



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA
DOUTORADO EM FÍSICA

MARCIO MACIEL BASTOS

ANÁLISE TOPOLÓGICA DE REDES Q-EXPONENCIAIS ALEATÓRIAS

FORTALEZA

2024

MARCIO MACIEL BASTOS

ANÁLISE TOPOLÓGICA DE REDES Q-EXPONENCIAIS ALEATÓRIAS

Tese apresentada ao Programa de Pós-Graduação em Física do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Física. Área de Concentração: FÍSICA DA MATÉRIA CONDENSADA.

Orientador: Prof. Dr. José Soares de Andrade Júnior.

FORTALEZA

2024

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

B329a Bastos, Marcio Maciel Bastos.

Análise Topológica de Redes q-Exponenciais Aleatórias / Marcio Maciel Bastos Bastos. – 2024.
113 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Física, Fortaleza, 2024.

Orientação: Prof. Dr. José Soares de Andrade Júnior.

1. q-Exponencial. 2. Redes. 3. Grafos. 4. Probabilidades. 5. Percolação. I. Título.

CDD 530

MARCIO MACIEL BASTOS

ANÁLISE TOPOLÓGICA DE REDES Q-EXPONENCIAIS ALEATÓRIAS

Tese apresentada ao Programa de Pós-Graduação em Física do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Física. Área de Concentração: FÍSICA DA MATÉRIA CONDENSADA.

Aprovada em: 08 de Fevereiro de 2024

BANCA EXAMINADORA

Prof. Dr. José Soares de Andrade
Júnior (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. André Auto Moreira
Universidade Federal do Ceará (UFC)

Prof. Dr. Cesar I. N. Sampaio Filho
Universidade Federal do Ceará (UFC)

Prof. Dr. Erneson Alves de Oliveira
Universidade de Fortaleza (UNIFOR)

À minha família, por sua capacidade de acreditar em mim e investir em mim. Mãe, seu cuidado e dedicação foi que deram, em alguns momentos, a esperança para seguir. Pai, sua presença significou segurança e certeza de que não estou sozinho nessa caminhada.

AGRADECIMENTOS

À minha querida esposa Sara Bastos, que foi a base sólida em todos os momentos desta trajetória. Seu apoio incondicional, compreensão e amor foram essenciais para que eu pudesse seguir em frente, tomar decisões importantes e enfrentar os desafios com determinação. Sua presença constante e encorajamento foram verdadeiras fontes de inspiração e força.

Ao meu orientador, professor José Soares de Andrade Júnior, cuja confiança no meu potencial e estímulo constante foram cruciais para o meu desenvolvimento e progresso. Sua orientação sábia e apoio contínuo me permitiram explorar novos horizontes e ultrapassar os limites do meu conhecimento e habilidades.

Ao professor André Auto Moreira, sou grato por ter iluminado os caminhos do conhecimento com clareza, transformando conceitos que antes me pareciam abstratos em entendimentos concretos e aplicáveis. Sua habilidade em tornar o complexo acessível mudou profundamente a minha maneira de ver e compreender o nosso campo de estudo.

E ao professor César Ivan Nunes Sampaio Filho, minha gratidão é imensa por sua paciência inigualável e dedicação em me ensinar as nuances da apresentação de dados de forma clara, coerente e padronizada. Seu comprometimento com a excelência e sua disposição em compartilhar seu conhecimento foram fundamentais para aprimorar minha expressão e comunicação científica.

A cada um de vocês, meu sincero agradecimento por fazerem parte desta jornada, contribuindo de maneira única e especial para o meu crescimento pessoal e acadêmico. Esta conquista não seria possível sem o valioso suporte e orientação que recebi ao longo deste caminho.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

“Desde os menores grãos de areia até as maiores montanhas, desde o menor inseto até o maior animal, desde o corpo inanimado e sem vida até as estrelas cujos movimentos percorremos, desde as aparências visíveis até as essências ocultas de coisas: tudo está sujeito a leis eternas e imutáveis.”

(Leonard Euler)

RESUMO

Esta tese explora a relevância e as propriedades fundamentais de redes com distribuição de grau q -exponencial, uma temática que permeia uma variedade de domínios, incluindo física de plasmas, econometria, biofísica e mais. A congruência entre experimentos, análises numéricas e modelos teóricos com a distribuição q -exponencial é notavelmente observada, particularmente em redes complexas tanto empíricas quanto modeladas. Ao contrário de abordagens prévias, que geralmente se concentravam em modelos de crescimento com associação preferencial, este trabalho se utiliza do modelo de configuração para gerar redes aleatórias. Esta metodologia permite uma avaliação mais acurada das propriedades inerentes a estas redes, desvinculadas de processos de crescimento específicos. A distribuição q -exponencial é caracterizada por seus dois parâmetros, q e λ , e serve como uma extensão generalizada da tradicional distribuição de lei de potência. A tese discute como a distribuição transita entre um comportamento de lei de potência para grandes valores de k (grau do nó) e uma distribuição de patamar para valores menores de k , com o parâmetro λ^{-1} desempenhando um papel crucial nessa transição. Este estudo destaca que variações do comportamento padrão de lei de potência em graus menores podem induzir alterações estruturais significativas nas redes. Tais mudanças têm o potencial de afetar profundamente as características e os processos fundamentais nas redes. As redes geradas sob o modelo de configuração revelam propriedades topológicas distintas, diretamente derivadas da natureza q -exponencial de sua distribuição de grau. Investiga-se detalhadamente características como assortatividade, caminho mínimo médio e robustez a falhas aleatórias e ataques direcionados. Os resultados obtidos sugerem que as redes q -exponenciais oferecem maior robustez em comparação às redes livres de escala convencionais, particularmente em cenários de ataques maliciosos. Além disso, a análise da decomposição de k -core nas redes q -exponenciais revela uma estrutura núcleo mais extensa e robusta do que aquela encontrada em redes de lei de potência pura.

Palavras-chave: q -Exponencial; Redes; Grafos; Probabilidades; Percolação.

ABSTRACT

This thesis explores the relevance and fundamental properties of networks with q-exponential degree distribution, a theme that pervades a variety of domains, including plasma physics, econometrics, biophysics, and more. The congruence between experiments, numerical analyses, and theoretical models with the q-exponential distribution is notably observed, particularly in both empirical and modeled complex networks. Unlike previous approaches that typically focused on growth models with preferential attachment, this work employs the configuration model to generate random networks. This methodology allows for a more accurate assessment of the inherent properties of these networks, detached from specific growth processes. The q-exponential distribution is characterized by its two parameters, q and λ , and serves as a generalized extension of the traditional power-law distribution. The thesis discusses how the distribution transitions from a power-law behavior for large values of k (node degree) to a plateau distribution for smaller values of k , with the parameter λ^{-1} playing a crucial role in this transition. This study highlights that deviations from the standard power-law behavior at smaller degrees can induce significant structural changes in networks. Such changes have the potential to profoundly affect the characteristics and fundamental processes within the networks. Networks generated under the configuration model reveal distinct topological properties, directly derived from the q-exponential nature of their degree distribution. Characteristics such as assortativity, shortest average path, and robustness against random failures and directed attacks are investigated in detail. The results suggest that q-exponential networks offer greater robustness compared to conventional scale-free networks, particularly in scenarios of malicious attacks. Furthermore, the analysis of the k-core decomposition reveals a more extensive and robust core structure in q-exponential networks than in pure power-law networks.

Keywords: q-Exponential; Networks; Graphs; Probabilities; Percolation.

LISTA DE FIGURAS

- Figura 1 – O grafo que consiste em seis vértices (A, B, C, D, E, F) e sete arestas que os conectam. A matriz de adjacência correspondente é a matriz 6x6 definida em (2.2), onde um '1' indica a presença de uma aresta entre dois nós e um '0' indica a ausência dela. 19
- Figura 2 – Dois Grafos Isomorfos: ambos os grafos têm quatro vértices e cinco arestas, com a mesma estrutura de conexão. O "Grafo 1" tem um arranjo onde um dos vértices está no centro conectado a todos os outros três vértices, enquanto no "Grafo 2", os vértices estão dispostos de forma que cada um esteja conectado a dois outros vértices, formando uma estrutura quadrada. Apesar de suas disposições diferentes, eles são isomorfos, pois há uma correspondência um-para-um entre seus vértices e arestas que preserva a estrutura de conexão. 22
- Figura 3 – Exemplo de Contração de Aresta em um Grafo: o "Grafo Original" contém quatro vértices (A, B, C, D) e cinco arestas. A contração é realizada na aresta que conecta os vértices A e B. No "Grafo Após Contração", os vértices A e B são fundidos em um único vértice (AB), e a aresta que os conectava é removida. Outras conexões são mantidas, refletindo as propriedades estruturais básicas do grafo original. 35
- Figura 4 – Neste grafo, o vértice "C" é destacado (preenchido em cinza) para indicar sua posição como o ponto de maior centralidade de intermediação. O vértice "C" está conectado a todos os outros vértices do grafo, atuando como um intermediário crucial no fluxo entre eles. Isso ilustra visualmente que a centralidade de intermediação de "C" é maior, pois ele se situa em muitos dos caminhos mais curtos entre os pares de vértices no grafo. Centralidade de cada vértice: A (0.1), B (0.1), C (0.6), D (0.1), E (0.05), F (0.05). 42

- Figura 5 – Esta figura ilustra uma pequena rede que exemplifica a presença de um 4-core e, simultaneamente, três 4-componentes distintos. A rede, como um todo, representa um único 4-core, pois cada um de seus vértices mantém conexão com pelo menos outros quatro vértices. Contudo, dentro desta mesma rede, identificam-se três 4-componentes independentes, o que é evidenciado pelos círculos sombreados em cinza. Este exemplo demonstra claramente que os conceitos de k -cores e k -componentes são distintos. É importante observar também que cada um dos 4-componentes constitui um clique dentro do grafo, reforçando a ideia de que um clique representa um caso especial de densidade de conexão entre os vértices. 55
- Figura 6 – Esta figura compara as funções de densidade de probabilidade de três distribuições: a distribuição q -exponencial com parâmetros $\lambda = 0.01$ e $q = 1.4$, a lei de potência com parâmetro $\gamma = 2.5$, e a exponencial com parâmetro $\lambda = 0.01$. É interessante notar que a distribuição q -exponencial é versátil: ao fazer q tender a 1, ela assume um perfil semelhante à distribuição exponencial com $\lambda = 0.01$. Alternativamente, quando λ se aproxima do infinito, a q -exponencial se assemelha mais à lei de potência com parâmetro $\gamma = 2.5$. . 76
- Figura 7 – Distribuição de grau de redes q -exponenciais (símbolos) comparada com a distribuição esperada da equação (4.1) (linhas contínuas) para $q = 1,4$ e para $\lambda = 0.01$ (estrelas azuis), $\lambda = 0.1$ (triângulos verdes), $\lambda = 1$ (quadrados vermelhos) e $\lambda = 100$ (círculos pretos). O detalhe mostra uma comparação entre as distribuições de grau de uma q -exponencial com $\lambda = 100$ e uma distribuição puramente livre de escala (linha tracejada preta) com $q = 1.4$ ($\lambda = 2.5$). Esses resultados são obtidos para redes com tamanho $N = 500000$ pela média de 100 amostras. Como pode ser visto, para pequenos valores de λ , as distribuições atingem um platô em graus pequenos e o regime de lei de potência se torna maior com o aumento de λ 82

