 (componentes motoras que solo se activan cuando se necesita pulsar el botón). También la componente IC 5 tiene una correlación más débil dado que se configuró con una componente de ruido de mayor amplitud.

Si se analiza la matriz de *Distance Correlation* se percibe un comportamiento similar, solo que los valores parecen más extremos ya que el rango solo va de [0,1]. Esto produce que la correlación de la DMN (IC 6) aparezca bastante alta pero no ofrezca información sobre su signo; simplemente indica que existe una fuerte asociación (como ocurriría con *Mutual Information*). Otro aspecto reseñable es la detección levemente superior de correlación entre las componentes de CSF (ICs 12 y 13, donde $IC_{13} = IC_{12} * IC_{12}$) que introdujimos al comienzo de este anexo. Lamentablemente, la visualización enmascara esta capacidad que tiene *Distance Correlation* para detectar las correlaciones no lineales frente a la matriz de correlación de Pearson, pero si observamos sus valores analíticos tenemos que la distancia de correlaciones es 0.448 y la correlación de Pearson es $-1.20e^{-16}$.

Nota: la representación de las matrices se ha realizado con Python. Los índices en Python empiezan en 0 mientras que la nomenclatura que se seguirá durante toda la memoria es que el primer nodo es el 1, no el 0. De este modo, la componente IC 6 (DMN) de la memoria se corresponde con el nodo 5 en Python y así sucesivamente. Esta consideración será necesaria solo durante la búsqueda de parámetros óptimos (que se realiza en Python); en la detección de comunidades (Matlab) se trabajará con la misma nomenclatura que la memoria.

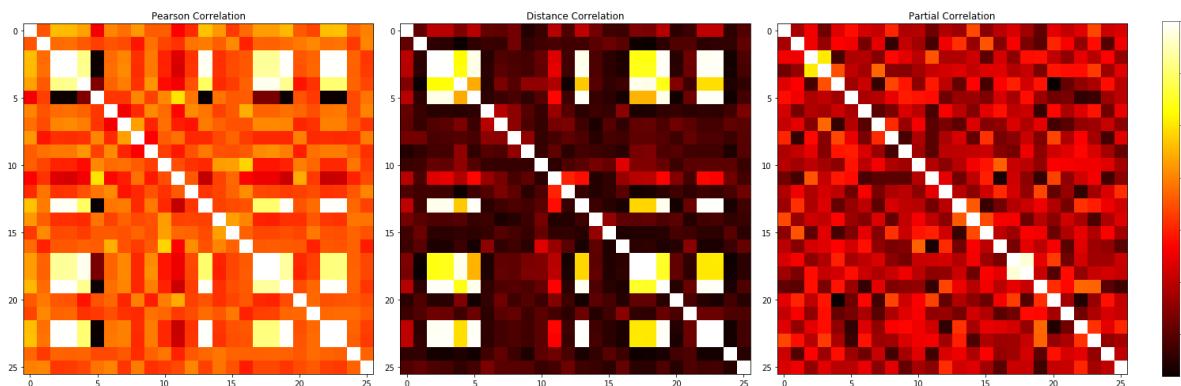


Fig. 9. Datos simulados: A la izquierda, la matriz de correlaciones. En el centro, la matriz de conectividad formada a partir de distance correlation. A la derecha, la matriz de conectividad formada a partir de partial correlation.

En cuanto a la correlación parcial, la explicación de porqué no aparece nada correlacionado es sencilla. Dadas dos variables aleatorias A y B pertenecientes a un conjunto de variables aleatorias M, la correlación parcial realiza un modelo de regresión donde "elimina" toda la información que hay en común en B y el resto del conjunto M (sin A) y lo mismo realiza con A. Así, la correlación se realiza sobre los residuos de los dos modelos lineales.

En este caso, la información que hay contenida en cada una de las variables correlacionadas es prácticamente la misma. Así, al realizar la correlación parcial entre, por ejemplo, las componentes 3 y 4, el residuo que se obtiene al "eliminar" la información de las componentes 5,6,14,18,19,20,23 y 24 es prácticamente nula, por lo que no existe correlación parcial.

De hecho, existe un pequeño grupo que corresponde a las componentes motoras 18 y 19, que solo se activaban en condiciones especiales, lo que conlleva que el resto de la matriz no contenga esa información. Esto implica que la correlación parcial sea utilizada solo bajo situaciones donde se conozcan realmente bien los datos y pueda decidirse qué elementos incluir en la regresión parcial y cuáles no para evitar tener una red no conexa.

En cuanto a *Distance Correlation* y a la correlación de Pearson, a partir de los datos simulados, preferimos trabajar con la correlación de Pearson normal. A pesar de que *Distance Correlation* detecta las relaciones no lineales, la fuerza de la asociación sigue siendo débil o difícil de detectar frente al ruido o relaciones poco significativas. Además, no ofrece información sobre el signo de la asociación, algo que puede ser de relevancia y sí lo ofrece la correlación de Pearson (a un coste computacional inferior).

- ***Sliding-window* y ventanas temporales.**

Por razones de espacio, en esta memoria solo se visualizará la evolución completa de la mejor combinación de parámetros encontrada, que corresponde a un tamaño de ventana $W = 20$ TRs y un *overlapping* de 0.8. No obstante, se explicarán las observaciones realizadas para cada caso. El resto de figuras pueden obtenerse en la parte final del *Anexo III. Estudio de parámetros óptimos*, al que se hará referencia en los comentarios. También es recomendable tener presente la figura 5 para seguir la evolución temporal del estímulo.

- **$W=10$**

Este tamaño de ventana es capaz de captar de forma considerable las transiciones entre ON-OFF (véase la transición de OFF-ON para $W=10$ y overlapping=0.8 en *From 80 to 90 seg.* u ON-OFF en *From 104 to 114 seg.*). Cuanto mayor sea el *overlapping*, esta transición aparecerá en un mayor número de ventanas seguidas (ya que el desplazamiento entre una ventana y la siguiente es muy pequeño).

Sin embargo, al ser la ventana temporal tan corta, la estimación tiene menos fiabilidad; no se distinguen los estados de ON-OFF (solo las transiciones) y existe un mayor número de componentes que aparecen como fuertemente relacionadas cuando realmente solo son ruido (véase *From 90 to 100* ó *From 144 to 154 seg.*). Esto hace que sea indistinguible cuando se encuentra en periodo ON o en período OFF. Los resultados son similares aunque se aumente el tamaño de *overlapping*.

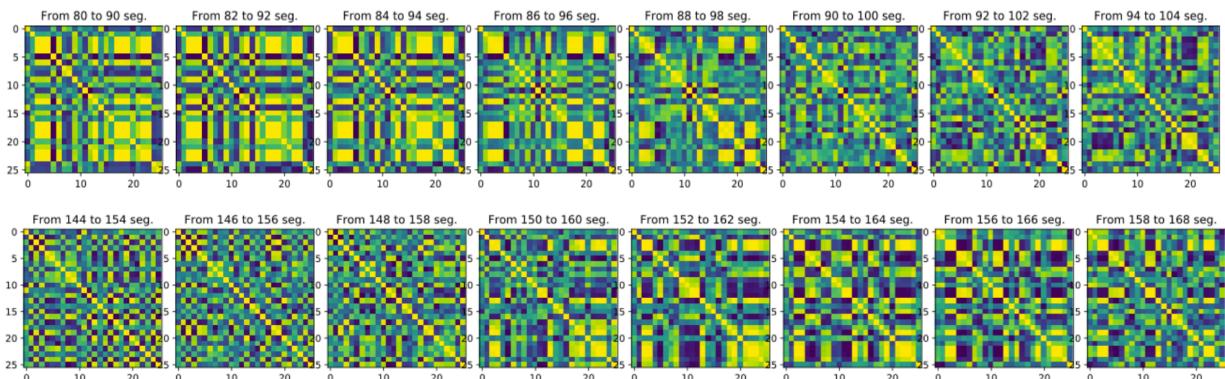


Fig. 10. Dos extractos de la evolución temporal para $W=10$ y overlapping = 0.8

- **$W=20$**

En este caso, el tamaño de ventana coincide con el tiempo que la tarea permanece en ON (sin contar el período de transición). Gracias a esto, sí que es posible detectar diferencias evidentes entre los estados de ON y OFF. Sin embargo, a simple vista podría parecer que la duración de ON es mucho mayor que la de OFF. Esto se debe a que las transiciones son capturadas por varias ventanas sucesivas (más cuanto mayor sea el *overlapping*), pareciendo así que el estado en ON es más largo.

No obstante, existe un patrón que permite detectar con buena aproximación cuándo comienza cada estado: 1) Siempre existen 2 matrices de conectividad completamente dispersas entre estados de activación (véase por ejemplo *From 212 to 232 seg.* y *From 216 to 236 seg.*) y 2) la matriz de conectividad se vuelve más *definida* en los estados de ON respecto de las ventanas de transición, esto es, los nodos de interés aumentan su correlación y la correlación entre el resto de nodos ruidosos disminuye, como si la evolución temporal alcanzase un máximo en estado de ON desde el que comienza a bajar cuando vuelve a contener la transición de bajada (ON-OFF).

Gracias a estas matrices *dispersas* (dispersas en el sentido de que no existen grandes correlaciones ni se identifican grupos de nodos activados frente a otros) entre activaciones se puede definir con precisión en qué instante comienza y termina cada período. La transición de OFF-ON se dará alrededor del instante final de la última matriz dispersa previa a la activación (por ejemplo, en *From 64 to 84 seg.*, se puede decir que la activación ocurre próxima a 84) y la transición de ON-OFF vendrá dado por el instante inicial de la primera ventana dispersa tras el período ON (en *From 112 to 132 seg.*, el período de OFF comenzará en 112 seg. aprox.). Véase

por ejemplo el proceso desde *From 168 to 188 seg.*, donde comienza a capturar la transición de OFF-ON, llegando a su máximo en *From 184 to 204 seg.* y disminuyendo en la siguiente ventana *From 188 to 208 seg.* porque vuelve a capturar la transición entre estados, en este caso de ON-OFF.

De hecho, este nivel de correlación respecto del fondo ruidoso es lo que permite a su vez detectar dos posibles estados de activación: cuando existe actividad motora o no. Compárense las matrices de conectividad cuando no hay actividad motora (correlaciones menores) de *From 120 to 140 seg.* a *From 156 to 176 seg.* con la activación motora (siguiente estado en ON correspondiente al 4 estimulo) de *From 168 to 188 seg.* a *From 208 to 228 seg.*

Estas propiedades se han encontrado en $W=20$ para todos los solapamientos, siendo *overlapping = 0.8* el que mejor permite detectar esta transición y donde las 2 matrices dispersas de estado OFF que hay entre cada estado de ON se mantienen más evidentes.

