



UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

Propriedades estatísticas do espectro de redes cerebrais funcionais

DISSERTAÇÃO DE MESTRADO

Matheus Rodrigues Xavier

São Cristóvão/SE

Julho/2024

UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

Matheus Rodrigues Xavier

Propriedades estatísticas do espectro de redes cerebrais funcionais

Dissertação apresentada ao Programa de Pós-Graduação em Física da Universidade Federal de Sergipe como requisito parcial para a obtenção do título de Mestre em Física.

Orientador: Prof. Dr. Gerson Cortês Duarte Filho

São Cristóvão/SE

Julho/2024

UNIVERSIDADE FEDERAL DE SERGIPE
Programa de Pós-Graduação em Física

**“Propriedades estatísticas do espectro de redes
cerebrais funcionais”**

por

Matheus Rodrigues Xavier

Dissertação de mestrado **APROVADA** em sessão pública no décimo nono dia de julho do ano dois mil e vinte e quatro perante a banca examinadora constituída pelos seguintes professores:

Documento assinado digitalmente
 GERSON CORTES DUARTE FILHO
Data: 19/07/2024 12:54:29-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Gerson Cortês Duarte Filho
(DFI-UFS)

Documento assinado digitalmente
 MARCOS VINICIUS DOS SANTOS REZENDE
Data: 21/07/2024 21:06:03-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Marcos Vinicius dos Santos Rezende
(DFI-UFS)

Documento assinado digitalmente
 ANTONIO MURILO SANTOS MACEDO
Data: 22/07/2024 09:31:37-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Antonio Murilo Santos Macêdo
(DF-UFPE)

Agradecimentos

Ao professor Dr. Gerson Cortês Duarte Filho pela orientação e ensinamentos durante toda a colaboração e realização deste trabalho, sempre muito solícito e paciente.

À minha família, por todo o apoio e incentivo, em especial à minha mãe, Naide Rodrigues Mendes, e ao meu pai, Valter Xavier Mendes.

Aos meus amigos que estiveram ao meu lado durante toda essa jornada, sempre me apoiando e ajudando em todos os momentos difíceis, em especial Junio, Klivan, Flávio, Lauro, Arthur, Erison, William e Sandro.

Agradeço também à Universidade Federal de Sergipe, ao Programa de Pós-Graduação em Física e a todo seu corpo docente de alta qualidade, que demonstraram estar totalmente comprometidos com a qualidade do ensino, assim como à CAPES, por ter tornado este trabalho possível.

Resumo

O cérebro é uma das redes mais complexas conhecidas, composto por uma enorme quantidade de neurônios e conexões, capazes de realizar inúmeras sinapses, organizados anatomicamente em diferentes escalas espaciais e interativos em diversas escalas temporais, formando uma estrutura de arquitetura complexa. Compreender o funcionamento desse sistema é essencial para o direcionamento de estudos em neurociência, oferecendo novas perspectivas para o tratamento de distúrbios cerebrais que ainda carecem de terapias eficazes. Devido aos avanços teóricos em matemática e física estatística, com modelos cada vez mais precisos para a descrição de dados reais, juntamente com avanços tecnológicos em técnicas de análise computacional de dados e exames de mapeamento cerebral, tornou-se possível abordar o conectoma, definido como a matriz de conectividades funcionais das regiões anatômicas do cérebro, permitindo assim o estudo do cérebro como uma rede complexa. Essa abordagem possibilita, por exemplo, a aplicação da Teoria de Matrizes Aleatórias (TMA) para estudar a estatística espectral desse tipo de sistema. Neste trabalho, analisamos um conjunto de dados de conexões funcionais cerebrais obtidos via imagem por ressonância magnética funcional (fMRI) de indivíduos em estado de repouso. Estudamos a distribuição de espaçamentos consecutivos (DEC), assim como a distribuição das razões entre espaçamentos consecutivos (DREC) do espectro das matrizes de adjacência dos dados. Utilizamos a distribuição de Brody para ajustar a DEC e propusemos uma expressão semi-empírica para a DREC, a fim de avaliar o grau de correlação dos autovalores no conjunto de matrizes de adjacência. Essas distribuições semi-empíricas dependem de um parâmetro β que varia continuamente entre zero, quando a estatística espectral é descrita pela distribuição de Poisson, e um, quando é descrita pelo Ensemble Gaussiano Ortogonal (GOE) da TMA. Os ajustes realizados forneceram estimativas de $\beta = 0,89$ para a DEC e $\beta = 0,87$ para a DREC, sugerindo uma estatística espectral intermediária entre a distribuição de Poisson e o GOE.

Palavras-chave: Redes cerebrais. Conectividade funcional. Teoria de matrizes aleatórias. Distribuição de Brody.

Abstract

The brain is one of the most complex networks known, composed of an enormous number of neurons and connections capable of performing countless synapses, anatomically organized across different spatial scales and interactive across various temporal scales, forming a structure of complex architecture. Understanding the functioning of this system is essential for guiding neuroscience studies, offering new perspectives for the treatment of brain disorders that still lack effective therapies. Due to theoretical advances in mathematics and statistical physics, with increasingly precise models for the description of real data, along with technological advances in computational data analysis techniques and brain mapping exams, it has become possible to approach the connectome, defined as the matrix of functional connectivities of the anatomical regions of the brain, thus allowing the brain to be studied as a complex network. This approach enables, for example, the application of Random Matrix Theory (RMT) to study the spectral statistics of this type of system. In this work, we analyzed a dataset of functional brain connections obtained by functional magnetic resonance imaging (fMRI) of individuals in a resting state. We studied the distribution of consecutive spacings (DCS), as well as the distribution of ratios between consecutive spacings (DRCS) of the adjacency matrix spectra of the data. We used the Brody distribution to fit the DCS and proposed a semi-empirical expression for the DRCS to assess the degree of correlation of the eigenvalues in the ensemble of adjacency matrices. These semi-empirical distributions depend on a parameter β that varies continuously between zero, when the spectral statistics are described by the Poisson distribution, and one, when described by the Gaussian Orthogonal Ensemble (GOE) of RMT. The fits performed provided estimates of $\beta = 0.89$ for the DCS and $\beta = 0.87$ for the DRCS, suggesting an intermediate spectral statistic between the Poisson distribution and the GOE.

Keywords: Brain networks. Functional connectivity. Random matrix theory. Brody distribution.

Lista de ilustrações

- Figura 1 – (a) Uma matriz de conectividade para um cérebro representando a conectividade de 32 áreas corticais. (b) Esta matriz está organizada de modo que os elementos coloridos representem uma projeção da região listada na coluna para a região listada na linha. O tamanho dos pontos em cada elemento da matriz é proporcional à distância de projeção e as cores mais escuras indicam uma intensidade média de conectividade relatada mais forte. (c) As localizações anatômicas das áreas listadas na matriz em (b), cores mais escuras identificam regiões com maior conectividade total com o restante da rede. 17
- Figura 2 – (a) Rede direcionada, ponderada; (b) Rede direcionada, binária; (c) Rede não direcionada, ponderada (d); Rede não direcionada, binária. . . 24
- Figura 3 – A organização modular de um grafo cerebral via fMRI em uma configuração topológica. 30
- Figura 4 – Tractografia via MRI de difusão usada para construir uma matriz de conectividade ponderada e não direcionada, onde cada elemento quantifica a densidade das projeções axonais entre um par específico de regiões. Essa rede pode ser ilustrada como um grafo cerebral projetado usando um layout definido algoritmicamente para enfatizar aspectos topológicos da organização da rede. 31
- Figura 5 – (a) e (b), representam distribuições semicircular e triangular das densidades espectrais para a média em ensemble de redes aleatórias e livres de escala, respectivamente, onde λ representa os autovalores e $\rho(\lambda)$ é a densidade espectral. Todas as redes têm uma dimensão $N = 2000$ e grau médio $k = 20$. Essas figuras foram plotadas com uma média de 10 realizações aleatórias das redes. 34
- Figura 6 – (a) e (b) correspondem a distribuição de espaçamento $P(s)$, mostrando que essas distribuições seguem as estatísticas GOE. Os histogramas são resultados numéricos e a curva representa a distribuição de Brody ajustada, com seu parâmetro β indicado. Esses resultados para as mesmas redes da Figura 5 34

Figura 7 – (a)–(d) mostram as densidades espectrais $\rho(\lambda)$ e (e)–(h) mostram as distribuições de espaçamento correspondentes $P(s)$, para $P = 0$, $P = 5 \times 10^{-5}$, $P = 2 \times 10^{-4}$ e $P = 5 \times 10^{-4}$, respectivamente. Os histogramas são dados numéricos e as curvas representam o ajuste dos dados pela distribuição de Brody (2.4). Todas as redes têm dimensão $N = 2000$ nó, um grau médio $k = 40$ por nó e os dados são a média de 10 realizações aleatórias do processo de reconexão.	37
Figura 8 – Visualização reduzida de 4 matrizes de conectividades, escolhidas aleatoriamente, representadas como Data Frame no Python	39
Figura 9 – (a) Uma rede de conectividade funcional baseada em correlação, medida com fMRI de um cérebro humano. Cada elemento nesta matriz tem um valor diferente de zero. (b) A mesma matriz de (a) após a aplicação de um limiar para reter apenas os 20% dos pesos mais intensos. (c) A rede após a binarização.	43
Figura 10 – Visualização de uma matriz adjacência obtida a partir de uma matriz conectividade.	44
Figura 11 – Histograma da distribuição de autovalores degenerados.	45
Figura 12 – Histograma da distribuição de autovalores não degenerados.	46
Figura 13 – Histograma da distribuição média em ensemble de autovalores.	51
Figura 14 – Histograma da distribuição média em ensemble de autovalores desdobrados.	52
Figura 15 – Histograma da distribuição média em ensemble de espaçamentos consecutivos entre autovalores não desdobrados.	53
Figura 16 – Histograma da distribuição média em ensemble de espaçamentos consecutivos entre autovalores desdobrados	54
Figura 17 – Dados fitados no histograma da média em ensemble da distribuição entre espaçamentos consecutivos de autovalores desdobrados pela distribuição de Brody. A Curva <i>Best fit</i> representa a distribuição de Brody com os valores de β e α_p que melhor se ajustam aos dados, indicados na Tabela 3.	55
Figura 18 – Histograma da distribuição dos valores de β para cada membro do ensemble na DEC.	56
Figura 19 – Os 10 melhores ajustes individuais da DEC levando em consideração os maiores valores de R^2 , com os valores de β e α_p listados na Tabela 5.	57
Figura 20 – Histograma da distribuição média em ensemble de razões entre espaçamentos consecutivos de autovalores não desdobrados	58
Figura 21 – Ajuste do histograma da média em ensemble da distribuição de razões entre espaçamentos consecutivos de autovalores desdobrados pela distribuição de Brody, com o valor de β informado na Tabela 6.	59

Figura 22 – Histograma da distribuição os valores de β para cada membro do ensemble considerando a DREC.	60
Figura 23 – Ajustes realizados nos histogramas sobrepostos das razões entre espaçamentos consecutivos de autovalores não desdobrados, pela distribuição semi-empírica equação (3.8) utilizando os valores de β listados na Tabela 8.	61

Lista de tabelas

Tabela 1	– Relação entre quantidades de autovalores não degenerados e a quantidade de matrizes, no intervalo de matrizes que possuem entre 90 e 100 elementos.	46
Tabela 2	– Valores da constante de normalização Z_β	48
Tabela 3	– Medidas estatísticas sobre valores dos parâmetros do modelo da distribuição de Brody, melhor ajustado aos dados, para a DEC pela equação (3.8).	52
Tabela 4	– Os parâmetros de Brody e R^2 estimados a partir das distribuições de β , α_p e R^2 obtidas pelos ajustes individuais de cada membro do ensemble, para a DEC.	53
Tabela 5	– 10 melhores ajustes dos parâmetros β e α_p para a DEC, considerando os maiores valores de R^2	55
Tabela 6	– Parâmetro de Brody que melhor ajusta os dados, da DREC obtida pela média em ensemble.	58
Tabela 7	– Medidas estatísticas associadas a distribuição dos valores de β apresentados na Figura 22.	58
Tabela 8	– 10 melhores valores do parâmetro β para a DREC.	60
Tabela 9	– Valores estimados para o parâmetro β para DEC e DREC, por média em ensemble e análise individual de membros do ensemble.	62

Lista de abreviaturas e siglas

fMRI	Imagem por Ressonância Magnética Funcional
MRI	Imagem por Ressonância Magnética
TMA	Teoria de Matrizes Aleatórias
DEC	Distribuição de Espaçamentos Consecutivos
GOE	Ensemble Ortogonal Gaussiano
GUE	Ensemble Unitário Gaussiano
GSE	Ensemble Simplético Gaussiano
DREC	Distribuição de Razões entre Espaçamentos Consecutivos

Sumário

1	Introdução	12
1.1	Redes cerebrais	12
1.2	Conectoma	15
2	Fundamentos Teóricos	20
2.1	Sistemas Complexos	20
2.1.1	Teoria dos grafos	23
2.1.2	Grafos Cerebrais	26
2.2	Teoria de Matrizes Aleatórias	31
2.2.1	Distribuição de Brody	36
3	Metodologia	38
3.1	Matrizes de Conectividade	38
3.2	Matrizes de Adjacência	40
3.2.1	Diagonal Principal	41
3.2.2	Filtragem	41
3.2.3	Binarização	42
3.3	Análise espectral das matrizes adjacência	43
3.4	Distribuições Estatísticas	47
3.5	Análise de Regressão	48
4	Resultados e Discussões	50
4.1	Distribuição de Autovalores	50
4.2	Distribuição de Espaçamentos Consecutivos	50
4.3	Distribuição de Razões entre Espaçamentos Consecutivos	54
4.4	Análise sobre os valores do parâmetro β	59
5	Conclusões	63
	Referências	65

1 Introdução

1.1 Redes cerebrais

O cérebro humano é uma das redes mais complexas conhecidas pelo homem, compreendendo cerca de 100 bilhões (10^{11}) neurônios conectados por cerca de 100 trilhões (10^{14}) de sinapses. Essas conexões são responsáveis por diferentes processos cognitivos e estão anatomicamente organizadas em uma estrutura de múltiplas escalas espaciais, além de serem funcionalmente interativas em múltiplas escalas de tempo. Esse sistema é o hardware biológico do qual originam-se todos os nossos pensamentos, sentimentos e comportamentos [Fornito, Zalesky e Bullmore 2016, Geerligs, Cam-CAN e Henson 2016].

Distúrbios clínicos das redes cerebrais humanas, como esquizofrenia, autismo, Alzheimer, TDAH, entre outros, estão entre os problemas de saúde mais difíceis de tratar atualmente. A investigação da conectividade dentro dessas redes, a compreensão dos aspectos da cognição e dos transtornos neurais têm sido preocupações centrais na neurociência há muito tempo. Recentemente, isso impulsionou uma era de iniciativas em larga escala e colaborações para mapear as redes cerebrais de maneira mais abrangente e detalhada. Tais análises de redes tornaram-se uma ferramenta importante para caracterizar diferenças individuais relacionadas à função cognitiva, idade e saúde mental [Fornito, Zalesky e Bullmore 2016, Riedl et al. 2015, Wijk, Stam e Daffertshofer 2010, Weissenbacher et al. 2009, Geerligs, Cam-CAN e Henson 2016].

As redes cerebrais são bem adaptadas para integrar funcionalmente os processos neuronais locais que sustentam a cognição, possuindo uma arquitetura que engloba múltiplas escalas espaciais, desde a microescala de células individuais e sinapses à macroescala de sistemas cognitivos inteiros. Essa arquitetura também é encontrada em outras redes complexas, por exemplo, na organização multiescala de redes sociais, que vão desde relações interpessoais e pequenos grupos sociais, até grandes comunidades e organizações políticas globais. Em sistemas multiescala, os níveis são interdependentes, com padrões em cada nível dependendo de processos em níveis inferiores e superiores. As descrições em microescala das sinapses são essenciais para entender a base da arquitetura somática e axodendrítica. A arquitetura somática é o design arquitetônico moldado pela compreensão e incorporação das respostas somáticas dos indivíduos, explorando como as experiências corporais e sensoriais podem influenciar e ser influenciadas pelo ambiente vivido [Anderwald, Grond e Pérez 2021, Gálvez 2019]. Por outro lado, a arquitetura axodendrítica está relacionada à estrutura e função das conexões neurais no cérebro, referindo-se à estrutura e organização das conexões entre axônios e dendritos em neurônios [Kandel et al. 2013]. Em contraste, as macroescalas descrevem regiões ou áreas do cérebro e suas vias de fibras associadas.

Esta escala é especialmente útil para correlacionar a anatomia com comportamentos e funções específicas e pode ser abordada com as técnicas atuais de neuroimagem.

Para entender o funcionamento do cérebro, precisamos estudá-lo em múltiplas escalas, identificando redes que conectam células em populações, organizando grupos de células em regiões funcionais, integrando regiões em sistemas e conectando o cérebro ao corpo em um organismo completo. Nessa organização, nenhum nível é privilegiado sobre os demais. Somente por meio de interações de rede multiescala moléculas e células podem dar origem a comportamento e a cognição. O conhecimento sobre interações de rede através de múltiplos níveis de organização é crucial para uma compreensão mais completa do cérebro como um sistema integrado [Sporns 2010, Silva 2022, Makris et al. 2023].

A motivação para estudar o cérebro humano como uma rede complexa é que essa abordagem pode fornecer informações fundamentais sobre como seus elementos se organizam em padrões dinâmicos, ampliando o entendimento em comparação ao estudo de elementos individuais de forma isolada [Sporns 2010].

As relações entre estrutura e função são um princípio fundamental em muitos sistemas complexos que ocorrem naturalmente. No entanto, a pesquisa em neurociência de redes sugere a existência de uma ligação imperfeita entre a conectividade estrutural e a conectividade funcional no cérebro [Suárez et al. 2020]. O funcionamento complexo dessas redes continua sendo um grande desafio para a neurociência. Portanto, é essencial procurar compreender a natureza das redes cerebrais, realizando um mapeamento detalhado da relação estrutura-função, para que possamos investigar o impacto dos danos cerebrais na cognição e avaliar os efeitos das intervenções [Silva 2022].

A interdependência de regiões cerebrais é descrita pela conectividade funcional [Cordes et al. 2000]. O estudo da conectividade cerebral já abriu novos caminhos experimentais e teóricos em várias áreas da neurociência, atuando na neuroanatomia, neurodesenvolvimento, eletrofisiologia, na imagem cerebral funcional e na base neural da cognição. Recentemente, em muitas doenças cognitivas e afetivas foram observados padrões anatômicos ou funcionais de conectividade alterados, definindo as síndromes de desconectividade. Alguns trabalhos mostram que certos tipos de traumas neurais causam alterações generalizadas na conectividade funcional, em comparação a indivíduos saudáveis [Derks et al. 2017], pois se descobriu os padrões de conectividade funcional se adaptam ao comportamento ou a cognição, e são interrompidos em pacientes que sofrem de transtornos psiquiátricos específicos [Sporns 2010, Silva 2022, Riedl et al. 2015, Derks et al. 2017].

A patologia funcional do cérebro pode ocorrer e se espalhar pela rede, de acordo com mecanismos relacionados à ela, particularmente fazendo uso dos *hubs*. Os *hubs* são geralmente caracterizados por uma alta taxa de transferência de informações, refletida no cérebro por alta conectividade ou centralidade da região. Eles são cruciais para o funcionamento cerebral e são intensamente utilizados em um cérebro saudável. Quando

ocorre uma doença neurológica, as regiões de *hub* serão eventualmente implicadas, porque todas as rotas de comunicação passam por essas áreas [Derks et al. 2017].

A análise da estrutura organizacional e conectividade das redes cerebrais esclarece bem diversos problemas relacionados à função integrativa do cérebro. Os sistemas nervosos, compostos por elementos neurais interconectados, são moldados por princípios arquitetônicos explorados pela ciência das redes. Neurônios individuais geram respostas fisiológicas complexas através de interações em rede. Características sensoriais distintas são integradas em estados perceptuais e cognitivos por processos distribuídos no cérebro. Em repouso cognitivo, o cérebro exibe atividade neural dinâmica, moldada por redes estruturais. Mudanças sensoriais ou cognitivas ativam padrões específicos de rede. Traumas e doenças causam *déficits* neurológicos, com danos à rede informando sobre a disfunção e recuperação. Variações individuais em cognição são ligadas a redes cerebrais, e o desenvolvimento modifica conexões anatômicas e capacidades cognitivas. Cérebro e corpo são acoplados dinamicamente, com redes cerebrais modulando a dinâmica interna através de movimentos corporais.

Essas e outras questões fundamentais da neurociência podem ser abordadas de forma eficaz pela perspectiva das redes complexas [Sporns 2010]. A relação entre estrutura e função é um conceito fundamental para o entendimento das propriedades de sistemas de diversas naturezas, assim como, interrupções estruturais se relacionam com perdas funcionais [Suárez et al. 2020].

O mapeamento das redes cerebrais é realizado por identificação de flutuações de sinal síncronos, ou conectividade funcional, entre regiões cerebrais quando há atividade neuronal. As três principais técnicas para a identificação das estrutura de redes cerebrais, obtidas por vias de imagens por ressonância magnética (MRI) em seres humanos, são: a MRI ponderada por difusão, que pode ser usada para estimar a integridade dos tratos de substância branca entre regiões de interesse; a imagem por ressonância magnética funcional (fMRI), que mede a conectividade dentro de um indivíduo, quantificando correlações entre séries temporais de atividade neuronal entre diferentes regiões temporais; e a MRI estrutural, a partir da qual a covariância entre regiões de interesse de uma propriedade do tecido, como volume ou espessura da substância cinzenta, pode ser examinada entre indivíduos, o que pode refletir mudanças maturacionais sincronizadas em regiões do cérebro anatomicamente conectadas [Riedl et al. 2015, Geerligs, Cam-CAN e Henson 2016, Santos et al. 2019].

Ao longo deste trabalho, utilizamos dados obtidos pela técnica de fMRI de indivíduos em estado de repouso para extrair os dados de conectividade cerebral a serem analisados. Alguns estudos mostram que o córtex pré-frontal ventromedial e o córtex cingulado posterior exibem paradoxalmente altos níveis de atividade durante o estado de repouso e reduções de atividade durante tarefas cognitivas voltadas para o meio externo [Esposito et al. 2006]. Técnicas de mapeamento cerebral, como a fMRI, permitem o estudo da atividade

neural espontânea e das interações entre redes cerebrais em repouso, já que essas redes exibem maior atividade metabólica em repouso do que durante a realização de tarefas cognitivas voltadas para o exterior [Uddin et al. 2009]. Em indivíduos que estão realizando tarefas específicas, a fMRI pode mapear áreas onde o fluxo sanguíneo cerebral regional aumenta de maneira sincronizada com a atividade realizada [Cordes et al. 2000].

No entanto, é importante notar que esse mapeamento apresenta algumas limitações: embora a conectividade funcional indique correlações entre sinais neuronais, apenas a conectividade efetiva descreve a influência exercida por um sistema neural sobre outro [Riedl et al. 2015].

1.2 Conectoma

O comportamento de um ser humano depende da atividade coordenada de uma enorme quantidade de neurônios interconectados. Determinar a conectividade desses neurônios, por meio de métodos anatômicos e eletrofisiológicos combinados, sempre foi um grande objetivo da neurociência. As análises experimentais para a cognição humana foram significativamente aprimoradas com a chegada da neuroimagem funcional, um conjunto de técnicas que pode ser aplicado para estudar uma ampla gama de funções cognitivas, com resolução espacial e temporal cada vez maior. A interpretação dos dados de neuroimagem é limitada devido à falta de informações sobre as redes que geram os padrões de ativação observados.

Um quadro teórico para entender a cognição como um fenômeno de rede baseia-se em dois princípios do córtex cerebral: segregação funcional e integração funcional. Segregação funcional refere-se à organização cerebral em múltiplos níveis, indicando que diferentes regiões do cérebro são especializadas em realizar funções específicas. Já a integração funcional refere-se à capacidade do cérebro de combinar e coordenar a atividade dessas diferentes regiões especializadas, assim gerando comportamentos integrados. As teorias de rede da cognição destacam a natureza distribuída e dinâmica das relações entre estrutura e função no cérebro. Para essas análises, é necessário um mapeamento detalhado das conexões entre neurônios por meio de um modelo estrutural abrangente [Bargmann e Marder 2013, Sporns, Tononi e Kötter 2005, Tononi, Sporns e Edelman 1994, Sporns 2013].