- Figura 8 – Dependência do grau médio $\langle k \rangle$ no parâmetro λ para diferentes valores de q . Estas curvas correspondem a $q = 1.4$ (círculos pretos), $q = 1.33$ (quadrados vermelhos) e $q = 1.25$ (estrelas azuis). Para valores de $\lambda < 1$, o grau médio segue $\langle k \rangle \sim \lambda^{-1}$, conforme esperado para distribuições q -exponenciais. No limite de lei de potência pura ($\lambda \gg 1$), o grau médio satura em um valor independente de λ 83
- Figura 9 – Dependência do grau médio do vizinho mais próximo k_{nn} dos vértices no grau k para redes com tamanho $N = 10^7$ nós, $q = 1.4$, e $\lambda = 100$ (círculo preto), $\lambda = 1$ (quadrado vermelho), $\lambda = 0.1$ (triângulo verde) e $\lambda = 0.01$ (estrela azul). Os patamares de cada curva correspondem ao valor $\langle k^2 \rangle / \langle k \rangle$ (Catanzaro *et al.*, 2005). Todas as curvas apontam para um comportamento intrinsecamente dissortativo para valores suficientemente grandes de k 83
- Figura 10 – Dependência do tamanho no comprimento médio do menor caminho $\langle l \rangle$ para $q = 1.33$ e para $\lambda = 0.01$ (estrelas azuis), $\lambda = 0.1$ (triângulos verdes), $\lambda = 1$ (quadrados vermelhos) e $\lambda = 1$ (círculos pretos). Os resultados são obtidos através da média de 105, 104, 104, 104, 104, 103 e 103 amostras de tamanho $N = 5000, 10000, 20000, 40000, 80000, 160000$ e 320000 , respectivamente. 85
- Figura 11 – O pré-fator α como função de q para $\lambda = 100$ (círculos pretos), $\lambda = 1$ (quadrados vermelhos) e $\lambda = 0.1$ (triângulos verdes). Cada ponto é obtido de um ajuste de mínimos quadrados aos dados, $\alpha \sim \log_{10} N$, com barras de erro menores que os símbolos. As linhas contínuas representam guias visuais. . . . 85
- Figura 12 – A densidade de nós no maior aglomerado $S(f)$ como função da fração f de nós removidos para falhas aleatórias em redes q -exponenciais com $q = 1.25$, $\lambda = 0.01$ (estrelas azuis), $\lambda = 0.1$ (triângulos verdes), $\lambda = 1$ (quadrados vermelhos) e $\lambda = 100$ (círculos pretos). Estes resultados correspondem a redes de tamanho $N = 320000$ por média de 200 amostras. As linhas contínuas para $\lambda = 0.01, 0.1, 1$ representam guias visuais, enquanto para $\lambda = 100$, a linha sólida preta corresponde ao resultado para ataques aleatórios em uma rede puramente livre de escala. O detalhe mostra a fração crítica f_c determinada pelos critérios de Molloy-Reed (Molloy e Reed, 1995). 86

Figura 13 – (a) A densidade de nós no maior *cluster* $S(f)$ como função da fração f de nós removidos para ataques maliciosos direcionados pelo maior grau, para $q = 1.4$ e $\lambda = 0.01$ (estrelas azuis), $\lambda = 0.1$ (triângulos verdes), $\lambda = 1$ (quadrados vermelhos) e $\lambda = 100$ (círculos pretos). As linhas contínuas para $\lambda = 0.01, 0.1$ e 1 representam guias para o olho, enquanto para $\lambda = 100$ a linha contínua preta corresponde ao resultado para ataques maliciosos em uma rede puramente livre de escala. Esses resultados correspondem a redes de tamanho $N = 320000$ pela média de 200 amostras. (b) A densidade $S(f)$ como função da fração f de redes geradas com $q = 1.4$ e submetidas a ataques maliciosos direcionados pelo maior grau. Os resultados correspondem a redes livres de escala com $k_{min} = 2$ e $\langle k \rangle = 6$ (círculos pretos), $k_{min} = 8$ e $\langle k \rangle = 24$ (estrelas azuis) e para redes q-exponenciais com $\lambda = 10, k_{min} = 2$ e $\langle k \rangle = 25$ (quadrados vermelhos). Eles foram obtidos com redes de tamanho $N = 10^6$ pela média de 100 amostras. 88

Figura 14 – A fração crítica f_c para ataques maliciosos como função de q para $\lambda = 0.1$ (triângulos verdes), $\lambda = 1$ (quadrados vermelhos) e $\lambda = 100$ (círculos pretos). Os resultados são obtidos para redes de tamanho $N = 500000$ pela média de 2000 amostras. A fração a partir da qual o maior *cluster* não obedece ao critério de Molloy-Reed é a fração crítica f_c . As redes q-exponenciais com valores menores de λ são claramente mais robustas, já que uma maior fração de nós precisa ser removida para atingir o ponto crítico. As linhas contínuas representam guias para o olho. 89

Figura 15 – (a) k-core mais alto e (b) massa do k-core mais alto versus N para $q = 1.4$ e $\lambda = 0.01$ (estrelas azuis), $\lambda = 0.1$ (triângulos verdes), $\lambda = 1$ (quadrados vermelhos) e $\lambda = 100$ (círculos pretos). Os resultados são obtidos pela média de 105, 104, 104, 104, 104, 103 e 103 amostras de tamanho $N = 5000, 10000, 20000, 40000, 80000, 160000$ e 320000 , respectivamente. 91

Figura 16 – Representação do Conjunto de Cantor em quatro iterações. Cada linha representa uma iteração no processo de construção do conjunto de Cantor, com intervalos centrais removidos progressivamente para criar um padrão fractal. 106

LISTA DE ABREVIATURAS E SIGLAS

CDF *Cumulative Distribution Function*

PMF *Probability Mass Function*

LISTA DE SÍMBOLOS

| | |
|---------------------------------|---|
| G | Grafo |
| V | Conjunto de vértices de um grafo |
| E | Conjunto de arestas de um grafo |
| $G(V, E)$ | Grafo formado pelo conjuntos V e E |
| $[V]^2$ | Conjunto de todos os pares de vértices de um grafo. |
| $L(v)$ | Lista de vértices adjacentes ao vértice $v \in V$ |
| A | Matriz de adjacência de um grafo |
| A_{ij} | Entrada ij da matriz de adjacência A |
| $ G $ ou $G(V)$ | Número de vértices do grafo G (ordem de G) |
| $\ G\ $ ou $G(E)$ | Número de arestas do grafo G (tamanho de G) |
| \overline{G} | Grafo complementar de G |
| $N_G(v)$ ou $N(v)$ | Conjunto de todos os vértices adjacentes a $v \in V$ |
| $d(v)$ ou k_v | Grau de um vértice $v \in V$ |
| $\delta(G)$ | Menor grau de G |
| $\Delta(G)$ | Maior grau de G |
| $d(G)$ ou $\langle k \rangle$ | Grau médio do grafo não direcionado G |
| $d_e(G)$ ou $\langle k \rangle$ | Grau médio de entrada do grafo direcionado G |
| $d_s(G)$ ou $\langle k \rangle$ | Grau médio de saída do grafo direcionado G |
| ρ | Densidade do grafo |
| $\kappa(G)$ | Nível máximo de conectividade que o grafo G |
| $\lambda(G)$ | Conectividade por arestas do grafo G |
| $C_D(v)$ | Centralidade de grau de $v \in V$ |
| $C_A(v)$ | Centralidade de autovetor de $v \in V$ |
| $C_K(v)$ | Centralidade de Katz de $v \in V$ |
| ℓ_{ij} | Menor distância entre os nós i e j |
| ℓ_i | Distância média entre o nó i e os demais nós de G |

| | |
|------------------------|--|
| $\langle \ell \rangle$ | Caminho geodésico médio |
| C_i | Centralidade de proximidade do nó i |
| $C_B(v)$ | Centralidade de proximidade de $v \in V$ |
| C | Coefficiente de agrupamento |
| δ_{ij} | Delta de Kronecker para os elementos i e j |
| S_q | Entropia generalizada de Tsallis |
| \mathbb{R} | Conjunto dos Números Reais |
| W | Número total de estados microscópicos do sistema |
| p_i | Probabilidade associada ao i -ésimo estado do sistema |
| q | Parâmetro livre da Entropia generalizada |
| Z_q | Função de partição generalizada |
| λ | Parâmetro das distribuições exponencial e q -exponencial |
| $e_q(x)$ | Função q -exponencial |
| $T(x)$ | $e_q(x)^{(q-1)}$ |
| Ω | Espaço amostral |
| \mathcal{A} | σ -álgebra de eventos |
| \mathbb{P} | Medida de Probabilidade em \mathcal{A} |
| $\mathit{supp}(f)$ | Suporte da função f |

SUMÁRIO

| | | |
|-----|--|-----|
| 1 | INTRODUÇÃO | 16 |
| 2 | FUNDAMENTAÇÃO TEÓRICA | 19 |
| 2.1 | Teoria dos Grafos | 19 |
| 2.2 | Percolação | 48 |
| 2.3 | Redes Complexas | 56 |
| 2.4 | Mecânica estatística de redes complexas | 70 |
| 3 | METODOLOGIA | 78 |
| 3.1 | Geração dos grafos | 78 |
| 3.2 | Implementação | 79 |
| 4 | REDES ALEATÓRIAS COM DISTRIBUIÇÃO Q -EXPONENCIAL . . | 81 |
| 4.1 | Resultados e discussão | 82 |
| 5 | CONCLUSÕES | 92 |
| | REFERÊNCIAS | 93 |
| | APÊNDICE A –PROBABILIDADE E VARIÁVEIS ALEATÓRIAS . . . | 97 |
| | APÊNDICE B –MOMENTOS DE UMA DISTRIBUIÇÃO DE PROBA- BILIDADE | 103 |
| | APÊNDICE C –FRACTAIS | 105 |
| | APÊNDICE D –GERAÇÃO DE NÚMEROS ALEATÓRIOS SEGUNDO A DISTRIBUIÇÃO Q -EXPONENCIAL | 111 |

1 INTRODUÇÃO

Em diversas disciplinas científicas—abrangendo desde turbulência (Beck *et al.*, 2001), difusão anômala em plasmas (Liu e Goree, 2008), estatísticas de raios cósmicos (Beck, 2004), econometria (Souza *et al.*, 2006), (Borges, 2004) e biofísica (Upadhyaya *et al.*, 2001)—existem experimentos, resultados numéricos e modelos teóricos que se alinham notavelmente bem com q -exponenciais. Este padrão é particularmente prevalente no domínio dos sistemas complexos, encapsulado no estudo da mecânica estatística, fractais e teoria matemática dos grafos.

Pesquisas anteriores exploraram o campo das redes complexas, com ênfase na análise de vértices que seguem diversas distribuições de grau, como expostos por (Brito *et al.*, 2019) e (Albert e Barabási, 2002). Esses estudos adotaram predominantemente o método do Modelo de Associação Preferencial (*Preferential Attachment*). Contudo, observa-se uma lacuna significativa: embora algumas redes geradas por esse método apresentem distribuições de graus de vértices assemelhando-se à q -exponencial, falta uma comprovação rigorosa que confirme a aderência efetiva a essa distribuição específica. Neste contexto, a presente tese visa estabelecer um marco distintivo ao construir e analisar minuciosamente redes aleatórias caracterizadas por terem seus vértices inequivocamente regidos por uma distribuição de grau q -exponencial, empregando para isso o modelo de configuração.

A distribuição q -exponencial, definida pela função (1.1), oferece um quadro único para a análise de redes complexas. Para graus de vértices grandes ($k \gg \lambda^{-1}$), a distribuição se aproxima de um decaimento de lei de potência caracterizado por $k^{-\gamma}$, com $\gamma = 1/(q - 1)$. Este aspecto adiciona uma camada adicional de complexidade para o entendimento dessas redes, e tem mostrado ter implicações significativas tanto para seu comportamento topológico quanto dinâmico.

$$p_q(k) = (2 - q)\lambda [1 - (1 - q)\lambda k]^{-1/(q-1)} \quad (1.1)$$

Dada a ubiquidade de sistemas complexos em fenômenos naturais e a relevância de seus modelos matemáticos, esta pesquisa é de particular interesse para físicos especializados em mecânica estatística, fractais e sistemas complexos, bem como para matemáticos na área de teoria dos grafos. A originalidade deste trabalho sublinha ainda mais sua importância: até onde

sabemos, este é o primeiro estudo abrangente de redes aleatórias geradas com uma distribuição de grau q -exponencial confirmada.

Por meio desta tese, pretendemos oferecer um arcabouço analítico robusto que possa não apenas avançar nossa compreensão teórica de sistemas complexos, mas também ter aplicações de grande alcance, incluindo a otimização de algoritmos e configurações de rede.

Motivação

Até o momento, não havia um estudo que caracterizasse definitivamente o comportamento de redes q -exponenciais e como o ajuste de seus parâmetros influenciaria a mudança de suas propriedades. A busca pela caracterização dessas redes é germinalmente abordada no artigo (Brito *et al.*, 2016), que apresenta um modelo matemático para o crescimento da rede. Este modelo busca superar as limitações dos modelos existentes, como o modelo de Barabási-Albert (BA). Em contraste com o modelo BA, que foca apenas na topologia, o novo modelo adota uma abordagem multidimensional, incorporando a distância euclidiana entre os nós, sua conectividade existente e a dimensionalidade do sistema.