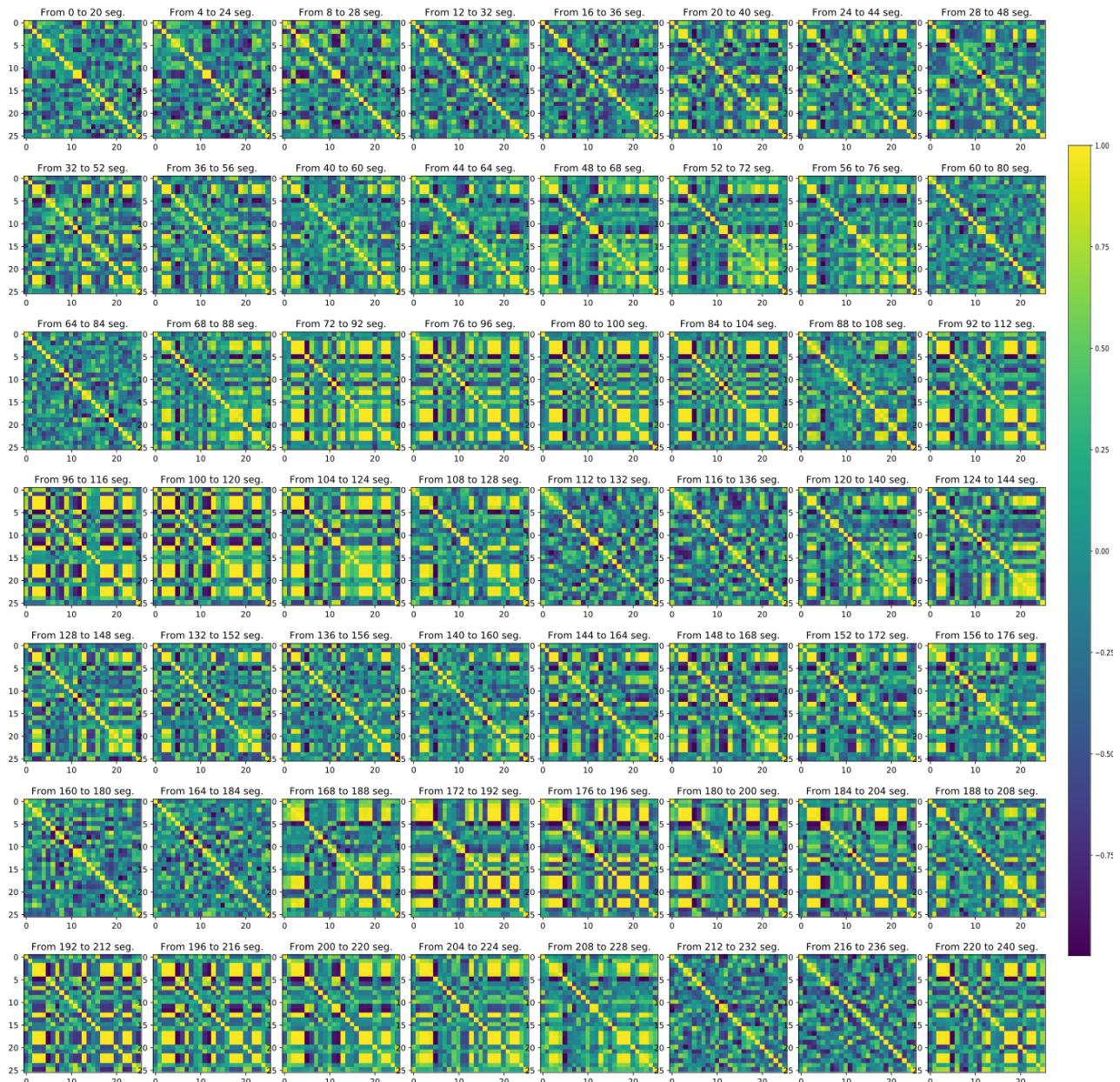


Fig. 11. Evolución temporal completa para $W=20$ y $overlapping=0.8$. Esta ha sido la configuración que mejor resultados ha ofrecido de las que se han probado.

- **W = 30**

En este punto empieza a notarse uno de los problemas que se verán acentuados con ventanas más grandes; dado que las ventanas son más largas que el estado de ON, existen solapamientos entre estados, lo que hace que siempre tengamos una matriz conectada en todas las ventanas y sea difícil detectar ON-OFF (no existe transición de matrices dispersas como con W=20). No obstante, con W=30 aún se puede parametrizar el solapamiento (cuanto mayor solapamiento, más gradual será el avance temporal y los cambios detectados) de forma que pueda intuirse cuándo son los cambios de estado. Por ejemplo, véase el cambio que existe para *overlapping*=0.8 entre la ventana *From 180 to 210 seg.*, que coincide justamente con el estado ON, respecto de las ventanas anteriores y posteriores a ella. La matriz de conectividad refleja de manera mucho más limpia la conectividad que existe entre los nodos de interés, mientras que entre el resto de nodos se mantiene más dispersa. También existe un cambio importante en la intensidad de las correlaciones respecto del estado OFF a la transición entre OFF-ON, aunque no llegan a distinguirse entre los bloques de activación motora y solo estímulo auditivo.

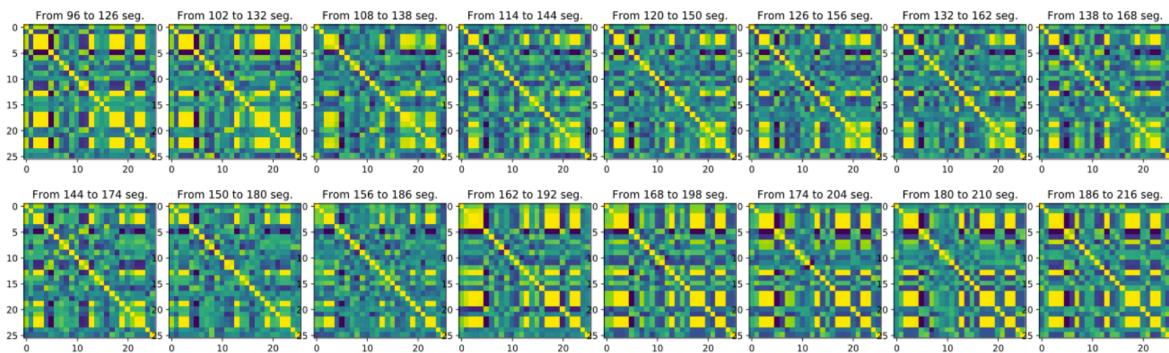


Fig. 12. Parte de la evolución temporal para W=30 y overlapping=0.8, la mejor configuración para este tamaño de ventana.

- **W = 40, W = 50**

Como se comentaba en W=30, las ventanas se hacen tan grandes respecto de la duración de cada estado que siempre capturan períodos de transición o activación, por lo que los nodos de interés aparecen siempre correlacionados y. Cuanto mayor sea W, más complicado será definir en qué estado se encuentra la red.

Con todo esto, se concluye que un tamaño de ventana aceptable estaría entre 20 y 30 seg. Es decir, el tamaño de ventana vendrá dado por la actividad que se deseé capturar. Si se desconoce esta información es muy probable que la ventana utilizada sea incapaz de capturar toda la información. En este punto podría ser útil utilizar técnicas con tamaño variable de ventana como las wavelets.

Es importante balancear el tamaño de ventana con un *overlapping* acorde. El *overlapping* permite detectar cambios y transiciones graduales en la red. Aunque esto puede ayudar a definir mejor cada estado, existen situaciones en las que vuelven los resultados más difusos por no encontrar variaciones definidas.

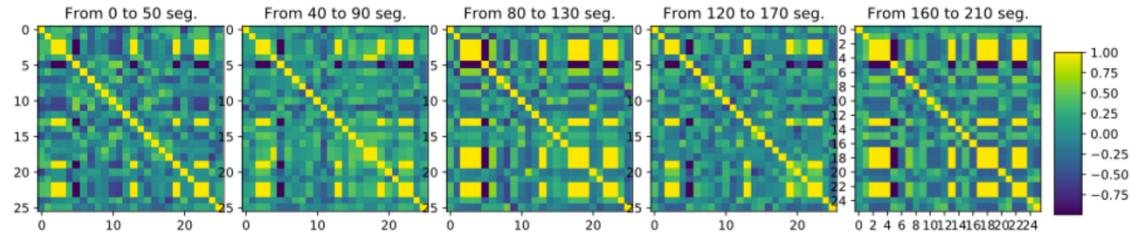


Fig. 13. $W=50$ y $overlapping=0.2$. El tamaño de ventana es tan grande que siempre parece conectada y no es posible diferenciar entre estados ON-OFF

- **Conclusión**

Para capturar los períodos de activación es necesario utilizar una ventana de tamaño similar a dicho período. Tamaños mucho menores no van a permitir discernir entre estados y ventanas temporales mucho más largas pueden solapar un estado de activación con otro, a menos que el tiempo en reposo sea mucho mayor que el tamaño de la ventana. Aún así, la ventana capturaría ruido de antes y después de la activación, con lo que estaría degradándose la información.

El *overlapping* permite graduar la forma en la que se detectan las transiciones. Aunque en ocasiones pueda ofrecer información redundante es una herramienta útil que evita cambios bruscos en la evolución temporal y permite detectar pequeñas fluctuaciones entre estados.

5.3. COMPARACIÓN SUJETOS SANOS vs. PACIENTES

Cabe decir que, dados los datos de los que se disponía para esta parte del análisis, no ha sido posible extraer un mismo número de nodos en todas las redes (cada una proviene de un estudio diferente). Por lo tanto, se ha optado por obtener un número similar de enlaces para todas las redes con el objetivo de hacer que los resultados sean comparables.

Así, se ha decidido mantener un threshold mínimo de 0.3 y un máximo de 0.7 para todos los casos y tener una asociación significativa entre nodos, reduciendo a unos 1000 el número de enlaces totales de la red (en caso contrario, la inmensa mayoría serían no significativos). Este hecho también ayudará a extraer mejor las propiedades de mundo pequeño en la red. Por otro lado, los grafos binarios estarán construidos sobre distintos umbrales de correlación, algo que debe tenerse en cuenta a la hora de interpretar los resultados.

En este apartado se caracterizarán las redes funcionales y anatómicas especificadas en el apartado 4.1. *Paradigma Experimental*. Para ello, se utilizarán métricas que permite cuantificar propiedades como la integración y la segregación de la red, la centralidad de distintas regiones, sus influencias sobre el resto de la red y sus implicaciones en la robustez del sistema (ver Anexo I. *Métricas y características de grafos*). Los resultados mostrados atienden las principales características globales que se manifiestan en la literatura como posibles biomarcadores que diferencien a sujetos sanos de pacientes de Alzheimer [105,106], ADHD [107,108,109] o Autismo [110,111]. Aún así, como se apreciará, la mayor parte de los razonamientos están aún por demostrar clínicamente, ya que existe gran controversia de resultados en la literatura, incluso arrojando conclusiones opuestas en algunos de ellos.

El código utilizado se encuentra en el Anexo IV. *Análisis anatómico-funcional* para redes binarias no dirigidas.

A. Distribución del grado (degree distribution)

El grado de un nodo refleja biológicamente de forma más o menos directa su importancia en la red, mientras que la distribución del grado de todos los nodos de la red ofrece pistas sobre su desarrollo y resiliencia.

Generalmente, el cerebro sigue una distribución heterogénea o *broad-scale*. Esto quiere decir que tiene mayor probabilidad que una red de Erdos-Renyi (ER) de contener *hubs* altamente conectados. Así, la distribución no llega a corresponderse con una ley de potencias, sino que se asemeja a una exponencial truncada o incluso puede tomar forma de campana.

En la fig. 14 se puede observar dicha apreciación en pacientes sanos de fMRI. En cuanto a la distribución en adolescentes con ADHD, este truncamiento de la distribución por la parte izquierda se acusa aún más, mostrando regiones que aumentan enormemente su grado quizás a causa de su hiperactivación. En cuanto al paciente con Alzheimer, por un lado tiene conectividad baja al ser una red generada a partir de una matriz de conectividad anatómica (DTI), pero también debido a las lesiones producidas en sustancia blanca por esta enfermedad. Los nodos con grados altos son menos y con grado más bajo (aunque esto puede estar influenciado por el hecho de que una es funcional y otra anatómica).

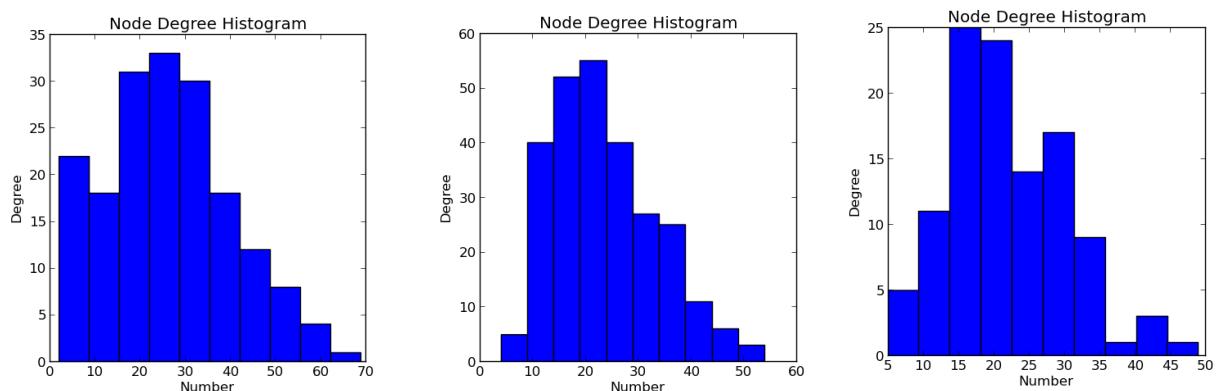


Fig.14. Distribución de grado de los nodos. Izquierda: DTI de paciente sano. Centro: fMRI de paciente con ADHD. Derecha: DTI de paciente con Alzheimer.

Otra de las diferencias que se pueden observar en la tabla 2 y en la fig. 14, entre los distintos tipos de redes es que el grado medio de las redes funcionales es muy superior a las anatómicas, básicamente porque el número de nodos y enlaces es mucho mayor.