Para entender o funcionamento de uma rede, é necessário conhecer seus elementos e suas interconexões, com essa intenção foi proposto por [Sporns, Tononi e Kötter 2005] estratégias de pesquisa voltadas para uma descrição estrutural abrangente da rede de elementos e conexões que formam o cérebro. Esse conjunto de dados, denominado "conectoma", sendo o objeto de estudo de uma nova ciência chamada "conectômia", é considerado fundamental para a neurociência cognitiva e a neuropsicologia. As perspectivas apontam que a conectômia pode promover um aumento significativo sobre nossa compreensão de

como os estados funcionais do cérebro emergem de sua estrutura anatômica e fornecerá novos *insights* sobre os mecanismos de como a função cerebral é afetada quando detalhes dessa estrutura são interrompidos. Além disso, oferece recursos de neuroinformática unificada, invariável ao longo do tempo e prontamente disponível, que poderá ser utilizado em praticamente todas as áreas da neurociência experimental e teórica [Sporns, Tononi e Kötter 2005, HCP 2024].

Muitos novos projetos promovidos por instituições ao redor do mundo estão pesquisando aspectos de como o desenvolvimento cronológico, doenças neurais e outros fatores podem afetar as conexões em constante mudança no cérebro humano. Todas essas iniciativas contribuem para construção de conectomas cada vez mais precisos, capazes de fornecer cada vez mais informações para todas áreas da neurociência. Nesse contexto, a *Connectome Coordination Facility* (CCF) foi criada para ajudar a coordenar esses inúmeros projetos de pesquisa, disponibilizando dados, e assim facilitando a distribuição e disseminação de resultados [HCP 2024].

O objetivo central da conectômia é mapear de forma abrangente as conexões entre elementos neurais anatomicamente distribuídos. Esses elementos podem ser neurônios individuais, populações neuronais específicas ou regiões cerebrais em larga escala. O número de conexões possíveis entre esses elementos é grande, para uma rede de N nós, o número de conexões possíveis está na ordem de N^2 . Portanto, é necessário um método para representar essa grande quantidade de dados de forma sucinta e significativa.

Atualmente o estudo sobre redes cerebrais utiliza fortemente o conceito de conectoma, para definir uma matriz que representa todas as conexões anatômicas possíveis entre pares de elementos neurais do cérebro Figura 1. Esse conceito inclui a matriz de conexões anatômicas entre grandes áreas cerebrais e entre neurônios individuais, bem como a matriz de interações funcionais, revelada pela análise de processos fisiológicos. Esses processos variam desde as lentas flutuações na oxigenação do sangue cerebral, medidas por fMRI, até as rápidas oscilações neuronais de alta frequência, detectáveis por eletrofisiologia. A quantificação, visualização e compreensão da organização da rede cerebral em diversas escalas de espaço e tempo são características fundamentais do campo da conectômia [Fornito, Zalesky e Bullmore 2016].

É comum representar a conectividade entre cada par de regiões cerebrais em uma rede como uma matriz bidimensional. Nesta matriz, cada linha e coluna corresponde a uma região diferente, e o elemento da matriz posicionado na interseção da linha i e coluna j codifica informações sobre a conexão entre as regiões i e j .

O conectoma consiste em dois principais aspectos de descrição para a definir sua arquitetura de rede: elementos e conexões neurais, que no contexto de matrizes podem ser interpretados como nós e arestas, respectivamente. Um nó geralmente é identificado com uma região do cérebro, frequentemente selecionada de maneira fundamentada no

Figura 1 – (a) Uma matriz de conectividade para um cérebro representando a conectividade de 32 áreas corticais. (b) Esta matriz está organizada de modo que os elementos coloridos representem uma projeção da região listada na coluna para a região listada na linha. O tamanho dos pontos em cada elemento da matriz é proporcional à distância de projeção e as cores mais escuras indicam uma intensidade média de conectividade relatada mais forte. (c) As localizações anatômicas das áreas listadas na matriz em (b), cores mais escuras identificam regiões com maior conectividade total com o restante da rede.

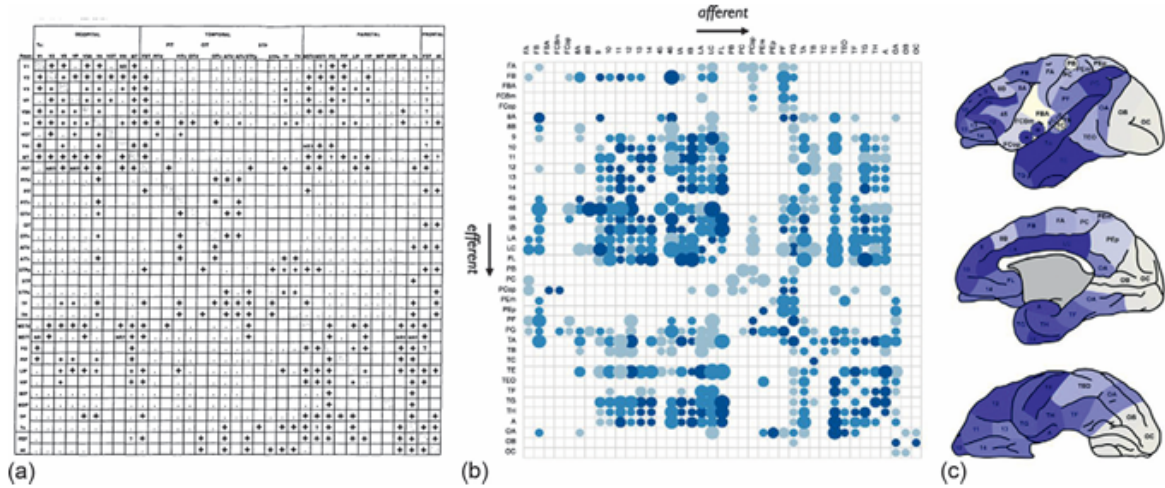


Imagem adaptada de [Fornito, Zalesky e Bullmore 2016]

conhecimento da anatomia e função cerebral, e pode variar em tamanho e especificidade desde a escala de um único neurônio microscópico até uma região cerebral macroscópica. Uma aresta representa uma conexão entre dois nós. Em uma rede estrutural, uma aresta representa uma conexão anatômica entre dois elementos cerebrais. Em uma rede funcional, uma aresta representa a associação estatística entre atividades registradas de elementos cerebrais separados. O conjunto de todos os N elementos constitui as colunas e linhas de uma matriz de conexão $N \times N$, cujas entradas a_{ij} representam conexões entre os elementos individuais i e j . A ausência confirmada de uma conexão é denotada por $a_{ij} = 0$, enquanto a presença confirmada de uma conexão, independentemente de sua intensidade ou características fisiológicas, resulta em $a_{ij} = 1$. Uma vez que uma conexão é confirmada, seu elemento de matriz não zero recebe entradas de dados adicionais catalogando uma variedade de parâmetros estruturais e fisiológicos, como direção da fibra, densidade da conexão, força, sinal (excitador/inibidor), atraso de condução, potenciais efeitos moduladores (dependência de voltagem), entre outros parâmetros de interesse. A união da matriz de conexão binária e dos dados fisiológicos específicos da conexão resulta em uma descrição estrutural que combina topologia de conexão e biofísica [Sporns, Tononi e Kötter 2005, Kopell et al. 2014, Fornito, Zalesky e Bullmore 2016].

A compilação do conectoma enfrenta dois desafios consideráveis não presentes em outras redes naturais ou tecnológicas. Primeiro, o cérebro humano é um órgão altamente complexo, com um grande número de componentes estruturalmente distintos, heterogêneos,

porém interconectados. Um segundo desafio é que os elementos estruturais básicos do cérebro humano, em termos de nós e conexões de rede, são difíceis de serem definidos. Diferentes tipos de descrições estruturais podem abordar pelo menos três níveis bastante distintos de organização. Nos extremos da escala estão o nível de neurônios individuais e sinapses (microescala) e o nível de regiões cerebrais anatomicamente distintas e vias inter-regionais (macroescala). Entre esses dois níveis está o nível de grupos ou populações neuronais (mesoscala) [Sporns, Tononi e Kötter 2005].

O surgimento relativamente recente do conectoma não deve ser entendido como uma falta de interesse anterior pelas redes cerebrais, pois muitos neurocientistas dos séculos XIX e XX como Ramón y Cajal, Golgi, Meynert, Wernicke, Flechsig, Brodmann, Meynert, Betz e Bevan-Lewis [Zilles 2018, Triarhou 2020, López-Muñoz, Boya e Alamo 2006], já tinham conhecimento de sua relevância para a compreensão dos sistemas nervosos. Estes e outros neurocientistas fundamentais fizeram descobertas expressivas e escreveram trabalhos conceituais duradouros [Cajal 1889, Brodmann 1909], que desde então têm sustentado a forma como pensamos sobre os sistemas nervosos.

Alguns fatores que impulsionaram a ascendência científica do conectoma foram o rápido crescimento na ciência das redes em geral. Desde a década de 1980, houve grandes desenvolvimentos conceituais na física estatística de redes complexas e aplicações cada vez mais amplas da ciência de redes para a análise e modelagem de *big data*. Foram encontradas novas maneiras de quantificar a complexidade topológica de grandes sistemas de agentes em interação, e foram observadas semelhanças impressionantes nas propriedades organizacionais de uma ampla gama de redes de mundo real.

Outro fator foi a evolução tecnológica dos métodos para medir e visualizar a organização do cérebro, em múltiplas escalas de resolução. Desde a década de 1990, houve progressos intensos na ciência da neuroimagem humana, especialmente usando a ressonância magnética para mapear redes anatômicas e funcionais do cérebro inteiro em escala macroscópica em voluntários saudáveis e pacientes com distúrbios neurológicos e psiquiátricos. Nos últimos 10 anos, também houve desenvolvimentos metodológicos espetaculares no rastreamento do trato, microscopia óptica, optogenética, registro multieletrodo, expressão gênica histológica e muitas outras técnicas de neurociência que agora podem ser usadas para mapear sistemas cerebrais em nível mesoscópico e escalas microscópicas, sob condições experimentais mais controladas e em uma gama mais ampla de espécies. A convergência destas duas tendências poderosas, os desenvolvimentos matemáticos e conceituais na ciência de redes complexas e a evolução das tecnologias para medir os sistemas nervosos motivaram o novo campo do conectoma [Fornito, Zalesky e Bullmore 2016].

Assim, conectoma se mostra como o substrato estrutural para compreender as funções cognitivas e interpretar estudos de neuroimagem, focando na descrição estrutural do cérebro [Sporns, Tononi e Kötter 2005]. No entanto, as conexões funcionais por si só

não fornecem informações cruciais sobre como o cérebro gera a cognição. É fundamental compreender não apenas o que está conectado, mas também como e em que direções as regiões cerebrais estão interligadas: que tipos de sinais elas transmitem e como esses sinais são processados como parte de uma operação computacional neural. Essa compreensão é essencial para entender como diferentes áreas cerebrais combinam suas operações individuais para sustentar a função cognitiva [Kopell et al. 2014].

As pesquisas sobre redes cerebrais são fundamentais para entender a complexidade do cérebro humano, um campo que ainda apresenta muitas questões em aberto. A compreensão desse assunto pode gerar novos conhecimentos tanto na área da saúde quanto no estudo de redes em geral. Utilizar o conectoma facilita a aplicação de métodos matemáticos e estatísticos, proporcionando uma compreensão mais profunda das dinâmicas cerebrais e abrindo novas possibilidades para tratamentos de distúrbios neurológicos. Dessa forma, esse é o nosso objeto de estudo.

Nosso objetivo no estudo de redes cerebrais é analisar um conjunto de dados de conectividade funcional neural, verificando as distribuições estatísticas de seu espectro de autovalores. Além disso, propomos modelos para descrever esses dados e avaliamos a qualidade desses modelos.

Nos próximos capítulos, abordaremos os conceitos teóricos aplicados ao estudo dessa rede. No Capítulo 2, discutimos os fundamentos teóricos. No Capítulo 3, apresentamos os procedimentos metodológicos utilizados ao conjunto de dados estudados. Os resultados das análises dos dados são apresentados e discutidos no Capítulo 4. Por fim, as conclusões deste trabalho são abordadas no Capítulo 5.

2 Fundamentos Teóricos

2.1 Sistemas Complexos

Existem muitos sistemas de interesse para os cientistas que são compostos de partes ou componentes individuais ligados de alguma forma. Como exemplos temos, a Internet, como um conjunto de computadores conectados por ligações de dados, as sociedades humanas, como grupos de pessoas conectadas por opiniões ou interesses comuns, interações sociais, entre outros tipos de sistemas. Muitos aspectos desses sistemas são dignos de estudos mais aprofundados tanto sobre a natureza dos elementos individuais e seus comportamentos em relação aos demais, como sobre o funcionamento, as dinâmicas de conexões e as tendências do sistema em si [Newman 2010].

Os sistemas complexos têm propriedades que não são totalmente aleatórias nem completamente regulares, mas demonstram características não triviais que são indicativas de uma organização mais elaborada ou complexa. O estudo de sistemas complexos ou redes complexas abrange diversos campos científicos, desde as ciências sociais até as ciências naturais. A ciência está cada vez mais focada em entender a estrutura, o comportamento e a evolução de sistemas complexos, como células, cérebros, ecossistemas, sociedades e economia. Para compreender esses sistemas, precisamos além do conhecimento de seus componentes elementares, também é necessário entender como esses componentes interagem e das propriedades particulares de suas interações. Todos estes são grandes sistemas, muitas vezes compostos por milhões de agentes interagindo entre si, e são representados por tipos de dados muito diversos. Nesse contexto, a física estatística tem sido bem-sucedida como uma poderosa ferramenta para modelar redes complexas.

A crescente disponibilidade de grandes conjuntos de dados e computadores cada vez mais poderosos torna mais fácil, analisar e modelar o comportamento de sistemas compostos de uma enorme quantidade de elementos interagentes. Assim, a ciência das redes explodiu na era da informação, graças à produção e armazenamento sem precedentes de dados sobre quase todas as atividades humanas. Isso ocorre porque as redes são uma maneira simples e eficaz de modelar uma grande variedade de sistemas tecnológicos, sociais, econômicos e biológicos que podem ser descritos como um conjunto de entidades (nós) com interações entre elas (*links*). Essas interações representam os graus de liberdade fundamentais da rede e podem ser de diferentes tipos, não direcionadas ou direcionadas, binárias ou ponderadas, dependendo da natureza do sistema e da resolução usada para descrevê-lo.

Esses sistemas complexos demonstram uma variedade de padrões característicos

devido ao acoplamento altamente estruturado e seletivo entre seus elementos, alcançados pela complexidade detalhada das conexões da rede, podendo ser descritos como redes com topologia complexa. A conectividade vem em muitas formas, por exemplo, interações moleculares, vias metabólicas, conexões sinápticas, associações semânticas, teias alimentares ecológicas, redes sociais, *hyperlinks* da *web* ou padrões de citação. Em todos os casos, a análise quantitativa da conectividade requer técnicas matemáticas e estatísticas sofisticadas [Fornito, Zalesky e Bullmore 2016, Sporns 2010, Barabási e Albert 1999, Cimini et al. 2019].

O conceito de rede tem ganhado reconhecimento como uma ferramenta fundamental para entender o comportamento dinâmico e as respostas de sistemas reais, um campo onde a teoria dos sistemas complexos encontra terreno fértil para suas aplicações [Aguiar e Bar-Yam 2005, Bandyopadhyay e Jalan 2007]. Inicialmente, redes de topologia complexa foram descritas pela teoria dos grafos aleatórios de Erdős e Rényi [Erdős e Rényi 1960]. No entanto, na época, devido à ausência de dados sobre grandes redes, as previsões dessa teoria foram pouco aplicadas a problemas do mundo real. À medida que foram desenvolvidos métodos para trabalhar com grandes conjuntos de dados e esses métodos foram amplamente aplicados em diferentes domínios, tornou-se claro que sistemas superficialmente diferentes podem expressar propriedades notavelmente gerais em termos de sua organização em rede [Barabási e Albert 1999, Fornito, Zalesky e Bullmore 2016, Caldas 2007, Albert e Barabási 2002, Newman 2003].

Esses progressos levaram ao surgimento de um campo interdisciplinar de ciência de redes, que utiliza métodos analíticos para modelar redes complexas e investigar princípios comuns ou quase universais de organização, função, crescimento e evolução das redes. A teoria dos grafos é o método principal entre esses métodos gerais [Fornito, Zalesky e Bullmore 2016].

O padrão de conexões em um determinado sistema pode ser representado como uma rede ou grafo, sendo uma estrutura composta por um conjunto de elementos, os vértices (nós) interligados por suas conexões (arestas). As arestas representam as linhas que conectam dois vértices, e podem ser direcionadas, quando as conexões entre os pares de vértices têm direção e sentido, possuindo diferentes valores para os pares de uma conexão, ou não-direcionada quando as conexões entre pares é a mesma para todos os pares de conexão, resultando em conexões sem direções.

O padrão particular de interações de tais redes pode ser um efeito do comportamento do sistema. Por exemplo, esse padrão de conexões entre computadores na Internet, interfere nas rotas que os dados percorrem na rede e a eficiência com que a rede transporta a informação. As conexões em uma rede social afetam como as pessoas absorvem, formam opiniões e obtêm notícias, assim como influenciam outros fenômenos menos óbvios, como a disseminação de doenças. A menos que saibamos algo sobre a estrutura dessas redes, não podemos esperar entender completamente como os sistemas correspondentes funcionam.

Uma rede é uma representação simplificada que transforma um sistema em uma estrutura abstrata, focando apenas nos padrões de conexão essenciais. Vértices e arestas na rede podem ser anotados com informações extras, como nomes ou intensidades, para adicionar mais detalhes. No entanto, muita informação é geralmente perdida ao reduzir um sistema completo a uma rede. Isso tem suas desvantagens, mas também oferece vantagens.

Dentre as vantagens da abordagem de redes complexas, podemos citar a capacidade de modelar sistemas de elementos interconectados de maneira não trivial, permitindo um estudo aprofundado de suas interações. Essa abordagem facilita a análise de estruturas, a identificação de pontos críticos e a detecção de padrões dentro de um sistema, o que pode indicar previsibilidades e pontos de intervenção. Além disso, é possível trabalhar com sistemas muito grandes, compostos por milhares ou até milhões de nós, e essa abordagem é aplicável a diversos tipos de sistemas, como físicos, biológicos, sociais e econômicos, entre outros.

No entanto, algumas desvantagens incluem o elevado custo computacional na análise ou simulação de redes muito grandes ou complexas, a dificuldade na interpretação de certos resultados e na tradução desses resultados em ações práticas para o sistema. Além disso, a abstração necessária para simplificar o sistema a uma rede de interações pode resultar na perda de detalhes importantes. Por fim, há desafios na validação de modelos que consigam descrever os dados reais com alta precisão [Newman 2010, Caldas 2007, Barabási 2002, Strogatz 1998, Barabási e Albert 1999, Albert e Barabási 2002].

A maioria das redes complexas não são conectadas aleatoriamente, elas seguem um comportamento de escala livre [Albert e Barabási 2002]. Em redes aleatórias livre de escala, a probabilidade de um nó ter k conexões, $P(k)$, segue uma lei de potência, uma propriedade que pode ser construída pela preferência por conexões aos nós mais conectados de novos nós adicionados à rede, onde novos nós são mais propensos a se conectar aos já altamente conectados. As propriedades dessas redes são frequentemente caracterizadas pela presença de alguns nós altamente conectados, os *hubs*, enquanto a maioria dos nós restantes têm poucas conexões.

Embora as redes livre de escala descrevam muitas propriedades estatísticas de redes biológicas, como as redes cerebrais, elas falham em considerar um aspecto importante: a modularidade exibida pela maioria dos sistemas complexos. A modularidade assume que a rede de interações pode ser dividida em sub-redes ou módulos. Cada módulo é composto por vários elementos que são mais interconectados entre si do que com o resto da rede. Sistemas modulares podem ser organizados em uma hierarquia estrutural, com múltiplos níveis de decomposição modular. Moléculas, organelas, células, tecidos, órgãos e organismos, famílias, comunidades, são exemplos dessa hierarquia de estruturas. Uma propriedade bastante usada para caracterizar redes modulares é o coeficiente de agrupamento, que representa, o grau em que os elementos adjacentes de um nó estão conectados entre si,

que é maior do que o dos modelos genéricos livre de escala [Aguilar e Bar-Yam 2005].

2.1.1 Teoria dos grafos

A teoria dos grafos é a área da matemática que estuda redes. Um grafo é uma representação matemática de uma rede do mundo real, de algum sistema composto de elementos interconectados. Embora a teoria dos grafos pertença à matemática pura e não tenha aplicações imediatas, o uso do formalismo teórico dos grafos é frequentemente usado para resolver problemas do mundo real [Sporns 2010].

Uma das primeiras aplicações da teoria dos grafos para resolver um problema real foi feita pelo matemático Leonhard Euler (1707-1783), em 1736 para solucionar um quebra-cabeça popular envolvendo sete pontes sobre o rio Pregel na cidade de Königsberg, na Prússia Oriental (atual cidade de Kaliningrado na Rússia). Euler encontrou uma solução geral que poderia ser aplicada a um arranjo arbitrário de pontes e massas de terra, e foi capaz de perceber que o problema poderia ser resolvido levando em consideração apenas a posição relativa de pontes e massas de terra e que a posição geográfica precisa ou a distância física não eram relevantes. Com isso, Euler foi creditado como fundador do campo ao qual ele se referiu como "geometria da posição" (*geometria situs*) e que agora é conhecido como teoria dos grafos.

Ao longo da história a teoria dos grafos se mostrou muito eficiente na resolução, interpretação e tratamento de questões que aparentemente não haviam como serem resolvidas, devido a sua flexibilidade de representações, podendo ser aplicada em sistemas de diversas naturezas e especificidades. Alguns outros exemplos de aplicações são: O físico Gustav Kirchhoff (1824-1887) representou redes complexas como grafos e formulou vários teoremas originais que estabeleceram a base da teoria de circuitos na engenharia elétrica; O matemático Arthur Cayley (1821-1895), aplicou conceitos teóricos dos grafos ao problema de enumerar isômeros químicos, moléculas que contêm o mesmo número de átomos ligados de maneiras diferentes, essa abordagem permitiu a caracterização completa de vários grupos de hidrocarbonetos como famílias de grafos químicos [Sporns 2010, Fornito, Zalesky e Bullmore 2016].

A topologia de um grafo define como os links entre os elementos do sistema são organizados dentro de uma rede de conexões, sem nenhum outro aspecto físico sendo informado, de forma que qualquer outras análises topológicas serão invariantes sob transformações contínua do sistema [Fornito, Zalesky e Bullmore 2016]. A topologia tem por objetivo descrever propriedades globais de um sistema, enquanto a geometria geralmente está associada às características locais do sistema, ambas procuram reproduzir de forma simplificada situações presentes no mundo real [Santos et al. 2019].

A topologia desenvolveu-se intensamente como um campo da matemática a partir

do final do século XIX, precedendo notáveis avanços na análise estatística de gráficos no século XX. Na década de 1950, este trabalho foi liderado por Paul Erdős e Alfred Rényi, que introduziram um modelo estatístico capaz de gerar gráficos aleatórios e para prever algumas de suas propriedades topológicas [Erdős e Rényi 1959]. Em um grafo de Erdős-Rényi, existem N nós e uma probabilidade uniforme p de cada aresta possível entre eles. Se p estiver próximo de um, o gráfico está fortemente conectado e se p estiver próximo de zero, o gráfico está fracamente conectado. Erdős e Rényi mostraram que muitas propriedades importantes desses gráficos, como o número médio de conexões anexadas a qualquer nó, poderiam ser previstos analiticamente a partir de seu modelo generativo [Fornito, Zalesky e Bullmore 2016].

Um grafo simples compreende um conjunto de nós e um conjunto de arestas. Os nós representam os elementos fundamentais do sistema, as arestas representam conexões entre pares de nós. As arestas podem ser não direcionadas ou direcionadas da origem ao destino [Sporns 2010]. Podemos ter 4 tipos de grafos para representar os diferentes tipos de redes complexas reais, formadas pelas combinação de duas propriedades:

- Redes direcionadas ou não direcionadas.
- Redes ponderadas ou binárias.

Portanto, temos os 4 tipos de redes, esquematizadas na Figura 2. Na representação gráfica, as setas representam a direção da conectividade, de modo que a ponta da seta aponta para o alvo de cada ligação. Nesses casos, na ponta da seta temos um valor não nulo e em sua base um valor igual à zero.

Figura 2 – (a) Rede direcionada, ponderada; (b) Rede direcionada, binária; (c) Rede não direcionada, ponderada (d); Rede não direcionada, binária.