Uma das características distintivas do modelo proposto em (Brito *et al.*, 2016) é a probabilidade de escolha de sítio, definida pela função (1.2) e a probabilidade de associação, denotada pela função (1.3), que determina se os novos nós (localizado no sítio escolhido anteriormente) será adicionado à rede existente. Os autores introduzem um parâmetro de controle, α_A/d , que tem implicações significativas para a adaptabilidade e eficiência do modelo. No limite onde $\alpha_A = 0$, o modelo converge para o modelo de Barabási-Albert. No entanto, os autores enfatizam que os dois modelos são fundamentalmente diferentes.

$$p(r) \propto \frac{1}{r^{d+\alpha_G}} \quad (\alpha_G > 0; d = 1, 2, 3, 4); \quad (1.2)$$

$$P_{ij} = \frac{k_i r_{ij}^{-\alpha_A}}{\sum_i k_i r_{ij}^{-\alpha_A}} \quad (1.3)$$

O artigo também aprofunda sua análise através de simulações em larga escala em múltiplas dimensões. Notavelmente, a distribuição de grau $p(k)$ da rede se mantém independente de α_G , uma descoberta que valida suas construções teóricas anteriores e enriquece nossa compreensão de como α_A/d afeta $p(k)$. Os parâmetros q e κ surgem como dependentes apenas

dessa razão, simplificando o modelo enquanto ampliam sua aplicabilidade universal. Isso foi confirmado por uma relação linear entre q e κ , tornando ambos os parâmetros críticos para entender o comportamento do sistema, de forma similar ao conceito de temperaturas críticas em termodinâmica. Perceba que o autor utiliza a variável $\kappa = \lambda^{-1}$ para parametrizar a distribuição de probabilidades q -exponencial. Embora o modelo seja matematicamente robusto e ofereça insights inovadores, há algumas ressalvas quanto à sua validação empírica. Os autores utilizam o teste Kolmogorov-Smirnov (K-S) para validar o ajuste de seu modelo q -exponencial. Embora popular, o teste K-S não é otimizado para distribuições de cauda pesada, frequentemente encontradas em redes complexas. Ademais, o autor remove dados escassos na cauda da distribuição, que podem conter informações cruciais para sua caracterização. Portanto, o teste K-S pode não captar adequadamente a semelhança entre as distribuições. Dado que o modelo proposto não é padrão, a escolha do teste K-S, sem mais justificativas, torna-se questionável. Além disso, as suposições que fundamentam o teste K-S podem não ser inteiramente válidas no contexto dessas redes, requerendo métodos adicionais para uma validação empírica mais robusta.

Assim, para caracterizar redes com distribuição q -exponencial, seria necessário fornecer uma prova rigorosa de que o grau dos vértices da rede, construída pelo método de *Preferential Attachment* proposto em (Brito *et al.*, 2016), de fato segue essa distribuição. No entanto, essa prova não é fornecida pelo autor. Uma alternativa para caracterizar redes com distribuição q -exponencial seria atribuir os graus dos vértices *a priori* de acordo com a distribuição q -exponencial e, em seguida, realizar a conexão mantendo os graus fixos. Essa modelo é conhecido como configuracional, proposto em (Newman, 2010). Essa abordagem é computacionalmente mais desafiadora, mas nos permite afirmar que as propriedades topológicas únicas observadas não são um artefato do processo de crescimento, mas são diretamente consequenciais à natureza q -exponencial da distribuição de grau em si. A presente tese buscou adotar essa última abordagem e explorar as propriedades que governam redes definitivamente q -exponenciais.

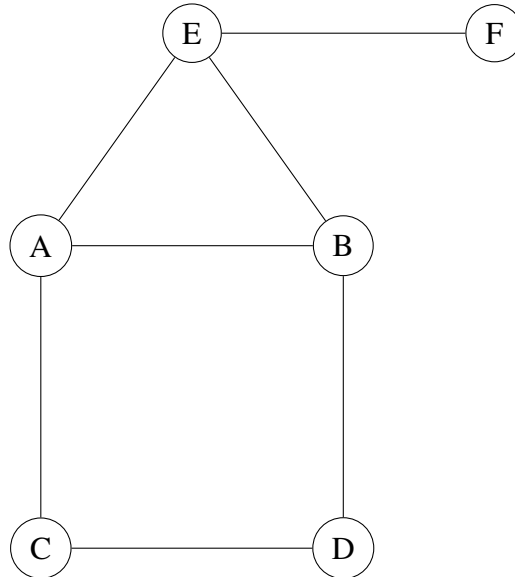
2 FUNDAMENTAÇÃO TEÓRICA

2.1 Teoria dos Grafos

Conceitos Básicos

Um grafo é definido como um par $G = (V, E)$, onde V é o conjunto de nós ou vértices e E é o conjunto de arestas, tal que $E \subseteq [V]^2$ (Anderson, 2001). Aqui, $[V]^2$ refere-se ao conjunto de todos os pares possíveis de vértices de V . É crucial destacar que, embora não exista restrição quanto a ambos os vértices de um par serem o mesmo, para as finalidades desta análise, imporemos a restrição de que os vértices de cada par sejam distintos nas discussões subsequentes. A representação matemática de grafos é crucial na análise estrutural e algorítmica de redes complexas. Uma representação visual comum de G utiliza um conjunto de pontos para representar V e linhas para representar E , conectando os respectivos pontos. Esta abordagem é amplamente adotada devido à sua capacidade de ilustrar propriedades topológicas com clareza, como ilustrado na Figura 1.

Figura 1 – O grafo que consiste em seis vértices (A, B, C, D, E, F) e sete arestas que os conectam. A matriz de adjacência correspondente é a matriz 6x6 definida em (2.2), onde um '1' indica a presença de uma aresta entre dois nós e um '0' indica a ausência dela.



Em uma abordagem estrutural, a lista de adjacência é uma representação prevalente. Dado um vértice $v \in V$, a lista de adjacência $L(v)$ contém todos os vértices adjacentes a v . Matematicamente, para uma aresta $(v, w) \in E$, tanto w está em $L(v)$ quanto v está em $L(w)$. Esta representação é espacialmente eficiente para grafos esparsos, uma vez que o espaço necessário é

proporcional ao número de arestas, a lista de adjacências definida em (2.1) representa o grafo da Figura 1.

$$L = \begin{bmatrix} \begin{bmatrix} B & C & E \end{bmatrix} \\ \begin{bmatrix} A & D & E \end{bmatrix} \\ \begin{bmatrix} A & D \end{bmatrix} \\ \begin{bmatrix} B & C \end{bmatrix} \\ \begin{bmatrix} A & B & F \end{bmatrix} \\ \begin{bmatrix} E \end{bmatrix} \end{bmatrix} \quad (2.1)$$

Por outro lado, a matriz de adjacência, denotada como A , é uma matriz $|V| \times |V|$ onde A_{ij} é 1 se existe uma aresta entre os vértices v_i e v_j , e 0 caso contrário, a matriz de adjacência definida em (2.2) representa o grafo da Figura 1. Para grafos ponderados, A_{ij} pode representar o peso da aresta. Embora A permita uma consulta rápida de adjacência, seu consumo de espaço é $O(|V|^2)$. Uma lista de arestas é simplesmente uma enumeração de todas as arestas, onde cada aresta $e \in E$ é representada por um par (v, w) . Em contraste, a matriz de incidência, para um grafo não direcionado, é uma matriz $|V| \times |E|$ onde cada coluna representa uma aresta e possui exatamente dois valores não nulos que indicam os vértices conectados por essa aresta. A escolha da representação apropriada é dependente do contexto. Por exemplo, algoritmos que frequentemente iteram sobre os vizinhos de um vértice beneficiar-se-iam da lista de adjacência devido à sua eficiência em tais operações (Newman, 2018). Dois grafos são considerados idênticos se tiverem os mesmos conjuntos de vértices e arestas, independentemente da forma como são desenhados.

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (2.2)$$

Se temos um grafo com um conjunto de vértices X , dizemos que se trata de um grafo em X (Anderson, 2001). Denotamos o conjunto de vértices de um grafo G como $V(G)$ e o

conjunto de arestas como $E(G)$, independentemente das letras específicas usadas para definir esses conjuntos. Pode-se também simplificar a notação e dizer que um vértice v pertence a G ao invés de $v \in V(G)$, e o mesmo vale para as arestas, podemos simplesmente dizer que a aresta e pertence a G ($e \in G \equiv e \in E(G)$).

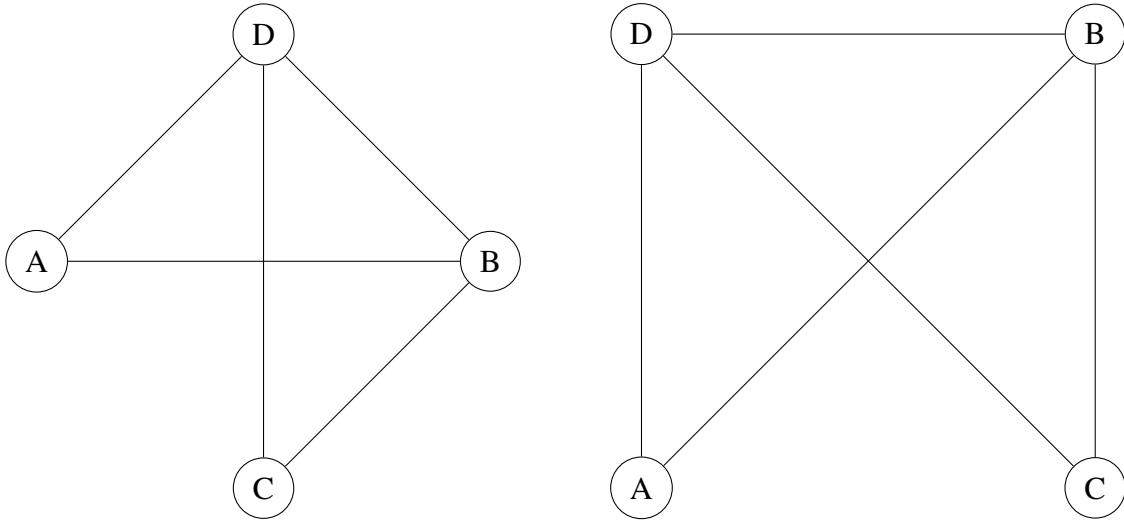
A ordem de um grafo G , indicada por $|G|$, é o número de vértices que o grafo contém. Analogamente, o tamanho de um grafo representa o número de arestas e é denotado por $\|G\|$. Um grafo vazio, que não possui vértices nem arestas, tem $\|G\| = |G| = 0$ e pode ser representado como $G(\emptyset, \emptyset)$ ou simplesmente \emptyset . Grafos vazios e grafos com apenas um vértice são chamados de grafos triviais. No grafo da Figura 1, temos $|G| = 6$ e $\|G\| = 7$.

Um vértice v é dito incidente a uma aresta e se v é um dos extremos dessa aresta. Os vértices x e y associados a uma aresta e são chamados de extremidades (ou pontas) da aresta, essa pode ser representada por xy . O conjunto de todas as arestas associadas a um vértice v é representado por $E(v)$. Se dois vértices x e y de G estiverem conectados por uma aresta, eles são chamados de adjacentes ou vizinhos. Se todos os vértices de G estiverem conectados entre si, o grafo é considerado completo e é denotado por K_n , onde n é a ordem do grafo. Inversamente, dois vértices que não são adjacentes são designados como independentes. De forma mais abrangente, afirmamos que um conjunto de vértices (ou arestas) é independente se, e somente se, nenhum de seus elementos for adjacente a qualquer outro elemento distinto desse conjunto.

Sejam $G = (V, E)$ e $G' = (V', E')$ dois grafos. Dizemos que G e G' são isomórficos ($G \simeq G'$), se existe uma bijeção $\varphi : V \rightarrow V'$ (isomorfismo), tal que $xy \in E \iff \varphi(x)\varphi(y) \in E'$, $\forall x, y \in V$. Num isomorfismo, as relações de adjacência e não adjacência são preservadas. Se $G = G'$, então dizemos que φ é um automorfismo, ou seja, um automorfismo é uma reordenação dos vértices do grafo de tal forma que a estrutura do grafo permanece inalterada, ou seja, um isomorfismo de um grafo com ele mesmo. Uma propriedade ou medida que é a mesma para dois grafos isomórficos é chamada de invariante de grafo. Em outras palavras, é uma característica que não muda quando o grafo é submetido a um isomorfismo. Os invariantes de grafos ajudam a identificar ou distinguir grafos, ou a verificar se dois grafos são isomórficos. Podemos citar a ordem de um grafo ($V(G)$) e o seu tamanho ($E(G)$). A Figura(2) ilustra um exemplo de isomorfismo entre grafos.

Definimos a união de dois grafos disjuntos como $G \cup G' \triangleq (V \cup V', E \cup E')$ e a interseção de dois grafos como $G \cap G' \triangleq (V \cap V', E \cap E')$. Se V' é um subconjunto de V e E' é um subconjunto de E , então G' é chamado de subgrafo de G , representado por $G' \subseteq G$.

Figura 2 – Dois Grafos Isomorfos: ambos os grafos têm quatro vértices e cinco arestas, com a mesma estrutura de conexão. O "Grafo 1" tem um arranjo onde um dos vértices está no centro conectado a todos os outros três vértices, enquanto no "Grafo 2", os vértices estão dispostos de forma que cada um esteja conectado a dois outros vértices, formando uma estrutura quadrada. Apesar de suas disposições diferentes, eles são isomorfos, pois há uma correspondência um-para-um entre seus vértices e arestas que preserva a estrutura de conexão.