B. Resiliencia:

El grado también se puede interpretar como una medida de resiliencia a lesiones. Por ejemplo, distribuciones en ley de potencias suelen ser resistentes a deterioros aleatorios graduales pero altamente vulnerables a los *hubs* centrales. Así, la distribución cerebral en forma de exponencial truncada, al tener más nodos conectados a más nodos (comparado con la distribución en ley de potencias), permite una integración mayor y mejora la resistencia a lesiones en nodos centrales.

Si se vuelve a analizar el caso del paciente con Alzheimer, (aparte de tener el menor grado medio de todas las redes) vemos que ha perdido ese truncamiento por la parte izquierda de la

distribución (que era lo que permitía tener más nodos conectados a más nodos), llegando a asimilarse más a una distribución en ley de potencias. Esto podría asociarse a la perdida de *hubs* cognitivos que expliquen el deterioro mental y la mayor fragilidad a lesiones (conforme más avanzada es la enfermedad, más se notan los cambios, menos resiliente es la red).

C. Intermediación (Betweenness centrality):

Aquellas regiones que se comportan como *hubs* en la red suelen estar asociadas a las regiones que más interactúan con el resto, es decir, que facilitan la integración de la red y potencian la resistencia a lesiones. Un nodo con alto coeficiente de intermediación equivaldría a aquellas regiones que están envueltas en múltiples procesos, es decir, aquellas por las que pasan un mayor número de caminos más cortos. Así, los coeficientes de centralidad más altos se encuentran en las redes anatómicas por su construcción jerárquica y su condición de enlace físico.

Como se puede observar en la fig.15, en un análisis anatómico, aquellas regiones del caudado que conectan ambos hemisferios del córtex (mitad central inferior) tienen los mayores coeficientes de intermediación, ya que gran parte de la información pasa a través de ellos, mientras que las regiones exteriores (a las que solo llegan enlaces pero no salen) tienen coeficientes de intermediación extremadamente bajos.

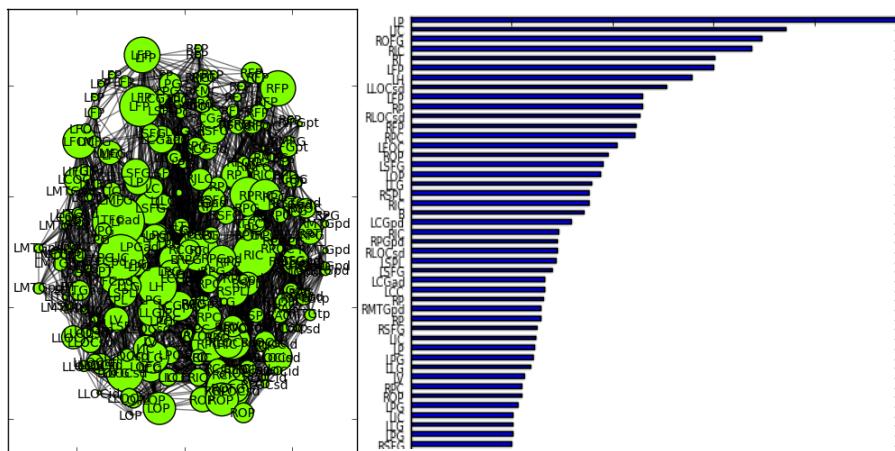


Fig.15. Intermediación. DTI de sujeto con Autismo.

D. Coeficiente de participación:

Otra medida interesante en redes cerebrales es el **coeficiente de participación**, que evalúa la diversidad de conexiones intermodulares de los nodos, donde por “diversidad” se entiende a las múltiples posibles combinaciones de enlaces que pueden darse entre dos nodos: nodo poco conectado con nodo muy conectado, nodo que une varias comunidades con nodo que solo está conectado con nodos de su propia comunidad, etc. Por ejemplo, nodos con alto grado de enlaces intra-módulo pero bajo coeficiente de participación, conocidos como *hubs* provinciales, favorecen la segregación modular. Por otro lado, nodos con alto grado de participación suelen favorecer la integración intermodular y son conocidos como *hubs* conectores. En general, no se han encontrado grandes diferencias en este parámetro. Quizá la diferencia más relevante se diese entre los estudios funcionales de sujeto sano y paciente con ADHD.

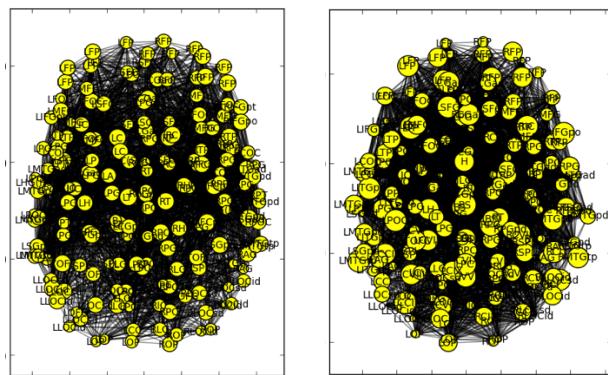


Fig.16. Coef. Participación. Izquierda: fMRI sujeto sano. Derecha: fMRI paciente con ADHD

Interpretación personal anatómico-funcional de la fig 16.: es un proceso algo confuso. Por un lado, los nodos centrales en redes anatómicas permiten una mayor integración de la red, lo que facilita que existan conexiones funcionales entre regiones anatómicamente no conectadas. Sin embargo, la generación de estos enlaces funcionales ayudan a que se “reparta” la actividad entre más nodos y más lejanos, hacen que los nodos centrales sean menos prominentes y, por lo tanto, reducen la sensibilidad a medidas de centralidad en redes funcionales. En estado de tranquilidad (resting-state), no existe una demanda clara de actividad en el sujeto sano y la DMN está repartida en diversas regiones cerebrales, por lo que el coeficiente de participación en la red es muy homogéneo. Sin embargo, en el paciente con ADHD se observa una ligera superioridad del coeficiente de participación en las las regiones anatómicas de procesamiento visual (zona posterior) y de procesamiento mental (frontal), posibles síntomas de distracción en sujetos con ADHD en estado basal.

E. Camino característico más corto, L (Shortest Path Length):

Otro concepto importante en análisis de grafos es el de *path* o camino. Se basa en las secuencias de nodos y enlaces que representan posibles caminos de transmisión de la información. Representa una medida simplificada de la eficiencia global en la transmisión de información de la red, es decir, una medida de integración. Así, cuanto más corto sea un camino entre dos nodos cualquiera de la red, mayor integración habrá en la red y más fácil y rápida será la transmisión de información. En este sentido, dada su connotación de propiedad física de la red, interesa conocer su valor en redes anatómicas más que en redes funcionales.

Dado que en los sujetos de Alzheimer ha sido el único caso en el que se han encontrado nodos no conectados, hemos decidido trabajar con la **eficiencia global** como métrica de integración, ya que es proporcional a $1/L$. El *characteristic path length* L es igual promedio de todos los caminos mínimos entre todos los pares de nodos; cuando hay nodos no conectados, $L = \infty$, por lo que $1/L = \text{eficiencia global} = 0$, valor que sí es computable y con el que se puede trabajar y comparar redes (véase Anexo I. Métricas y Características). Otra posible forma de trabajar con esta situación sería obtener la componente gigante de la red y extraer sus características, asumiendo que serían los descriptores globales de la red.

No obstante, esta desconexión es ya en sí un posible resultado: dado que el estudio de Alzheimer es sobre datos anatómicos, una región desconectada puede significar una lesión en sustancia blanca, asociada al deterioro neurodegenerativo que caracteriza esta enfermedad. Una interpretación clínica de las regiones afectadas se escapa del conocimiento del autor de este trabajo, pero llama la atención lo rápido que decrece dicha eficiencia global, es decir, las pocas regiones que favorecen esa integración (transmisión de información) comparado con un sujeto

normal. También es recomendable estudiar la eficiencia a nivel local (por regiones), como en la figura 17, ya que observar simplemente el valor promedio de la eficiencia global (tabla 2) puede enmascarar comportamientos como este decrecimiento de la eficiencia local tan acusado.

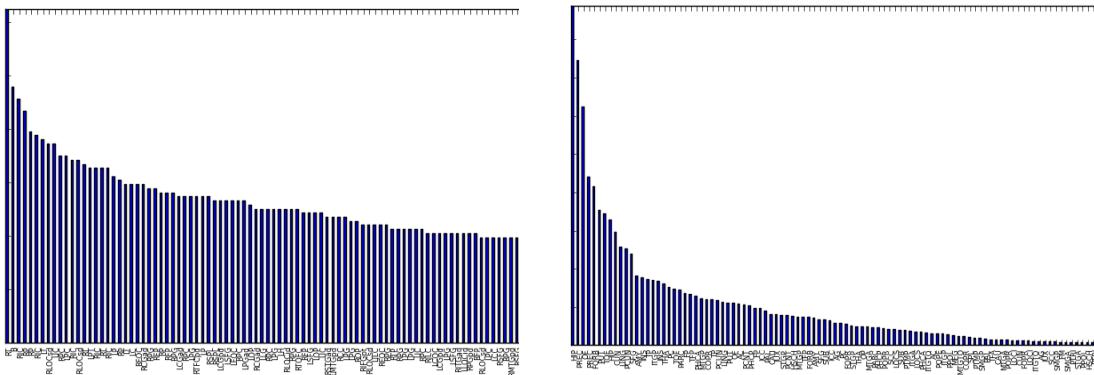


Fig.17. Eficiencia local de cada región cerebral (nodo) de: Izquierda: DTI de sujeto sano. Derecha: DTI paciente con Alzheimer.

F. Coeficiente de clustering, C (clustering coefficient):

En este contexto, puede interpretarse como una medida de la eficiencia de la conectividad local y de segregación. El coeficiente de *clustering* global puede verse influido por nodos con muy alto o, como casi siempre, muy bajo coeficiente de *clustering*. Existe un parámetro denominado transitividad, que está normalizado para no sufrir este problema y también está generalizado para grafos ponderados.

Generalmente, un mayor número de triángulos (*clusters*) en redes cerebrales suele estar asociado a un mayor nivel de segregación en la red. En este trabajo se ha encontrado que el coeficiente de *clustering* para las redes funcionales es significativamente mayor que en las anatómicas, algo que parece concordar con la topología de este tipo de redes que están mucho más densamente conectadas y con una estructura modular (*clusters*) mayor.

En cuanto a las patologías, todas las enfermedades han reportado un coeficiente de *clustering* mayor que los sujetos sanos. Dado que dos de las enfermedades, ADHD y Autismo, no son degenerativas y se han estudiado en *resting-state*, parece tener sentido que existan regiones más segregadas en estado de reposo sin afectar a la integración de la red. Por ejemplo, los sujetos con autismo tienen mayor capacidad de evadirse en sus pensamientos, mostrando complejos patrones de hiperactividad dentro de la corteza frontal (segregación) pero hipoconectividad entre el frontal y el resto del cerebro. En cuanto al DTI de Alzheimer, esta mayor segregación sí que puede corresponder a un decrecimiento de la integración a nivel anatómico, asociada a depósitos locales de beta-amiloido en esas regiones.

G. Small-world σ :

A partir de las dos métricas anteriores (*Clustering* y *Shortest Path Length*), se puede definir la propiedad de mundo pequeño en todas las redes cerebrales. Por ejemplo, comparando fMRI de pacientes con autismo con la red de Erdos-Renyi se tiene que:

$$\sigma = \frac{C/C_r}{L/L_r} = \frac{\gamma}{\lambda} = \frac{0.599/0.196}{1.96/1.8} = 3.395 > 1 \quad (7)$$

donde C es el coeficiente de *clustering* de la red del paciente con autismo, C_r el coeficiente de *clustering* del modelo nulo de ER, L es el camino característico más corto de la red del paciente con autismo y L_r el análogo para el modelo nulo de ER. Se debe cumplir (y se cumple) que $C > C_r$ y que $L = L_r$ aprox., de forma que $\sigma = \frac{C}{C_r} > 1$. Si no, podríamos categorizar falsamente como redes de mundo pequeño ($\sigma > 1$) que tengan alta segregación ($\frac{C}{C_r} > 1$) y baja integración ($L \ll L_r$).

H. Modularidad y comunidades:

La modularidad representa una métrica de segregación más sofisticada que el *clustering*. No solo ofrece una medida global (*hubs*), si no que encuentra patrones locales de conectividad (comunidades). En *resting-state*, existen diferentes redes co-activadas simultáneamente, siendo la *Default Mode Network* la más relevante de todas ellas.