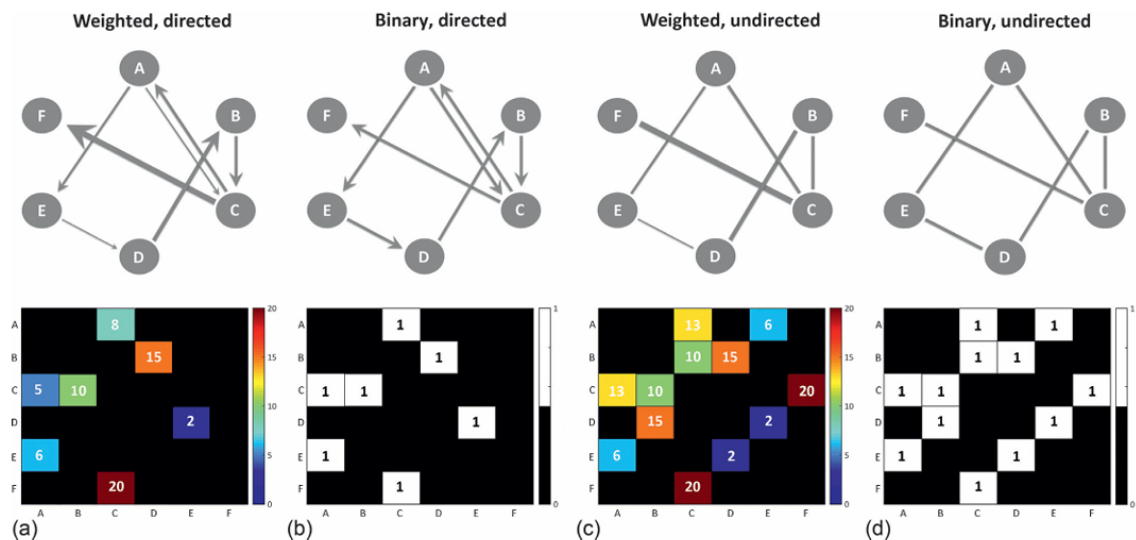


Imagem adaptada de [Fornito, Zalesky e Bullmore 2016]

Nem todas as conexões entre elementos neurais são iguais. Alguns pares de elementos neurais podem formar mais sinapses entre si, por meio de comunicações de feixes axonais mais densos ou terem mais fibras mielinizadas que facilitam a transmissão de sinais de forma mais rápida. Essas variações podem ser descritas por diferenças no peso da conectividade. A maioria dos métodos para medir a conectividade cerebral fornece algum índice de peso de conectividade, que para qualquer par de nós i e j , é denotado como w_{ij} . Tal informação pode ser usada em uma matriz de conectividade cerebral atribuindo os elementos da matriz com esses pesos de conexão, de forma que $C_{ij} = w_{ij}$, onde C_{ij} são os elementos da matriz. O resultado é um grafo ou rede ponderada. Nas representações gráficas, é comum representar diferenças no peso da conectividade entre pares de nós como variações na espessura das arestas, essa situação também pode ser vista na Figura 2. Nas representações matriciais, as variações de peso são frequentemente representadas usando uma escala de cores. O peso em grafos cerebrais pode ser interpretado como a intensidade de conectividade, quando maior o valor do peso mais forte deve ser a conexão entre um determinado par de nós.

Em grafos de redes não ponderadas, os pesos de conectividade podem mostrar uma ampla variação, dependendo do método usado para estimar a conectividade entre regiões cerebrais. Em muitos modelos de grafos os valores dos pesos em si não são de grande interesse, o que nos permite indexar a conectividade de forma binária admitindo apenas dois valores possíveis para w_{ij} , 1 ou 0, de modo que $w_{ij} = 0$ indica ausência de conectividade, entre os nós, e para $w_{ij} = 1$ indicando que os nós estão conectados. O processo de conversão de uma rede ponderada para uma rede não ponderada ocorre a partir da inserção de um limite superior para os valores de w_{ij} . Esse processo tem por objetivo simplificar e facilitar a interpretação da rede de forma geral. Redes binárias nos dizem onde as conexões estão na rede, mas não fornecem informações sobre variações no peso da conectividade entre diferentes nós da rede. Dado que os pesos das conexões nas redes cerebrais podem variar em várias ordens de magnitude, uma vez que todas as análises de redes cerebrais deveriam ser ponderadas. No entanto, percepções valiosas sobre a organização da rede podem muitas vezes ser obtidas considerando a topologia binária de uma matriz de conectividade. A maioria das medidas da teoria dos grafos usadas na análise de dados conectômicos foi desenvolvida para redes binárias.

Os grafos aleatórios gerados pelo modelo de Erdős-Rényi são exemplos da classe mais simples de grafos: grafos binários não direcionados. Eles são gráficos binários porque as arestas estão ausentes ou presentes ou equivalentemente, o peso da aresta é zero ou um. Eles são gráficos não direcionados porque as arestas conectam os nós simetricamente, não existe nenhuma diferença entre a origem e o destino de uma conexão. Desde então, os princípios da análise topológica foram estendidos a grafos mais sofisticados que incluem conectividade ponderada e direcionada, essas informações adicionais são particularmente significativas para caracterizar certos tipos de dados de redes cerebrais [Fornito, Zalesky e

Bullmore 2016].

No modelo de grafos aleatórios de Erdős e Rényi, quaisquer dois nós estão conectados aleatoriamente com uma probabilidade p . Este modelo assume que as interações entre os nós são aleatórias. Recentemente, com a disponibilidade de grandes mapas de redes do mundo real, verificou-se que o modelo de grafos aleatórios não é apropriado para estudar o comportamento dessas redes. Assim, muitos novos modelos foram introduzidos. Watts e Strogatz [Strogatz 1998] propuseram um modelo, popularmente conhecido como "rede de mundo pequeno", que possui propriedades de pequeno diâmetro e alta clusterização. Além disso, essa modelagem é muito esparsa: uma rede com um número muito reduzido de arestas, outra característica apresentada por muitas redes do mundo real.

O modelo livre de escala de Barabási-Albert e o modelo mundo pequeno de Watts-Strogatz contribuíram imensamente para a compreensão da evolução e do comportamento dos sistemas reais com estruturas de rede. Seguindo esses dois novos modelos, houve uma explosão de estudos na área de redes complexas que descrevem problemas de mundo real. Esses estudos mostram que as redes do mundo real possuem coexistência de aleatoriedade e regularidade.

Muitos estudos têm mostrado que as propriedades das redes ou grafos podem ser bem caracterizadas pelo espectro da matriz de adjacência associada A . Para um grafo binarizado, é definida da seguinte maneira: $A_{ij} = 1$, se os nós i e j estão conectados, e zero caso contrário. Em uma rede não direcionada, essa matriz é simétrica e, conseqüentemente, possui autovalores reais. Os autovalores fornecem informações sobre algumas propriedades topológicas básicas da rede subjacente. A distribuição dos autovalores de uma matriz com uma probabilidade finita p segue a lei semicircular de Wigner no limite $p \rightarrow 1$. Para valores muito pequenos de p que correspondem à matriz aleatória esparsa, também se obtém uma lei semicircular com vários picos em diferentes autovalores [Jalan e Bandyopadhyay 2007].

2.1.2 Grafos Cerebrais

A teoria dos grafos tem desempenhado um papel fundamental para a compreensão da estrutura e função de sistemas complexos. Os sistemas nervosos são sem dúvida complexos, assim a teoria dos grafos também pode ser útil para a neurociência. É importante ressaltar que representações de redes cerebrais baseadas em gráficos (grafos cerebrais), podem ser facilmente construídas a partir de matrizes de conectividade cerebrais, como mostra a Figura 1. Cada linha ou coluna que representa uma região diferente do cérebro na matriz é desenhada como um nó no gráfico, e os valores em cada elemento da matriz são desenhados como arestas. As representações matriciais e gráficas de uma rede são formalmente equivalentes, e grande parte da matemática da teoria dos grafos é aplicada através da análise matriciais. A teoria dos grafos foi aplicada para compreender as redes cerebrais e ela emergiu como uma ferramenta analítica poderosa para a conectomia [Fornito,

Zalesky e Bullmore 2016, Wijk, Stam e Daffertshofer 2010].

A teoria do neurônio, estabelecida pelos brilhantes estudos microscópicos e pensamentos teóricos de Santiago Ramón y Cajal (1852-1934), considerado o pai da neurociência moderna, no final do século XIX e início do século XX, preparou o cenário para que a teoria dos grafos mais tarde fizesse sentido como um modelo para sistemas nervosos. Ramón y Cajal e alguns outros cientistas, usando a então revolucionária técnica de impregnação de prata para visualizar os complexos processos de ramificação de neurônios individuais, afirmaram que os neurônios eram células discretas que entravam em contato entre si muito estreitamente por meio de junções sinápticas. Este modelo contradiz o principal paradigma alternativo, a teoria reticular defendida por Camille Golgi (1843—1926), que inventou o método de coloração neuronal utilizado por Ramón y Cajal. Em 1906, ambos compartilharam o Prêmio Nobel de medicina, mas só na década de 1950 é que a microscopia eletrônica finalmente conseguiu resolver a questão teórica a favor de Ramón y Cajal.

O modelo de Ramón y Cajal de neurônios discretos interconectados por sinapses é naturalmente adequado para uma representação teórica em grafos, em que os neurônios são representados por nós e as projeções axonais ou junções sinápticas são representadas por arestas. Essa aplicação só foi possível devido aos esforços de Ramón y Cajal, possibilitando a análise teórica dos grafos de sistemas neurais, que surgiram formalmente cerca de 100 anos depois [Fornito, Zalesky e Bullmore 2016].

Todos os quatro tipos de gráficos binários ou ponderados e não direcionados ou direcionados são importantes para descrever redes cerebrais. A diferença entre gráficos não direcionados e direcionados é especialmente importante, pois a maioria das medidas de gráficos são definidas e computadas de forma ligeiramente diferente para essas duas classes principais de gráficos. Porém, a maior parte do trabalho clássico em teoria dos grafos foi realizada para grafos binários não direcionados [Sporns 2010].

As redes anatômicas do cérebro são intrinsecamente direcionadas, pois cada projeção axonal se origina no corpo celular e termina em uma ou mais sinapses. Mapear a direcionalidade da conectividade neuronal requer o uso de métodos invasivos, como rastreamento de tratos ou microscopia eletrônica. Métodos não invasivos adequados para uso em organismos vivos, como a ressonância magnética por difusão (diffusion MRI), atualmente não permitem resolver a direção de uma projeção axonal. As redes resultantes de tais técnicas são, portanto, não direcionadas. Nas análises de conectividade funcional do cérebro e conectividade efetiva, a direcionalidade pode ser medida ou modelada usando métodos específicos. Em geral, no entanto, a maioria das análises usa o coeficiente de correlação ou medidas relacionadas no domínio do tempo ou no domínio da frequência para quantificar a conectividade funcional, resultando em uma rede não direcionada. A direcionalidade da conectividade cerebral proporciona uma caracterização mais rica da organização da rede do que a análise de redes não direcionadas. No entanto, os gráficos

não direcionados ainda podem oferecer informações importantes sobre a organização dos sistemas nervosos [Fornito, Zalesky e Bullmore 2016].

Todas as abordagens de grafos desenvolvidas neste trabalho são de redes binárias e não direcionadas. Embora aplicações mais especializadas da teoria dos grafos permitam o tratamento matemático de grafos que incluem arestas múltiplas e negativas, esses métodos ainda não foram amplamente aplicados na neurociência.

A matriz de adjacência permite a obtenção de uma das medidas mais fundamentais do grafo, o grau. Em um grafo não direcionado, o grau de um nó é o número de arestas conectadas a esse nó. Em grafos direcionados, o grau de entrada e o grau de saída correspondem ao número de arestas de entrada e saída, respectivamente. Em grafos ponderados, a soma de todos os pesos de aresta de um nó fornece a força do nó, que é analisada de forma semelhante ao grau do nó. Os graus de todos os nós juntos formam a distribuição de graus da rede, que mostra se a rede contém nós com graus aproximadamente iguais ou se os graus dos nós variam em um intervalo mais amplo. Os graus dos nós são fundamentais porque têm um impacto significativo na maioria das outras medidas, e a distribuição de graus pode ser altamente explicativa sobre a arquitetura de rede do grafo. Outra medida simples baseada no grau é a assortatividade, definida como o coeficiente de correlação para os graus de nós vizinhos. A assortatividade positiva indica que as arestas tendem a ligar nós de grau semelhante, enquanto a assortatividade negativa indica que nós de alto grau se conectam preferencialmente com nós de baixo grau. Em redes cerebrais, o grau do nó e a força do nó podem ser vistos simplesmente como uma medida de interação direta: nós de alto grau de conectividade podem interagir diretamente com um grande número de outros nós. Um nó com alto grau de entrada é influenciado por muitos outros nós, enquanto um nó com alto grau de saída tem muitos alvos funcionais potenciais. O equilíbrio do grau de entrada e do grau de saída do nó é uma indicação da maneira como o nó está inserido na rede geral; por exemplo, esse equilíbrio especifica se o nó envia ou recebe informações.

A matriz adjacência tem interpretações bastante diretas no contexto de redes cerebrais anatômicas. Os nós da rede representam elementos neurais, como células, populações de células ou regiões cerebrais, enquanto as bordas da rede representam conexões entre nós, como sinapses. A estrutura das matrizes de adjacência descreve os padrões de comunicações dentro da rede. A presença de uma borda ligando dois nós indica que os dois nós podem se comunicar diretamente. Caminhos de vários comprimentos registram possíveis maneiras pelas quais os sinais podem viajar indiretamente entre dois nós [Sporns 2010].

Desde a década de 1980, houve uma crescente busca sobre o entendimento das interações entre pares de séries temporais neurofisiológicas registradas simultaneamente em duas localizações anatômicas diferentes. A versão mais simples da análise consiste em estimar o coeficiente de correlação entre duas séries temporais, regiões cerebrais que

demonstram padrões correlacionados de mudança de sinal ao longo do tempo são então consideradas conectadas funcionalmente. Portanto, a conectividade funcional é definida como uma dependência estatística entre as séries temporais dos sinais neurofisiológicos medidos. Este conceito foi originalmente desenvolvido para a análise de trens de impulsos registrados de unidades individuais e foi então traduzido para a análise de dados de neuroimagem funcional humana por Karl Friston, Barry Horwitz, Randy McIntosh e outros. O conceito básico de que duas localizações podem ser consideradas funcionalmente conectadas se tiverem dinâmicas coerentes ou sincronizadas. Essas análises foram generalizadas para muitas modalidades diferentes de imagens neurofisiológicas, incluindo a fMRI, eletroencefalografia, magnetoencefalografia, registro de matrizes de multieletródos de potenciais de campo local, e tomografia por emissão de pósitrons [Fornito, Zalesky e Bullmore 2016].

O cérebro de mamíferos apresenta lentas flutuações no fluxo sanguíneo cerebral regional durante o estado de repouso, quando não está envolvido em nenhuma tarefa. Em regiões funcionalmente relacionadas, mesmo aquelas separadas por grandes distâncias, essas flutuações ocorrem de forma sincronizada. Essa sincronia sugere a existência de conexões neuronais que facilitam a atividade coordenada entre essas áreas. Estudos de fMRI identificam essas regiões do cérebro com base em suas flutuações sincronizadas. As fortes correlações na intensidade do sinal de fMRI refletem mudanças simultâneas no efeito BOLD (dependente dos níveis de oxigênio no sangue) em regiões funcionalmente conectadas. A partir das imagens adquiridas pela fMRI é possível mapear as mudanças na intensidade do sinal no cérebro, permitindo identificar o aumento do fluxo sanguíneo quando um estímulo é aplicado ou uma tarefa cognitiva é realizada. As variações na intensidade do sinal, correlacionadas temporalmente com a execução da tarefa, são usadas para localizar as alterações hemodinâmicas decorrentes da atividade ou do estímulo. Flutuações sincrônicas na intensidade do sinal no cérebro são usadas para examinar a força das conexões neurais entre diferentes regiões cerebrais enquanto elas não estão ativadas por uma tarefa específica [Cordes et al. 2000].

As primeiras análises de grafos das redes funcionais do cérebro humano foram baseadas em matrizes de conectividade funcional estimadas a partir de dados de fMRI e outros métodos, realizando a correlação ou coerência entre séries temporais registradas em diferentes locais do cérebro (nós), foi estimada para cada par possível de nós, e as correlações par-a-par foram escolhidas de forma arbitrária, assim definindo arestas binárias que constituem um grafo da rede funcional em larga escala. Esses trabalhos mostraram que as redes de conectividade funcional do cérebro humano reproduzem propriedades organizacionais semelhantes às que foram descobertas nas redes anatômicas de outros mamíferos, assim como em muitos outros sistemas complexos. Por exemplo, essas redes são redes de mundo pequeno, contêm *hubs*, possuem uma estrutura modular hierárquica e parecem ser restringidas para minimizar os custos de fiação, aproximadas pela distância

euclidiana de suas arestas, como mostra a Figura 3. De fato, a organização estrutural devido suas restrições espaciais e topologia são uma característica proeminente e hereditária das redes cerebrais humanas.

Figura 3 – A organização modular de um grafo cerebral via fMRI em uma configuração topológica.

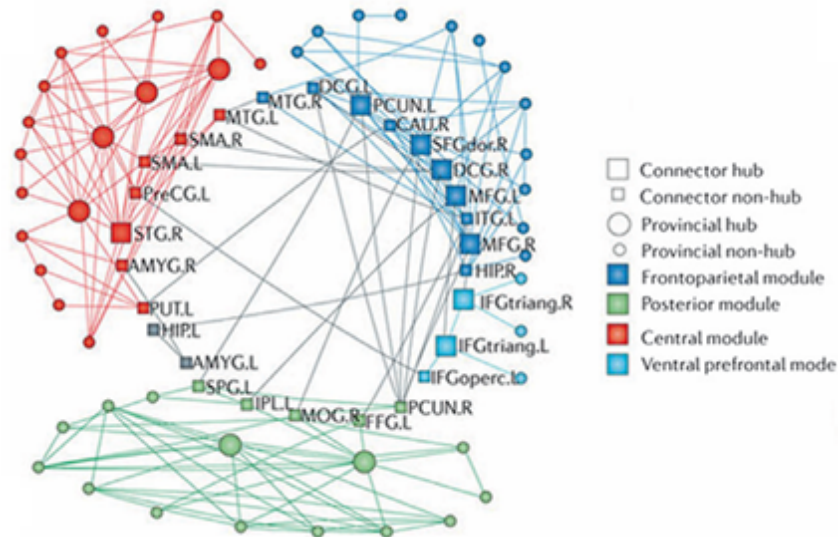


Imagem adaptada de [Fornito, Zalesky e Bullmore 2016]

As primeiras análises de grafos das redes anatômicas do cérebro humano foram baseadas na análise tractográfica de dados de MRI de difusão Figura 4 e na análise de covariância estrutural de dados de MRI convencional. Independentemente do método de construção, essas redes anatômicas do cérebro humano têm consistentemente mostrado as mesmas propriedades organizacionais não triviais descobertas em outros tipos de redes cerebrais.

Embora essa convergência entre grafos cerebrais construídos a partir de mapeamentos por MRI estrutural, de difusão e funcional sejam animadoras, é necessário ressaltar que conectividade funcional e conectividade estrutural são conceitos bastante diferentes de muitas maneiras. A conectividade anatômica pode ser definida de forma mais fundamental como uma projeção axonal de uma célula para outra, ou em escalas espaciais mais amplas por um trato de axônios projetando-se de uma área para outra. Tais projeções axonais são esperadas a mudar apenas lentamente ao longo do tempo.

A maioria dos estudos teóricos de grafos das redes de conectividade funcional analisa a atividade neurofisiológica coerente ao longo de um período prolongado usando uma métrica escalar, como um coeficiente de correlação, "calculando a média" das mudanças dinâmicas na configuração da rede. Dessa forma, a conectividade funcional está correlacionada, embora não seja idêntica, à conectividade anatômica. Em modelos computacionais, a convergência entre a conectividade funcional e a conectividade anatômica subjacente aumenta com o tempo sobre o qual a conectividade funcional é calculada. Portanto, quanto

Figura 4 – Tractografia via MRI de difusão usada para construir uma matriz de conectividade ponderada e não direcionada, onde cada elemento quantifica a densidade das projeções axonais entre um par específico de regiões. Essa rede pode ser ilustrada como um grafo cerebral projetado usando um layout definido algoritmicamente para enfatizar aspectos topológicos da organização da rede.

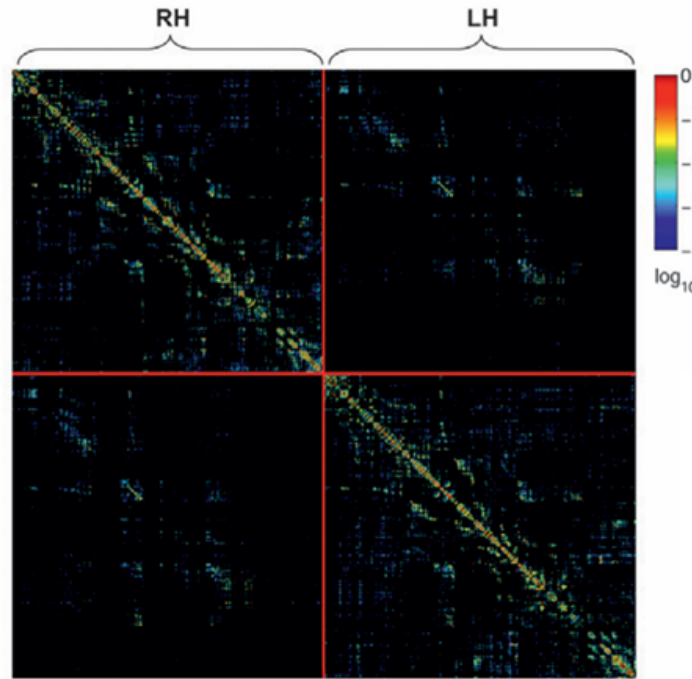


Imagem adaptada de [Fornito, Zalesky e Bullmore 2016]

mais longo o período de cálculo, mais a conectividade funcional pode se aproximar da conectividade estrutural. No entanto, é importante lembrar que conectividade estrutural e funcional são medidas diferentes e podem resultar em conectomas com diferentes valores para alguns parâmetros topológicos [Fornito, Zalesky e Bullmore 2016].

2.2 Teoria de Matrizes Aleatórias

As matrizes aleatórias surgiram na década de 1950, o com o físico e matemático Eugene Paul Wigner (1902-1995), que introduziu a Teoria das Matrizes Aleatórias (TMA) com o objetivo de descrever o espectro de energia de núcleos de átomos pesados. Wigner publicou um conjunto de trabalhos que chamou a atenção das comunidades de física e matemática, dos quais podemos destacar [Wigner 1951, Wigner 1955, Wigner 1958]. Ele percebeu que a distribuição de autovalores, determinada através da diagonalização dessas matrizes em que os elementos eram determinados aleatoriamente seguindo uma distribuição gaussiana, coincidia com a estatística das flutuações dos níveis de átomos pesados, obtidos experimentalmente. Devido a grande importância de seu trabalho ele foi laureado em 1963 por contribuições para a teoria do núcleo atômico e partículas elementares [Albert e Barabási 2002, Mehta 2004, Almeida e Souza 2009, Silva 2000, Atas et al. 2013].

Apesar da TMA ser inicialmente proposta para explicar as propriedades estatísticas dos espectros nucleares, ela também obteve previsões bem-sucedidas para as propriedades espectrais de diferentes sistemas complexos, como sistemas desordenados, sistemas caóticos quânticos, grandes átomos complexos [Beenakker 1997] e outros sistemas [Luo et al. 2006, Sarkar e Jalan 2016], respaldada por verificações numéricas e experimentais nas últimas décadas [Jalan e Bandyopadhyay 2007, Bandyopadhyay e Jalan 2007, Mehta 2004, Atas et al. 2013].

Os ensembles ortogonal gaussiano (GOE), unitário gaussiano (GUE) e simplético gaussiano (GSE) são fundamentais na TMA, frequentemente referidos como os ensembles clássicos dessa teoria, pois as propriedades estatísticas dos autovalores desses ensembles são bastante interessantes para análises utilizando essa metodologia, sujeitas apenas às restrições de simetria [Nieminen e Muche 2017, Sabri et al. 2014]. Esses ensembles correspondem a matrizes hermitianas aleatórias cujas entradas são distribuídas independentemente como variáveis aleatórias reais (GOE), complexas (GUE) ou quaterniônicas (GSE) [Mehta 2004, Atas et al. 2013].

As correlações de longo alcance dos autovalores são quantificadas utilizando estatísticas como a variância ou a rigidez espectral, enquanto as correlações de curto alcance são estudadas através da distribuição de espaçamentos consecutivos (DEC). A DEC fornece informações sobre a correlação entre autovalores adjacentes, mas não revela a correlação entre autovalores mais distantes [Nieminen e Muche 2017, Jalan e Bandyopadhyay 2007].

A universalidade da TMA significa que os ensembles de matrizes aleatórias descrevem os espectros de sistemas reais a um nível estatístico, e apenas em uma janela de energia local quando a densidade média de níveis é ajustada para a unidade. Diferentes modelos podem, de fato, apresentar densidades de níveis muito variadas e, para comparar as funções usuais de correlação espectral, como a DEC, é necessário realizar uma transformação chamada "*unfolding*" (desdobramento), que é uma forma de normalização. O procedimento de *unfolding* consiste em alterar os valores dos níveis verdadeiros, e_n , para níveis normalizados, $e'_n = N(e_n)$, onde $N(e)$ é o número médio de níveis menores que e , obtido por suavização ao longo de muitas realizações no caso de sistemas desordenados, ou por suavização local em uma janela de energia grande comparada ao espaçamento entre níveis, mas pequena comparada às variações de $N(e)$. O espectro desdobrado tem automaticamente um espaçamento médio entre níveis igual a um, e suas propriedades estatísticas podem ser comparadas diretamente com as da TMA. Quando a forma funcional de $N(e)$ é conhecida, ou quando há estatísticas suficientes disponíveis, o desdobramento é direto e facilmente implementado [Atas et al. 2013, Mehta 2004].