Grafo 1

Grafo 2

Adicionalmente, se U é um subconjunto arbitrário de vértices de V e as arestas entre eles são exatamente as arestas de G que ligam vértices em U , então $G[U]$ é o subgrafo *induzido* de G por U . Finalmente, se V' abrange todo G , ou seja, se $V' = V$, então $G' \subseteq G$ é um subgrafo gerador de G . Se U é um conjunto de vértices qualquer, então as notações $G - U$ e $G[V \setminus U]$ representam o grafo obtido de $G[V]$ após removermos os vértices $V \cup U$ e suas arestas incidentes. Analogamente, para $F \subseteq [V]^2$, escrevemos $G(V, E) - F \triangleq G(V, E \setminus F)$ e $G(V, E) + F \triangleq G(V, E \cup F)$.

Dizemos que um grafo $G(V, E)$ é maximal em relação a uma propriedade específica se ele satisfaz essa propriedade, mas ao adicionar qualquer aresta adicional xy (onde $x, y \in V$ e $xy \notin E$), ele deixa de satisfazê-la (Anderson, 2001). Em outras palavras, o grafo G é maximal se ele não pode ganhar mais arestas sem perder a propriedade em questão. Por contraste, um grafo é considerado minimal em relação às suas arestas com essa mesma propriedade se ele a possui, mas a remoção de qualquer aresta do grafo faz com que ele perca tal propriedade. Em outras palavras, enquanto um grafo maximal não pode ganhar mais arestas sem perder a propriedade em questão, um grafo minimal não pode perder nenhuma aresta sem que a propriedade seja comprometida. Por exemplo, em relação à conectividade, uma árvore é um grafo minimalmente conectado, pois a remoção de qualquer uma de suas arestas resultará em um grafo desconectado.

Se $G(V, E)$ e $G'(V', E')$ são dois grafos disjuntos, nós escrevemos $G * G'$ para denotar o grafo obtido por $G \cup G'$ acrescido de todas as arestas dos vértices V para os vértices V' . O

grafo complementar \overline{G} de $G(V, E)$ é o grafo $G + [V]^2 \setminus E$. O grafo linha $L(G)$ de um grafo $G(V, E)$ é construído associando a cada aresta de E um vértice em $L(G)$, e conectando dois vértices em $L(G)$ por uma aresta se suas arestas correspondentes em E forem adjacentes, ou seja, compartilharem um vértice comum. Esta transformação é útil em teoria dos grafos, pois transforma problemas sobre vértices do grafo original em problemas sobre suas arestas no grafo linha. Por exemplo, para um grafo G com vértices $\{a, b, c\}$ e arestas $\{(a, b), (b, c)\}$, seu grafo linha $L(G)$ terá dois vértices, representando as arestas (a, b) e (b, c) de G , respectivamente, e uma aresta conectando esses vértices, pois em G , as arestas (a, b) e (b, c) são adjacentes através do vértice b (Anderson, 2001).

Grafos dirigidos, também referidos como digrafos ou grafos direcionados, são estruturas nas quais as arestas possuem uma direção definida, partindo de um vértice e direcionando-se a outro (Newman, 2018). Formalmente, um grafo dirigido $G(V, E)$ é composto por um conjunto V de vértices e um conjunto E de arestas, em que cada aresta é representada por um par ordenado de vértices. Isto é, se temos uma aresta (u, v) em E , o vértice u é identificado como a origem e o vértice v como o destino da aresta. Gráficamente, essa relação é denotada por uma seta que se origina no vértice u e aponta para o vértice v (Anderson, 2001). Tais arestas são frequentemente denominadas *arestas dirigidas* ou arcos. Define-se um digrafo D como uma *orientação* do grafo não direcionado G se, e somente se, $V(D) = V(G)$ e cada aresta $e \in E(G)$ é transformada em uma aresta direcionada em $E(D)$, preservando os vértices originais. Dito de outra forma, uma aresta $e = xy \in E(G)$ pode ser representada no digrafo como uma aresta dirigida (x, y) ou (y, x) . Esse processo consiste, intuitivamente, na atribuição de uma direção a cada aresta do grafo original (Anderson, 2001). Grafos dirigidos, ou digrafos, são comumente utilizados para representar relações assimétricas, como situações que envolvem direções de fluxo, dependências temporais ou relações hierárquicas.

Em muitos contextos, as arestas de um grafo não são uniformes; elas podem possuir diferentes valores ou pesos. Quando isso ocorre, referimo-nos ao grafo como tendo *arestas ponderadas*. O peso de uma aresta pode simbolizar diversos aspectos, como a intensidade de uma conexão, a distância entre dois pontos ou o tempo necessário para se deslocar entre eles.

Grau de um vértice

Considere um grafo $G(V, E)$ não vazio. A vizinhança de um vértice v em G , denotada por $N_G(v)$ com $v \in V$, é o conjunto de todos os vértices adjacentes a v em G (Anderson, 2001).

Se não houver risco de confusão, podemos simplificar essa notação para $N(v)$. A vizinhança de um subconjunto de vértices $U \subseteq V$ é definida como $N(U) = \cup_{v \in U} N(v) \setminus U$, ou seja, é o conjunto de todos os vértices que são vizinhos de pelo menos um vértice em U , mas que não pertencem a U . O grau de um vértice é o tamanho do conjunto de seus vizinhos, denotado por $d(v)$ ou k_v . Especificamente, é o número de arestas incidentes ao vértice v . Assim, temos que $k_v = d(v) = |N(v)|$. Um vértice é considerado isolado se seu grau for zero, ou seja, se $k_v = d(v) = 0$. Uma observação crucial na teoria dos grafos é a afirmação de que a soma dos graus de todos os vértices no grafo não direcionado é exatamente duas vezes o número de arestas. Matematicamente, essa relação é expressa pela equação (2.3) (Anderson, 2001) que é conhecida como *Lema do Aperto de Mão*.

$$|E| = \frac{1}{2} \sum_{v \in V(G)} d(v) \quad (2.3)$$

O grau mínimo e o grau máximo de G são representados, respectivamente, por $\delta(G) = \min\{d(v) | v \in V\}$ e $\Delta(G) = \max\{d(v) | v \in V\}$. Um grafo é dito regular se todos os seus vértices têm o mesmo grau. O grau médio de um grafo, denotado por $\langle k \rangle$ ou $d(G)$, é a média aritmética dos graus de todos os vértices, e é expressa na equação (2.4) (Anderson, 2001).

$$d(G) = \frac{1}{|V|} \sum_{v \in V(G)} d(v) \quad (2.4)$$

É fácil perceber que $\delta(G) \leq d(G) \leq \Delta(G)$, onde a igualdade só é verdadeira em grafos regulares (todos os vértices têm o mesmo grau). O grau médio também pode ser interpretado como o número médio de arestas incidentes por vértice, e frequentemente é calculado diretamente a partir da razão entre o número total de arestas e o número total de vértices, ou seja, $\varepsilon(G) = \frac{|E|}{|V|}$. A relação entre $\varepsilon(G)$ e $d(G)$ é dada pela equação (2.5) (Anderson, 2001).

$$\begin{aligned} \varepsilon(G) &= \frac{|E|}{|V|} \\ &= \frac{\frac{1}{2} \sum_{v \in V} d(v)}{|V|} \\ &= \frac{1}{2} \left(\frac{1}{|V|} \sum_{v \in V} d(v) \right) \\ &= \frac{1}{2} d(G). \end{aligned} \quad (2.5)$$

Nos contextos em que estamos tratando exclusivamente de grafos simples, isto é, grafos que não apresentam auto-arestas nem múltiplas arestas paralelas entre dois vértices distintos, é possível definir a *densidade* de um grafo. Esta densidade, representada por ρ , indica a proporção de arestas presentes no grafo em comparação com o total de arestas que um grafo completo com o mesmo número de vértices poderia ter. Para um grafo completo e não direcionado, denotado por K_n , o total de arestas é determinado pela combinação de seus vértices em pares, expressa como $\binom{n}{2}$. Para um grafo G de ordem $n = |V(G)|$ e tamanho $m = |E(G)|$, a densidade ρ é especificada pela equação (2.6). Vale ressaltar que, dado que $0 \leq m \leq \binom{n}{2}$, a densidade de um grafo também pode ser entendida como a probabilidade de que, ao selecionarmos aleatoriamente dois vértices de G de maneira uniforme, estes estejam ligados por uma aresta (Newman, 2018).

$$\begin{aligned}
 \rho &= \frac{m}{\binom{n}{2}} \\
 &= \frac{m}{\frac{n!}{2!(n-2)!}} \\
 &= \frac{2m(n-2)!}{n!} \\
 &= \frac{2m}{n} \cdot \frac{1}{(n-1)} \\
 &= d(G) \cdot \frac{1}{n-1} \\
 &\approx \frac{d(G)}{n} \quad (\text{para } n \gg 1).
 \end{aligned} \tag{2.6}$$

A densidade de um grafo pode ser usada para classificar o grafo quanto à sua *compacidade* em termos de arestas. Quando estudamos o comportamento assintótico da densidade à medida que o número de vértices n cresce, podemos diferenciar grafos *densos* de grafos *esparsos*. Precisamente, se para um grafo G temos $\lim_{n \rightarrow \infty} \rho > 0$, então o grafo G é considerado *denso*. Por outro lado, se $\lim_{n \rightarrow \infty} \rho = 0$, então o grafo é caracterizado como *esparso*. Essa distinção implica que, em grafos esparsos, a função $\rho(n)$ cresce sublinearmente com n , enquanto em grafos densos o crescimento é linear ou supralinear (Newman, 2018).

É importante notar que essa classificação se baseia em um comportamento assintótico, e muitos dos grafos que examinamos, especialmente aqueles que modelam estruturas topológicas na natureza ou em sistemas reais, têm um número finito de vértices. Assim, embora não se possa tecnicamente classificar esses grafos como densos ou esparsos em um contexto formal,

informalmente, é comum referir-se a grafos finitos com valores pequenos de ρ como esparsos e com valores altos de ρ como densos (Newman, 2018).

Nos digrafos, é importante considerar os graus de entrada e saída de forma independente. Para um vértice $v \in V(G)$ no digrafo G , o grau de entrada, representado por $d_e(v)$, refere-se ao número de arcos que têm v como vértice terminal, enquanto o grau de saída, denotado por $d_s(v)$, é o número de arcos que têm v como vértice inicial. Vale notar que, para qualquer digrafo, a soma total dos graus de entrada de todos os vértices é sempre igual à soma total dos graus de saída, dado que cada arco tem exatamente um ponto inicial e um ponto terminal, conforme a relação (2.7).

$$|E(G)| = \sum_{v \in V(G)} d_e(v) = \sum_{v \in V(G)} d_s(v). \quad (2.7)$$

Dessa forma, o grau médio de entrada para o digrafo, $d_e(G)$, é idêntico ao grau médio de saída, $d_s(G)$. Assim, pode-se representar ambos os graus médios pelo simples termo $d(G)$, conforme demonstrado na equação (2.8).

$$\begin{aligned} d_e(G) &= \frac{1}{|V(G)|} \sum_{v \in V(G)} d_e(v) \\ &= \frac{1}{|V(G)|} \cdot |E(G)| \\ &= \frac{1}{|V(G)|} \sum_{v \in V(G)} d_s(v) \\ &= d_s(G). \end{aligned} \quad (2.8)$$

Conectividade

Um grafo G é dito *conectado* se, para todo par de vértices u e v pertencentes a $V(G)$, existe um caminho em G que os conecta (Anderson, 2001). Se tal caminho não existir, então u e v pertencem a *componentes conexas* distintas e o grafo é dito desconectado. Uma componente conexa de G é um subgrafo S no qual qualquer par de vértices dentro de S é conectado por um caminho, e nenhum vértice adicional $w \in V(G) \setminus V(S)$ pode ser adicionado a S sem violar essa propriedade. Assim, uma componente conexa é *maximal* em relação à conectividade do subgrafo induzido (Newman, 2018). Com essa definição, podemos facilmente enunciar a proposição (2.1.1).