Una métrica que ayuda a cuantificar estos ratios o patrones de co-activación es el coeficiente de participación comentado anteriormente, de forma que se definen nodos conectores (altamente conectados dentro de un módulo) y provinciales (aquellos que tienen baja conectividad intermodular). Así, se puede trazar qué regiones podrían tener un rol topológico más crítico en la transmisión de información en la red. Sin embargo, como se comentaba anteriormente y como se puede observar en la *tabla 2* esta métrica no ha sido especialmente relevante.

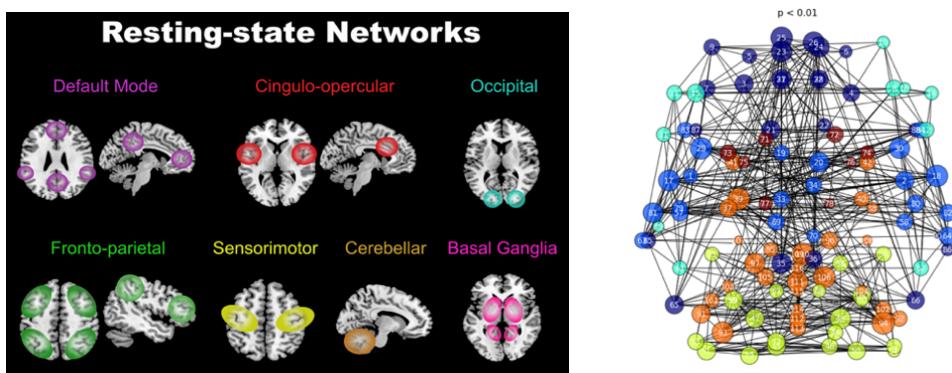


Fig.18. Detección de comunidades. Izquierda: redes funcionales co-activadas durante el resting-state. Derecha: Gold standard de comunidades obtenidas en resting-state

Sin embargo, sí ha sido posible localizar en sujetos sanos algunas de las redes que se muestran en la figura 18. En la figura 19 se revelan 3 de las redes principales de *resting-state* en sujeto sano: en azul, la *Default Mode Network*; en rojo, la red sensorimotora y, en amarillo, la fronto-pariental. Sin embargo, en el sujeto con autismo no puede identificarse de manera tan clara; existe un incremento de la importancia en la corteza frontal y una disminución de la DMN.

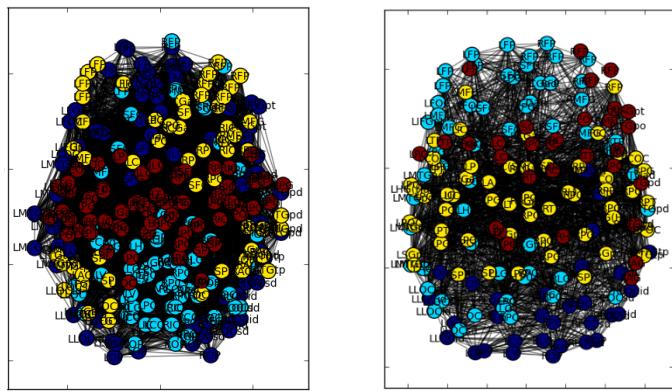


Fig.19. Detección de comunidades. Izquierda: fMRI de sujeto sano y redes resting-state. Derecha: fMRI de paciente con Autismo.

	Erdos-Renyi	Regular-Graph	Star-Graph	Sano_DTI	Alzheimer_DTI	Autismo_DTI	Autismo_fMRI	ADHD_fMRI	Sano_fMRI
Clustering	0.196564	0.000000	0.000000	0.551587	0.478278	0.427841	0.599802	0.621237	0.566202
Grado Medio	21.818182	3.618182	1.981818	22.010638	11.109091	12.469697	99.166667	75.715789	72.598870
Betweenness Centrality	0.007486	0.055556	0.009091	0.007207	0.015587	0.008094	0.002383	0.003194	0.003358
Closeness Centrality	0.553271	0.146311	0.506829	0.432862	0.377277	0.324983	0.616051	0.625227	0.630940
Global Efficiency	0.598638	0.208476	0.509091	0.486311	0.439152	0.366363	0.688285	0.700139	0.706225
Average Shortest Path	1.808507	7.000000	1.981818	2.340539	2.683403	3.120607	1.624409	1.600390	1.587635
Asortatividad	-0.001453	0.575822	-1.000000	0.097217	0.066331	0.281060	0.123351	0.023325	0.175237
Segregación(SW)	1.000000	0.000000	0.000000	2.806150	2.433197	2.176601	3.051440	3.160488	2.880501
Integración(SW)	1.000000	3.870596	1.095831	1.294183	1.483767	1.725516	0.898205	0.884923	0.877870

Tabla 2. Tabla comparativa del análisis entre sujetos sanos y pacientes con redes binarias.

5.4. DETECCIÓN DINÁMICA DE COMUNIDADES

La primera aproximación que más se adaptaba a los requisitos de [21], *Instant Optimal*, tiene uno de los procedimientos más sencillos y versátiles: analizar cada ventana temporal como si de un grafo independiente se tratara y evaluar la evolución que han seguido las comunidades y las características de la red a lo largo del tiempo.

Aunque, tras cada iteración de la detección de comunidades, la comunidad esté formada por los mismos nodos, el etiquetado en cada iteración no está regulado; esto es, que los nodos de la comunidad A en la primera iteración pueden recibir la etiqueta 1, en la siguiente la etiqueta 4, etc. Para evitar estas variaciones de etiquetado, se utilizará un método de *clustering* por consenso [70] y se comparará con los resultados ofrecidos utilizando simplemente la detección de comunidades de Louvain.

- **Comunidades y *Consensus clustering***

En la figura 20 se representa la estructura de comunidades (en distintos colores) para cada nodo (cada una de las filas) a lo largo de cada ventana temporal (cada ventana es una columna). Para seguir de manera más sencilla la evolución temporal, se recomienda tener presentes las figuras 5 y 11.

Analizando dicha figura detenidamente se puede observar cómo aquellos períodos donde hay activación motora (en magenta, *ON motor*) la uniformidad de las comunidades es mucho mayor que cuando solo hay estímulo auditivo (en línea discontinua naranja, ON). Este hecho se aprecia incluso en aquellos nodos que no son de interés (cuadros blancos 1 y 2), quizás producido porque los niveles de activación en *ON motor* son mucho mayores que durante el estímulo simplemente auditivo, lo que hace más evidentes las diferencias entre nodos de interés y nodos ruidosos. Otro detalle que puede apreciarse es, en el cuadro blanco 3, cómo se pierde la cohesión de los nodos motores (y también parte de la DAN) a las comunidades de los nodos activados cuando solo hay estímulo auditivo y no es necesario pulsar el botón (movimiento motor).

Sin embargo, incluso en las comunidades de los nodos de interés durante los períodos de *ON motor*, que es donde se hacen más evidentes, hay fluctuaciones entre nodos y/o una ventana y otra respecto a la etiqueta de la comunidad. Esto puede tener dos posibles razones: 1) que el algoritmo de *consensus clustering* no sea del todo eficiente, o 2) que esté relacionado con los problemas de estabilidad que tienen los algoritmos de *Instant Optimal*, comentados en el apartado 3.1.2.

Comparando los resultados con y sin *clustering*, para este experimento, no existen grandes diferencias, ya que las particiones iniciales eran correctas. En cierto modo, permite identificar mejor los períodos de transición y hace que las comunidades de los nodos que no son activados tengan una mayor variabilidad, algo que podría interpretarse como un posible indicador de ruido.

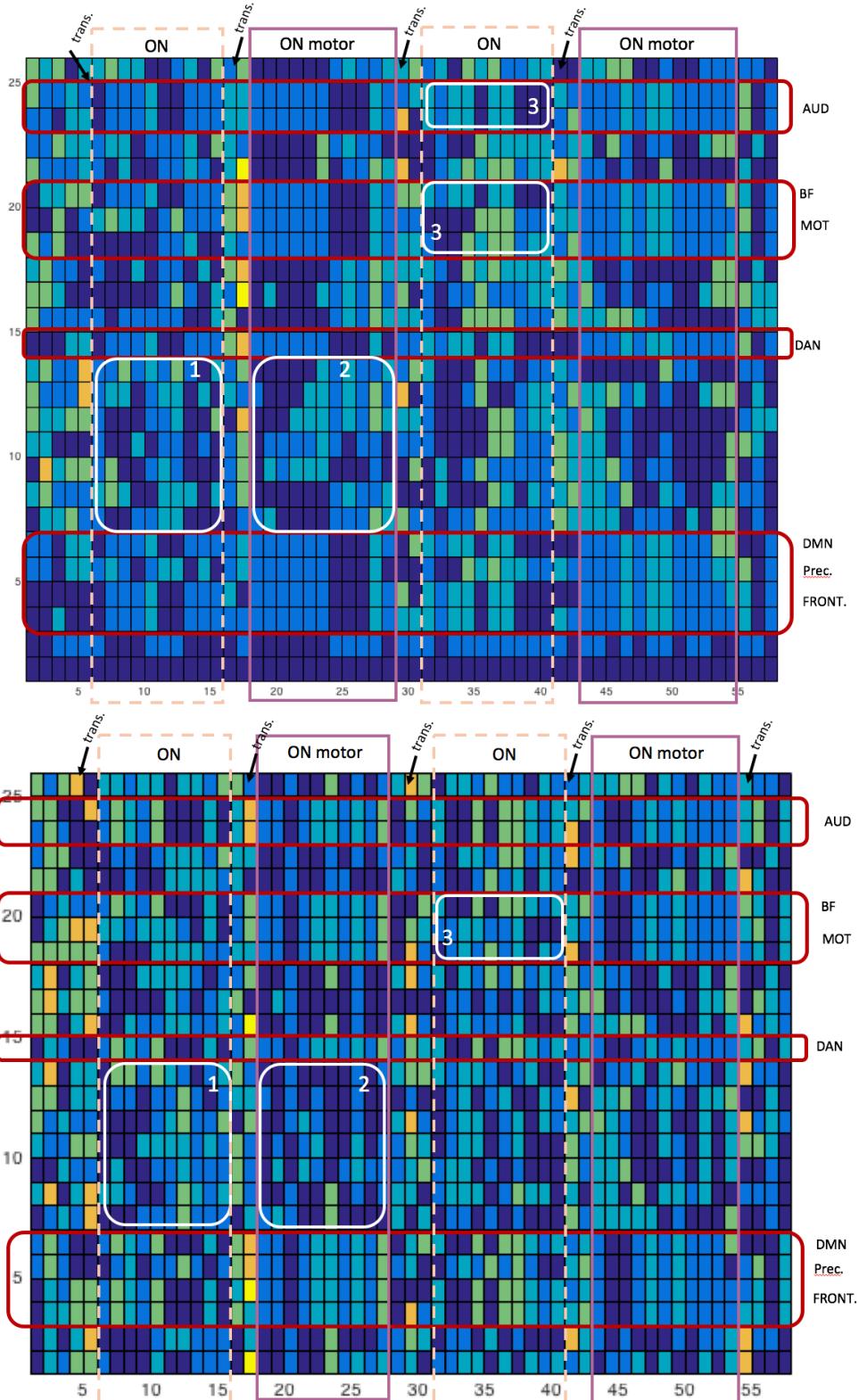


Fig. 20. Uniformidad en los resultados de detección de comunidades (arriba) y tras someterlo a un proceso de clustering de consenso (abajo). Cada fila representa un nodo y cada columna hace referencia a una ventana temporal. Cada color representa una comunidad. Los marcos rojos encuadran a los nodos de interés (Lobulos frontales y bifrontal, Precuneo, Default Mode Network, Dorsal Atención Network, componentes motores y auditivos), que coinciden con aquellos que pertenecen a la misma comunidad durante los períodos de ON. Las celdas amarillas aparecen de forma periódica y coinciden con las ventanas donde se inicia una transición de OFF a ON

- **Características de la red**

En la figura 21 se muestra la evolución temporal de algunas de las características de la red que han demostrado tener cierto comportamiento periódico, tanto a nivel local como global:

Características Globales:

Las dos primeras filas de la figura 21 muestran algunas de las propiedades globales más importantes de una red. La primera de ellas es la **modularidad**, el parámetro cuyo valor se intenta optimizar en la detección de comunidades realizada. Es quizás, de las 4 propiedades mostradas, la que tiene el comportamiento periódico más acorde con el diseño del experimento. Aunque las variaciones no son enormes, se puede observar cómo alcanza su valor máximo durante los estímulos y decrece durante el tiempo en reposo. Cabe recordar que, cuanto mayor Q, mayor cohesión hay entre los elementos de una misma comunidad (y menor con la del resto de comunidades) por lo que, al detectarse una estructura de comunidades durante los períodos de activación más definida, Q aumenta.