Uma estratégia muito útil para se trabalhar com redes complexas sob o arcabouço da TMA é estudar a DEC dos espectros de autovalores das matrizes de adjacência e Laplaciana de várias redes complexas de interesse. A DEC fornece a probabilidade de

encontrar autovalores vizinhos com um dado espaçamento e segue duas propriedades universais, dependendo das correlações subjacentes entre os autovalores.

- Para autovalores correlacionados, a DEC segue as estatísticas dos ensembles clássicos da TMA.
- Para autovalores não correlacionados, a DEC segue a estatística de Poisson.

Uma das principais vantagens da abordagem TMA é que, dependendo da natureza das correlações dos autovalores, pode-se separar a parte dependente do sistema da parte universal aleatória, que estão interligadas devido à complexidade do sistema [Bandyopadhyay e Jalan 2007]. As abordagens recentes da TMA aplicadas ao estudo de séries temporais de redes complexas do mundo real, de diversas naturezas, mostram que a DEC dessas redes apresenta uma boa concordância com as estatísticas universais do GOE da TMA. Esses resultados indicam que as estatísticas do GOE conseguem explicar com sucesso esses tipos de sistemas [Plerou et al. 2002].

Com a crescente disponibilidade de dados de redes complexas do mundo real, tornou-se possível realizar a análise espectral da matriz de adjacência dessas redes, bem como de redes modeladas com propriedades semelhantes. Essas análises mostram que a matriz de adjacência correspondente a uma rede aleatória não ponderada também segue a lei semicircular de Wigner, conforme ilustrado na Figura 5. Redes modeladas de mundo pequeno apresentam uma densidade espectral muito complexa, com muitos picos acentuados, enquanto a densidade espectral das redes modeladas livres de escala exibe uma distribuição chamada triangular. As DEC das matrizes de adjacência dessas redes foram estudadas extensivamente, e esses estudos mostraram que elas são bem descritas pela DEC das matrizes pertencentes ao ensemble GOE da TMA Figura 6 [Jalan e Bandyopadhyay 2007, Bandyopadhyay e Jalan 2007].

Uma matriz aleatória é uma matriz $A = (a_{ij})_{i,j=1}^N$, onde os elementos a_{ij} são escolhidos aleatoriamente, respeitando as simetrias do sistema e o interesse principal são os autovalores das matrizes. Frequentemente, exigimos que A seja auto-adjunta, garantindo que seus autovalores sejam reais [Speicher 2020].

Ao se analisar espectros de autovalores de matrizes aleatórias, é necessário considerar duas propriedades:

- Propriedades globais, como a densidade espectral ou a distribuição de autovalores $\rho(\lambda)$;
- Propriedades locais, como as flutuações dos autovalores ao redor de $\rho(\lambda)$.

Figura 5 – (a) e (b), representam distribuições semicircular e triangular das densidades espectrais para da média em ensemble de redes aleatórias e livres de escala, respectivamente, onde λ representa os autovalores e $\rho(\lambda)$ é a densidade espectral. Todas as redes têm uma dimensão $N = 2000$ e grau médio $k = 20$. Essas figuras foram plotadas com uma média de 10 realizações aleatórias das redes.

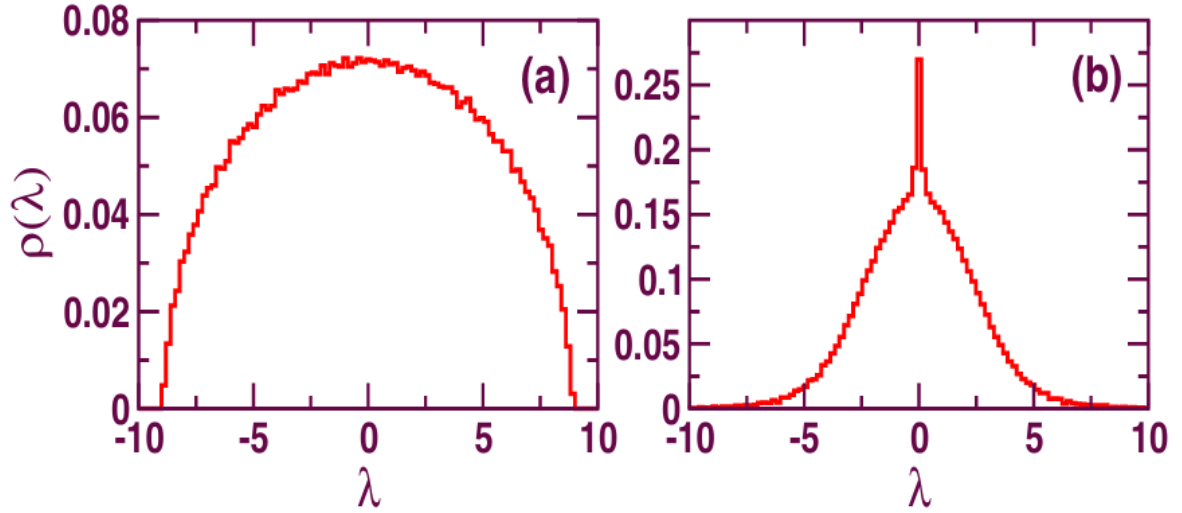


Imagem adaptada de [Bandyopadhyay e Jalan 2007]

Figura 6 – (a) e (b) correspondem a distribuição de espaçamento $P(s)$, mostrando que essas distribuições seguem as estatísticas GOE. Os histogramas são resultados numéricos e a curva representa a distribuição de Brody ajustada, com seu parâmetro β indicado. Esses resultados para as mesmas redes da Figura 5

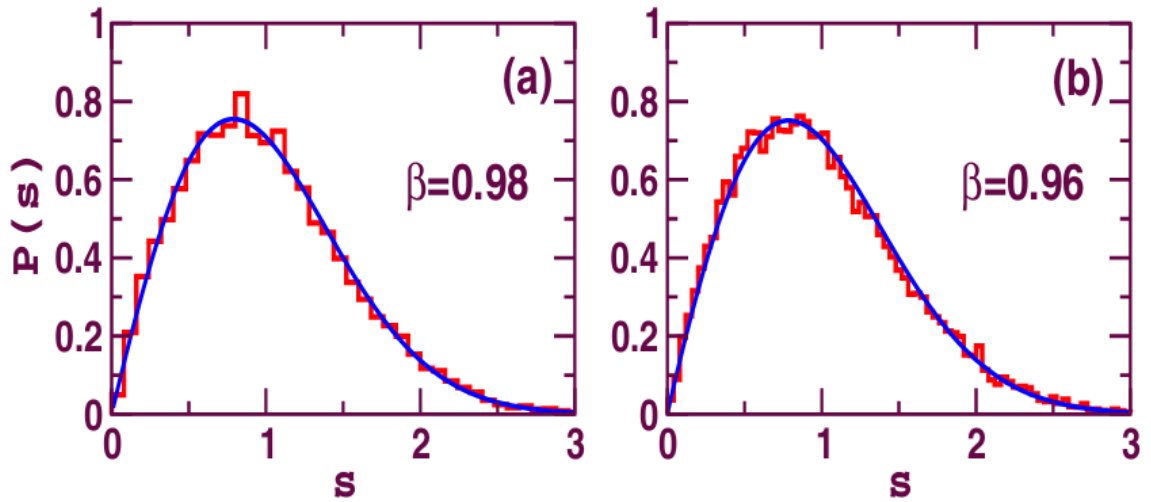


Imagem adaptada de [Bandyopadhyay e Jalan 2007]

Destas propriedades, as flutuações dos autovalores são as mais estudadas. Isso geralmente é obtido a partir da DEC dos autovalores. Os autovalores da rede são denotados por λ_n , $n = 1, \dots, N$, onde N é o tamanho da rede e $\lambda_n < \lambda_{n+1}$. Para obter propriedades universais das flutuações dos autovalores, é necessário remover os efeitos espúrios devido às variações da densidade espectral e trabalhar com densidade espectral constante em média. Portanto, é costume na TMA desdobrar os autovalores, assim como já foi discutido anteriormente,

por uma transformação $\bar{\lambda}_n = \bar{N}(\lambda_n)$, onde $\bar{N}(\lambda) = \int_{\lambda_{\min}}^{\lambda} \rho(\lambda') d\lambda'$ é a densidade média dos autovalores integrada do menor autovalor $\lambda_{\min}$ até o autovalor λ . Como a forma analítica para $\bar{N}(\lambda)$ não é conhecida, realizamos o desdobramento numericamente através de uma curva polinomial que melhor ajusta o espectro. Usando o espectro desdobrado, calculamos os espaçamentos consecutivos como:

$$s_n = \bar{\lambda}_{n+1} - \bar{\lambda}_n, \quad (2.1)$$

onde os autovalores são listados em ordem crescente, de forma que $\bar{\lambda}_{n+1} \geq \bar{\lambda}_n$, e por consequência $s_n \geq 0$.

Após o desdobramento, os espaçamentos médios consecutivos $\langle s \rangle$ tornam-se unitários, sendo independentes do sistema. Usando os espectros desdobrados, a DEC, $P(s)$, é definida como a distribuição de probabilidade desses s . No caso da estatística de Poisson,

$$P(s) = \exp(-s) \quad (2.2)$$

enquanto que para GOE, obedece a fórmula aproximada,

$$P(s) = \frac{\pi}{2} s \exp\left(-\frac{\pi s^2}{4}\right). \quad (2.3)$$

Para autovalores correlacionados, a DEC segue a fórmula de Wigner-Dyson da estatística GEO (2.3), que é uma propriedade exibida por matrizes aleatórias simétricas reais com elementos sendo números aleatórios obtidos de uma distribuição gaussiana. Por outro lado, para autovalores não correlacionados, a DEC segue a estatística de Poisson (2.2), que é uma propriedade mostrada por matrizes aleatórias que possuem elementos não nulos apenas ao longo de suas diagonais [Jalan e Bandyopadhyay 2007, Bandyopadhyay e Jalan 2007, Oganessian e Huse 2007, Atas et al. 2013, Sarkar e Jalan 2018].

Também são observados sistemas de transição, que não são bem explicados nem pela estatística de Poisson, nem pela GOE. Na TMA, modelos de matrizes que fazem a transição entre os ensembles clássicos: uma transição de GOE para GUE; ou modelos de matrizes que fazem a transição entre autovalores espaçados aleatoriamente e os conjuntos clássicos, ou seja, uma transição de Poisson para GOE têm sido investigados. As estatísticas de autovalores intermediários podem indicar alguma quebra de simetria subjacente ou uma mistura de autovalores de diferentes subespaços que possuem boas simetrias, particularmente para sistemas físicos. Uma possibilidade interessante é utilizar os modelos de transição da TMA como ponto de partida para se entender melhor esses casos intermediários [Nieminen e Muche 2017].

2.2.1 Distribuição de Brody

Em 1973, ao estudar a estatística de repulsão entre níveis de energia, Thomas A. Brody (1922-1988), propôs em seu artigo [Brody 1973] uma extensão da fórmula de Wigner para a DEC do GOE [Brody 1973].

Em redes com conexões aleatórias, mantendo o número de vértices ou arestas fixos, como por exemplo, as redes de mundo pequeno de Watts-Strogatz, podem ser construídas reestruturando as arestas de uma rede regular com uma probabilidade P . Para $P = 0$, a estrutura da rede regular permanece a mesma; Por outro lado, para $P = 1$ a rede se torna uma rede aleatória. Para valores intermediários de P , temos uma rede de mundo pequeno. Em redes regulares com $P = 0$, a DEC segue a estatística de Poisson, quando $P = 1$, a DEC segue a estatística GOE, e para $0 < P < 1$, devemos ter uma estatísticas intermediárias entre Poisson e GOE. Exemplos que representam a densidade espectral e a distribuição de espaçamento dessas redes, podem ser visto na Figura 7. Para os casos intermediários, a DEC não obedece exatamente as estatísticas de Poisson (2.2) ou GOE (2.3), podemos descreve-la pela distribuição de Brody,

$$P_\beta(s) = As^\beta \exp \left[-\alpha s^{\beta+1} \right], \quad (2.4)$$

onde A e α são determinados pelo parâmetro β da seguinte forma:

$$A = (1 + \beta)\alpha, \quad (2.5)$$

$$\alpha = \left[\Gamma \left(\frac{\beta + 2}{\beta + 1} \right) \right]^{\beta+1}, \quad (2.6)$$

a expressão (2.4) é uma equação semi-empírica caracterizada por um único parâmetro, β , conhecido como parâmetro de Brody. Os parâmetros A e α foram determinados pela normalização de $P_\beta(s)$, $\int_0^\infty P_\beta(s) ds = 1$ e fazendo a média de s igual a unidade, $\langle s \rangle = 1 = \int_0^\infty s P_\beta(s) ds$. À medida que β varia de 0 a 1, a distribuição de Brody muda suavemente de Poisson para GOE.

A DEC $P(s)$, muda de Poisson para GOE de forma suave, estabelecendo uma relação entre essa transição e o parâmetro de Brody. Esse parâmetro descreve um ajuste semi-empírico dos dados de espaçamento de autovalores estudados na TMA para modelar a transição de Poisson para GOE. A aplicação da distribuição de Brody permite ajustar os dados experimentais de sistemas intermediários.

Em outras palavras, ajustando as DEC de diferentes redes pela distribuição de Brody $P_\beta(s)$, somos capazes de estimar β e identificar se a distribuição de espaçamento de uma dada rede é Poisson, GOE ou uma distribuição intermediária entre as duas [Bandyopadhyay

Figura 7 – (a)–(d) mostram as densidades espectrais $\rho(\lambda)$ e (e)–(h) mostram as distribuições de espaçamento correspondentes $P(s)$, para $P = 0$, $P = 5 \times 10^{-5}$, $P = 2 \times 10^{-4}$ e $P = 5 \times 10^{-4}$, respectivamente. Os histogramas são dados numéricos e as curvas representam o ajuste dos dados pela distribuição de Brody (2.4). Todas as redes têm dimensão $N = 2000$ nó, um grau médio $k = 40$ por nó e os dados são a média de 10 realizações aleatórias do processo de reconexão.

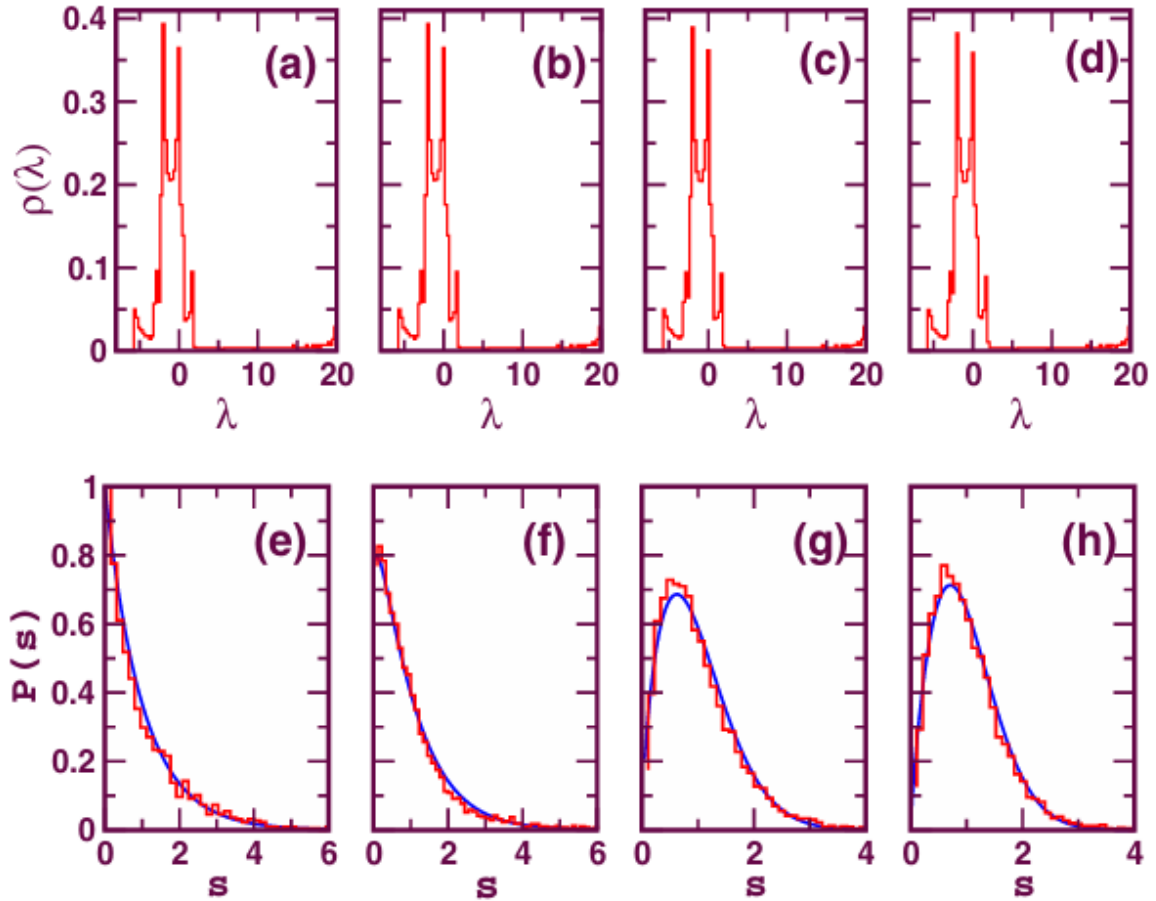


Imagem adaptada de [Bandyopadhyay e Jalan 2007]

e Jalan 2007, Jalan e Bandyopadhyay 2007, Brody 1973, Atas et al. 2013, Ganesan e Lakshmanan 1994, Sarkar e Jalan 2018].

3 Metodologia

3.1 Matrizes de Conectividade

Nosso trabalho foi desenvolvido com base em redes cerebrais funcionais, utilizando um conjunto de dados composto por dados de conectividade funcional mapeados por fMRI de indivíduos em estado de repouso, disponibilizados publicamente pelo projeto Human Connectome Project (HCP) (<https://www.humanconnectome.org/>) por meio de sua plataforma, The Connectome Coordination Facility, que distribui dados de pesquisa pública para uma série de estudos que se concentram nas conexões dentro do cérebro humano [Essen et al. 2012, Harms et al. 2018]; e também por dados fornecidos pelo VU University Medical Center, localizado em Amsterdã [Derks et al. 2017].

Os dados iniciais são um conjunto de 998 matrizes de conectividades funcionais cerebrais obtidas por fMRI. Todas as matrizes são quadradas $N \times N$, onde temos $N = 116$, indicando que cada cérebro foi dividido em 116 regiões de múltiplas escalas espaciais, sendo estas regiões os nós da rede, assim nossas matrizes são formadas.

Todas análises de dados e cálculos matemáticos foram realizados com o uso da linguagem de programação **Python** e suas bibliotecas, devido a sua versatilidade em lidar com grandes bases de dados e facilidade de programação dentro da plataforma **Google Colaboratory** ou **Google Colab**. Optamos por essa interface por ser uma plataforma gratuita que permite escrever e executar códigos em **Python** através de um navegador de internet, além da facilidade no armazenamento e compartilhamento de arquivos devido a sua integração com o **Google Drive**.

A matriz conectividade fornece uma descrição compacta da conectividade entre todos os pares de nós da rede. A matriz conectividade C para uma rede cerebral pode ser descrita matematicamente da seguinte forma:

$$C_{ij} = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1N} \\ C_{21} & C_{22} & \cdots & C_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ C_{N1} & C_{N2} & \cdots & C_{NN} \end{bmatrix}, \quad (3.1)$$

ou seja, matriz quadrada com N linhas e N colunas. Cada linha e coluna correspondentes representam um único nó da rede. Os elementos da matriz representam os valores de conectividade entre cada dois pares de nós [Fornito, Zalesky e Bullmore 2016]. Uma representação de nossas matrizes com os dados de conectividades reais podem ser vistos na Figura 8.

Figura 8 – Visualização reduzida de 4 matrizes de conectividades, escolhidas aleatoriamente, representadas como Data Frame no Python.

Matrix index: 251							Matrix index: 653						
	0	1	2	113	114	115		0	1	2	113	114	115
0	1.000000	0.831771	0.234380	0.275582	0.301047	0.103589	0	1.000000	0.833358	0.466566	0.638110	0.647326	-0.081818
1	0.831771	1.000000	0.035142	0.215839	0.143918	0.107088	1	0.833358	1.000000	0.452349	0.609813	0.612677	-0.090420
2	0.234380	0.035142	1.000000	0.059904	0.350221	0.053209	2	0.466566	0.452349	1.000000	0.391784	0.461750	0.032534
113	0.275582	0.215839	0.059904	1.000000	0.343333	0.238793	113	0.638110	0.609813	0.391784	1.000000	0.579374	-0.158980
114	0.301047	0.143918	0.350221	0.343333	1.000000	0.233230	114	0.647326	0.612677	0.461750	0.579374	1.000000	-0.061657
115	0.103589	0.107088	0.053209	0.238793	0.233230	1.000000	115	-0.081818	-0.090420	0.032534	-0.158980	-0.061657	1.000000
Matrix index: 871							Matrix index: 418						
	0	1	2	113	114	115		0	1	2	113	114	115
0	1.000000	0.743959	0.179560	0.295724	0.132900	-0.161679	0	1.000000	0.755131	0.411654	0.177194	0.208243	0.034424
1	0.743959	1.000000	0.131520	0.268960	0.117824	-0.112213	1	0.755131	1.000000	0.247233	0.133110	0.126646	0.060055
2	0.179560	0.131520	1.000000	0.135704	0.249802	0.024080	2	0.411654	0.247233	1.000000	0.023367	0.150040	0.043032
113	0.295724	0.268960	0.135704	1.000000	0.355107	-0.074456	113	0.177194	0.133110	0.023367	1.000000	0.282257	0.038147
114	0.132900	0.117824	0.249802	0.355107	1.000000	0.054055	114	0.208243	0.126646	0.150040	0.282257	1.000000	0.075202
115	-0.161679	-0.112213	0.024080	-0.074456	0.054055	1.000000	115	0.034424	0.060055	0.043032	0.038147	0.075202	1.000000

Fonte do Autor

Todas as 998 matrizes são da forma descrita na equação (3.1) e consequentemente apresentam as seguintes características:

- Os elementos da diagonal principal são todos iguais 1.

$$C_{ij} = 1, \quad \text{se } i = j \quad (3.2)$$

- Os elementos fora da diagonal C_{ij} , para $i \neq j$, são simétricos.

$$C_{ij} = C_{ji}. \quad (3.3)$$

Os elementos da diagonal principal podem ser interpretados como a representação da conectividade de cada nó consigo mesmo, mas, de forma mais geral, eles podem ser usados para codificar alguma propriedade intrínseca de cada nó. Em redes de conectividade funcional, a diagonal pode refletir variações na dinâmica local de cada região do cérebro. Na prática, a diagonal da matriz é pouco usada para esses fins dentro da neurociência. Em vez disso, as diferenças intrínsecas entre os nós da rede são comumente ignoradas e a diagonal é convencionalmente definida para um valor comum para todos os nós. De fato, muitas das medidas de teoria dos grafos que foram aplicadas a redes cerebrais ignoram a diagonal principal da matriz, muitas vezes assumindo que ela contém apenas valores zero. Nós também faremos essa suposição.

Os elementos fora da diagonal da matriz, representam a conectividade entre pares de elementos neurais distintos. Os valores contidos nesses elementos correspondem às medidas de conectividade entre pares, e sua amplitude dependerá do método particular

usado para estimar essa conectividade. Em geral, os valores desses elementos podem ser usados para representar a intensidade da conectividade entre cada par de regiões. Esses elementos apresentam uma simetria em torno da diagonal principal da matriz seguindo a equação (3.3). A simetria é justificada pela não direcionalidade da rede. De forma simples, a conectividade entre um par de nós quaisquer é a mesma no sentido oposto [Fornito, Zalesky e Bullmore 2016].

Valores negativos para conectividade em uma matriz conectividade indicam uma correlação negativa ou "anticorrelação". Enquanto correlações positivas fortes entre regiões indicam que essas regiões são tipicamente moduladas na mesma direção durante a realização de tarefas, anticorrelações entre redes indicam que as redes são moduladas temporalmente em direções opostas. Anticorrelações podem, portanto, revelar aspectos-chave da organização funcional do cérebro, pois se acredita que representem a conectomia e também a dicotomia entre o aumento da atividade cerebral em regiões que sustentem a execução de uma tarefa e diminuição da atividade cerebral em regiões envolvidas em processos não relacionados. Segundo [Fransson 2006] as anticorrelações representam uma "divisão do trabalho" entre redes com funções aparentemente opostas, assim como está sugerido em seu trabalho [Uddin et al. 2009].