Proposição 2.1.1. *Para qualquer grafo conectado G , é possível ordenar seus vértices como $v_1, v_2, \dots, v_n \in V(G)$ de tal forma que, ao se considerar sucessivos subgrafos induzidos, começando de v_1 e adicionando um vértice por vez, cada subgrafo G_i formado pelos vértices v_1, \dots, v_i será conectado.*

Demonstração. Para $i = 1$, escolhamos arbitrariamente um vértice v_1 de G . O subgrafo induzido G_1 é trivialmente conectado, pois consiste apenas do vértice v_1 . Agora suponha, por hipótese de indução, que tenhamos escolhido os vértices v_1, \dots, v_i de forma que o subgrafo induzido G_i seja conectado. Queremos mostrar que podemos escolher um vértice v_{i+1} de forma que G_{i+1} seja conectado. Dado que G é conectado e G_i não é o grafo todo, deve existir um vértice v em G mas não em G_i . Considerando um caminho P que ligue v a v_1 (esse caminho existe porque G é conectado), escolhamos v_{i+1} como o último vértice deste caminho que não está em G_i . Por definição, v_{i+1} tem um vizinho em G_i , garantindo que G_{i+1} seja conectado. \square

Para digrafos (grafos direcionados), distinguimos entre conectividade *fraca* e *forte*. A conectividade é dita *fraca* se, ao ignorar a direção dos arcos, o grafo equivalente é conectado. Por outro lado, a conectividade é *forte* se, para quaisquer componentes $A, B \subset G$, existe ao menos um caminho de A para B e um caminho de B para A . Nesse contexto, a *componente de entrada* (*in-component*) de um vértice é o conjunto de vértices que podem alcançá-lo através de um caminho direcionado, enquanto a *componente de saída* (*out-component*) é o conjunto de vértices que podem ser alcançados a partir dele (Newman, 2018).

Quando estudamos a *conectividade de um grafo* G , estamos interessados em quão resistente esse grafo é a desconexões. Uma maneira de medir isso é através do conceito de k -conectividade. Um grafo G é dito k -conectado se, para fazer com que qualquer par de vértices do grafo não esteja mais conectado, precisamos remover pelo menos k vértices (Anderson, 2001). Em outras palavras, o grafo possui uma certa resiliência que o impede de ser facilmente fragmentado em componentes menores. É imediato que todo grafo, desde que não esteja vazio, tem pelo menos a propriedade de ser 0-conectado, pois sem remover nenhum vértice, o grafo ainda está conectado consigo mesmo. Já nos grafos 1-conectados há um caminho entre qualquer par de vértices, mas, se removermos um único vértice em alguns casos, podemos fazer com que o grafo se divida em partes desconexas.

A notação $\kappa(G)$ indica o nível máximo de conectividade que o grafo G possui em relação aos seus vértices. Se $\kappa(G) = 3$, significa que o grafo G é robusto o suficiente para que não possamos separar quaisquer dois vértices removendo apenas 1 ou 2 vértices. No entanto,

existe algum conjunto de 3 vértices cuja remoção desconectaria o grafo. Portanto, o valor de $\kappa(G)$ nos mostra a quantidade mínima de vértices que precisamos remover para começar a fragmentar o grafo em componentes menores. O conjunto minimal de vértices cuja remoção aumenta o número de componentes conexas é chamado de **conjunto mínimo de corte**. Se tivermos um grafo que já é desconectado ou se o grafo representa apenas um ponto isolado, então $\kappa(G) = 0$.

A medida da conectividade ($\kappa(G)$) é limitada pela conectividade por arestas (número mínimo de arestas que devemos remover para desconectar o grafo G), representada por $\lambda(G)$. Adicionalmente, essa última não ultrapassa o grau mínimo de qualquer vértice no grafo, simbolizado por $\delta(G)$. Isto sugere que um grafo com uma conectividade marcante geralmente apresenta também um alto grau mínimo. Assim, de forma concisa, para um grafo G temos a relação de ordem definida na proposição (2.1.2). Para o grafo completo K_n , a conectividade é sempre $n - 1$, pois precisaríamos remover $n - 1$ vértices para desconectar qualquer par de vértices restantes. O valor de $\kappa(G)$ está correlacionado com o conceito de percolação, como veremos mais a frente (Anderson, 2001).

É possível flexibilizar a definição da medida de conectividade a fim de identificar grupos de vértices interligados com uma robustez caracterizada por um valor k menor ou igual a $\kappa(G)$. Neste contexto, introduzimos o conceito de k -componentes. Estes representam um subconjunto de vértices dentro de um grafo, onde cada vértice é interconectado com outro do mesmo grupo por, no mínimo, k caminhos independentes. Embora a explanação detalhada sobre caminhos independentes seja abordado mais adiante, essa definição preliminar sugere que os vértices dentro de um grupo não podem ser isolados pela remoção de menos de k vértices pertencentes a esse mesmo grupo (Newman, 2018).

Proposição 2.1.2. *Para qualquer grafo conectado G , é possível estabelecer a relação de ordem:*
 $\kappa(G) \leq \lambda(G) \leq \delta(G)$.

Demonstração. Dividiremos a demonstração da proposição em duas partes, uma para cada desigualdade:

(i) $\kappa(G) \leq \lambda(G)$:

Suponha que $\kappa(G) = k$, isso significa que, se removermos $k - 1$ vértices, o grafo ainda permanece conectado. Assim, o k -ésimo vértice, necessário para desconectar o grafo, deve estar ligado a outro componente do grafo através de, pelo menos, uma aresta. Se removermos esta aresta, também conseguimos desconectar o grafo. Dessa forma, para cada vértice que

removemos para desconectar o grafo, podemos, em vez disso, optar por remover uma aresta que o conecta a outro componente do grafo. O número de arestas que removemos dessa forma será, no máximo, o número de vértices que removemos. Em outras palavras, em um cenário ideal, poderíamos remover o mesmo número de arestas e vértices para desconectar o grafo. No entanto, na prática, muitas vezes precisamos remover menos arestas do que vértices. No exemplo acima, se removemos k -ésimo vértice teremos removido duas arestas, o que evidencia o fato de que $\kappa(G) \leq \lambda(G)$.

(ii): $\lambda(G) \leq \delta(G)$:

Seja v um vértice em G com grau mínimo, ou seja, v tem $\delta(G)$ vizinhos. Se removermos todas as arestas incidentes a v , desconectamos v do restante do grafo. Portanto, precisamos remover no máximo $\delta(G)$ arestas para fazer isso. Assim, a conectividade por arestas, $\lambda(G)$, que é o número mínimo de arestas que precisamos remover para desconectar o grafo, é no máximo $\delta(G)$. Logo, $\lambda(G) \leq \delta(G)$.

Combinando a **(i)** com a **(ii)**, obtemos a relação procurada: $\kappa(G) \leq \lambda(G) \leq \delta(G)$. Perceba que se $\kappa(G)$ ou $\lambda(G)$ são grandes, isso implica que $\delta(G)$ também deve ser grande, conforme afirmamos anteriormente. \square

Um grafo G é caracterizado como ℓ -aresta-conectado quando, ao removermos menos de ℓ arestas de sua estrutura, ele continua conectado, contanto que o número total de vértices no grafo seja maior que um. A medida $\lambda(G)$ representa a maior quantidade de arestas que podemos retirar sem desconectar o grafo, sendo denominada a conectividade por arestas de G . É fundamental reconhecer que se um grafo estiver desconectado de início, então sua conectividade por arestas, representada por $\lambda(G)$, será igual a 0 (Anderson, 2001).

Vale ressaltar que um alto grau mínimo não é garantia de uma conectividade robusta, nem especificamente de uma forte conectividade por arestas. Existem certos grafos que, embora possuam um grau mínimo consideravelmente elevado, possuem conectividade limitada. Como ilustração, consideremos o Grafo de Duas Partes. Este grafo é composto por dois subconjuntos, A e B . O subconjunto A é formado por k vértices, todos interligados entre si, enquanto B segue o mesmo padrão. A única ligação entre A e B é uma aresta solitária que conecta um vértice de A a um de B . Assim, cada vértice dentro de um desses subconjuntos está conectado a exatamente $k - 1$ outros vértices. Quando k é um número elevado, $\delta(G)$, o grau mínimo, também é alto. No entanto, tanto a conectividade, $\kappa(G)$, quanto a conectividade por aresta, $\lambda(G)$, são igual a 1, visto que a remoção de apenas uma aresta já é suficiente para desconectar o grafo. Este cenário

evidencia que, mesmo com um grau mínimo expressivo, um grafo pode ter baixa conectividade e conectividade por aresta. Se temos dois subconjuntos de vértices, A e B em um grafo, e um conjunto X composto por vértices ou arestas, dizemos que X separa A e B se qualquer caminho de A para B passar por um elemento de X . Se alguns vértices pertencem tanto a A quanto a B , eles também estarão em X . De forma mais ampla, se X consegue separar dois vértices que não estão em X , dizemos que X separa o grafo. Um vértice que tem o poder de separar dois outros vértices de sua mesma componente é chamado de *vértice de corte*. Já uma aresta que, quando removida, separa seus vértices extremos é chamada de *ponte*. Assim, as pontes em um grafo são exatamente aquelas arestas que não pertencem a nenhum ciclo.

Teorema 2.1.1. *Todo grafo com grau médio ($\langle k \rangle$) maior ou igual a $4k$, possui um subgrafo k -conectado (Mader, 1972).*

Demonstração. Considerando um grafo $G = (V, E)$ com $|V| = n$ (número de vértices) e $|E| = m$ (número de arestas). Para $k \in \{0, 1\}$, a afirmação é trivial, uma vez que um grafo sem vértices ou com um único vértice não possui arestas e, portanto, não tem conectividade. Portanto, a afirmação é imediatamente verdadeira nesses casos. Considerando agora os casos não triviais ($k \geq 2$), a condição $\langle k \rangle \geq 4k$ implica em duas premissas mais fortes que utilizaremos para realizar a prova do teorema, são elas:

(i) $n \geq 2k - 1$.

(ii) $m \geq (2k - 3)(n - k + 1) + 1$.

Dado que $n - 1 \geq \Delta(G) \geq \langle k \rangle$, inferimos que $n \geq 4k + 1$. Como $k \geq 2 \rightarrow n \geq 2k - 1$, correspondendo à primeira premissa. Usando o *Lema do Aperto de Mão*, temos $m = \frac{1}{2} \langle k \rangle n \geq \frac{1}{2} (4k)n = 2kn$. Considere o polinômio $f(k) = -2k^2 + 5k - 2$, que tem raízes $\frac{1}{2}$ e 2 . Como é uma função côncava, $f(k) \leq 0$ para todo $k \geq 2$. Dado que $n > 0$, temos $f(k) \leq 0$ para todo $k \geq 2 - 3n \leq 0$ para todo $k \geq 2$. Portanto, $m \geq 2kn \geq 2kn + [f(k) - 2n] = (2k - 3)(n - k + 1) + 1$. Assim, temos a segunda premissa. Se $n = 2k - 1$, temos $k = \frac{1}{2}(n + 1)$ e, aplicando $k = \frac{1}{2}(n + 1)$

em (ii), desenvolvemos a igualdade (2.9):

$$\begin{aligned}
m &\geq (2k-3)(n-k+1) + 1 \\
&= 2\frac{1}{2}(n+1) - 3[n+1 - \frac{1}{2}(n+1)] + 1 \\
&= (n-2)\frac{1}{2}(n+1) + 1 \\
&= \frac{1}{2}(n^2 - 2n + n - 2 + 2) \\
&= \frac{1}{2}(n^2 - n) \\
&= \frac{1}{2}n(n-1).
\end{aligned} \tag{2.9}$$

Assim, $m \geq \frac{1}{2}n(n-1)$, mas $m \leq \frac{1}{2}n(n-1)$, concluímos que $m = \frac{1}{2}n(n-1)$ e portanto $G = K_n$. Como $n = 2k - 1 = k + (k - 1)$ e $k \geq 2$, concluímos que $n \geq k + (2 - 1) = k + 1$ e portanto $G = K_n \supseteq K_{k+1}$, confirmando nossa afirmação para esse caso. Considerando os demais casos, ou seja, um grafo G com $n \geq 2k$ vértices e grau mínimo $\delta(G)$. Podemos assumir que todos os vértices de G têm grau pelo menos $2k - 2$, ou seja, $\delta(G) \geq 2k - 2$. Caso contrário, se existir um vértice v em G tal que $d(v) \leq 2k - 3$, podemos remover esse vértice para obter o subgrafo $G - v$ com menos vértices tal que $\delta(G) \geq 2k - 2$. Esse passo pode ser repetido recursivamente até que a condição alvo seja atendida. Assim, cada vértice no grafo resultante possui pelo menos $2k - 2$ vizinhos. Se G já for k -conectado, a afirmação se mantém e nada mais precisa ser feito. Caso contrário, podemos decompor G em dois subgrafos G_1 e G_2 de tal maneira que $|G_1 \cap G_2| < k$, onde ambos $|G_1|$ e $|G_2|$ são menores que n e ao menos um dos dois subgrafos é k -conectado. Como não existem arestas entre $G_1 - G_2$ e $G_2 - G_1$, e sabendo que cada vértice nesses subgrafos tem pelo menos $2k - 2$ vizinhos, concluímos que $|G_1|, |G_2| \geq 2k - 1$, portanto a premissa (i) é verdadeira. Suponha por absurdo que a premissa (ii) é falsa ($|G_i| \leq (2k - 3)(|G_i| - k + 1)$ para $i = 1, 2$). Dessa forma, podemos desenvolver a igualdade (2.10).

$$\begin{aligned}
m &< ||G_1|| + ||G_2|| \\
&\leq (2k-3)(|G_1| - k + 1) + (2k-3)(|G_2| - k + 1) \\
&\leq (2k-3)(|G_1| + |G_2| - 2k + 2) \\
&\leq (2k-3)(|G_1 \cup G_2| + |G_1 \cap G_2| - 2k + 2) \\
&\leq (2k-3)(n + (k-1) - 2k + 2) \\
&\leq (2k-3)(n - k + 1),
\end{aligned} \tag{2.10}$$

Assim, concluímos que $m < (2k - 3)(n - k + 1)$, contradizendo a premissa (ii) que afirma que $m \geq (2k - 3)(n - k + 1) + 1$. Essa contradição finaliza a prova. \square

A conectividade, por outro lado, pode ser relacionada a um par de vértices, representando o grau de acoplamento entre esses vértices. Nesse contexto, dizemos que a conectividade entre um par de vértices é definida pelo número de caminhos independentes entre esses dois vértices. Perceba a conectividade pode variar conforme contamos caminhos independentes por vértice ou independentes por arestas, e portanto, dizemos conectividade por vértices e por arestas respectivamente.