Relacionado con esta cohesión entre grupos se define la **transitividad T** (también llamado coeficiente de clustering), la cual también ha mostrado cierto comportamiento oscilatorio, coincidiendo las pendientes decrecientes con los períodos de activación y recuperándose durante los períodos en reposo. La transitividad era una medida normalizada del *clustering*, revelando la existencia de comunidades altamente conectadas. En este tipo de redes se ha encontrado que cuanto mayor grado de transitividad, mayor segregación de la red; cuando la red esté en periodo de estimulación y aparezcan los *clusters* de nodos activados, este grado de segregación aumenta dado que hay una estructura de comunidades más definida que durante el período en reposo. Por otro lado, la transitividad también se define como la propiedad de que si A está conectado con B y B está conectado con C, es muy probable que A y C estén conectados. Los nodos de interés fueron diseñados para tener una alta correlación entre ellos durante los períodos de activación. De hecho, el coeficiente de *clustering* global (al igual que T) está basado en una medida de cuántos de estos *loops* A-B-C hay en la red; en ON, se forman varios de estos *loops* entre los nodos de interés.

Otros de los parámetros mostrados son el **camino característico más corto** (*Characteristic Path Length*) y la Eficiencia Global (relacionado inversamente con el camino más corto, mirar valores de los ejes), que son medidas de integración o, dicho de otro modo, cómo de fácil es transmitir la información entre un nodo y otro de la red. Cuanto menor sea el camino que debe seguir la información para llegar al destino, menos posibilidades de error habrá y más eficiente será la red. Así, los mínimos de L se localizan al final de los períodos de activación, que coincide con los máximos de eficiencia global.

Dado que la red es totalmente conexa (todos los nodos están conectados entre sí), existen métricas que se han mantenido constantes y no han podido aportar nada de información, como la densidad de la red o el grado de cada nodo. Por otro lado, parámetros como la asortatividad han mostrado resultados no interpretables, algo similar a lo encontrado en el estudio entre sujetos sanos y pacientes del apartado 4.3. Los resultados obtenidos para estos y otros parámetros como la intermediación pueden extraerse del Anexo V. *Detección dinámica de comunidades*.

Por último, debe tenerse en cuenta que las ventanas de ON en la figura 21 son orientativas, dado que el eje de tiempos está representado en ventanas y las ventanas tenían solapamiento temporal, por lo que no coinciden de forma exacta con el tiempo de estímulo.

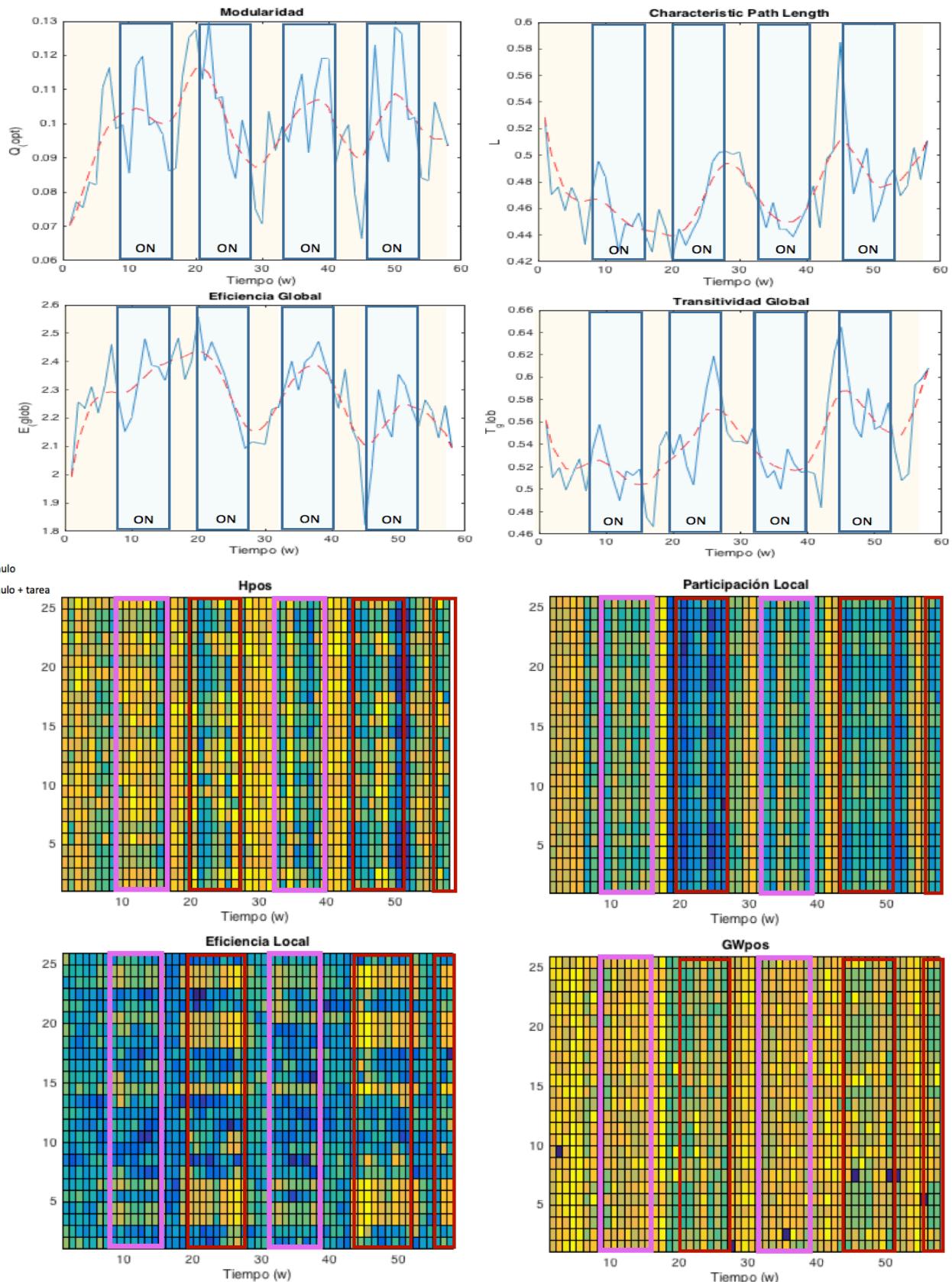


Fig. 21. Evolución a lo largo de las ventanas temporales a nivel global: Modularidad (fila 1, izquierda), Characteristic Path Length (fila 1, derecha), Eficiencia global (fila 2, izquierda), Transitividad Global (fila 2, derecha). A nivel local: Hpos (fila 3, izquierda), Participación local (fila 3, derecha), Eficiencia local (fila 4, izquierda), GWpos (fila 4, derecha).

Características Locales:

Siguiendo con la figura 21, aunque existen ciertos patrones en las matrices de evolución temporal de las propiedades locales, estos aparecen menos definidos que en la detección de comunidades. Por ejemplo, se identifica perfectamente cómo la **eficiencia local** de los nodos de interés aumenta durante los bloques de estimulación; en los casos en los que hay movimiento motor, los nodos relacionados con la tarea se añaden a los anteriores y la eficiencia local de todos ellos es considerablemente mayor que en los bloques de solo estimulación. Como se puede ver en el *Anexo I. Propiedades de un grafo*, la eficiencia local de un nodo es la eficiencia global promedio del subgrafo generado por sus vecinos; esto permite añadir algún matiz a la interpretación de la eficiencia local respecto de la global. En este caso puede caracterizar cómo de bien un nodo intercambia información con sus vecinos (y cómo de resistente es la red a si este falla o se elimina). Dado que los nodos de interés tienen una alta correlación (son la misma señal con algo de ruido añadido), la eficiencia en los períodos de activación crece enormemente (sobre todo si se compara con el resto de señales ruidosas). Todos forman un gran *cluster* con el mismo comportamiento; si falla un enlace, el nodo sigue teniendo una fuerte conexión a otros nodos de interés (propiedad de transitividad) y si falla el nodo, la información se mantiene en la red gracias a los otros nodos. Durante los períodos de activación, la red se vuelve más resistente a fallos.

Un comportamiento similar se encuentra también con la **participación** de cada nodo, que disminuye durante los períodos de activación y aún más cuando hay movimiento. En cambio, la diferencia entre los nodos de interés y el resto de nodos no es tan evidente.

La ponderación del **coeficiente de participación** por la importancia de las conexiones da lugar al **Gateway** (GWpos). Es decir, si un nodo es la única conexión entre una comunidad y otra, será una conexión “importante” para esa comunicación intermodular y tendrá un alto coeficiente de Gateway. En este caso, como todos están conectados con todos, los coeficientes GW son muy homogéneos y solo varían levemente durante los períodos de activación.

Hpos define el **coeficiente de diversidad** obtenido a partir de las conexiones positivas (en este caso son todas positivas). Este tiene un significado análogo a los anteriores, solo que evalúa la diversidad de conexiones intermodulares a partir de la entropía de Shannon [107], lo que parece ayudar a diferenciar entre estados, aunque no así entre nodos de interés y nodos aleatorios. Aún así, se sigue identificando relativamente con facilidad los mismos períodos de mayor activación; compárese los bloques *estímulo + tarea* en rojo respecto de solo *estímulo* (en rosa).

Otras dos medidas de centralidad de cada nodo son el coeficiente **Z** y los **autovalores** o **eigenvalues**, mostrados en la figura 22. En ellos se puede ver cómo los nodos de interés tienen valores muy similares entre ellos y diferentes del resto de nodos de la red. Z evalúa cómo de *fuerte* es la conexión entre los nodos de la misma comunidad, por lo que parece lógico que los nodos de interés, que pertenecen de la misma forma a la comunidad, tengan un mismo coeficiente Z. Los autovalores consideran como nodos centrales de la red a aquellos nodos con alto grado que a su vez están conectados a otros nodos de alto grado; como la fuerza de las conexiones eran más altas en los nodos de interés, estos nodos tienen los autovalores más altos de la red. Por último, nótese que todos estos coeficientes se han definido tras la detección de comunidades, pues es necesario tener primero la estructura de comunidades definida.

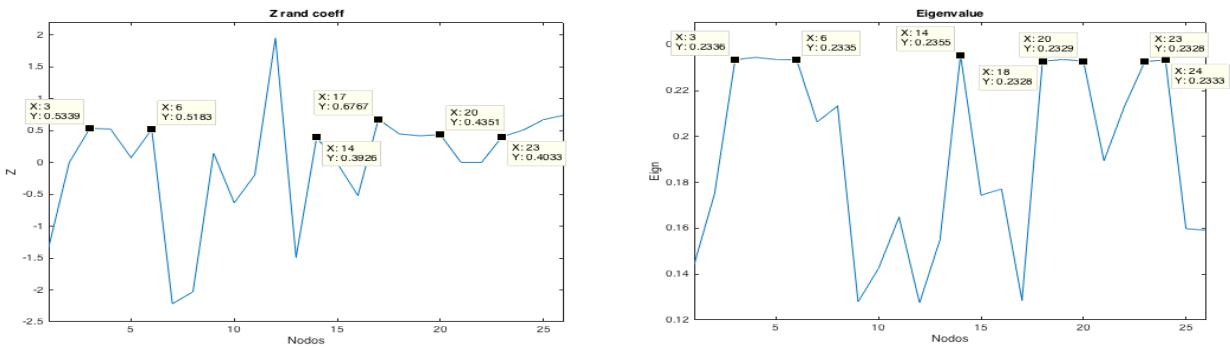


Fig. 22. Medidas de centralidad de los nodos. A la izquierda, el coeficiente Z. A la derecha, los autovalores.

- **Comparación con modelos nulos temporales y nodales**

El mismo procedimiento puede llevarse a cabo para los modelos nulos comentados en el apartado 3.1.2. El modelo de Girvan-Newman ha sido utilizado como R_{ij} durante la detección de comunidades. En este caso, dado que lo que se quiere examinar es el comportamiento dinámico, se mostrarán dos modelos nulos temporales: el primero de ellos proviene de la aleatorización de la serie temporal de cada nodo antes de generar las ventanas temporales y las matrices de conectividad; el segundo de ellos es una reordenación aleatoria de las ventanas temporales de la matriz de conectividad. Los resultados para la detección de comunidades en dichos modelos son los mostrados en la figura 23.