3.2 Matrizes de Adjacência

A matriz de conectividade, assim como qualquer outro grafo $N \times N$, pode ser representada como uma matriz de adjacência, A , com elementos de matriz A_{ij} definidos da seguinte forma

$$A_{ij} = \begin{cases} 1, & \text{se } i \text{ e } j \text{ estiverem conectados} \\ 0, & \text{se } i \text{ e } j \text{ não estiverem conectados} \end{cases}. \quad (3.4)$$

Na teoria dos grafos, dois nós que estão diretamente conectados por uma aresta são considerados adjacentes ou vizinhos. A matriz adjacência, portanto, define o padrão de adjacências par a par entre os nós [Albert e Barabási 2002, Fornito, Zalesky e Bullmore 2016].

Para a construção de nossas matrizes de adjacência, a partir das matrizes de conectividade via fMRI, realizamos as seguintes etapas, sobre cada uma das 998 matrizes iniciais:

- Atribuir valores nulos para os elementos da diagonal principal.
- Manter apenas um percentual dos elementos não nulos definidos através de um limiar.

- Binarização da matriz resultante.

3.2.1 Diagonal Principal

Para os elementos da diagonal principal de cada uma das matrizes atribuímos valores nulos, pois a conexão de ponto cerebral com ele mesmo não deve existir. Matematicamente essa interpretação é descrita como:

$$C_{ij} = 0, \quad \text{para } i = j. \quad (3.5)$$

3.2.2 Filtragem

Algumas redes cerebrais, como redes de conectividade funcional baseadas em correlações de séries temporais, contêm um valor diferente de zero em todos os elementos fora da diagonal da matriz, alguns desses muito pequenos. Essas redes são ditas como totalmente conectadas, pois existe um link não nulo entre cada par de regiões. Porém, não sabemos se de fato essas conexões existem ou são apenas ruído do processo de medição das correlações. De forma geral, sabemos que o cérebro não é uma rede totalmente conectada. Por exemplo, os 302 neurônios que compõem o sistema nervoso central são ligados por cerca de 5600 sinapses químicas e junções comunicantes [White et al. 1986], representando aproximadamente 6% do total possível de conexões, uma vez que poderiam existir $N \times (N - 1)$, para $N = 302$, um total de 90.902 conexões. A conectividade em conectomas de menor resolução tende a ser maior, embora mais densas, ainda estão longe de indicarem serem totalmente conectadas [Markov et al. 2013]. Essas estimativas sugerem que qualquer método de reconstrução de rede cerebral que produza uma matriz de conectividade totalmente conectada, em que cada elemento neural esteja conectado a todos os outros, provavelmente incluirá uma porcentagem substancial de conexões espúrias. De fato, a maioria dos métodos para quantificar a conectividade cerebral, seja estrutural ou funcional, está associada a ruído de medição, complicando a distinção entre conexões reais e não reais. Assim, um pequeno valor diferente de zero na matriz de conectividade pode refletir ruído de medição, em vez da presença de uma conexão real.

Para resolver este problema, na conectômia frequentemente aplicamos algum tipo de processamento ou filtragem na matriz de conectividade antes da análise, com o objetivo de transformar uma matriz ponderada em uma matriz binária, para isso podemos escolher um limiar para reduzir a influência de conexões de baixa intensidade e potencialmente espúrias na topologia da rede. O limiar funciona como filtro atribuindo valores nulos para quaisquer elementos de matrizes menores que um determinado valor escolhido como valor do limiar. Portanto, a matriz de adjacência pode ser entendida como uma filtragem da matriz conectividade [Fornito, Zalesky e Bullmore 2016].

O processo de limiar funciona da seguinte maneira, a partir de um valor de corte τ :

$$A_{ij} = \begin{cases} C_{ij}, & \text{se } C_{ij} > \tau \\ 0, & \text{se } C_{ij} \leq \tau \end{cases}. \quad (3.6)$$

O valor de τ determina quais conexões devem ser retidas na matriz de adjacência. Essa abordagem ajuda a eliminar conexões espúrias, mantendo apenas aquelas com peso de conectividade acima de um certo valor. Isso nos permite focar nas conexões mais fortes e provavelmente reais na análise da rede. A aplicação do limiar é útil por maximizar a separação entre sinal e ruído, entre conexões supostamente reais e possíveis erros de medida, enfatizando as propriedades topológicas das arestas que são menos propensas a refletir erros de medição. Também pode melhorar a eficiência computacional, já que a maioria das métricas de rede são calculadas mais rapidamente em redes com menos arestas [Fornito, Zalesky e Bullmore 2016, Newman 2010].

Uma das principais desvantagens da utilização de um limiar é o fato de que não existe uma regra clara para a escolha desse valor específico, de modo que a escolha do valor de τ fica sujeita a escolha do pesquisador por meio de tentativas e ajustes experimentais. Outro problema causado pela aplicação do limiar é que ele pode gerar diferenças nas densidades de conexão de redes individuais dentro de uma população mais ampla. Por exemplo, se um limiar for aplicado a uma amostra de redes de conectividade funcional humana, as matrizes de adjacência de indivíduos com altos níveis médios de conectividade terão um número maior de arestas sobrevivendo a esse limiar do que as matrizes de pessoas com baixa conectividade média. Essa variação introduz um grande fator de confusão, pois a maioria das medidas de rede é sensível ao número de arestas na rede.

Uma possível solução a este problema é variar adaptativamente o valor de τ para alcançar o mesmo número de arestas em todas as redes. Esse método tem limitações que podem complicar as análises, por isso sua escolha deve ser cuidadosa e seus resultados devem ser bem observados e interpretados, para um melhor ajuste em seu valor, uma vez que essa escolha ainda é algo de caráter empírico [Fornito, Zalesky e Bullmore 2016, Wijk, Stam e Daffertshofer 2010]. Neste trabalho escolhemos um limiar de 10%, dessa forma, apenas os 10% maiores valores de C_{ij} sobrevivem a essa etapa.

3.2.3 Binarização

Após a aplicação de um limiar à matriz de conectividade, podemos realizar uma binarização sobre os elementos restantes da matriz A_{ij} , de tal forma:

$$A_{ij} = \begin{cases} 1, & \text{se } C_{ij} > \tau \\ 0, & \text{se } C_{ij} \leq \tau \end{cases}. \quad (3.7)$$

Concluídas essas etapas obtemos nossas matrizes de adjacência A , onde seus elementos A_{ij} apresentam apenas dois valores possíveis: 1, indicando que os nós estão conectados, ou 0, na ausência de conexão entre nós. As análises dessas matrizes estão principalmente preocupadas em entender os padrões topológicos das conexões entre os nós, independentemente das variações em seus pesos. Focar apenas em quais conexões estão presentes e onde elas estão posicionadas pode fornecer informações sobre o plano arquitetônico básico de um conectoma [Fornito, Zalesky e Bullmore 2016, Newman 2010, Bandyopadhyay e Jalan 2007, Aguiar e Bar-Yam 2005]. Uma representação visual das etapas utilizadas na obtenção da matriz adjacência pode ser vista na Figura 9.

Figura 9 – (a) Uma rede de conectividade funcional baseada em correlação, medida com fMRI de um cérebro humano. Cada elemento nesta matriz tem um valor diferente de zero. (b) A mesma matriz de (a) após a aplicação de um limiar para reter apenas os 20% dos pesos mais intensos. (c) A rede após a binarização.

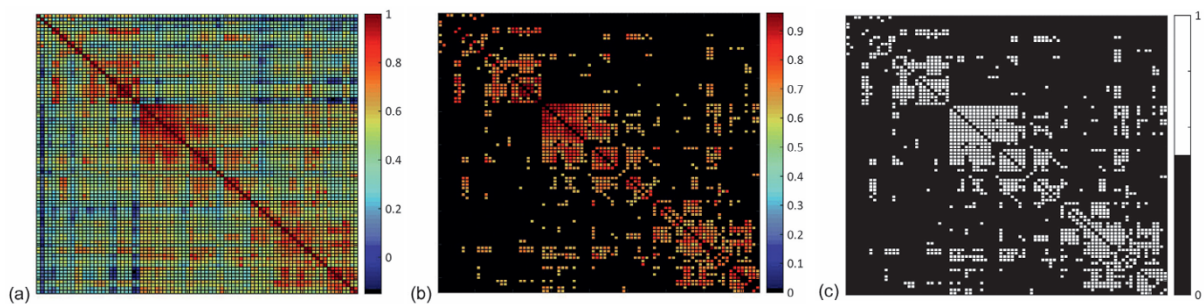


Imagem adaptada de [Fornito, Zalesky e Bullmore 2016]

As matrizes de adjacência foram obtidas no `Python`, utilizamos as bibliotecas básicas para importação e manipulação das matrizes de conectividade dos 998 cérebros e implementamos uma rotina para executar as três etapas descritas anteriormente para obter nossas matrizes de adjacência. Também fizemos o uso das bibliotecas `networkx` para manipular os grafos e `numpy` para operações matriciais. Como dito anteriormente, utilizamos um limiar de 10%, de modo que as matrizes resultantes possuem 10% das conexões originais. Na Figura 10, temos a visualização de uma das nossas matrizes de adjacência, escolhida aleatoriamente, obtida a partir de uma matriz conectividade.

3.3 Análise espectral das matrizes adjacência

As propriedades das redes complexas são bem caracterizadas pelo espectro das matrizes de adjacência associadas, onde a matriz adjacência A representa a topologia do sistema, indicando quais elementos estão interconectados [Bandyopadhyay e Jalan 2007, Aguiar e Bar-Yam 2005]. O espectro de autovalores de redes de sistemas complexos contém informações sobre as características estruturais das redes subjacentes e também fornece *insights* sobre o comportamento dinâmico e a estabilidade dos sistemas correspondentes. Os espectros dessas redes demonstram características específicas, de acordo com as

Figura 10 – Visualização de uma matriz adjacência obtida a partir de uma matriz conectividade.

	0	1	2	3	4	5	6	7	8	9	...	106	107	108	109	110	111	112	113	114	115
0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0
1	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	1.0	0.0	0.0	1.0	1.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	...	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
111	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0
112	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
113	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
114	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
115	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

116 rows × 116 columns

Fonte do Autor

características estruturais das redes e modelos subjacentes, indicando que os espectros de redes podem ser usados como uma ferramenta prática para classificar e entender diferentes sistemas do mundo real representados como redes complexas [Sarkar e Jalan 2018].

Para uma rede não direcionada, ela é simétrica e, conseqüentemente, possui autovalores reais. Esse espectro fornece informações sobre algumas propriedades topológicas básicas das redes subjacentes, como a densidade espectral da matriz de adjacência de uma rede aleatória e também seguem a lei semicircular de Wigner [Bandyopadhyay e Jalan 2007].

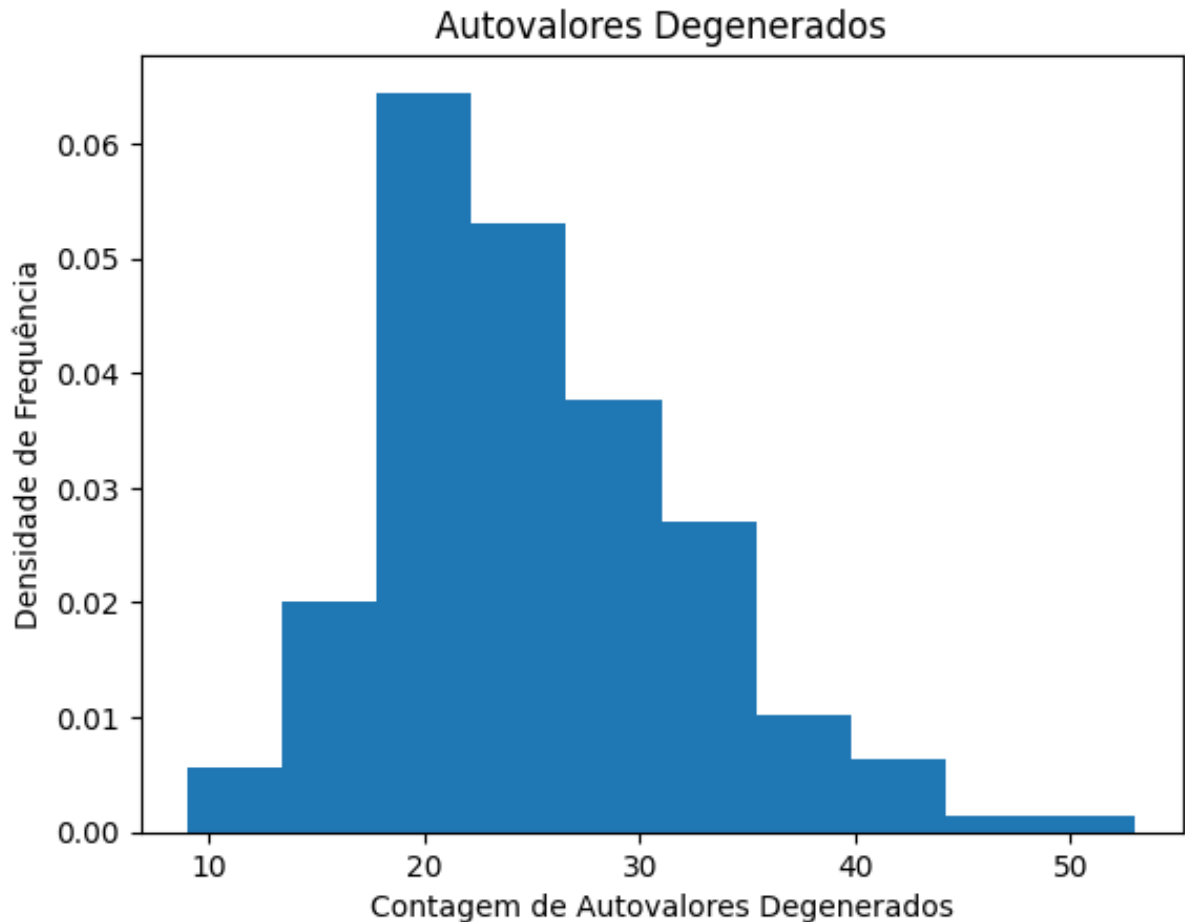
O espectro de cada uma das matrizes consideradas nesse trabalho foi determinado pela sua diagonalização exata, utilizando a biblioteca **numpy**. Ao analisar as listas de autovalores de cada uma das matrizes, percebemos uma quantidade significativa de valores nulos ou aproximadamente nulos. A aparição desses valores em nossas matrizes já era algo esperado, pois essa é uma característica de redes complexas do mundo real [Cimini et al. 2019]. Esses valores podem impactar nossas análises estatísticas futuras. Devido a esse fato, fizemos uma breve observação estatística das quantidades de valores nulos ou aproximadamente nulos em cada uma das matrizes.

Portanto, nossas análises começaram por uma contagem de autovalores duplicados ou aproximadamente duplicados, definindo uma tolerância de 10^{-6} como intervalo para que dois autovalores sejam considerados duplicados. Autovalores duplicados representam a degenerescência de um sistema. Observa-se que uma característica de redes complexas de mundo real é que a maioria dos autovalores degenerados são autovalores nulos.

Por meio do histograma, mostrado na Figura 11, podemos visualizar a distribuição

da quantidade de autovalores degenerados em todas as lista de autovalores associados a cada matriz.

Figura 11 – Histograma da distribuição de autovalores degenerados.

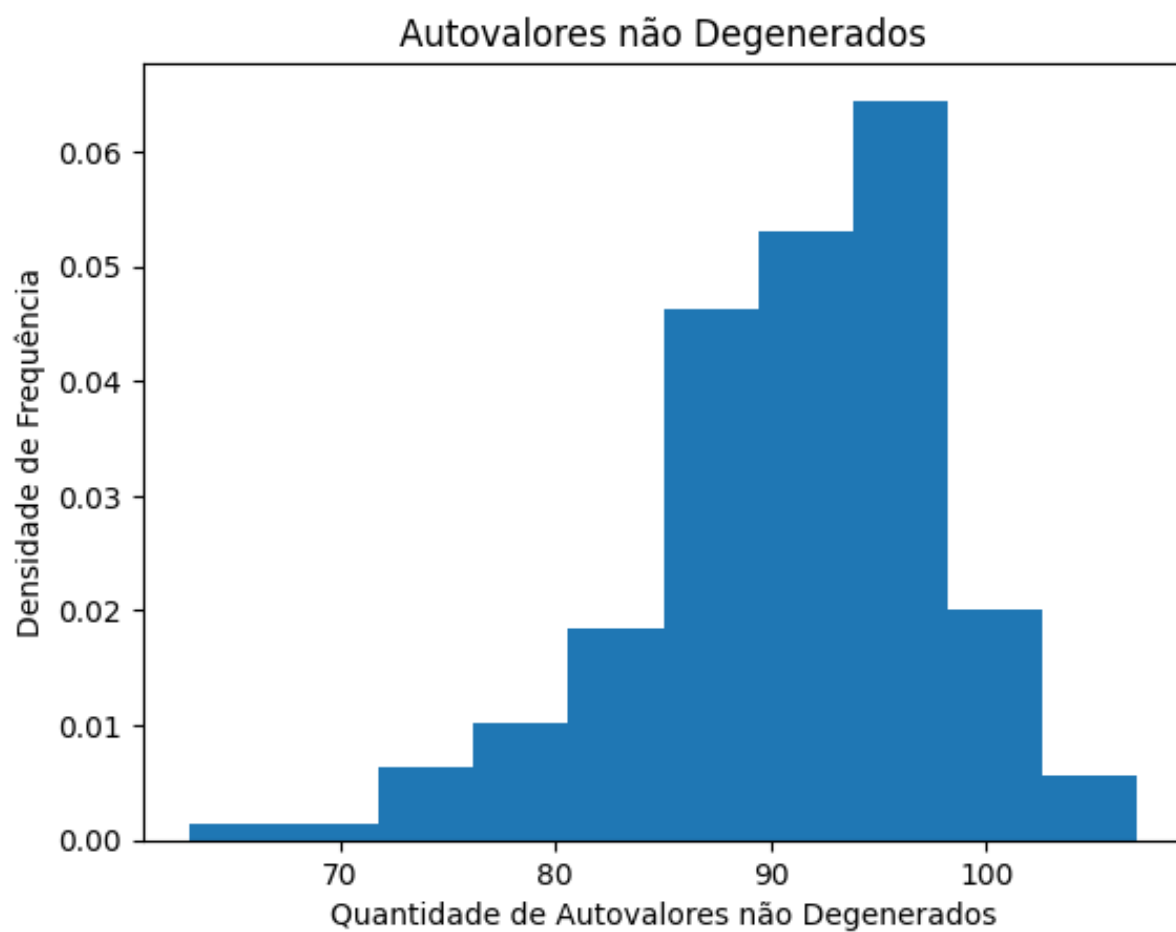


Fonte do Autor

Após essa verificação, iniciamos o próximo passo que foi a eliminação dos autovalores degenerados. Como consequência, o número de elementos em cada uma das listas de autovalores foi alterado. Inicialmente, tínhamos 998 listas contendo 116 autovalores cada, mas após a eliminação dos autovalores degenerados, essas quantidades foram modificadas. Essa alteração tem implicações para os cálculos estatísticos futuros, pois precisamos que todas as listas tenham a mesma quantidade de elementos para podermos realizar médias em ensemble. Para ajustar essa condição, verificamos a quantidade de autovalores não degenerados em todas as listas após o processo de eliminação da degenerescência. As quantidades de autovalores não degenerados em todas as lista pode ser visto por meio do histograma apresentado na Figura 12. Como pode ser visto no histograma, maioria das listas possuem entre 90 e 100 elementos. Portanto, analisamos esse intervalo com mais detalhes na Tabela 1.

Com as quantidades mostradas na Tabela 1, fica claro o a maioria das matrizes possuem elementos no intervalo entre 90 e 98 autovalores não degenerados. Esse intervalo

Figura 12 – Histograma da distribuição de autovalores não degenerados.



Fonte do Autor

Tabela 1 – Relação entre quantidades de autovalores não degenerados e a quantidade de matrizes, no intervalo de matrizes que possuem entre 90 e 100 elementos.

Quantidade de autovalores não degenerados	Quantidade de matrizes
90	53
91	61
92	62
93	57
94	57
95	73
96	57
97	48
98	48
99	35
100	19

Fonte do Autor

contempla um total de 516 matrizes, representando aproximadamente 52% das 998 matrizes iniciais.

Após determinar o subconjunto de matrizes com aproximadamente o mesmo número

de autovalores não degenerados, fizemos a uniformização das quantidades de elementos. Como todas as listas devem ter o mesmo número de elementos, executamos uma rotina sobre todas as listas que elimina valores além dos 90 primeiros, onde os valores estão ordenados de forma crescente em cada uma das listas.

Ao fim dessa análise, chegamos ao subconjunto uniformizado de listas de autovalores não degenerados. Nosso conjunto é formado por 516 listas de 90 elementos. Esse conjunto será o nosso ensemble.

3.4 Distribuições Estatísticas

Além da análise da DEC do nosso ensemble, também analisamos a distribuição de razões entre espaçamentos consecutivos (DREC), uma quantidade adimensional proposta por [Oganesyan e Huse 2007], para caracterizar as correlações entre os autovalores.

Para a obtenção da DEC, podemos utilizar a distribuição de Brody, definida pela equação (2.4), ou uma forma alternativa, introduzindo o parâmetro α_p , que é introduzido para que o espaçamento médio $\langle s \rangle$ não seja mais fixado em um. O modelo matemático para o ajuste dos dados de espaçamentos de níveis utilizado nessa dissertação é apresentado na equação abaixo

$$P_\beta(s) = A \left(\frac{\alpha_p}{\alpha} \right) s^\beta \exp \left[-\alpha_p s^{\beta+1} \right], \quad (3.8)$$

onde α_p e β são parâmetros do modelo, aos quais temos o objetivo de determinar. A e α são parâmetros definidos pelas equações, 2.5 e 2.6, respectivamente, ambos determinados pelo parâmetro β .

Para o estudo da DREC, realizamos o cálculo das razões entre espaçamentos consecutivos, definidos da seguinte forma:

$$r_n = \frac{\min(s_n, s_{n+1})}{\max(s_n, s_{n+1})}. \quad (3.9)$$

Para cada par consecutivo de espaçamentos (s_n, s_{n+1}) , a razão r é calculada como a razão do menor pelo maior valor entre os dois. Isso garante que r será sempre um valor no intervalo $0 \leq r_n \leq 1$.

Essa quantidade tem a vantagem de não precisar de *unfolding*, pois as razões dos espaçamentos consecutivos de autovalores são independentes da densidade local de estados. Essa distribuição proporciona uma comparação mais transparente com experimentos quando comparados à tradicional DEC da TMA. Tal fato, pode ser visto em muitos trabalhos recentes, que utilizam essa quantidade em diferentes contextos de sistemas de muitos corpos. Em alguns trabalhos, mostraram que a DREC $P(r)$ produziu resultados

mais assertivos que a DEC $P(s)$ [Oganesyan e Huse 2007, Santos e Rigol 2010, Atas et al. 2013].

Utilizamos a equação a seguir para descrever a DREC

$$P_\beta(r) = \frac{(r + r^2)^\beta}{[1 + r^{\beta+1} + (1 + r)^{\beta+1}]^{\frac{3\beta+2}{\beta+1}}}. \quad (3.10)$$

A expressão (3.10) [Duarte-Filho et al. 2024] foi inspirada na expressão aproximada proposta por [Atas et al. 2013], para DREC dos ensembles clássicos de TMA,

$$P_\beta(r) = \frac{1}{Z_\beta} \frac{(r + r^2)^\beta}{(1 + r + r^2)^{1+\frac{3}{2}\beta}}, \quad (3.11)$$

onde Z_β é uma constante de normalização que depende do parâmetro β dos ensembles clássicos de TMA. Os valores de Z_β , estão indicados na Tabela 2.

Tabela 2 – Valores da constante de normalização Z_β .

	$\beta = 1$	$\beta = 2$	$\beta = 4$
Z_β	$\frac{8}{27}$	$\frac{4}{81} \frac{\pi}{\sqrt{3}}$	$\frac{4}{729} \frac{\pi}{\sqrt{3}}$

Tabela adaptada de [Atas et al. 2013]

A equação (3.11) foi obtida a partir de expressões para a distribuição de probabilidade da razão entre dois espaçamentos consecutivos de níveis para os ensembles clássicos da TMA [Atas et al. 2013].