Um conjunto de corte refere-se a um conjunto de vértices ou arestas cuja remoção resulta em um grafo desconectado. Em outras palavras, após a remoção desses elementos, o grafo não é mais conectado. É uma generalização dos conceitos de vértice de corte e aresta de corte, que se referem à remoção de um único vértice ou aresta, respectivamente, que leva à desconexão do grafo. O teorema de Menger (Menger, 1927) relaciona a cardinalidade do conjunto corte e o número de caminhos independentes, conforme mostrado em 2.1.2. Além disso, O teorema oferece uma maneira de entender e quantificar a robustez de um grafo à remoção de vértices ou arestas.

Teorema 2.1.2. *Seja G um grafo e sejam u e v dois vértices não adjacentes em G . O número mínimo de vértices que precisam ser removidos de G para desconectar u e v é igual ao número máximo de caminhos independentes entre u e v em G .*

Demonstração. Dividiremos o teorema em duas afirmações e iremos provar cada uma delas de forma independente:

Afirmação 1:

Para qualquer conjunto S de vértices que separe u e v , existem pelo menos $|S|$ caminhos independentes de u a v .

Afirmação 2: Para qualquer conjunto P de caminhos independentes de u a v , qualquer conjunto de vértices que separe u e v tem pelo menos $|P|$ vértices.

Prova da Afirmação 1:

Seja S um conjunto de vértices que separa u e v . Para cada vértice w em S , escolha um caminho P_w de u a v que passe por w . Seja Q o subcaminho de P_w de u até w . Como S separa u e v , nenhum caminho em Q pode passar por um vértice em S que não seja w . Assim, todos

esses caminhos são independentes.

Prova da Afirmação 2:

Suponha, por contradição, que existe uma coleção P de k caminhos independentes de u a v e um conjunto S com menos de k vértices que separa u e v . Considere um subgrafo mínimo H de G que contém u , v , todos os caminhos em P , e nenhum outro caminho de u a v . Os vértices de grau 1 em H são u e v , e a remoção de S de H desconecta u e v . Como H é mínimo, ele não tem outros caminhos de u a v . Isso é uma contradição, pois H contém k caminhos independentes de u a v .

Combinando ambas as afirmações, concluímos a prova do Teorema de Menger. \square

No contexto topológico, um conjunto é dito ser de primeira categoria (ou do primeiro tipo) em um espaço se ele pode ser expresso como uma união contável de conjuntos fechados em nenhum lugar densos. Um conjunto é dito ser de segunda categoria se ele não é de primeira categoria. Para grafos infinitos, a validade do Teorema de Menger está relacionada à topologia desses grafos. Um grafo infinito que satisfaça a propriedade da primeira categoria (no sentido da topologia) satisfará uma versão do Teorema de Menger. Em particular, um grafo infinito é dito satisfazer a propriedade da primeira categoria se, para cada par de vértices distintos, a quantidade de caminhos independentes entre eles é um número finito ou um conjunto que é do primeiro tipo (ou seja, é uma contagem de conjuntos que são fechados e em nenhum lugar densos).

Encontrar todos os caminhos independentes em um grafo é $\#P$ -completo, o que significa que não é apenas NP-difícil, mas está em uma classe de problemas que contam soluções e são geralmente considerados ainda mais difíceis do que os problemas NP-completos. Por outro lado, encontrar o tamanho do corte mínimo entre dois vértices específicos pode ser realizado em tempo polinomial como o algoritmo de *Ford-Fulkerson* ou o algoritmo de *Edmonds-Karp* (Cormen *et al.*, 2022). No entanto, o Teorema de Menger fornece uma relação entre o número máximo de caminhos independentes e a cardinalidade do corte mínimo.

Passeios, Caminhos e Ciclos

Dentro da teoria dos grafos, um *passeio* é uma sequência de vértices onde alguns vértices podem ser revisitados. Mais formalmente, um passeio P é representado como $P = x_0, x_1, \dots, x_k$. Em um digrafo, o passeio deve seguir a direção dos arcos. O *tamanho* de um passeio é o número de arestas que o compõem ($\|P\|$) (Newman, 2018).

Um *caminho* é um tipo especial de passeio, onde todos os vértices são distintos. Matematicamente, um caminho em um grafo é descrito por $P(V, E)$, onde $V = \{x_0, x_1, \dots, x_k\}$ e $E = \{x_0x_1, x_1x_2, \dots, x_{k-1}x_k\}$, com a condição de que todos os vértices x_i sejam distintos (Anderson, 2001). A **distância** entre dois vértices em um grafo, denotada por $d(u, v)$, é definida pelo número mínimo de arestas em um caminho que liga os vértices u e v . Assim, o **caminho mais curto** ou **caminho geodésico** entre dois vértices u e v é aquele que minimiza esta distância. Em um grafo ponderado, a distância é determinada pela soma dos pesos das arestas no caminho mais curto. O **diâmetro** de um grafo conectado, denotado por D , representa a maior distância entre quaisquer dois vértices de G , como definido na equação (2.11).

Dado um grafo $G = (V, E)$, onde V representa o conjunto de vértices e E o conjunto de arestas, dois ou mais caminhos em G são considerados *independentes* se eles não compartilham certos elementos específicos. Sejam P_1 e P_2 dois caminhos em G . Eles são considerados *caminhos independentes por aresta* se as arestas que compõem P_1 e P_2 são disjuntas, ou seja, $P_1 \cap P_2 = \emptyset$ quando consideramos apenas as arestas. Por outro lado, eles são considerados *caminhos independentes por vértice* se os vértices que compõem P_1 e P_2 são disjuntos, excetuando-se possivelmente os vértices iniciais e finais. É importante notar que dois caminhos que são independentes por aresta nem sempre são independentes por vértice. Por exemplo, dois caminhos podem compartilhar um vértice, mas não uma aresta. No entanto, se dois caminhos são independentes por vértice, eles também são independentes por aresta (Newman, 2018).

$$D = \max_{u, v \in V(G)} d(u, v). \quad (2.11)$$

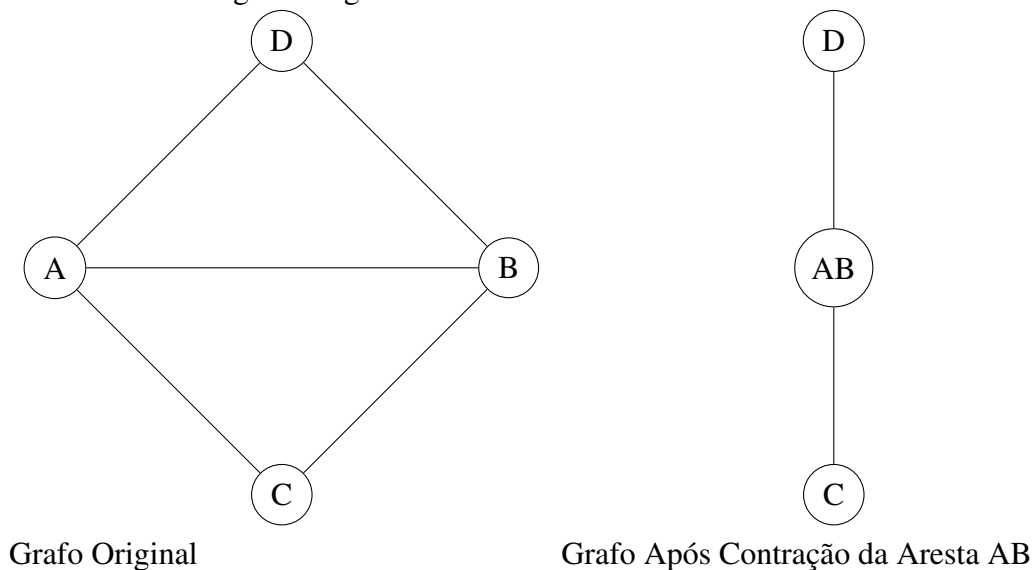
Finalmente, um *ciclo* é um caminho que começa e termina no mesmo vértice. Portanto, se $P = x_0, x_1, \dots, x_k$ e $C = x_0, x_1, \dots, x_k, x_0$, então C é um ciclo. A *cintura* de um grafo G , denotada por $g(G)$, é o tamanho do menor ciclo em G . Em contraste, a *circunferência* de G é o tamanho do maior ciclo. Para um grafo acíclico, a cintura é infinita (∞) e a circunferência é zero. Em relação a um ciclo C que é um subgrafo induzido de G , uma aresta $e \in E(G)$ que conecta dois vértices de C , mas não é parte de C , é chamada *corda* (Anderson, 2001).

Contrações e menores

Na teoria dos grafos, a *contração* de uma aresta é uma transformação fundamental que consiste em fundir os dois vértices conectados por essa aresta, resultando em um único

vértice e , conseqüentemente, na remoção da aresta que os conectava inicialmente (Anderson, 2001). Caso haja outras arestas no grafo original que ligassem esses dois vértices, elas também são eliminadas durante essa operação para garantir que o grafo resultante não contenha laços. A operação inversa da contração é a subdivisão, que envolve dividir um vértice em dois, conectados por uma nova aresta. Essas transformações, tanto a contração quanto a subdivisão, são essenciais na teoria dos grafos (Anderson, 2001), pois, mesmo que alterem a aparência do grafo, mantêm suas propriedades estruturais básicas intactas. Ao aplicar essas operações, pode-se estudar a essência e as características subjacentes dos grafos, independentemente de suas configurações específicas (Anderson, 2001). As contrações e suas operações inversas são exemplos de operações relacionadas ao conceito de *homeomorfismo* em grafos. A Figura 3 ilustra um exemplo de contração em uma das arestas do grafo.

Figura 3 – Exemplo de Contração de Aresta em um Grafo: o "Grafo Original" contém quatro vértices (A, B, C, D) e cinco arestas. A contração é realizada na aresta que conecta os vértices A e B. No "Grafo Após Contração", os vértices A e B são fundidos em um único vértice (AB), e a aresta que os conectava é removida. Outras conexões são mantidas, refletindo as propriedades estruturais básicas do grafo original.



Quando se diz que um grafo H é um *menor* de outro grafo G , estamos nos referindo a uma relação específica de derivabilidade entre os dois grafos (Anderson, 2001). Especificamente, H pode ser obtido a partir de G ao realizar uma série ordenada de operações que incluem: (1) a remoção de vértices selecionados e todas as arestas conectadas a eles e (2) a contração de arestas escolhidas. Assim, se é possível transformar G em H ao suprimir certos vértices e arestas e ao contrair outras arestas específicas sem realizar outras modificações, então podemos afirmar com confiança que H é um menor de G . Esta relação de menor é fundamental para compreender e

caracterizar a estrutura intrínseca e as propriedades de diferentes grafos.

As propriedades hereditárias desempenham um papel fundamental ao identificar características que são transmitidas dos grafos para seus subgrafos induzidos (Anderson, 2001). Quando dizemos que um grafo possui uma propriedade hereditária, significa que, se o grafo original possui essa propriedade, todos os seus subgrafos induzidos – criados pela remoção de vértices e suas arestas associadas – também a possuirão. Relacionado a isso, o conceito de *grafos proibidos* é central para a caracterização de certas propriedades dos grafos (Anderson, 2001). Um grafo proibido é um grafo que não pode ser um subgrafo ou um menor de um grafo que possui uma propriedade específica. Em outras palavras, a presença de um grafo proibido dentro de um grafo maior indica que o grafo maior não pode ter a propriedade em questão (Anderson, 2001).