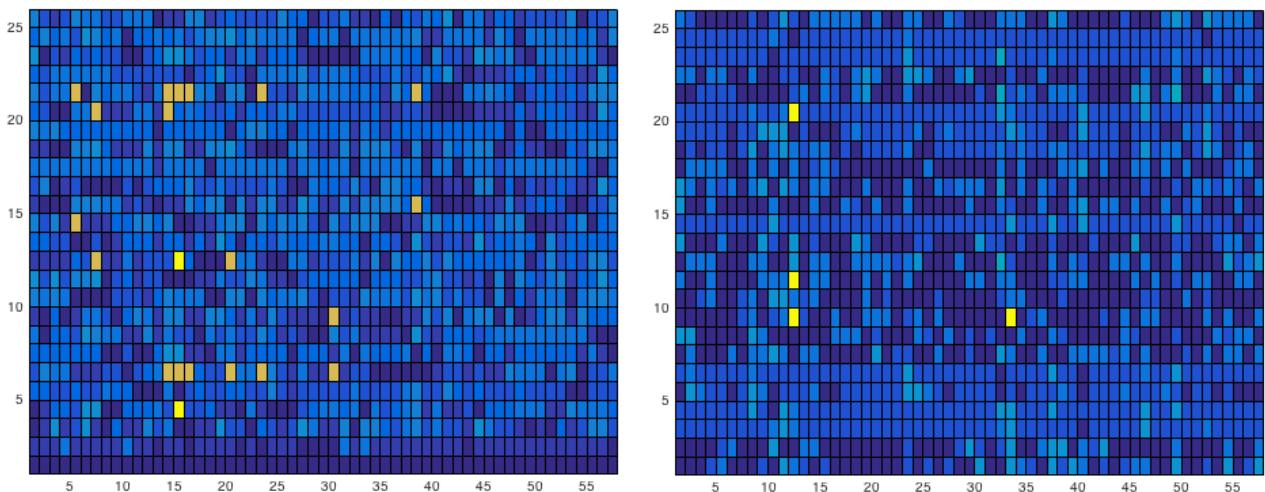


Fig. 23. Comunidades en modelos nulos temporales. A la izquierda, modelos nulos sobre las series temporales. A la derecha, modelo nulo sobre las ventanas temporales de la matriz de conectividad.

Si la aleatorización se produce antes de generar las ventanas y las matrices de conectividad, es como si se aleatorizase toda la información, de modo que no se preserva ninguna estructura de comunidades. De hecho, ni siquiera se mantiene el orden de los pulsos. Véase en la figura 24 el aspecto que tendrían las series temporales de las ICs 6 (DMN) y 18 (componente motora) comparadas con cualquiera de las otras señales aleatorias:

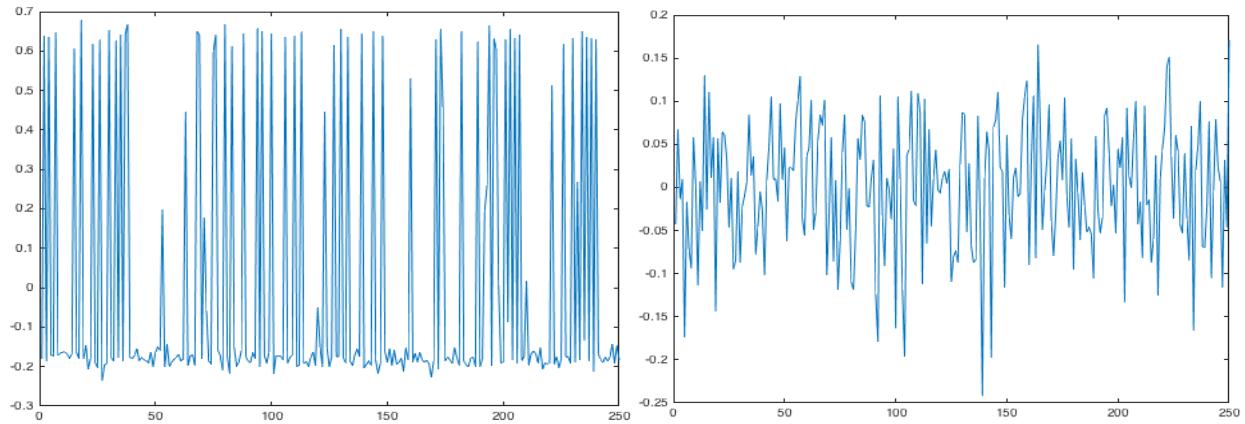


Fig. 24. Modelo nulo de series temporales. A la izquierda, la serie temporal de la componente motora 18. A la derecha, la serie temporal de la componente ruidosa IC 13.

La componente 18 sigue manteniendo valores de amplitud mucho mayor que el ruido, asociados a esos períodos de activación, pero están desordenados (deberían formar los pulsos asociados a los bloques de activación) y se pierde toda la información, lo que produce que no se pueda encontrar ninguna estructura de comunidades. En cuanto a la componente IC13, tras la aleatorización, sigue siendo una señal ruidosa que no proporciona ninguna información.

Respecto al segundo modelo temporal, sí que se pueden identificar perfectamente cuáles son los nodos de interés, dado que al reordenar las ventanas temporales de la matriz de conectividad no se pierde la información sobre las relaciones entre nodos. Sin embargo, lo que no puede obtenerse son los períodos de activación y reposo, ni siquiera analizando el comportamiento de las componentes motoras por separado.

Con estos dos resultados se concluye que las comunidades obtenidas en el apartado anterior no son aleatorias sino fruto de la dinámica de las series temporales, que han sido capturadas por el algoritmo.

Las mismas conclusiones pueden extraerse también si se analizan el resto de propiedades de la red como se hizo en el apartado anterior (véase figura 25).

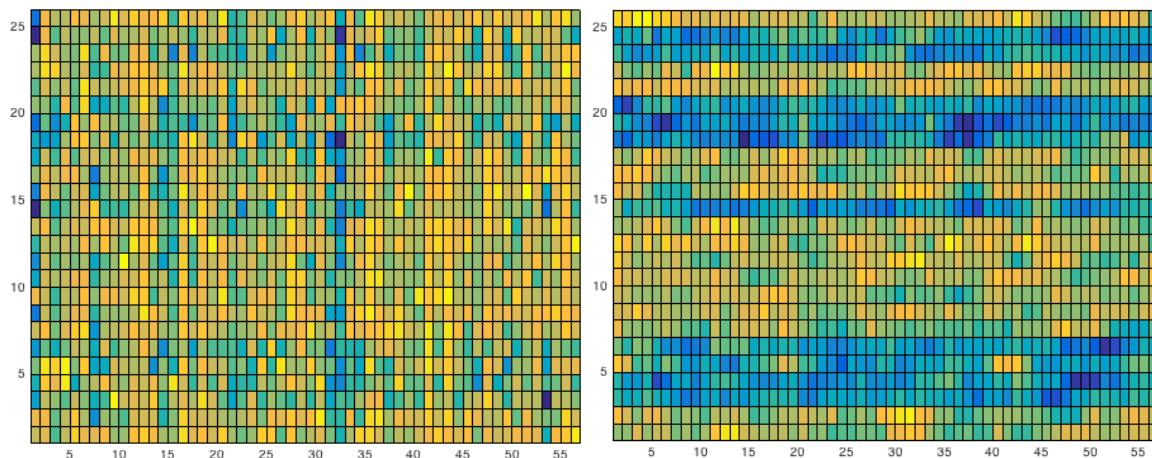


Fig. 25. Eficiencia local de los modelos nulos. A la izquierda, modelos nulos sobre las series temporales. A la derecha, modelo nulo sobre las ventanas temporales de la matriz de conectividad.

6. CONCLUSIONES Y TRABAJO FUTURO

El trabajo se ha centrado en un campo específico de aplicación para la teoría de grafos, la neurociencia computacional, donde se han encontrado multitud de aspectos por formalizar y que pueden influir en los resultados: desde algunos generales como la construcción del grafo o la definición de nodo y enlace hasta las metodologías o estrategias empleadas para aprovechar los grafos como un entorno multimodal (donde se combinen redes anatómicas y funcionales) para el análisis dinámico de redes cerebrales. Este proceso puede causar la pérdida de información o el sesgo de los resultados en el análisis de redes cerebrales, lo que conlleva que la reproducibilidad de los resultados en *network neuroscience* sea aún una dirección en la que trabajar.

No obstante, se han logrado extraer algunas conclusiones sobre la topología cerebral, al igual que se han encontrado diferencias entre distintas patologías como el Alzheimer. La mayoría de estas conclusiones han sido reportadas en la bibliografía anteriormente, lo que puede servir como validación para las técnicas utilizadas. Algunas de ellas son:

- Las redes funcionales están más densamente conectadas que las anatómicas, ya que no tienen restricciones físicas y esto les permite generar multitud de enlaces entre regiones distantes.
- La dependencia de la actividad cerebral al estado anímico, el proceso cognitivo y múltiples de factores ambientales pueden hacer que las comparaciones a nivel funcional entre sujetos tengan una mayor variabilidad. Sin embargo, las redes anatómicas se mantienen similares si se comparan individuos con grupos de individuos.
- Dado el grado de modularidad, la organización jerárquica, interdependencia entre estos módulos, la distribución del grado y la alta eficiencia global (*bajo path length*) tanto en el análisis funcional como en el anatómico, se podría pensar que el desarrollo genético ha conducido a generar una red que maximiza la capacidad de transferir información en poco tiempo, algo así como una red óptima. Sin embargo, la alta densidad de conexiones (*alto coste topológico*) indica que no se ha optimizado en función de reducir el *wiring cost* (aunque sí que seguiría siendo menor que en una red de ER). Una posible justificación estaría basada en la resiliencia a lesiones y enfermedades neurodegenerativas, de modo que si un *hub* se ve dañado, la red puede encontrar otro camino útil para seguir funcionando. De hecho, reducciones en la eficiencia de la red suele estar asociada a lesiones de sustancia blanca, lo que aumenta la vulnerabilidad de la red a lesiones (como en Esclerosis Múltiple o el Alzheimer).

Aún así, cabe destacar que en muchas de las propiedades de la red no se han detectado diferencias significativas entre sujetos con distintas patologías. Por otro lado, existen preguntas que se mantienen abiertas, como definir si tener una correlación mayor que lo normal es algo positivo o negativo a nivel cognitivo (como ocurre en autismo, pero también en gente con mayor capacidad intelectual). Cabe notar también que las edades de los sujetos eran diferentes (ADHD son adolescentes y en Alzheimer son ancianos), aunque los parámetros de adquisición, procesamiento e incluso atlas eran iguales o similares.

Otro punto clave del trabajo ha sido el estudio sobre el comportamiento dinámico de la actividad en la red. La definición de relaciones y operaciones entre comunidades a lo largo del tiempo es un nuevo marco conceptual y todas las aproximaciones actuales tienen ventajas e

inconvenientes que dependen de la aplicación. Algunos puntos que se pueden destacar respecto a la detección dinámica de comunidades son:

- Los algoritmos de detección de comunidades han evolucionado y consiguen dar resultados robustos con unos recursos computacionales factibles. Existe la posibilidad de enfocar esta detección dinámica a partir de técnicas estáticas y de manera eficiente, como la técnica de *Instant Optimal*, que reduce el número de parámetros y permiten paralelizar todo el cómputo. No obstante, es necesario tener en cuenta algunas consideraciones, como el uso de métodos de consenso que permitan obtener particiones verosímiles y la validación estadística de los resultados con modelos nulos.
- La técnica de *sliding-window* para caracterizar el comportamiento dinámico de una red es tremadamente dependiente de la ventana temporal que se defina. No obstante, esta dependencia puede suavizarse hasta cierto punto (para tamaños de ventana próximos al óptimo) a través del *overlapping*. Sin embargo, en situaciones en las que se desconozca el *ground-truth* o directamente no exista un paradigma (por ejemplo, en *resting-state*), la resolución temporal sigue siendo un parámetro sin solución definida.

Con todo esto, la contribución de este trabajo no ha estado centrada en la implementación de nuevas metodologías, sino en la revisión y evaluación detallada de algunas de las existentes:

1. Revisión actualizada de la aplicación de teoría de grafos al estudio de redes cerebrales anatómicas y funcionales.
2. Estudio de los principales parámetros de configuración en la construcción de un grafo. Generalmente, como se puede observar en la bibliografía, suele estudiarse únicamente uno de los parámetros, mientras que en este trabajo se ofrece una visión global de cómo afecta cada uno de ellos para un mismo caso de estudio. Estos parámetros han sido: métricas de conectividad, uso de redes ponderadas o binarias, umbral en la construcción de redes binarias, utilización de valores negativos de correlación, parametrización de análisis dinámicos a través de *sliding-window*, generación de modelos nulos y el uso de datos simulados de fMRI que permitan conocer el *ground-truth* del experimento.
3. Integración en código abierto y en plataformas interactivas que incluye todos estos experimentos y técnicas. El autor desconoce de algún repositorio actual que ofrezca esta posibilidad (de nuevo, vuelven a estar centrados en algún tipo de análisis concreto).
4. Validación de algunos resultados reportados en la bibliografía, lo que refuerza la utilización de teoría de grafos para la caracterización de redes cerebrales.