A distribuição (3.10) é simétrica assim como (3.11), a inversão r_n e $1/r_n$, de modo que $P_\beta(r)$ deva satisfazer a seguinte equação:

$$P_\beta(r) = \frac{1}{r^2} P_\beta\left(\frac{1}{r}\right). \quad (3.12)$$

3.5 Análise de Regressão

Devido à crescente disponibilidade de computadores mais poderosos, a análise de grandes quantidades de dados (*big data*) tornou-se cada vez mais viável e precisa. Esse fato tem impulsionado o estudo da ciência de dados em redes do mundo real, possibilitando análises cada vez mais abrangentes sobre a estrutura e a dinâmica dessas redes, facilitando muito o trabalho de pesquisadores de diversas áreas. Nesse contexto, a análise de regressão é uma ferramenta amplamente utilizada para estimar a relação entre uma resposta observada (variável dependente) e fatores (variáveis independentes) que podem estar relacionados à resposta. A análise de regressão é empregada para ajustar uma equação aos dados observados. Após propor um modelo matemático que explique a distribuição dos dados, é

necessário verificar quão bem os dados são ajustados pela equação, a fim de fazer previsões com um modelo bem ajustado [Hahn 1973, Sarstedt e Mooi 2014, Jalan e Bandyopadhyay 2009, Rahman e Muktadir 2021].

Para verificar se os dados reais analisados estão bem ajustados aos modelos que estamos propondo para descrevê-los utilizamos o coeficiente R^2 . O coeficiente de determinação R^2 é uma medida estatística que indica a proporção da variância presente nos dados em um modelo de regressão. Em outras palavras, ele quantifica a relação de ajuste entre os dados e o modelo proposto [Zhang 2016, Hahn 1973].

O valor de R^2 varia entre 0 e 1, também podendo ser expresso em porcentagem, é uma medida estatística que avalia a proporção da variabilidade na variável dependente que é explicada pela variabilidade nas variáveis independentes em um modelo de regressão. Ele fornece uma indicação de quão bem os dados observados são replicados pelo modelo, baseado na proporção da variação total dos dados que o modelo explica. Quanto mais próximo R^2 está de 1, mais o modelo está ajustado aos dados. Outra observação em relação ao valor de R^2 é que sempre é um valor positivo, pois R^2 é o valor do coeficiente de correlação R elevado ao quadrado. Nesse trabalho, o método usado para a obtenção dos parâmetros que melhor ajustam o modelo aos dados e respectivo coeficiente R^2 foi feito resolvendo o problema não-linear de mínimos quadrados utilizando o pacote `lmfit` [lmfit 2024] do `python` para ajustar os dados com os modelos para DEC e DREC. Esse pacote fornece uma interface de alto nível para problemas de otimização não linear e ajuste de curva.

4 Resultados e Discussões

4.1 Distribuição de Autovalores

Os resultados apresentados nesse capítulo dizem respeito ao ensemble formado por 516 listas, todas contendo 90 autovalores sem degenerescência, associados às matrizes de adjacência construídas a partir das matrizes de conectividade funcional obtidas das correlações séries temporais de redes cerebrais medidas via fMRI.

Para visualizarmos o comportamento das distribuição de autovalores, DEC e DREC, iremos *plotar* os histogramas das médias em ensemble para cada uma dessas distribuições, utilizando um padrão de 10 *bins*. As médias em ensemble dessas distribuições são obtidas organizando os valores de interesse (por exemplo, os autovalores) de todos os membros do ensemble em ordem crescente. Em seguida, calcula-se a média do menor autovalor, depois a média do segundo menor autovalor, e assim por diante, para todos os autovalores. A distribuição média em ensemble de autovalores pode ser visto pelo histograma apresentado na Figura 13. É possível perceber que a distribuição de autovalores não possui um caráter homogêneo.

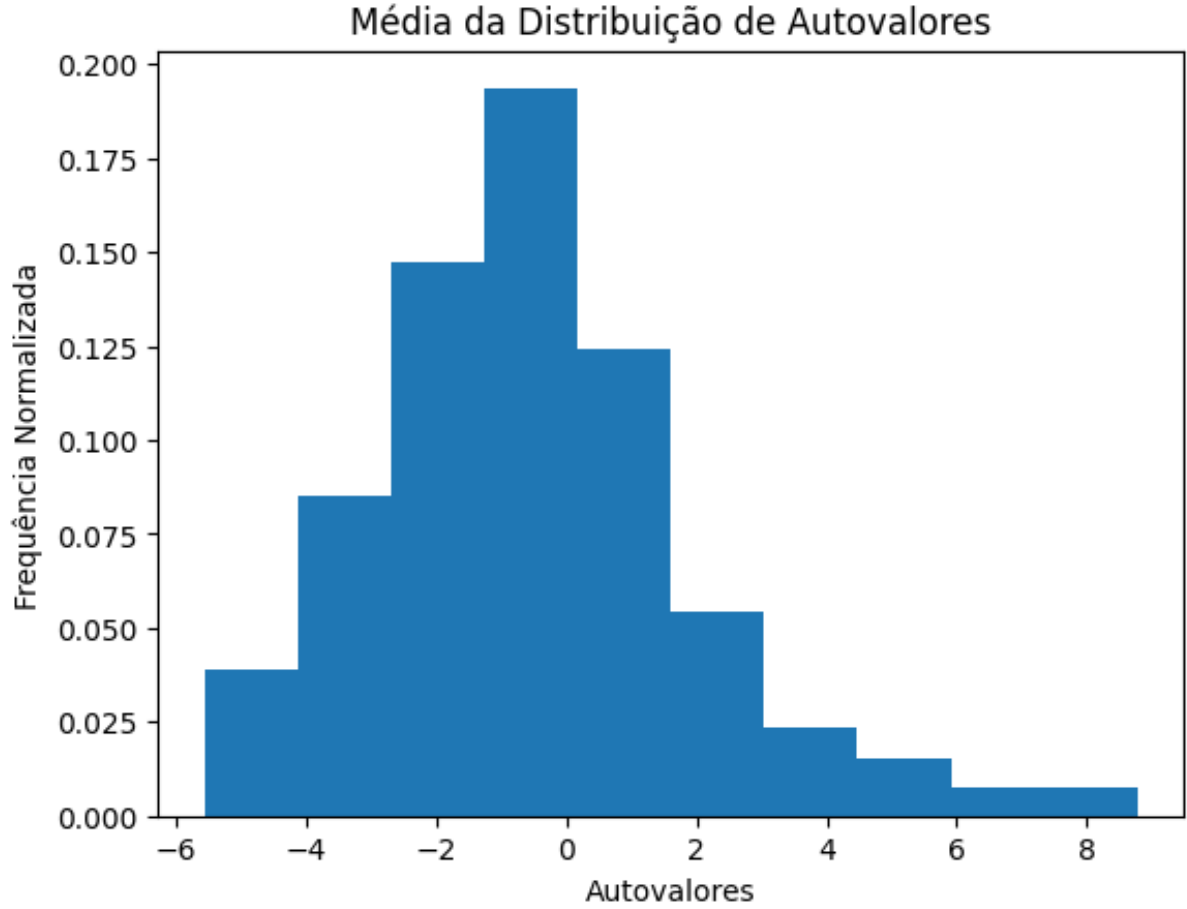
Para o cálculo dos espaçamentos entre autovalores consecutivos, é necessário que os autovalores passem por um processo de desdobramento (*unfolding*), que pode ser visto como um processo de normalização, assim como explicado no capítulo anterior (Capítulo 3). Dessa forma, a distribuição de autovalores desdobrados será mais uniforme que a original.

Para realizar o *unfolding* dos autovalores, utilizamos a biblioteca `scipy.stats`, fazendo uso da função `rv_histogram`, que permite tratar um histograma como uma distribuição de probabilidade contínua, a partir da qual podemos calcular a função densidade de probabilidade (PDF). Desta forma, obtemos um histograma normalizado, ou seja, a área total sob o histograma será igual a 1. A distribuição de autovalores desdobrados é mostrada na Figura 14. Este conjunto de autovalores desdobrado irá produzir resultados mais informativos que o conjunto original para a DEC.

4.2 Distribuição de Espaçamentos Consecutivos

Primeiramente, fizemos uso dos autovalores não desdobrados para realizamos o cálculo dos espaçamentos consecutivos, com o objetivo de fazer um comparativo com o caso no qual o conjunto de autovalores foram submetidos ao processo de *unfolding*. Essa distribuição pode ser vista na Figura 15. O histograma revela uma distribuição concentrada em valores próximos a zero, deixando claro que analisar a DEC sem efetuar o processo

Figura 13 – Histograma da distribuição média em ensemble de autovalores.



Fonte do Autor

de *unfolding* não representa bem as propriedades espectrais dos dados. Utilizando os autovalores desdobrados, calculamos os espaçamentos consecutivos, pela equação (2.1), resultando na DEC apresentada pela Figura 16. Podemos perceber que esta distribuição é melhor distribuída em relação ao caso anterior mostrado na Figura 15.

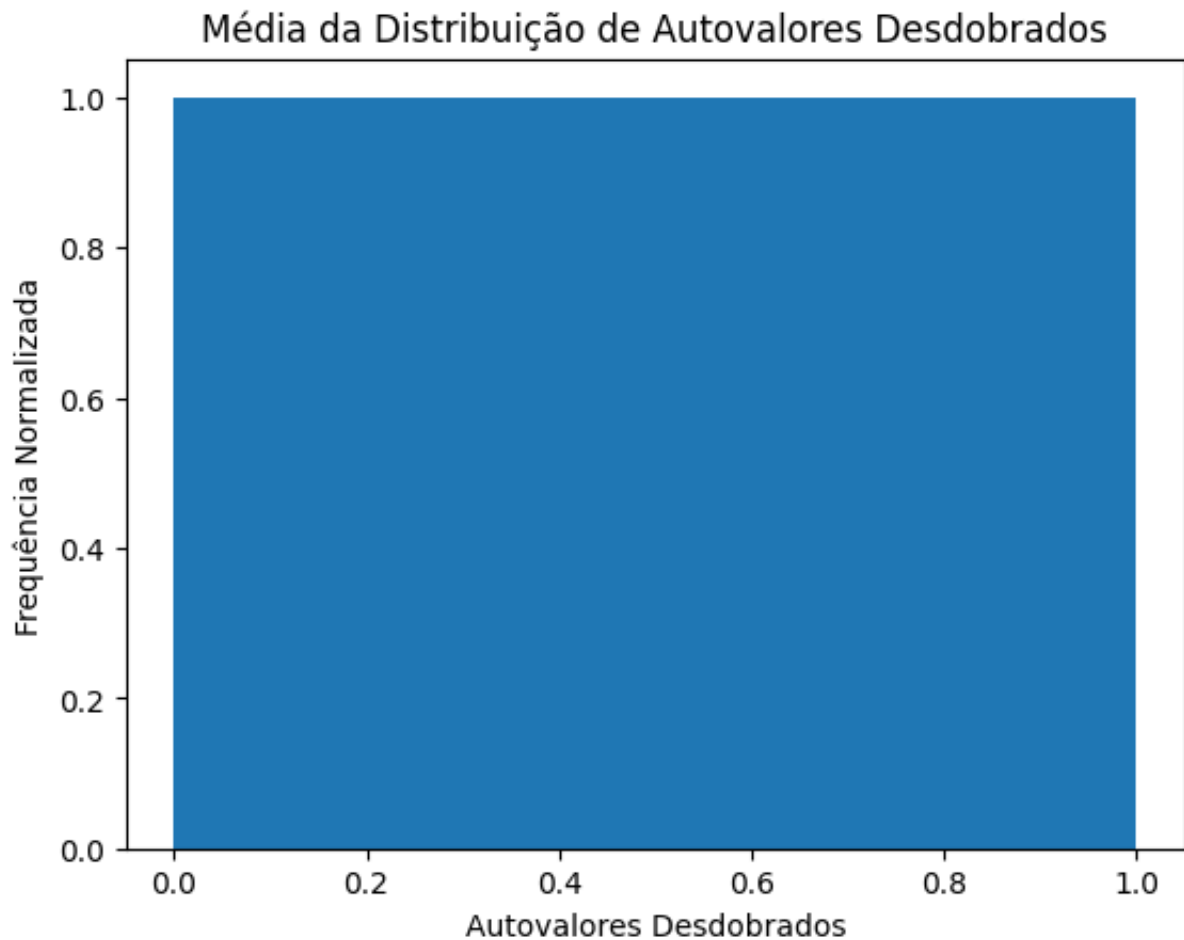
Após obtermos a DEC da média de ensemble, vamos verificar quão bem a distribuição de Brody (3.8) ajusta nossos dados. Para tanto, transformamos o histograma Figura 16 em um conjunto de pontos cujas abscissas são as distâncias dos pontos médios de cada *bin* em relação à origem

$$a_{n+1} = \frac{(B_{n+1} - B_n)}{2} + B_n; n = 0, 1, 2, \dots, \quad (4.1)$$

onde B_n e B_{n+1} são as extremidades inferior e superior, respectivamente, do $(n + 1)$ -ésimo *bin*. As ordenadas são as frequências normalizadas associadas a cada *bin*.

Após essa etapa, definimos uma função interna em nosso código que define o modelo que será ajustado aos dados. Essa função calcula a diferença entre os dados obtidos e a distribuição de Brody (3.8). Os parâmetros α_p e β da distribuição de Brody são

Figura 14 – Histograma da distribuição média em ensemble de autovalores desdobrados.



Fonte do Autor

determinados por um processo de minimização, por meio da função interna `Minimizer` do pacote `lmfit`. Ao executarmos esse processo, obtemos a minimização do erro entre o modelo que melhor ajusta os parâmetros aos dados. Também calculamos o coeficiente de determinação R^2 , para indicar o quão bem ajustado está o modelo aos dados. Os resultados dos ajustes do modelo que melhor descreve os dados estão apresentados na Tabela 3 e pela Figura 17.

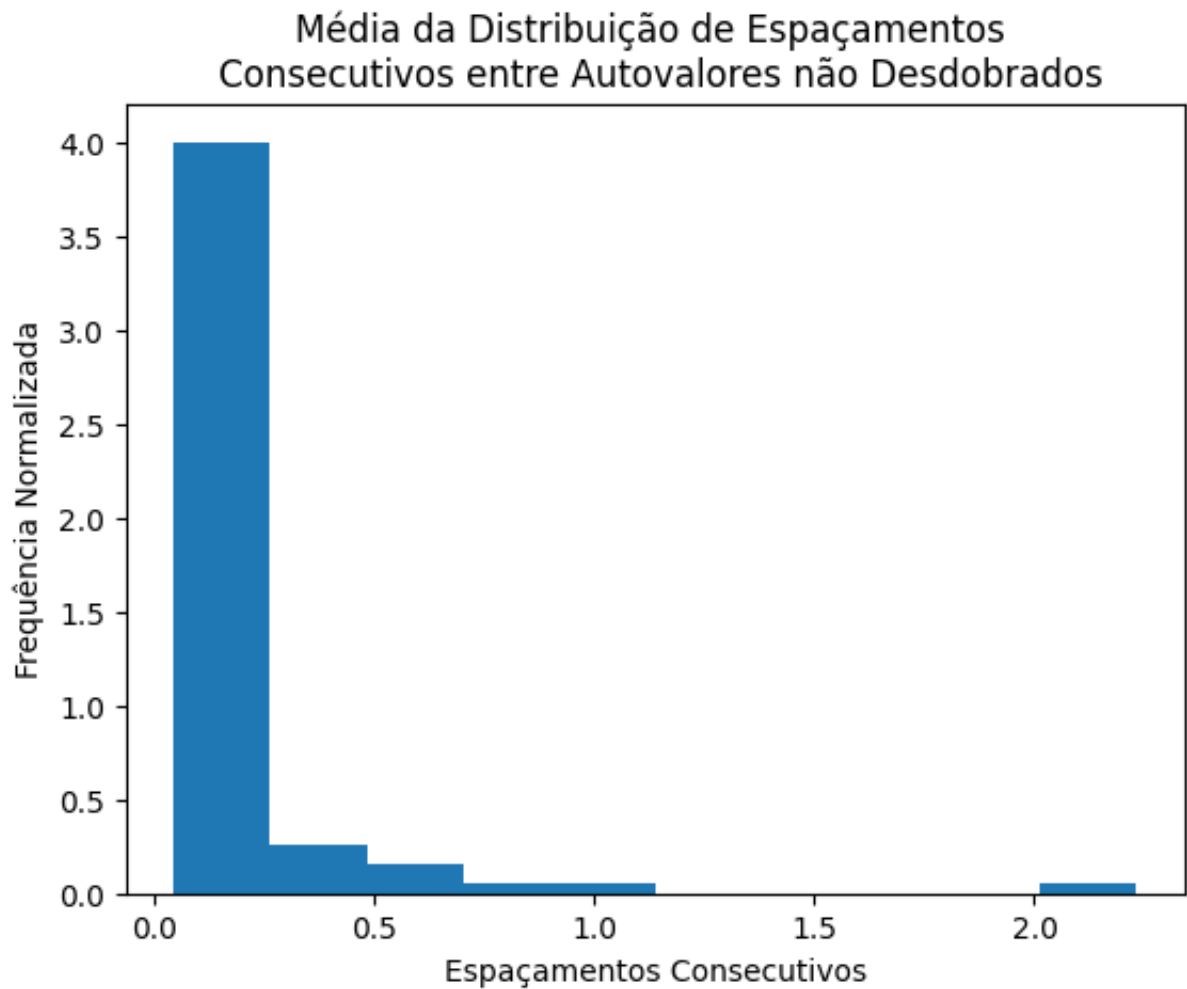
Tabela 3 – Medidas estatísticas sobre valores dos parâmetros do modelo da distribuição de Brody, melhor ajustado aos dados, para a DEC pela equação (3.8).

Parâmetro	Média do Valor Ajustado	Desvio Padrão	Incerteza Aproximada
β	0,89	$\pm 0,04$	4%
α_p	3785	± 612	16%

Fonte do Autor

Além dos dados expostos na Tabela 3, também calculamos a correlação entre β e α_p , de valor $+0,9878$, indicando uma alta correlação positiva entre esses parâmetros. Assim como, o valor do coeficiente de determinação $R^2 = 0,992$, esse valor de R^2 encontrado é próximo de 1, demonstrando que o modelo faz um bom ajuste dos dados.

Figura 15 – Histograma da distribuição média em ensemble de espaçamentos consecutivos entre autovalores não desdobrados.



Fonte do Autor

Além da análise da média em ensemble da DEC, ajustamos a distribuição de Brody para as DEC de cada membro individualmente do ensemble. Esse procedimento nos permitiu analisar, por exemplo, a distribuição dos valores de β estimados para os membros do ensemble. Os resultados dessa análise estão expostos na Figura 18 e na Tabela 4.

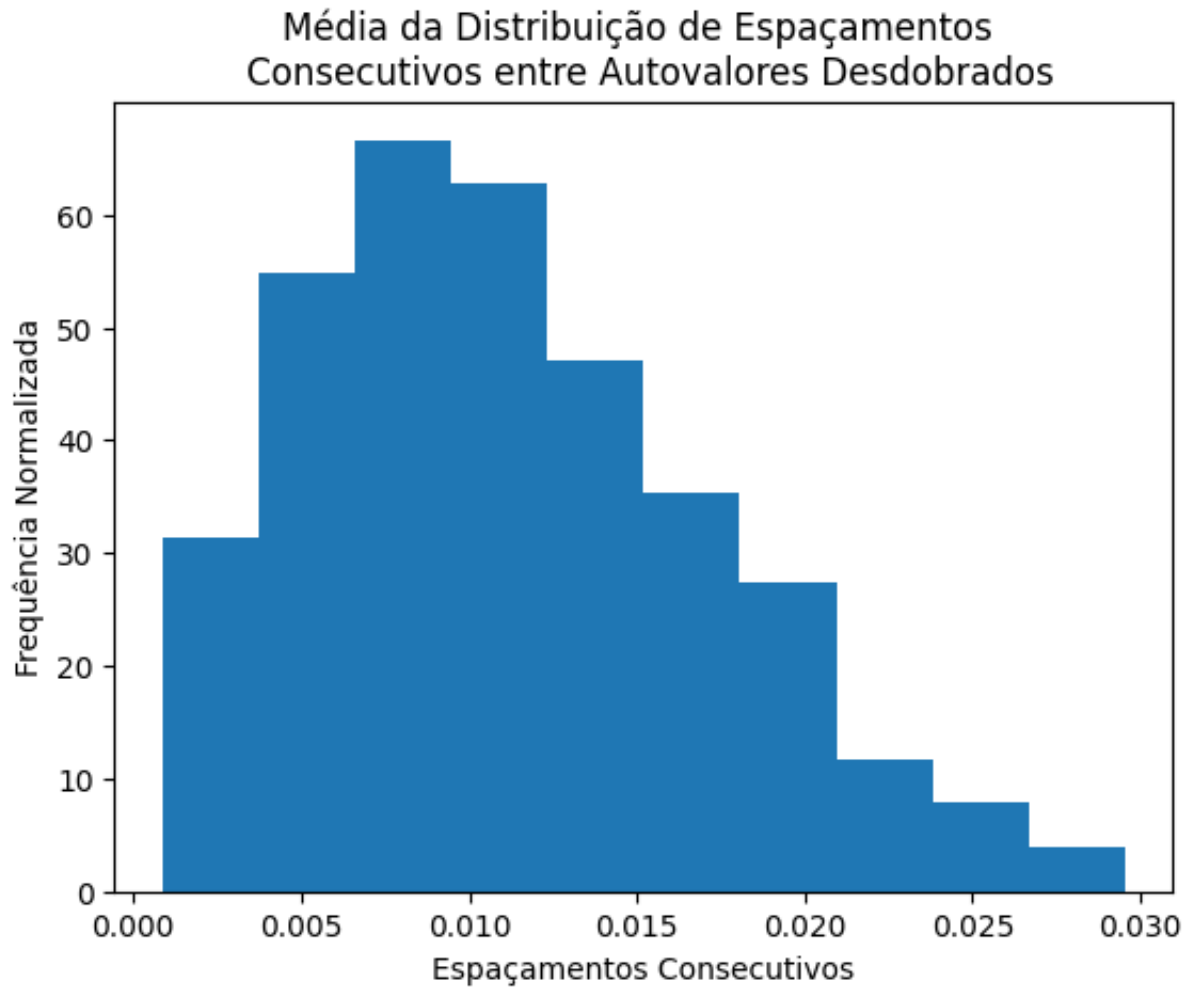
Tabela 4 – Os parâmetros de Brody e R^2 estimados a partir das distribuições de β , α_p e R^2 obtidas pelos ajustes individuais de cada membro do ensemble, para a DEC.

Parâmetro	Média	Desvio Padrão
β	0,9	$\pm 0,2$
α_p	3903	± 2003
R^2	0,812	$\pm 0,132$

Fonte do Autor

Os resultados das análises individuais apresentados pela Tabela 4 demonstram concordância com os resultados da média em ensemble (Tabela 3), mas é importante verificar que os resultados da análise individual possuem uma dispersão maior, em relação

Figura 16 – Histograma da distribuição média em ensemble de espaçamentos consecutivos entre autovalores desdobrados



Fonte do Autor

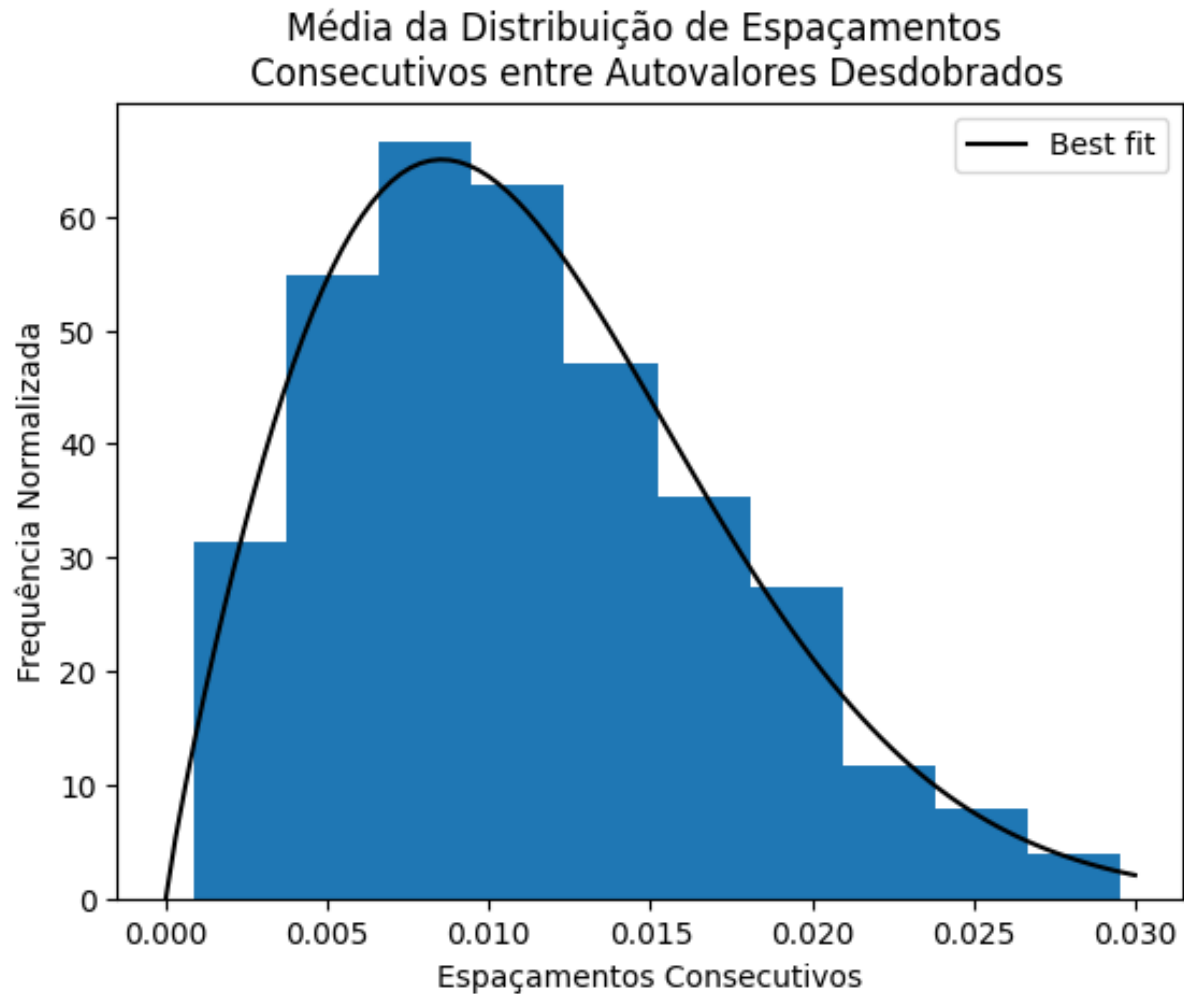
ao caso da média.