Um resultado famoso relacionado a menores de grafo é o Teorema de Robertson-Seymour (Robertson e Seymour, 1983), também conhecido como Teorema do Menor Gráfico, que estabelece que, para qualquer propriedade de fechamento hereditário (propriedades que são herdadas por subgrafos), existe um conjunto finito de menores de grafos proibidos, de forma que um grafo tem a propriedade se e somente se não contém nenhum desses menores proibidos (Anderson, 2001). Um exemplo clássico relacionado ao teorema de Robertson-Seymour é a caracterização de grafos planares. O Teorema de Kuratowski (Kuratowski e Whyburn, 1930) afirma que um grafo é planar (pode ser desenhado em um plano sem que suas arestas se cruzem) se e somente se ele não contém um subgrafo que seja homeomorfo a K_5 ou $K_{3,3}$.

Métricas de Centralidade

Métricas de centralidade desempenham um papel fundamental na análise de grafos, oferecendo uma abordagem quantitativa para identificar nós de maior relevância ou influência em uma rede. Diferentes métricas capturam distintas noções de importância, variando conforme o contexto e a estrutura do grafo. A centralidade avalia a relevância de um nó com base em critérios como sua posição, conectividade e interações com outros nós. No caso de grafos direcionados, podem-se considerar centralidades de grau de entrada e de saída. A compreensão da centralidade é vital para analisar a estrutura e o fluxo de informações em uma rede. Nós com alta centralidade de intermediação, por exemplo, podem funcionar como pontes ou controladores de acesso em uma rede. Em redes de infraestrutura, reconhecer nós de alta centralidade é essencial para identificar pontos vulneráveis e otimizar a resiliência da rede. No âmbito epidemiológico, nós com

elevada centralidade são mais suscetíveis a contrair e disseminar doenças, orientando estratégias de vacinação. Já em ambientes empresariais ou estratégicos, a análise de centralidade pode informar decisões sobre alocação de recursos ou maximização da disseminação de informações.

Centralidade de Grau

No contexto da análise de grafos, talvez a medida de centralidade mais simples para um nó seja o seu grau, que se refere ao número de arestas conectadas a ele. Na literatura de teoria dos grafos, o grau é por vezes denominado de **centralidade de grau**, para enfatizar seu uso como uma medida de centralidade. Em grafos direcionados, os nós possuem tanto um grau de entrada (in-degree) quanto um grau de saída (out-degree), e ambos podem ser valiosos como medidas de centralidade, dependendo do contexto em análise. Apesar de ser uma métrica de centralidade simples, a centralidade de grau pode ser extremamente reveladora. Em um grafo social, por exemplo, é razoável supor que indivíduos que têm muitos amigos ou conhecidos possam ter mais influência, mais acesso à informação ou mais prestígio do que aqueles com menos conexões. Um exemplo não-social é o uso de contagens de citações na avaliação de artigos científicos. O número de citações que um artigo recebe, que corresponde ao seu in-degree no grafo direcionado de citações, fornece uma medida quantitativa da influência do artigo e é amplamente usado para avaliar o impacto da pesquisa científica. Além disso, em modelos epidemiológicos, a centralidade de grau pode ser usada para identificar os indivíduos ou locais mais prováveis de serem pontos iniciais de disseminação de uma doença, ou para identificar indivíduos que, se vacinados, poderiam interromper mais eficazmente a disseminação de uma doença (Newman, 2018).

Formalmente, para um grafo $G = (V, E)$, a centralidade de grau $C_D(v)$ de um vértice v é definida como $C_D(v) = d(v)$. Para grafos direcionados, também podemos definir o in-degree e o out-degree para cada vértice. A centralidade de grau é frequentemente normalizada dividindo-se pelo valor máximo possível, que é $|V| - 1$. Assim, a centralidade de grau normalizada é definida pela equação (2.12).

$$C_D^{norm}(v) = \frac{d(v)}{|V| - 1}. \quad (2.12)$$

Centralidade de Autovetor

Suponha que cada nó $v \in V$ de um grafo $G = (V, E)$ comece com um valor de centralidade $C_A(v)$. Em cada passo iterativo, um nó distribui uma fração de sua centralidade atual entre seus vizinhos. Após diversas iterações, os valores convergem, resultando na **centralidade de autovetor** (*Eigenvector centrality*). Essa relação é expressa na Equação (2.13), onde $N(v)$ denota o conjunto de vizinhos do nó v e $d(v)$ representa o seu grau. Podemos começar com um vetor inicial arbitrário de centralidades (por exemplo, todas sendo 1) e iterar usando a fórmula até que a variação nos valores de centralidade entre iterações consecutivas esteja abaixo de um limiar definido.

$$C_A(v) = \frac{1}{d(v)} \sum_{u \in N(v)} C_A(u) \quad (2.13)$$

A definição iterativa sublinha o conceito de que a importância (ou centralidade) de um nó é intrinsecamente ligada à importância de seus vizinhos. Uma visão alternativa para a centralidade de autovetor é fornecida pelo estudo da matriz de adjacência A do grafo. A centralidade de autovetor é o autovetor correspondente ao maior autovalor de A . Denotando o autovetor de centralidade por \mathbf{C}_A e o maior autovalor por λ , temos que a relação definida pela equação (2.14). Esta relação matricial reitera o conceito de que a centralidade de um nó é proporcional à soma das centralidades de seus vizinhos, em harmonia com a definição iterativa. Ambas as perspectivas, seja a iterativa ou a matricial, destacam o papel crucial dos vizinhos de um nó na determinação de sua centralidade dentro do grafo.

$$A\mathbf{C}_A = \lambda\mathbf{C}_A \quad (2.14)$$

Centralidade de Katz

A **Centralidade de Katz** é uma métrica de centralidade em redes que leva em consideração a proximidade dos nós. Diferente de outras métricas que podem se basear puramente no grau de um nó ou nos caminhos mais curtos até outros nós, a **centralidade de Katz** considera todos os caminhos que conectam um nó a todos os outros nós na rede, dando menos peso a caminhos mais longos de acordo com um fator de atenuação (Katz, 1953). Formalmente, a centralidade de Katz de um nó é calculada através da equação 2.15.

$$C_K(v) = \sum_{l=1}^{\infty} \sum_{u \in V} \alpha^l A_{uv}^l \quad (2.15)$$

Onde:

- $C_K(v)$ é a centralidade do nó v .
- A_{uv}^l é o número de caminhos de comprimento l que conectam o nó u ao nó v . Obtemos A^l ao elevar a matriz de adjacência à potência l .
- α é um fator de atenuação (com $0 < \alpha < 1$) que dá menos peso a caminhos mais longos.

A Centralidade de Katz também pode ser expressa em relação à matriz de adjacência do grafo. Dado um grafo com matriz de adjacência A , a centralidade de Katz de todos os nós do grafo pode ser obtida pela equação (2.16).

$$\mathbf{C}_K = (\mathbf{I} - \alpha A)^{-1} \mathbf{1} \quad (2.16)$$

Onde:

- \mathbf{C}_K é um vetor coluna de elementos $C_K(v)$ da centralidade de Katz do vértice v .
- I é a matriz identidade do mesmo tamanho que A .
- α é o mesmo fator de atenuação já descrito, com $0 < \alpha < \frac{1}{\lambda_{max}}$ onde λ_{max} é o maior autovalor de A .
- $\mathbf{1}$ é um vetor coluna de uns, do mesmo tamanho que \mathbf{C}_K .

Essa formulação matricial fornece uma maneira rápida e eficiente de calcular a centralidade de Katz para todos os nós do grafo de uma só vez, especialmente para grafos grandes, utilizando métodos numéricos. Além disso, relacionando a definição original de Katz com a matriz de adjacência, podemos ver o papel crucial que a estrutura global do grafo desempenha na determinação da centralidade de cada vértice.

Intuitivamente, a ideia é que nós que estão próximos de muitos outros nós (independentemente de terem muitas conexões diretas ou não) terão uma alta centralidade de Katz. O fator α garante que caminhos mais longos sejam ponderados menos do que caminhos mais curtos. O valor de α deve ser escolhido de modo que seja menor que o inverso do maior autovalor da matriz de adjacência do grafo, para garantir a convergência da métrica. A centralidade de Katz é especialmente útil em situações em que caminhos indiretos entre nós (mesmo aqueles que são mais longos) são considerados importantes para a influência ou relevância de um vértice.

Centralidade de Proximidade (closeness centrality)

A centralidade de proximidade (*closeness centrality*) é uma métrica fundamental na análise de grafos que mede o quão próximo um nó está de todos os outros nós no grafo. Ela fundamenta-se na ideia de que os nós centrais são aqueles que podem interagir mais rapidamente com todos os outros nós do grafo, tornando-se, assim, influentes ou acessíveis. Diferentemente de outras métricas baseadas em conceitos matriciais, como os autovetores, a centralidade de proximidade é baseada em caminhos mais curtos em grafos. Em termos simples, esta métrica calcula a média das distâncias dos caminhos mais curtos entre um nó e todos os outros nós do grafo. Assim, um nó com uma alta centralidade de proximidade tem, em média, uma distância menor até os outros nós (Newman, 2018). Matematicamente, se definirmos $\ell_{i,j}$ como a menor distância entre os nós i e j , a média das menores distâncias do nó i para todos os outros nós no grafo é dada pela equação (2.17)

$$\ell_i = \frac{1}{n} \sum_j \ell_{i,j} \quad (2.17)$$

Este valor, ℓ_i , assume valores menores para nós que estão, em média, mais próximos dos outros nós. Como tal, a centralidade de proximidade é frequentemente definida como o inverso de ℓ_i , conforme a equação (2.18).

$$C_i = \frac{n}{\sum_j \ell_{i,j}} \quad (2.18)$$

Essa métrica é uma ferramenta poderosa em muitas aplicações. Por exemplo, na análise de grafos de atores de cinema, ela tem sido usada para classificar atores com base em suas colaborações em filmes (Newman, 2018). É importante notar que há nuances ao calcular a centralidade de proximidade em grafos com múltiplos componentes. Em grafos desconexas, a distância entre nós de componentes diferentes é considerada infinita. Uma abordagem para contornar essa limitação é redefinir a centralidade em termos da média harmônica das distâncias entre os nós (Newman, 2018), como definido pela equação (2.19).

$$C'_i = \frac{n}{\sum_j \frac{1}{\ell_{i,j}}} \quad (2.19)$$

Esta definição dá mais peso aos nós próximos e é menos sensível a nós distantes, refletindo a ideia de que distâncias curtas são mais cruciais em muitos cenários práticos. A centralidade de proximidade é uma métrica naturalmente intuitiva e tem sido amplamente utilizada em estudos de grafos sociais e outras análises de grafos para destacar a importância relativa dos nós em diferentes contextos.

Centralidade de Intermediação (Betweenness Centrality)

A **centralidade de intermediação** (*betweenness centrality*) é uma medida que captura até que ponto um vértice se situa em caminhos entre outros vértices dentro de um grafo. A ideia fundamental é considerar o quanto um vértice atua como uma ponte ou intermediário no fluxo de informações, dados, mensagens ou outro tráfego dentro de um grafo. Para um grafo não direcionado onde existe no máximo um caminho mais curto entre os vértices, a representação matemática da centralidade de intermediação é dada pela equação (2.20). No entanto, se existirem múltiplos caminhos mais curtos entre os vértices, para cada caminho é atribuído a um peso. A fórmula geral, nesse caso, deve ser dada pela equação (2.21), onde n_{st}^i representa o número de caminhos mais curtos do vértice s para o vértice t que passam pelo vértice i e g_{st} é o número total de caminhos mais curtos de s para t .

$$C_B(i) = \sum_{st} n_{st}^i \quad (2.20)$$

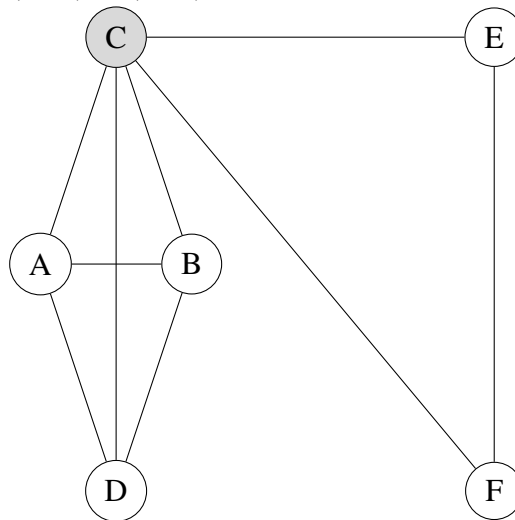
$$C_B(i) = \sum_{st} \frac{n_{st}^i}{g_{st}} \quad (2.21)$$

Para normalizar os valores de centralidade de intermediação, dividimos pelo número total de pares de nós ordenados (n^2), resultando em uma fração - que fica entre 0 e 1 - dos caminhos que passam por um determinado vértice, conforme a equação (2.22). Tais valores normalizados oferecem uma comparação mais abrangente entre nós em várias redes. (Newman, 2018). A figura 2.1 ilustra a importância de um vértice com elevada centralidade de intermediação para a integridade do grafo.

$$C_B(i) = \frac{1}{n^2} \sum_{st} \frac{n_{st}^i}{g_{st}} \quad (2.22)$$

Embora o conceito básico gire em torno dos caminhos mais curtos, existem variações da centralidade de intermediação. Algumas consideram caminhos além dos mais curtos, reconhecendo que o tráfego do grafo no mundo real pode não seguir sempre a rota mais direta. A *centralidade de intermediação de fluxo* é uma variante da centralidade de intermediação tradicional, considerando caminhos independentes das arestas entre os pares de vértices, ao invés de se restringir apenas aos caminhos mais curtos (Newman, 2018). Se houver mais de uma possível escolha de caminhos independentes entre um par de vértices, a contribuição para a centralidade de intermediação de um determinado vértice para esse par é definida como o máximo entre todas as escolhas. Essa métrica busca capturar a ideia de que o tráfego em um grafo pode seguir diferentes rotas, e não apenas o caminho mais direto. Matematicamente, pode ser calculado pela equação 4, onde n_{st}^i representa o número de caminhos independentes das arestas entre os vértices s e t que passam pelo vértice i e g_{st} é o número total de caminhos independentes das arestas entre s e t .