6.1. TRABAJO FUTURO

El estudio de las redes cerebrales es aún un campo emergente y queda mucho por hacer. Para terminar la memoria, se proponen algunas posibles líneas de trabajo futuro generales; algunas de ellas se corresponden con los retos actuales en análisis de grafos en neurociencia y otros son propuestos a modo personal tras la experiencia de este trabajo. Para finalizar, tras estas propuestas generales, se propone de forma más detallada la ampliación de un método de *Cross-Time Dynamic Detection*, el de Vaiana [91], que podría ser implementada junto a la codificación temporal en *spikes* propuesta en [21].

- **Propuestas generales**

- a) Una de las líneas más inmediatas y necesarias es continuar profundizando y mejorando la definición de la red. Por ejemplo, si existen regiones con diversas tareas ¿deberían representarse todos los nodos y enlaces por igual? ¿Si una región es más grande debería tener varios nodos o pesos más grandes? ¿tiene sentido biológico que signifiquen todas lo mismo?

Otro punto importante es la investigación sobre la métrica utilizada para definir los enlaces entre nodos. Métodos como *cross-correlation* puede ser interesante para estudiar el “flujo” o direccionalidad de la información. Sin embargo, proporciona resultados más complejos de manipular que una correlación (el resultado es una serie temporal) y sigue teniendo el inconveniente de que solo captura relaciones no lineales. Empiezan a aparecer alternativas similares a *distance correlation* como medida de dependencia (lineal o no lineal), como el Criterio de Independencia de Hilbert-Schmidt [108], que trabaja con *kernels* en lugar de solo con la distancia euclídea. También se podrían probar las otras métricas propuestas en la literatura (*wavelets* y FFT, etc.), o incluso investigar el uso de cónulas como herramienta para detectar dependencias en sistemas dinámicos [109].

- b) Una de las ramas que mayor interés suscita es la integración de datos multimodales para lograr entender mejor los mecanismos neuronales. La anatomía puede influir en los procesos cognitivos y viceversa, por lo que este tipo de estudios multimodales podría revelar más información sobre las interacciones topológicas con la geometría cerebral, al igual que sobre la variabilidad entre- e intra-sujetos.

Ligado a esto, resalta el hecho de que la detección de comunidades en redes cerebrales se ha construido sobre el principio de no solapamiento con el algoritmo de Louvain. Dada la naturaleza del cerebro, es evidente la necesidad de investigar nuevas líneas que permitan solapamiento en las comunidades, como el uso de lógica difusa [110].

- c) Una vez desarrolladas las herramientas para la caracterización del comportamiento dinámico de las redes, surge un nuevo objetivo: la predicción de dicho comportamiento. En este punto, técnicas de *Machine Learning* basadas en reconocimiento de patrones parecen ser una de las opciones más prometedoras. Con ello, surgen preguntas como si es posible predecir el comportamiento o la capacidad cognitiva a partir de cambios estructurales o predecir enfermedades y su progresión. Por último, si esto fuera posible, se abre otra pregunta: si somos capaces de predecirlo, ¿somos capaces de actuar de alguna forma para que se modifique ese patrón con beneficios terapéuticos (origen de la enfermedad o progresión de la misma)? Esto reforzaría al que se propone como el modelo de medicina del futuro: el de la prevención personalizada.

- **Modelos aplicables a *spike-trains***

Sorprendentemente, el parámetro ω que parecería jugar un rol importante en la detección de comunidades en redes temporales es comúnmente fijado a un valor constante e igual para todos los nodos. De hecho, esta complejidad para definir los valores de los parámetros de resolución y de generar un modelo nulo apropiado durante la optimización de Q ha provocado que muy recientemente se hayan publicado trabajos que proponen alternativas a este método [88,89], aunque cómo seleccionar los valores óptimos sigue siendo objeto de investigación. De entre todos ellos, la tesis de Michael Vaiana y las publicaciones asociadas a ella [76,90] resultan de especial interés para este trabajo.

En la primera de ellas [91] se identifica una cota superior para la resolución temporal o acoplamiento entre capas (ω) que limita la habilidad para detectar ciertos cambios en la modularidad de la estructura (concretamente, la unión de comunidades). Además, esta cota superior refuerza el principio de que para redes temporales (concretamente, en las redes cerebrales funcionales) construidas a base de métricas de correlaciones donde los pesos de los enlaces tienen como cota superior 1, usar un modelo nulo uniforme U en lugar del de Newman-Girvan es preferible para detectar la evolución de las comunidades (ya que el límite de detección Ω del modelo nulo con distribución uniforme es mayor que el de Newman-Girvan, $\Omega_U > \Omega_{NG}$), como se sugería en [88].

En [90] formulan la base teórica que permite variar ω durante el proceso de optimización de Q con el objetivo de mejorar la sensibilidad en la detección de comunidades frente a cambios dinámicos en redes temporales. Este método selecciona los pesos de los enlaces entre capas basado en medidas de lo que denominan auto-similitud temporal (*temporal self-similarity*) $s(i_t, i_{t+1})$. De este modo, si la métrica de auto-similitud decrece entre dos instantes temporales o capas, su acoplamiento ω también lo hará.

El algoritmo seguiría un proceso como:

1. Inicializar $B_{it} = w_{global}$ para todos los nodos de todas las capas.
2. Calcular nodo por nodo y capa por capa la función $s(i_t, i_{t+1})$ y definir un valor ρ proporcional a la similitud.
3. Actualizar $B_{it} = \rho w_{global}$, donde $\rho \leq 1$.
4. Repetir los pasos 2 y 3 hasta actualizar toda la red.

La función s puede ser cualquiera que resulte de interés. En [90] utilizan simplemente una comparación entre el *rate of spikes* de una neurona (nodo) entre un momento y el siguiente. Esta conceptualización de nodo=neurona y actividad=*spikes* es similar a la plataforma propuesta por el alumno donde la función de auto-similitud s aplicada equivaldría al *Spike-Timing Dependant Plasticity* (STDP) en [21]. Sin embargo, STDP es una métrica más compleja que, aparte de capturar el ratio de *spikes* producidos por cada neurona, tiene en cuenta el instante preciso de disparo (con resolución temporal igual a la resolución de adquisición) que ha demostrado ser una propiedad biológica determinante.

De este modo, se propone como trabajo futuro la ampliación del trabajo de Vaiana en [90] con las métricas funcionales propuestas en [21] con el objetivo de evaluar de forma simultánea:

- Nuevas definiciones de la red y la experimentación de nuevas métricas como STDP.
- La implementación de un entorno multimodal.
- Un estudio más profundo tanto de las dinámicas de la red.

Referencias

1. Newman, MEJ. Networks: An Introduction. Oxford University Press; 2010.
2. Jbabdi S, Sotiroopoulos SN, Haber SN, Van Essen DC, Behrens TE. Measuring macroscopic brain connections in vivo. *Nat Neurosci*. 2015; 18:1546–1555. [PubMed: 26505566]
3. Oh SW, et al. A mesoscale connectome of the mouse brain. *Nature*. 2014; 508:207–214. [PubMed: 24695228]
4. Hamel EJ, Grewe BF, Parker JG, Schnitzer MJ. Cellular level brain imaging in behaving mammals: an engineering approach. *Neuron*. 2015; 86:140–159. [PubMed: 25856491]
5. Keller PJ, Ahrens MB. Visualizing whole-brain activity and development at the single-cell level using light-sheet microscopy. *Neuron*. 2015; 85:462–483. [PubMed: 25654253]
6. Sporns, O, Tononi, G, Kötter, R (2005) The human connectome: A structural description of the human brain. *PLoS Comput. Biol.* 1, 245-251.
7. Rosazza, C.; Minati, L. (2011). "Resting-state brain networks: literature review and clinical applications". *Neurol Sci*. 32 (5): 773–785.
8. Friston, K (2009). "Causal Modelling and Brain Connectivity in Functional Magnetic Resonance Imaging. [Editorial Material]". *PLOS Biology*. 7 (2): 220–225
9. Huang H and Ding M. Linking functional connectivity and structural connectivity quantitatively: a comparison of methods. 2016. *Brain Connectivity*. 6(2): 99-108
10. Goulas A, et al. Comparative analysis of the macroscale structural connectivity in the macaque and human brain. *PLoS Comput Biol*. 2014; 10:e1003529. [PubMed: 24676052]
11. Li L, et al. Mapping putative hubs in human, chimpanzee and rhesus macaque connectomes via diffusion tractography. *Neuroimage*. 2013; 80:462–474. [PubMed: 23603286]
12. Mišić B, Sporns O. From regions to connections and networks: new bridges between brain and behavior. *Curr Opin Neurobiol*. 2016; 40:1–7. [PubMed: 27209150]
13. Fornito A, Bullmore ET. Connectomic intermediate phenotypes for psychiatric disorders. *Front Psychiatry*. 2012; 3:32. [PubMed: 22529823]
14. Vidal M, Cusick ME, Barabási AL. Interactome networks and human disease. *Cell*. 2011; 144:986–998. [PubMed: 21414488]
15. Crossley NA, et al. Cognitive relevance of the community structure of the human brain functional coactivation network. *Proc Natl Acad Sci USA*. 2013; 110:11583–11588. [PubMed: 23798414]
16. Raj A, Kuceyeski A, Weiner M. A network diffusion model of disease progression in dementia. *Neuron*. 2012; 73:1204–1215. [PubMed: 22445347]
17. Fornito A, Zalesky A, Pantelis C, Bullmore ET. Schizophrenia, neuroimaging and connectomics. *Neuroimage*. 2012; 62:2296–2314. [PubMed: 22387165]
18. de la Torre-Ubieta L, Won H, Stein JL, Geschwind DH. Advancing the understanding of autism disease mechanisms through genetics. *Nat Med*. 2016; 22:345–361. [PubMed: 27050589]
19. Lewis PM, Thomson RH, Rosenfeld JV, Fitzgerald PB. Brain neuromodulation techniques: a review. *Neuroscientist*. 2016; 22:406–421. [PubMed: 27130839]
20. Johnson MD, et al. Neuromodulation for brain disorders: challenges and opportunities. *IEEE Trans Biomed Eng*. 2013; 60:610–624. [PubMed: 23380851]
21. https://github.com/josepm/TM_TECI
22. Hutchison RM, et al. Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage*. 2013; 80:360–378. [PubMed: 23707587].
23. Liu, S et al. Multimodal neuroimaging computing: a review of the applications in neuropsychiatric disorders. *Brain Informatics*. 2015; 2. 3. 167-180.
24. Calhoun VD, Miller R, Pearlson G, Adali T. The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery. *Neuron*. 2014; 84:262–274. [PubMed: 25374354]
25. Holme P. Temporal networks. *Phys Rep*. 2012; 519:97–125.
26. Kivelä M, et al. Multilayer networks. *J Compl Netw*. 2014; 2:203–271.