A seguir, apresentamos uma tabela com os melhores ajustes individuais levando em consideração os maiores valores de R^2 . Os 10 melhores pares de parâmetros β e α_p estão listados na Tabela 5. Esse melhores ajustes são apresentados na Figura 19. Esse histograma revela que a DEC consegue alcançar um ajuste de alta qualidade, pois as curvas estão bem alinhadas umas com as outras.

4.3 Distribuição de Razões entre Espaçamentos Consecutivos

Para a DREC, realizamos uma análise similar à feita para a DEC. Iniciamos calculando as razões entre espaçamentos consecutivos de autovalores não desdobrados, por meio da equação (3.9). A DREC pode ser vista na Figura 20. Como podemos visualizar

Figura 17 – Dados fitados no histograma da média em ensemble da distribuição entre espaçamentos consecutivos de autovalores desdobrados pela distribuição de Brody. A Curva *Best fit* representa a distribuição de Brody com os valores de β e α_p que melhor se ajustam aos dados, indicados na Tabela 3.



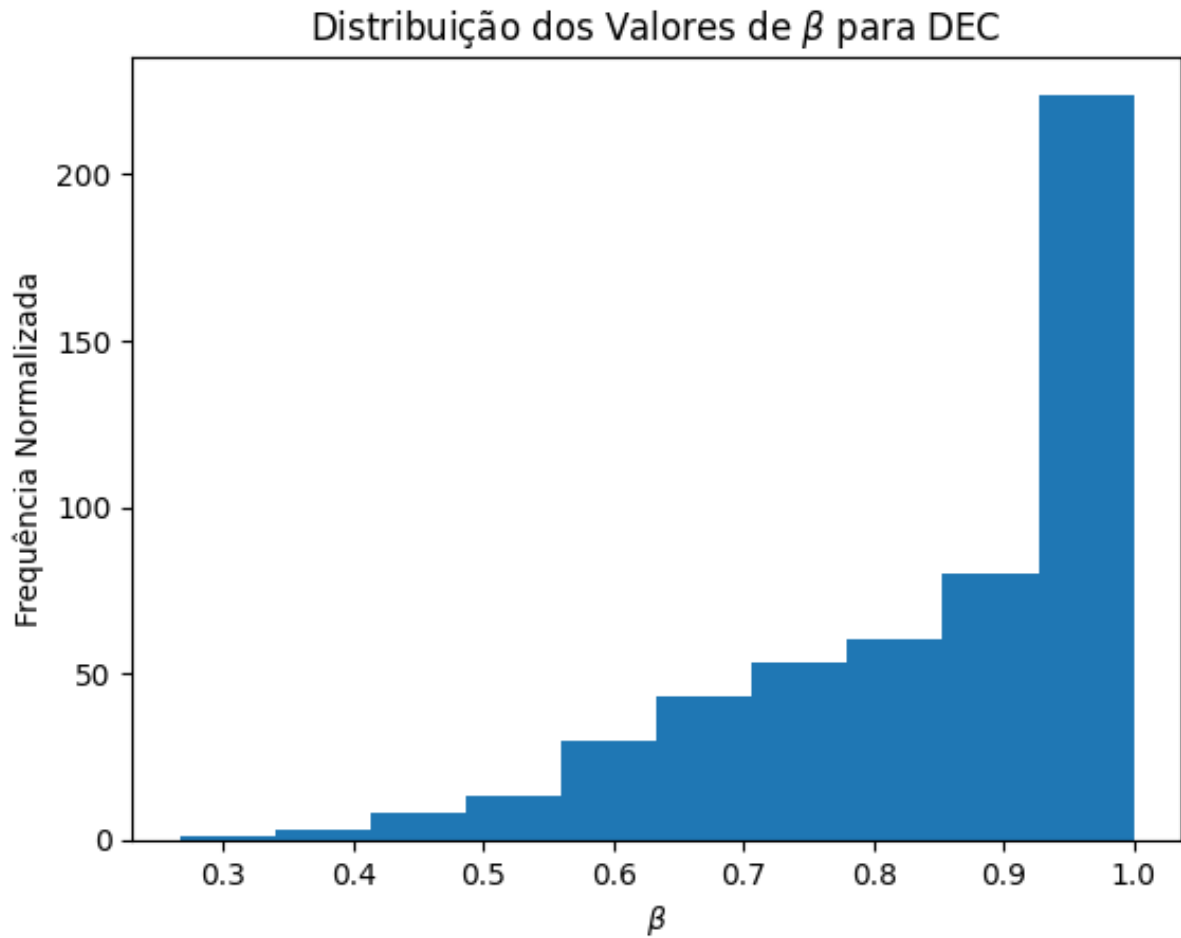
Fonte do Autor

Tabela 5 – 10 melhores ajustes dos parâmetros β e α_p para a DEC, considerando os maiores valores de R^2 .

Ordem	R^2	β	α_p
1º	0,993	0,9	5153
2º	0,992	0,7	1922
3º	0,990	0,9	5018
4º	0,986	0,9	4107
5º	0,983	0,9	4169
6º	0,983	1,0	3453
7º	0,982	1,0	5919
8º	0,982	1,0	4950
9º	0,982	0,9	4777
10º	0,978	0,9	3931

Fonte do Autor

Figura 18 – Histograma da distribuição dos valores de β para cada membro do ensemble na DEC.



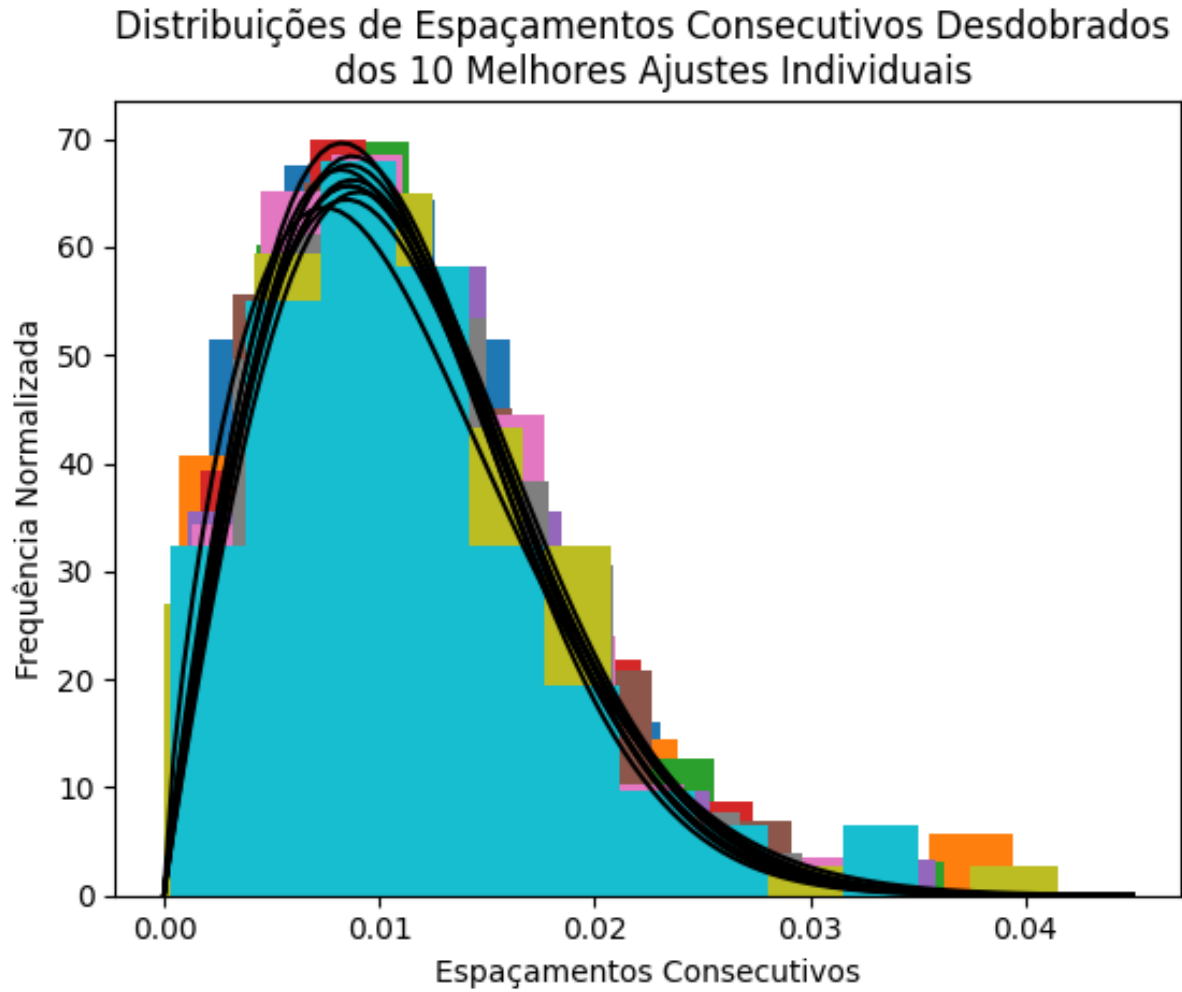
Fonte do Autor

a DREC apresenta uma distribuição mais uniforme em relação a DEC da Figura 15, ao utilizamos autovalores não desdobrados.

Em seguida, estudamos o modelo que melhor ajusta os dados da DREC, executando os passos descritos a seguir sobre todas as listas de autovalores não desdobrados, listadas em ordem crescente:

- Calcular as razões entre espaçamentos consecutivos de autovalores;
- Preparar os dados para ajustar ao modelo, obtendo os pontos a serem ajustados a partir do histograma Figura 20;
- Calcular os ponto médios de cada *bin*, para fazer o ajuste dos dados.
- Ajustar a distribuição semi-empírica proposta para a razão entre espaçamentos consecutivos, equação (3.10), aos dados através do processo de minimização utilizando o `Minimizer`.

Figura 19 – Os 10 melhores ajustes individuais da DEC levando em consideração os maiores valores de R^2 , com os valores de β e α_p listados na Tabela 5.



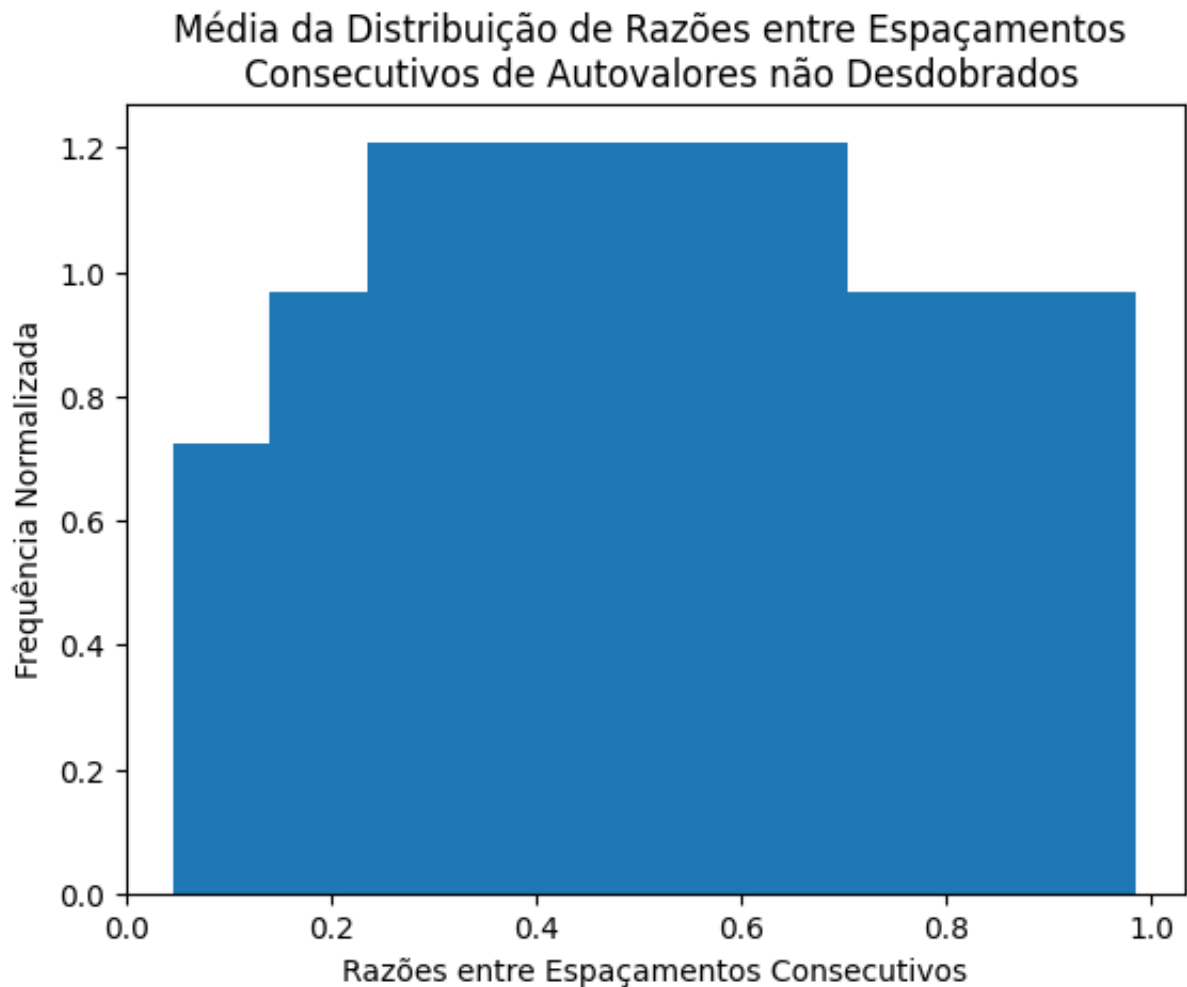
- Calcular o coeficiente de determinação R^2 , para avaliar a qualidade do ajuste.

Após executar esses passos, obtemos o parâmetro β melhor ajustado, assim como o respectivo R^2 . Fizemos duas análises para verificar como o modelo proposto *fit*a os dados. Analisamos o *fit* sobre a média em ensemble da distribuição e dos *fits* de cada membro individualmente do ensemble.

Os resultados da análise da DREC obtida da média em ensemble são apresentados na Figura 21 e na Tabela 6. Os resultados apresentados mostram que nosso modelo ajusta razoavelmente bem os dados, como pode ser visto na Figura 21. O R^2 obtido desse ajuste foi $R^2 = 0,87$.

Analisando a DREC de cada membro do ensemble, obtemos novamente uma distribuição dos valores de β Figura 22, de onde podemos estimar a média de β e outras quantidades estatísticas apresentadas na Tabela 7. O resultado exibido está de acordo com o apresentado no caso da média em ensemble Tabela 6, porém com uma maior dispersão.

Figura 20 – Histograma da distribuição média em ensemble de razões entre espaçamentos consecutivos de autovalores não desdobrados



Fonte do Autor

Tabela 6 – Parâmetro de Brody que melhor ajusta os dados, da DREC obtida pela média em ensemble.

Parâmetro	Valor Ajustado	Desvio Padrão	Incerteza Aproximada
β	0,87	$\pm 0,04$	5%

Fonte do Autor

O valor do coeficiente de determinação associados a essa medida é de $R^2 = 0,889$.

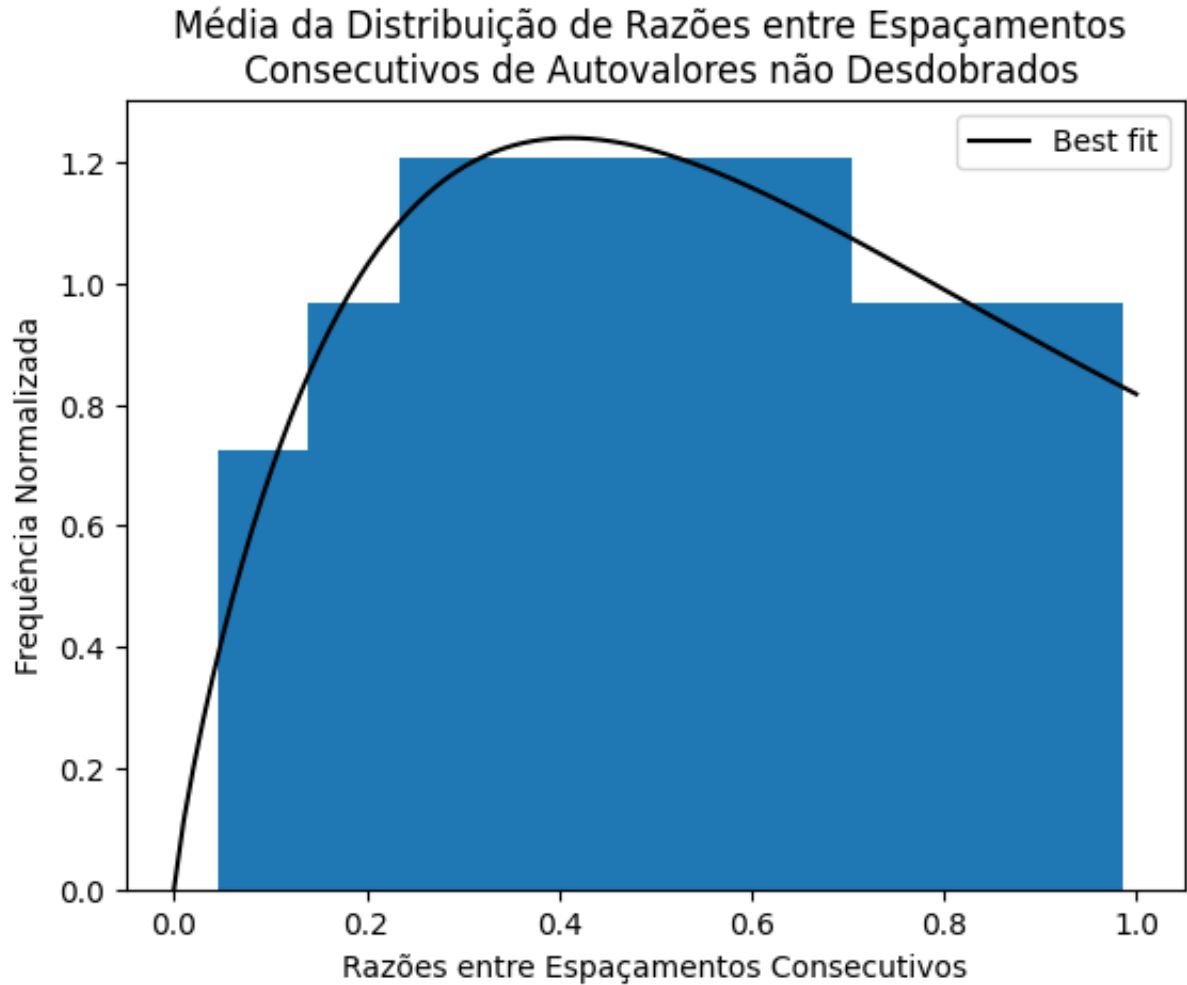
Tabela 7 – Medidas estatísticas associadas a distribuição dos valores de β apresentados na Figura 22.

Parâmetro	Média do Valor Ajustado	Desvio Padrão
β	0,8	$\pm 0,2$

Fonte do Autor

Assim, como fizemos no caso da DEC, listamos os 10 valores do parâmetro β associados com os mais altos valores do coeficiente de determinação R^2 , para os melhores ajuste aos dados. Esse valores estão listados na Tabela 8. Os valores dessa tabela foram

Figura 21 – Ajuste do histograma da média em ensemble da distribuição de razões entre espaçamentos consecutivos de autovalores desdobrados pela distribuição de Brody, com o valor de β informado na Tabela 6.



Fonte do Autor

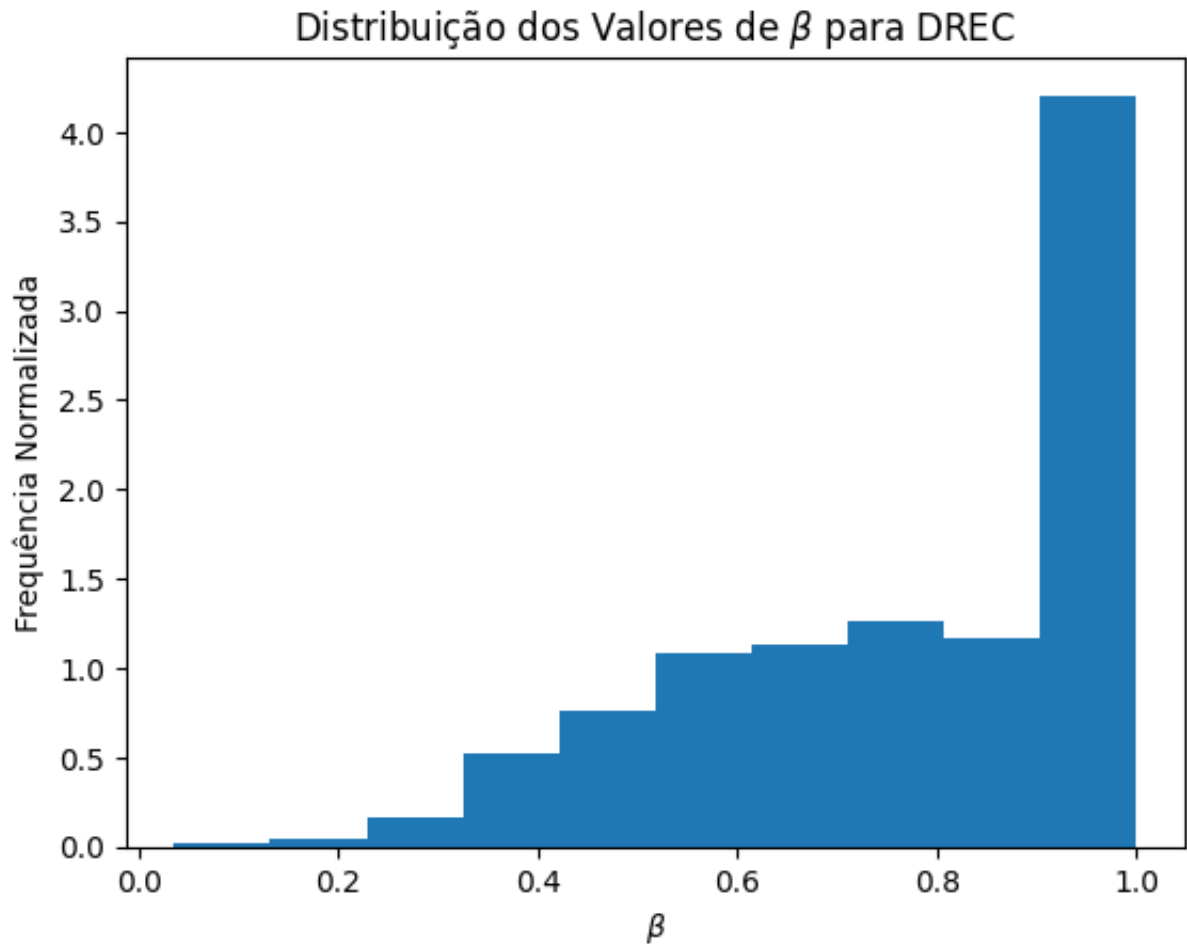
utilizados para ajustar os histogramas sobrepostos apresentados na Figura 23. O histograma mostra curvas que se ajustam aos dados, porém não estão tão bem alinhadas. Esse resultado é claramente inferior ao da DEC Figura 19.

4.4 Análise sobre os valores do parâmetro β

Em nossos resultados para DEC e DREC, o valor do parâmetro β foi estimado de duas maneiras. Primeiramente, por meio dos ajustes à média em ensemble para cada uma das distribuições e, em seguida, pela verificação de cada membro separadamente, depois calculada uma média desses resultados. O valor desse parâmetro é o principal parâmetro de ajuste dos dados aos modelos que os descrevem, portanto, essa quantidade é de grande importância para este estudo.

Para sintetizar os resultados anteriores e facilitar as comparações entre as diferentes

Figura 22 – Histograma da distribuição os valores de β para cada membro do ensemble considerando a DREC.



Fonte do Autor

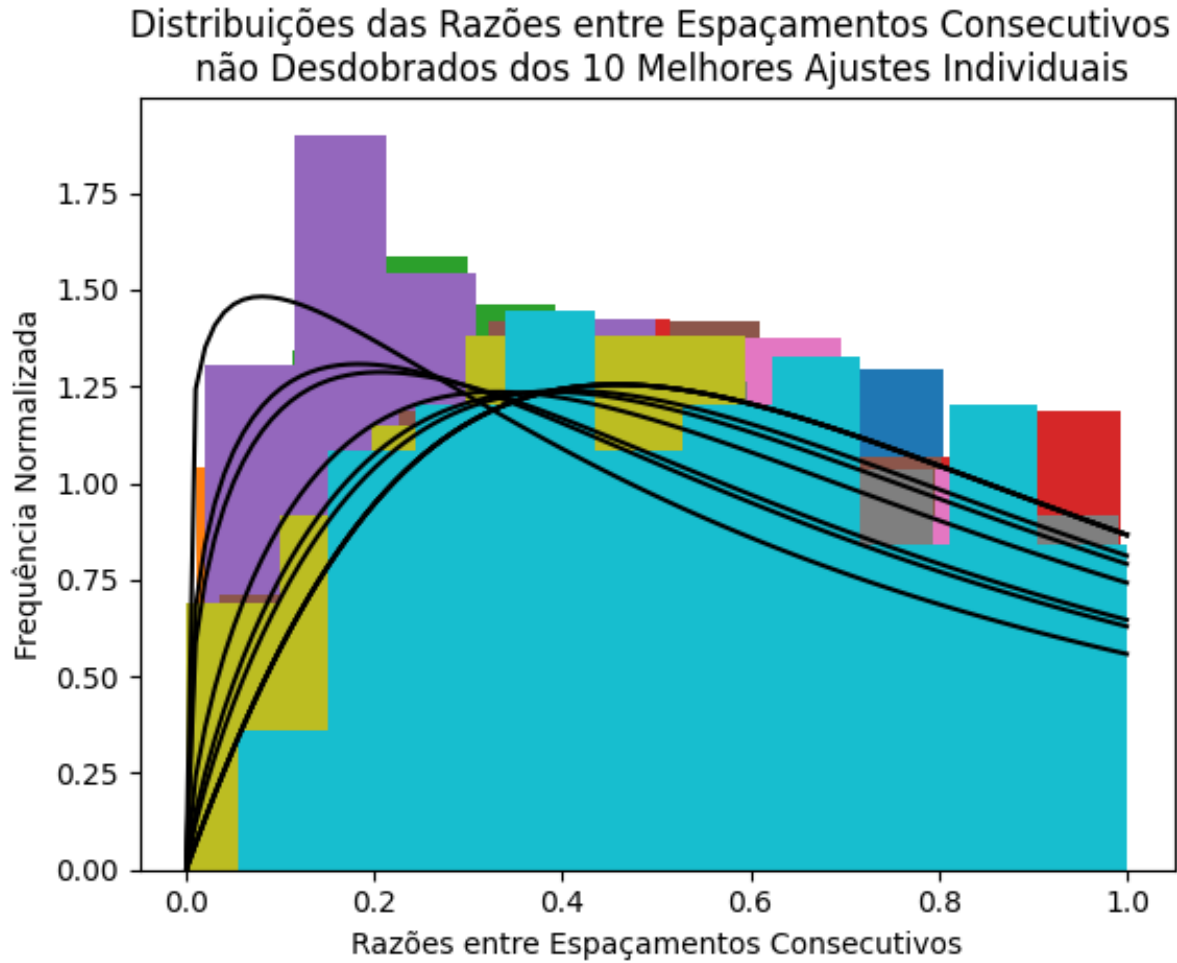
Tabela 8 – 10 melhores valores do parâmetro β para a DREC.

Ordem	R^2	β
1º	0,764	1,0
2º	0,758	0,4
3º	0,715	0,3
4º	0,706	1,0
5º	0,700	0,1
6º	0,696	0,8
7º	0,690	1,0
8º	0,687	0,8
9º	0,681	0,6
10º	0,676	1,0

Fonte do Autor

abordagens propostas, apresentamos a Tabela 9. Nessa tabela, β_s e $\bar{\beta}_s$ representam o valor de β obtidos pela DEC para a média em ensemble e para a análise individual dos membros do ensemble, respectivamente, assim como β_r e $\bar{\beta}_r$ para a DREC. Todos os resultados

Figura 23 – Ajustes realizados nos histogramas sobrepostos das razões entre espaçamentos consecutivos de autovalores não desdobrados, pela distribuição semi-empírica equação (3.8) utilizando os valores de β listados na Tabela 8.



Fonte do Autor

alcançados concordam estatisticamente, contudo, as análises das médias em ensemble β_s e β_r mostraram-se mais precisas. O comportamento das distribuições observadas em separadamente apresentam uma variabilidade considerável o que deixa impreciso estimar β que ajuste bem o modelo à distribuição dos dados. A Tabela 9 também indica que o valor do parâmetro tende a ser próximo a 1, sugerindo que ambas as distribuições seguem uma estatística intermediária, porém mais próxima de um GOE.

Tabela 9 – Valores estimados para o parâmetro β para DEC e DREC, por média em ensemble e análise individual de membros do ensemble.

	Valor Estimado	Desvio Padrão
β_s	0,89	$\pm 0,04$
$\bar{\beta}_s$	0,9	$\pm 0,2$
β_r	0,87	$\pm 0,04$
$\bar{\beta}_r$	0,8	$\pm 0,2$

Fonte do Autor

5 Conclusões

O estudo de redes cerebrais é, sem dúvida, desafiador, pois essas redes possuem uma enorme complexidade em muitos aspectos organizacionais, tanto em sua estrutura anatômica quanto em sua funcionalidade. Abordar esse sistema como uma rede de conexões de seus elementos se mostra um excelente caminho para simplificar e possibilitar manipulações estatísticas e matemáticas, facilitando assim o entendimento dessas redes. Isso pode proporcionar novos direcionamentos para questões ainda em aberto na neurociência.

Nesse sentido, usar a definição de conectoma para representar todas as conexões anatomicamente possíveis em uma rede e produzir sua matriz de conectividade permite a aplicação bem-sucedida de ferramentas algébricas. Isso viabiliza a implementação da análise espectral de matrizes de adjacência pela Teoria de Matrizes Aleatórias (TMA) nesse sistema. No entanto, é importante ressaltar que definir nós e arestas em uma rede com essas características é uma tarefa complexa.

Para evitar que conexões espúrias decorrentes do ruído no processo de medição dos sinais sejam levadas em conta na análise espectral subsequente, utilizamos um limiar de 10%. Portanto, as matrizes analisadas possuem apenas 10% das conexões mais fortes provenientes das matrizes iniciais. Após essa aplicação, realizamos uma redução do conjunto inicial de matrizes em aproximadamente 50%. As redes complexas do mundo real possuem muita degenerescência, principalmente de autovalores nulos, e, portanto, optamos por considerar apenas matrizes com 90 a 98 autovalores não degenerados. O tratamento realizado resultou em um subconjunto com aproximadamente 50% do número total de matrizes que tínhamos no início dessa análise. Esse subconjunto mostrou-se bem representativo em relação ao conjunto original e permitiu o cálculo das quantidades subsequentes. O ensemble obtido foi influenciado por essas escolhas empíricas. Investigações adicionais são necessárias para avaliar o uso de outros valores de limiar, bem como a influência da degenerescência do espectro em redes do mundo real.

As matrizes de adjacência foram construídas com sucesso, sendo capazes de refletir as propriedades estatísticas das redes por meio de seu espectro de autovalores sem degenerescência, explicando melhor as relações de conexões em comparação com a matriz de conectividade original.

Ao fim de todas as análises propostas neste trabalho, com base nos resultados obtidos, podemos concluir, que os autovalores associados a matriz adjacência dos dados estudados apresentam uma correlação significativa, com uma estatística muito mais próxima do GOE da TMA em comparação à estatística de Poisson. Todas as estimativas para o cálculo do parâmetro β , tanto para a DEC pela distribuição de Brody (3.8), como para a

DREC pelo modelo proposto (3.10), indicam valores próximos a um.

Os modelos utilizados para descrever a DEC e a DREC se ajustam bem aos dados, porém as análises sobre as média em ensemble tiveram resultados superiores ao caso individual de membros do ensemble, indicado que os valores obtidos por essa abordagem são as melhores estimativas ao valores de β para cada das distribuições. Assim temos $\beta_s = 0,89$ para DEC e $\beta_r = 0,87$ para DREC, como mostra a Tabela 9, como os valores que produzem o melhor ajuste aos dados.

Para a DEC, os resultados foram, em geral, mais precisos em comparação à DREC. Os valores do parâmetro β concordam em ambas as distribuições analisadas, tanto na análise da média em ensemble quanto pela distribuição de β obtida das análises individuais dos membros do ensemble. No entanto, a distribuição de β obtida do ajuste da DREC para cada membro do ensemble observou-se uma maior dispersão desse parâmetro em relação a média. Esse fato pode ser justificado pela dimensão da rede estudada ou por alguma outra característica dessa rede do mundo real.

Como perspectivas futuras, podemos indicar o estudo dos modelos trabalhados, tanto em termos de suas descrições matemáticas quanto pelo ajuste a dados de outras redes complexas de mundo real e problemas de matéria condensada. Além disso, novos modelos para as DEC e DREC que descrevam a fenomenologia dessas redes de forma mais precisa, assim como outras abordagens para o ensemble analisado, podem ser explorados.

Esperamos que as técnicas desenvolvidas e analisadas neste trabalho contribuam para novos estudos tanto na área de redes cerebrais quanto em outras ciências de redes.

Referências

AGUIAR, M. A. M. de; BAR-YAM, Y. Spectral analysis and the dynamic response of complex networks. *Physical Review E*, v. 71, Janeiro 2005. Citado 3 vezes nas páginas 21, 23 e 43.

ALBERT, R.; BARABÁSI, A.-L. Statistical mechanics of complex networks. *REVIEWS OF MODERN PHYSICS*, Indiana, v. 74, p. 47–97, Janeiro 2002. Citado 4 vezes nas páginas 21, 22, 31 e 40.

ALMEIDA, F. O. S.; SOUZA, A. M. C. Estudos sobre o ensemble de wishart de matrizes aleatórias. *Scientia Plena*, v. 5, 2009. Citado na página 31.

ANDERWALD, R.; GROND, L.; PÉREZ, M. A. G. Getting dizzy: A conversation between the artistic research of dizziness and somatic architecture. *The Journal of Somaesthetics*, v. 7, p. 59–82, 2021. Citado na página 12.

ATAS, Y. Y. et al. Distribution of the ratio of consecutive level spacings in random matrix ensembles. *Physical Review Letters*, v. 110, Fevereiro 2013. Citado 5 vezes nas páginas 31, 32, 35, 36 e 48.

BANDYOPADHYAY, J. N.; JALAN, S. Universality in complex networks: Random matrix analysis. *Physical Review E*, v. 76, Agosto 2007. Citado 9 vezes nas páginas 21, 32, 33, 34, 35, 36, 37, 43 e 44.

BARABÁSI, A.-L. *Linked-the new science of networks*. 1st. ed. [S.l.]: Perseus Books Group, 2002. Citado na página 22.

BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. *Science*, v. 286, p. 509–512, Outubro 1999. Citado 2 vezes nas páginas 21 e 22.

BARGMANN, C. I.; MARDER, E. From the connectome to brain function. *Nature Methods*, Maio 2013. Disponível em: <<https://www.nature.com/articles/nmeth.2451>>. Citado na página 15.

BEENAKKER, C. W. J. Random-matrix theory of quantum transport. *Reviews of Modern Physics*, v. 69, p. 731–808, 1997. Citado na página 32.

BRODMANN, K. *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. [s.n.], 1909. Disponível em: <<https://wellcomecollection.org/works/vrnkxtj>>. Citado na página 18.

BRODY, T. A. A statistical measure for the repulsion of energy levels. *Lettere Al Nuovo Cimento*, v. 7, N. 12, p. 482–484, Março 1973. Citado na página 36.

CAJAL, S. R. y. *Manual de histología normal y técnica micrográfica*. Valencia : Libreria de Pascual Aguilar, 1889. Disponível em: <<https://archive.org/details/b21779995/page/n41/mode/2up>>. Citado na página 18.

- CALDAS, B. R. da S. *Modelo do Voto da Maioria em rede direcionada de "Small-World-Voronoi-Delaunay"*. Dissertação (Mestrado) — Universidade Federal do Piauí, CCN, Programa de Pós-Graduação em Física, Teresina, 2007. Citado 2 vezes nas páginas 21 e 22.
- CIMINI, G. et al. The statistical physics of real-world networks. *Nature Reviews Physics*, v. 1, p. 58–71, 2019. Citado 2 vezes nas páginas 21 e 44.
- CORDES, D. et al. Mapping functionally related regions of brain with functional connectivity mr imaging. *American Journal of Neuroradiology*, v. 21, n. 9, p. 1636–1644, 2000. ISSN 0195-6108. Disponível em: <<https://www.ajnr.org/content/21/9/1636>>. Citado 3 vezes nas páginas 13, 15 e 29.
- DERKS, J. et al. Connectomic profile and clinical phenotype in newly diagnosed glioma patients. *NeuroImage: Clinical*, v. 14, p. 87–96, Janeiro 2017. Citado 3 vezes nas páginas 13, 14 e 38.
- DUARTE-FILHO, G. C. et al. General distribution for ratio of spacing levels of many-body hamiltonians. 2024. A ser publicado. Citado na página 48.
- ERDŐS, P.; RÉNYI, A. On random graphs. *Publicationes Mathematicae Debrecen*, v. 6, p. 290–297, 1959. Citado na página 24.
- ERDŐS, P.; RÉNYI, A. On the evolution of random graphs. *Mathematical Institute of the Hungarian Academy of Sciences*, 1960. Citado na página 21.
- ESPOSITO, F. et al. Independent component model of the default-mode brain function: Assessing the impact of active thinking. *Brain Research Bulletin*, v. 70, p. 263–269, 2006. Citado na página 14.
- ESSEN, D. C. V. et al. The human connectome project: A data acquisition perspective. *NeuroImage*, v. 62, p. 2222–2231, 2012. Citado na página 38.
- FORNITO, A.; ZALESKY, A.; BULLMORE, E. T. *Fundamentals of brain network analysis*. [S.l.]: Elsevier, 2016. ISBN 978-0-12-407908-3. Citado 19 vezes nas páginas 12, 16, 17, 18, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 38, 40, 41, 42 e 43.
- FRANSSON, P. How default is the default mode of brain function? further evidence from intrinsic bold signal fluctuations. *Neuropsychologia*, v. 44, p. 2836–2845, Agosto 2006. Citado na página 40.
- GANESAN, K.; LAKSHMANAN, M. Applicability of brody distribution in the study of quantum chaos of the hydrogen atom in a generalized van der waals potential. *Journal of Physics B: Atomic, Molecular and Optical Physics*, v. 27, p. 2809—2820, 1994. Citado na página 36.
- GEERLIGS, L.; CAM-CAN; HENSON, R. N. Functional connectivity and structural covariance between regions of interest can be measured more accurately using multivariate distance correlation. *NeuroImage*, v. 135, p. 16–31, 2016. Citado 2 vezes nas páginas 12 e 14.
- GÁLVEZ, M. A. *Espacio somático. Cuerpos múltiples*. [S.l.]: Ediciones Asimétricas, 2019. ISBN 9788417905163. Citado na página 12.

- HAHN, G. J. The coefficient of determination exposed! Outubro 1973. Citado na página 49.
- HARMS, M. P. et al. Extending the human connectome project across ages: Imaging protocols for the lifespan development and aging projects. *NeuroImage*, v. 183, p. 972–984, 2018. Citado na página 38.
- HCP. *Human Connectome Project*. 2024. <<https://www.humanconnectome.org/>>. Acessado em 17 de junho de 2024. Citado na página 16.
- JALAN, S.; BANDYOPADHYAY, J. N. Random matrix analysis of complex networks. *Physical Review E*, v. 76, Outubro 2007. Citado 5 vezes nas páginas 26, 32, 33, 35 e 36.
- JALAN, S.; BANDYOPADHYAY, J. N. Randomness of random networks: A random matrix analysis. *EPL*, v. 87, 2009. Citado na página 49.
- KANDEL, E. R. et al. *Principles of Neural Science*. 5th. ed. [S.l.]: McGraw Hill, 2013. ISBN 978-0071390118. Citado na página 12.
- KOPELL, N. J. et al. Beyond the connectome: The dynamome. *Neuron*, Setembro 2014. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4169213/>>. Citado 2 vezes nas páginas 17 e 19.
- LMFIT. *Non-Linear Least-Squares Minimization and Curve-Fitting for Python*. 2024. <<https://lmfit.github.io/lmfit-py/>>. Acessado em 02 de julho de 2024. Citado na página 49.
- LUO, F. et al. Application of random matrix theory to microarray data for discovering functional gene modules. *Physical Review E*, v. 73, 2006. Citado na página 32.
- LÓPEZ-MUÑOZ, F.; BOYA, J.; ALAMO, C. Neuron theory, the cornerstone of neuroscience, on the centenary of the nobel prize award to santiago ramón y cajal. *Brain Research Bulletin*, v. 70, p. 391–405, 2006. Citado na página 18.
- MAKRIS, N. et al. A proposed human structural brain connectivity matrix in the center for morphometric analysis harvard-oxford atlas framework: A historical perspective and future direction for enhancing the precision of human structural connectivity with a novel neuroanatomical typology. *Developmental Neuroscience*, Março 2023. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10526721/>>. Citado na página 13.
- MARKOV, N. T. et al. The role of long-range connections on the specificity of the macaque interareal cortical network. *Proceedings of the National Academy of Sciences*, v. 110, p. 5187–5192, 2013. Citado na página 41.
- MEHTA, M. L. *Random Matrices*. Third edition. [S.l.]: Academic Press, Elsevier, 2004. (Pure and Applied Mathematics 142). ISBN 978-0-12-088409-4. Citado 2 vezes nas páginas 31 e 32.
- NEWMAN, M. The structure and function of complex networks. *SIAM REVIEW*, v. 45, p. 167–256, 2003. Citado na página 21.
- NEWMAN, M. *Networks: An Introduction*. [S.l.]: Oxford University Press, 2010. ISBN 0199206651, 9780199206650. Citado 4 vezes nas páginas 20, 22, 42 e 43.

- NIEMINEN, J. M.; MUCHE, L. A random matrix model whose eigenvalue spacings are closely described by the brody distribution. *Acta Physica Polonica Series B*, v. 48, N. 4, p. 765–772, Abril 2017. Citado 2 vezes nas páginas 32 e 35.
- OGANESYAN, V.; HUSE, D. A. Localization of interacting fermions at high temperature. *Physical Review B*, v. 75, 2007. Citado 3 vezes nas páginas 35, 47 e 48.
- PLEROU, V. et al. Random matrix approach to cross correlations in financial data. *Physical Review E*, v. 65, 2002. Citado na página 33.
- RAHMAN, A.; MUKTADIR, G. Spss: An imperative quantitative data analysis tool for social science research. *International Journal of Research and Innovation in Social Science*, v. 5, 2021. Citado na página 49.
- RIEDL, V. et al. Metabolic connectivity mapping reveals effective connectivity in the resting human brain. *Proceedings of the National Academy of Sciences*, Julho 2015. Citado 4 vezes nas páginas 12, 13, 14 e 15.
- SABRI, H. et al. Generalization of brody distribution for statistical investigation. random matrices: Theory and applications. v. 3, 2014. Citado na página 32.
- SANTOS, F. A. N. et al. Topological phase transitions in functional brain networks. *Physical Review E*, v. 100, Setembro 2019. Citado 2 vezes nas páginas 14 e 23.
- SANTOS, L. F.; RIGOL, M. Localization and the effects of symmetries in the thermalization properties of one-dimensional quantum systems. *Physical Review E*, v. 82, Setembro 2010. Citado na página 48.
- SARKAR, C.; JALAN, S. Randomness and structure in collaboration networks: A random matrix analysis. *IEEE Transactions on Computational Social Systems*, v. 3, p. 132—138, 2016. Citado na página 32.
- SARKAR, C.; JALAN, S. Spectral properties of complex networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 2018. Citado 3 vezes nas páginas 35, 36 e 44.
- SARSTEDT, M.; MOOI, E. *A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics*. Second edition. [S.l.]: Springer, 2014. Citado na página 49.
- SILVA, P. H. R. da. *Brain structural and functional networks: from connections to cognition. A study applied to Cerebral Small Vessel Disease*. Tese (Doutorado) — Universidade de São Paulo, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Programa de Pós-Graduação em Física Aplicada à Medicina e Biologia, 2022. Citado na página 13.
- SILVA, R. da. *Distribuição de autovalores de matrizes aleatórias*. Dissertação (Mestrado) — Universidade de São Paulo, Instituto de Física, São Paulo, 2000. Citado na página 31.
- SPEICHER, R. *Random Matrices*. 2020. Disponível em: <<https://arxiv.org/pdf/2009.05157>>. Lecture notes, Saarland University, Faculty of Mathematics and Computer Science, Department of Mathematics. Citado na página 33.

- SPORNS, O. *Networks of the Brain*. 1. ed. [S.l.]: The MIT Press, 2010. ISBN 0262014696,9780262014694. Citado 7 vezes nas páginas 13, 14, 21, 23, 24, 27 e 28.
- SPORNS, O. Network attributes for segregation and integration in the human brain. *Current Opinion in Neurobiology*, v. 23, p. 162—171, 2013. Citado na página 15.
- SPORNS, O.; TONONI, G.; KÖTTER, R. The human connectome: A structural description of the human brain. *PLoS Computational Biology*, 2005. Disponível em: [<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1239902/>](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1239902/). Citado 4 vezes nas páginas 15, 16, 17 e 18.
- STROGATZ, D. J. W. . S. H. Collective dynamics of ‘small-world’ networks. *Nature*, p. 440–442, Junho 1998. Citado 2 vezes nas páginas 22 e 26.
- SUÁREZ, L. E. et al. Linking structure and function in macroscale brain networks. *Trends in Cognitive Sciences*, v. 24, p. 302–315, 2020. Citado 2 vezes nas páginas 13 e 14.
- TONONI, G.; SPORNS, O.; EDELMAN, G. M. A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences*, v. 91, p. 5033–5037, 1994. Citado na página 15.
- TRIARHOU, L. C. Pre-brodmann pioneers of cortical cytoarchitectonics i: Theodor meynert, vladimir betz and william bevan-lewis. *Brain Structure and Function*, v. 226, p. 49–67, 2020. Citado na página 18.
- UDDIN, L. Q. et al. Functional connectivity of default mode network components: Correlation, anticorrelation, and causality. *Proceedings of the National Academy of Sciences*, v. 30, p. 625–637, Janeiro 2009. Citado 2 vezes nas páginas 15 e 40.
- WEISSENBACHER, A. et al. Correlations and anticorrelations in resting-state functional connectivity mri: A quantitative comparison of preprocessing strategies. *NeuroImage*, v. 47, p. 1408–1416, Outubro 2009. Disponível em: <https://www.sciencedirect.com/science/article/pii/S105381190900487X>. Citado na página 12.
- WHITE, J. G. et al. The structure of the nervous system of the nematode caenorhabditis elegans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, v. 314, p. 1–340, 1986. Citado na página 41.
- WIGNER, E. P. On the statistical distribution of the widths and spacings of nuclear resonance levels. *Mathematical Proceedings of the Cambridge Philosophical Society*, v. 47, p. 790–798, 1951. Citado na página 31.
- WIGNER, E. P. Characteristic vectors of bordered matrices with infinite dimensions. *Annals of Mathematics*, Mathematics Department, Princeton University, v. 62, p. 548–564, Novembro 1955. Citado na página 31.
- WIGNER, E. P. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, Mathematics Department, Princeton University, v. 67, p. 325–327, Novembro 1958. Citado na página 31.
- WIJK, B. C. M. van; STAM, C. J.; DAFFERTSHOFER, A. Comparing brain networks of different size and connectivity density using graph theory. *PLoS ONE*, v. 5, p. 1–13, Outubro 2010. Citado 3 vezes nas páginas 12, 26 e 42.

ZHANG, D. A coefficient of determination for generalized linear models. *The American Statistician*, v. 71, p. 310–316, 2016. Citado na página 49.

ZILLES, K. Brodmann: a pioneer of human brain mapping—his impact on concepts of cortical organization. *Brain*, v. 141, p. 3262–3278, 2018. Citado na página 18.