27. Vaiana M, Muldoon S. Multilayer brain networks. 2017. Quantitative Biology.
28. Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*. 2009; 10:186–198. [PubMed: 19190637]
29. Fornito, A., Zalesky, A., Bullmore, E. *Fundamentals of Brain Network Analysis*. Academic Press; 2016.
30. Sporns, O. *Networks of the Brain*. MIT Press; 2010. 7. Cunningham JP, Yu BM.
31. Medaglia JD, Lynall ME, Bassett DS. Cognitive network neuroscience. *J Cogn Neurosci*. 2015; 27:1471–1491. [PubMed: 25803596]
32. Mucha PJ, Richardson T, Macon K, Porter MA, Onnela JP. Community structure in time-dependent, multiscale, and multiplex networks. *Science*. 2010; 328:876–878. [PubMed: 20466926]
33. Betzel RF, et al. Generative models of the human connectome. *Neuroimage*. 2016; 124(Pt A): 1054–1064. [PubMed: 26427642]
34. Deco G, Jirsa VK, Robinson PA, Breakspear M, Friston K. The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS Comput Biol*. 2008; 4:e1000092. [PubMed: 18769680]
35. Betzel RF, Gu S, Medaglia JD, Pasqualetti F, Bassett DS. Optimally controlling the human connectome: the role of network topology. *Sci Rep*. 2016; 6:30770. [PubMed: 27468904].
36. Fornito A, Zalesky A, Breakspear M. Graph analysis of the human connectome: promise, progress, and pitfalls. *Neuroimage*. 2013; 80:426–444. [PubMed: 23643999]
37. Sporns O. Contributions and challenges for network models in cognitive neuroscience. *Nat Neurosci*. 2014; 17:652–660. [PubMed: 24686784]
38. Zalesky A, Fornito A, Bullmore ET. Network-based statistic: identifying differences in brain networks. *Neuroimage*. 2010; 53:1197–1207. [PubMed: 20600983]
39. Biswal B, Zerrin Yetkin VM, Haughton JS, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. 1995. *Magnetic resonance in medicine* 34 (4), 537-541
40. Smith, Stephen M; Andersson, Jesper; Auerbach, Edward J.; Beckmann, Christian F; Bijsterbosch, Janine; Douaud, Gwenaëlle; Duff, Eugene; Feinberg, David A; Griffanti, Ludovica (2013-10-15). "Resting-state fMRI in the Human Connectome Project". *NeuroImage*. 80: 144–168.
41. Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, Xu J, Jbabdi S, Webster M, Polimeni JR, Van Essen DC, Jenkinson M; WU-Minn HCP Consortium. The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage*, 2013;20:105-24.
42. Hayasaka, S., and Laurienti, P.J. Comparison of characteristics between region- and voxel-based network analysis in resting-state fMRI data (2010). *NeuroImage*, 50(2), 499-508.
43. Power JD, et al. Functional network organization of the human brain. *Neuron*. 2011; 72:665–678. [PubMed: 22099467]
44. Cole MW, Bassett DS, Power JD, Braver TS, Petersen SE. Intrinsic and task-evoked network architectures of the human brain. *Neuron*. 2014; 83:238–251. [PubMed: 24991964]
45. Strotzer M. One century of brain mapping using Brodmann áreas. *Klin Neuroradiol*. 2009. 19(3):179-86
46. de Paula, DR et al. A method for independent component graph analysis of resting-state fMRI. *Brain and Behavior*. 2017; 7(3)
47. Zalesky A, Fornito A, Bullmore E. On the use of correlation as a measure of network connectivity. *Neuroimage*. 2012; 60:2096–2106. [PubMed: 22343126]
48. Baba, Kunihiro; Ritei Shibata; Masaaki Sibuya (2004). "Partial correlation and conditional correlation as measures of conditional independence". *Australian and New Zealand Journal of Statistics*. 46 (4): 657–664. doi:10.1111/j.1467-842X.2004.00360.x

49. Hindriks R, Adhikari MH, Murayama Y, Ganzetti M, Mantini D, Logothetis NK, Deco G. Can sliding-window correlations reveal dynamic functional connectivity in resting-state fMRI? *Neuroimage*. 2016; 127:242–256.
50. Honey CJ, Kötter R, Breakspear M, Sporns O. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proc Natl Acad Sci USA*. 2007; 104:10240–10245. [PubMed: 17548818]
51. Betzel, RF., et al. Functional brain modules reconfigure at multiple scales across the human lifespan. 2015. Preprint at <https://arxiv.org/abs/1510.08045>
52. Richiardi J, et al. BRAIN NETWORKS. Correlated gene expression supports synchronous activity in brain networks. *Science*. 2015; 348:1241–1244. [PubMed: 26068849]
53. Thomas C, et al. Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. *Proc Natl Acad Sci USA*. 2014; 111:16574–16579. [PubMed: 25368179]
54. Jones DK, Knösche TR, Turner R. White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. *Neuroimage*. 2013; 73:239–254. [PubMed: 22846632]
55. Edward T. Bullmore, Danielle S. Bassett (2011). “Brain Graphs: Graphical Models of the Human Brain Connectome”. *Annual Review of Clinical Psychology*, Vol. 7
56. M.E. Newman and M. Girvan "Finding and evaluating community structure in networks". *Physical review E* 69.2 (2004): 026113
57. García J, Ashourvan A, Muldoon S, Vettel J, Bassett, DS. Applications of community detection techniques to brain graphs: algorithmic considerations and implications for neural function. 2018. *Proceedings of the IEEE*. 2016;5 846-867.
58. Changsong Zhou, Lucia Zemanová, Gorka Zamora, Claus C. Hilgetag, and Jürgen Kurths. Hierarchical organization unveiled by functional connectivity in complex brain networks. 2006. *Phys. Rev. Lett.* 97, 238103
59. Calhoun VD, Miller R, Pearlson G, Adali T. The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery. *Neuron*. 2014; 84:262–274. [PubMed: 25374354]
60. Goñi J, et al. Resting-brain functional connectivity predicted by analytic measures of network communication. *Proc Natl Acad Sci USA*. 2014; 111:833–838. [PubMed: 24379387]
61. Fortunato, S. Community detection in graphs. 2010. *Physic reports*. 486:3-5; 75-174
62. Newman, MEJ. Communities, modules and large-scale structure in networks. 2012. *Nature Physics* v.8, pages 25–31 (2012).
63. Benjamin H. Good, Yves-Alexandre de Montjoye, and Aaron Clauset. Performance of modularity maximization in practical contexts. 2010. *Phys. Rev. E* 81, 046106
64. Fortunato S, Barthélémy M. Resolution limit in community detection. 2007. *PNAS*. 104(1) 36-41
65. Brandes, U., Delling, D, Gaertler, M., Görke, R., Hoefer, M., Nikoloski, Z., and Wagner, D. On Modularity – NP-Completeness and Beyond (2007)]
66. Blondel, V.D., Guillaume, J.-L., Lambiotte, R. and Lefebvre, E. Fast unfolding of communities in large networks (2008). *Journal of statistical mechanics: theory and experiment*
67. Fast unfolding of communities in large networks, Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, Etienne Lefebvre, *Journal of Statistical Mechanics: Theory and Experiment* 2008 (10), P10008 (12pp) doi: 10.1088/1742-5468/2008/10/P10008. ArXiv: <https://arxiv.org/abs/0803.0476>
68. Jeub, L.G.S., Bazzi, M., Jutla, I.S. and Mucha, P.J. A generalized Louvain method for community detection implemented in MATLAB (2011-2017)
69. Lancichinetti, A., and Fortunato, S. Consensus clustering in complex networks (2012). *Scientific reports*, 2, 336
70. Jeub, L.G., Sporns, O. and Fortunato, S. Multiresolution consensus clustering in networks (2018). *Scientific reports*, 8(1), 3259

71. Palla G, Barabasi AL, Vicsek T. Quantifying social group evolution. *Nature* (2007). 446(7136):664-667
72. Cazabet R, Takeda H, Hamakasi M, Amblard F. Using dynamic community detection to identify trends in user-generated content (2012). *Soc. Netw. Anal. Min.* 2: 361-371
73. Chen W, Liu Z, Sun X, Wang Y (2010) A game-theoretic framework to identify overlapping communities in social networks. *Data Min Knowl Discov* 21(2): 224–240
74. Dhouioui Z, Akaichi J. (2014). Tracking dynamic community evolution in social networks. *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)* Asonam (2014), 764–770
75. Danielle S. Bassett, Mason A. Porter, Nicholas F. Wymbs, Scott T. Grafton, Jean M. Carlson and Peter J. Mucha. Robust detection of dynamic community structure in networks. (2013). *Chaos*, 23
76. Vaiana, M. and Muldoon, S. Resolution limits for detecting community changes in multilayer networks. *arXiv preprint arXiv:1803.03597*
77. Karrer B and Newman MEJ. Stochastic blockmodels and community structure in networks. 2011. *Phys. Rev. E* 83, 016107
78. Honey C, Sporns O, Cammoun L, Gigandet X, Thiran JP, et al. 2009. Predicting human resting state functional connectivity from structural connectivity. *Proc. Natl. Acad. Sci. USA* 106:2035–40
79. Bassett, DS., Bullmore, ET. Small-world brain networks revisited. *The Neuroscientist*. 2016. <http://dx.doi.org/10.1177%2F1073858416667720>
80. Betzel, RF., et al. The modular organization of human anatomical brain networks: accounting for the cost of wiring. *Network Neurosci*. 2017. http://doi.org/10.1162/NETN_a_00002
81. Sporns O, Betzel RF. Modular brain networks. *Annu Rev Psychol*. 2016; 67:613–640.
82. van den Heuvel MP, Sporns O. Network hubs in the human brain. *Trends Cogn Sci*. 2013; 17:683– 696.
83. van den Heuvel MP, Sporns O. Rich-club organization of the human connectome. *J Neurosci*. 2011; 31:15775–15786.
84. Nigam S, et al. Rich-club organization in effective connectivity among cortical neurons. *J Neurosci*. 2016; 36:670–684.
85. Dean Prichard and James Theiler. Generating surrogate data for time series with several simultaneously measured variables. 1994. *Phys. Rev. Lett.* 73, 951
86. Maslov S and Sneppen K. Specificity and stability in topology of protein networks. 2002. *Science*. Vol. 296, Issue 5569, pp. 910-913
87. Traud AL, Kelsic ED, Mucha PJ and Porter MA. Comparing Community structure to characteristics in online collegiate social networks. 2011. *SIAM Rev*. 53(3) 526-543.
88. Bazzi, M., Porter, M. A., Williams, S., McDonald, M., Fenn, D. J. and Howison, S.D. Community detection in temporal multilayer networks, with an application to correlation networks. (2016). *Multiscale model. Simul.*
89. Weir, W.H., Emmons, S., Gibson, R., Taylor, D. and Mucha, P.J. Post-processing partitions to identify domains of modularity optimization (2017). *Algorithms*. 10(3), 93
90. Vaiana, M. and Muldoon, S. Optimizing state change detection in functional temporal networks through dynamic community detection. (2018)- Submitted
91. Vaiana, M. and Muldoon, S. Resolution limits for detecting community changes in multilayer networks. *arXiv preprint arXiv:1803.03597*
92. Erdős, P.; Rényi, A. (1960). "The Evolution of Random Graphs". *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 5: 17–61
93. Damoiseaux, J. S.; Rombouts, S.; Barkhof, F.; Scheltens, P.; Stam, C. J.; Smith, S. M.; et al. (2006). "Consistent resting-state networks across healthy subjects". *Proceedings of the National Academy of Sciences of the United States of America*. 103 (37): 13848–13853
94. Huettel, S. A.; Song, A. W.; McCarthy, G. (2009), Functional Magnetic Resonance Imaging (2 ed.), Massachusetts: Sinauer

95. <http://umcd.humanconnectomeproject.org/>
96. Welvaert M and Rosseel Y. A review of fMRI simulation studies. 2014. PLOS One 9(7).
97. Friston K. And Henson R. Convolution Models for fMRI. 2006.
98. paper de direccionalidad de neuroscience frontiers 2018
99. Korhonen, O., Saarimaki, H., Glerean, E., Sams, M., and Saramaki, J. Consistency of regions of interest as nodes of fMRI functional brain networks (2017). Network Neuroscience, 1(3), 254-274
100. Ryyppö, E., Glerean, E., Brattico, E., Saramäki, J., and Korhonen, Onerva. Regions of interest as nodes of dynamic functional networks (2018). Network Neuroscience (In press)
101. Erhardt EB, Allen EA, Wei Y, Eichele T, Calhoun VD. SimTB, a simulation toolbox for fMRI data under a model of spatiotemporal separability. Neuroimage. 2012 Feb 15;59(4):4160-7
102. G. J. Szekely; M. L. Rizzo; N. K. Bakirov (2007), "Measuring and testing independence by correlation of distances", *Annals of Statistics*, 35(6): 2769–2794
103. <https://networkx.github.io/>
104. <https://sites.google.com/site/bctnet/>
105. Jon del Etoile et al. (2017). Graph Theory and Brain Connectivity in Alzheimer's Disease. *The Neuroscientist Review*
106. David J. Philips et al. (2015). Graph theoretic analysis of structural connectivity across the spectrum of Alzheimer's disease: The importance of graph creation methods. *Neuroimage Clinical*. 377-390
107. Shannon CE (1948) Bell Syst Tech J 27, 379-423
108. Gretton, A. et al. (2005) Measuring Statistical Dependence with Hilbert-Schmidt Norms
109. Ayyad, C; Porcu, E (2008), «Inferencia y modelización mediante cópulas», Universidad Jaume
110. Detecting overlapping temporal community structure in time-evolving networks