Figura 4 – Neste grafo, o vértice "C" é destacado (preenchido em cinza) para indicar sua posição como o ponto de maior centralidade de intermediação. O vértice "C" está conectado a todos os outros vértices do grafo, atuando como um intermediário crucial no fluxo entre eles. Isso ilustra visualmente que a centralidade de intermediação de "C" é maior, pois ele se situa em muitos dos caminhos mais curtos entre os pares de vértices no grafo. Centralidade de cada vértice: A (0.1), B (0.1), C (0.6), D (0.1), E (0.05), F (0.05).



$$C_B(i) = \sum_{st} \max\left(\frac{n_{st}^i}{g_{st}}\right) \quad (2.23)$$

A *centralidade de intermediação de passeio aleatório* é uma métrica que imagina que as mensagens ou o tráfego em um grafo realizam caminhadas aleatórias entre todos os

possíveis pontos de partida e destino. Aqui, a centralidade de intermediação é definida como o número médio de tais mensagens ou unidades de tráfego que passam por cada vértice. Em vez de focar nos caminhos mais curtos ou diretos, essa métrica considera a ideia de que o tráfego em um grafo pode não ter uma direção clara e pode se mover de forma errática até atingir seu destino (Newman, 2018). Matematicamente, a *centralidade de intermediação de passeio aleatório* pode ser formulada usando uma representação de passeios aleatórios w_{st}^i entre os vértices s e t que passam pelo vértice i , onde g_{st} é o número total de passeios aleatórios entre s e t , conforme a equação (2.24).

$$C_B(i) = \sum_{st} \left(\frac{w_{st}^i}{g_{st}} \right) \quad (2.24)$$

Conceitos de poder, frequentemente discutidos em sociologia, estão entrelaçados com a centralidade de intermediação. Vértices com maior centralidade de intermediação podem alavancar sua posição para controlar o fluxo de informações, possivelmente obtendo uma vantagem em cenários em que isso se traduz em benefícios tangíveis, como em relacionamentos comerciais (Newman, 2018).

O cálculo da centralidade de intermediação (*betweenness centrality*) é uma tarefa intensiva em termos computacionais. Para um grafo não-dirigido e sem pesos com V vértices e E arestas, a ordem de complexidade do algoritmo mais eficiente conhecido é $O(VE)$, usando uma abordagem baseada em busca em largura (BFS, *Breadth-First Search*). Já para grafos ponderados (aqueles cujas arestas possuem peso), a ordem de complexidade é $O(VE + V^2 \log V)$ usando algoritmos de caminho mais curto, como o algoritmo de Dijkstra. Estas complexidades resultam da necessidade de calcular os caminhos mais curtos entre todos os pares de vértices no grafo, e depois, para cada par, distribuir a centralidade de intermediação proporcionalmente entre todos os vértices que fazem parte dos caminhos mais curtos.

Coefficiente de agrupamento (Cluster coefficient)

Uma relação R em um conjunto A é classificada como transitiva se, para quaisquer elementos x , y e z pertencentes a A , a relação entre x e y (indicada como xRy) e entre y e z (indicada como yRz) implica necessariamente em uma relação direta entre x e z (ou seja, xRz). Formalmente, esta propriedade é expressa pela seguinte fórmula (2.25) (Fraleigh, 2003).

$$\forall x, y, z \in A, (xRy \wedge yRz) \rightarrow xRz \quad (2.25)$$

Por exemplo, consideremos a relação de "ser maior que" em um conjunto numérico. Esta relação é transitiva pois, se um número a é maior que um número b , e b é maior que um número c , então é garantido que a será maior que c . A transitividade tem um papel significativo teoria dos grafos, especialmente nas não direcionados, exemplificada pela ideia de que "o amigo do meu amigo também é meu amigo". Neste contexto, se um nó u está conectado a um nó v , e v está conectado a w , então, sob a condição de transitividade, u estaria igualmente conectado a w . Contudo, a transitividade perfeita é rara na maioria das redes, ocorrendo somente em casos onde cada componente é um clique completo. Em redes como as sociais, observa-se mais frequentemente uma forma parcial de transitividade, onde a conexão entre u e w é mais provável se ambos têm um conhecimento comum em v . O coeficiente de agrupamento C quantifica esta ideia de transitividade. É definido como a fração de caminhos de comprimento dois no grafo que formam triângulos fechados. Matematicamente, é expresso pela equação (2.26). O valor de C pode variar de 0 (nenhum triângulo fechado) a 1 (transitividade perfeita). Uma forma alternativa de calcular C leva em conta os triângulos no grafo é definido pela equação (2.27), onde o fator de seis é devido ao fato de que cada triângulo contém seis caminhos de comprimento dois (Newman, 2018).

$$C = \frac{\text{(número de caminhos fechados de comprimento dois)}}{\text{(número de caminhos de comprimento dois)}} \quad (2.26)$$

$$C = \frac{\text{(número de triângulos)} \times 6}{\text{(número de caminhos de comprimento dois)}} \quad (2.27)$$

Para calcular o coeficiente de agrupamento local de um nó individual i , contamos quantos dos seus pares de vizinhos estão conectados e dividimos esse número pelo total de pares de vizinhos, que é $\frac{1}{2}k_i(k_i - 1)$, sendo k_i o grau do vértice. O coeficiente de agrupamento para toda a rede é obtido estendendo este cálculo a todos os vértices, considerando cada par de vizinhos u, v com $u < v$ e verificando se eles estão conectados. O total de conexões é dividido pelo número de triplas conectadas, conforme a equação (2.28). O tempo de execução desse algoritmo depende da distribuição dos graus na rede. Especificamente, ele é afetado pelo número

de nós n , o grau médio e o quadrado médio do grau, tornando o cálculo mais complexo do que simplesmente calcular os graus dos nós (Newman, 2018).

$$\sum_i \frac{1}{2} k_i (k_i - 1) \quad (2.28)$$

Assortatividade

A **assortatividade**, também chamada de *homofilia*, em grafos refere-se à propensão dos nós se conectarem com outros que possuem características semelhantes. Em um grafo $G(V, E)$, podemos avaliar a assortatividade observando a distribuição das características dos vizinhos de cada nó. Para características categóricas não ordenadas, a assortatividade indica a tendência dos nós associarem-se com base na similaridade dessas características. Já para características numéricas, observamos como os valores desses atributos influenciam as conexões entre os nós. Em contraste, a **disassortatividade** representa a tendência de nós se associarem a outros com características diferentes. A avaliação rigorosa da assortatividade depende tanto da natureza das características quanto dos métodos empregados para examinar as associações entre os vértices (Newman, 2018).

Assortatividade por características sem relação de ordem.

Considere que os vértices de G são classificados de acordo com uma característica que assume um conjunto finito de possíveis valores, denotados por $1, \dots, N$. Esses valores são apenas descritivos e não têm uma ordem específica. Seja g_i o grupo, classe ou tipo do nó i . Para um grafo ser considerado assortativo, uma proporção significativa das arestas deve conectar nós do mesmo tipo. Vamos definir um quantificador matemático para essa tendência. Primeiro, a quantidade total de arestas que conectam nós do mesmo tipo é dada pela equação (2.29). Onde A_{ij} é a matriz de adjacência do grafo e $\delta_{g_i g_j}$ é o delta de Kronecker, que é 1 se $g_i = g_j$ e 0 caso contrário (Newman, 2018).

$$E_{total} = \frac{1}{2} \sum_{ij \in E} A_{ij} \delta_{g_i g_j} \quad (2.29)$$

Para determinar o que esperaríamos ver em um grafo aleatório, consideramos que a probabilidade do outro extremo de uma aresta que está conectada ao nó i também estar conectada

ao nó j é $\frac{k_j}{2m-1}$, onde k_j é o grau do nó j e m é o número total de arestas no grafo. Assim, a equação (2.30) define o número esperado de arestas entre os nós i e j . Então, podemos definir o número esperado de arestas entre nós do mesmo tipo através da equação (2.31).

$$\langle E_{ij} \rangle = \frac{k_i k_j}{2m-1} \quad (2.30)$$

$$\langle E \rangle = \frac{1}{2} \sum_{ij} \langle E_{ij} \rangle \delta_{g_i g_j} = \frac{1}{2} \sum_{ij} \left(\frac{k_i k_j}{2m-1} \right) \delta_{g_i g_j} \approx \frac{1}{2} \sum_{ij} \frac{k_i k_j}{2m} \delta_{g_i g_j} \quad (\text{para } m \gg 1) \quad (2.31)$$

A modularidade, denotada por Q , é uma métrica crucial em redes complexas que quantifica a estrutura de comunidades dentro de um grafo. Esta medida avalia a densidade de conexões entre os vértices dentro de comunidades em comparação com a densidade que seria esperada em um cenário de conexões aleatórias. Matematicamente, a modularidade é expressa como a diferença normalizada entre o número observado e o esperado de arestas conectando nós do mesmo tipo, dividida pelo número total de arestas na rede. O valor de Q varia, sendo estritamente menor que 1. Um valor positivo indica que há mais arestas entre nós do mesmo tipo do que seria esperado ao acaso, refletindo uma estrutura de rede com comunidades bem definidas. Por outro lado, valores negativos de modularidade ocorrem quando há menos arestas interconectando nós semelhantes do que o esperado, sugerindo uma estrutura de rede menos clara ou até mesmo uma disposição aleatória de conexões. Além disso, a modularidade está intimamente relacionada ao conceito de assortatividade, que se refere à tendência de vértices na rede se conectarem com outros que possuem características semelhantes, como o número de conexões. Em contextos práticos, uma alta modularidade em uma rede assortativa sugere a presença de grupos coesos, onde vértices com propriedades semelhantes tendem a formar comunidades densamente interligadas. A modularidade Q pode ser reformulada utilizando as variáveis e_r e a_r , conforme descrito nas equações (2.32) e (2.33), respectivamente. Aqui, e_r representa a fração de arestas que conectam nós do mesmo tipo r , enquanto a_r denota a fração de extremidades de arestas ligadas a nós desse tipo. Esta reformulação é possível graças à propriedade $\delta_{g_i g_j} = \sum_r \delta_{g_i, r} \delta_{g_j, r}$, que nos permite expressar Q através da equação (2.34). Esta abordagem é particularmente útil em situações onde dispomos de uma lista de arestas e conhecemos os tipos dos nós em suas extremidades, mas não temos informações explícitas sobre os graus dos vértices (Newman, 2018).

$$e_r = \frac{1}{2m} \sum_{ij} A_{ij} \delta_{g_i,r} \delta_{g_j,r} \quad (2.32)$$

$$a_r = \frac{1}{2m} \sum_i k_i \delta_{g_i,r} \quad (2.33)$$

$$\begin{aligned} Q &= \frac{1}{m} (E_{total} - \langle E \rangle) = \frac{1}{m} \left[\left(\frac{1}{2} \sum_{ij} A_{ij} \delta_{g_i g_j} - \frac{1}{2} \sum_{ij} \frac{k_i k_j}{2m} \delta_{g_i g_j} \right) \right] = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta_{g_i g_j} \\ &= \sum_r (e_r - a_r^2) \end{aligned} \quad (2.34)$$

Assortividade por características com relação de ordem

Podemos ter assortividade em uma rede de acordo com características como idade ou renda, cujos valores seguem uma ordem particular. É possível dizer quando dois nós são aproximadamente iguais, de acordo com essa característica. Se os nós da rede com valores similares de uma característica escalar tendem a se conectar mais frequentemente do que aqueles com valores diferentes, então a rede é considerada assortativa de acordo com essa característica. Podemos também falar de grafos cujos vértices possuem características vetoriais, e assim, a similaridade pode ser medida por uma métrica apropriada (Newman, 2018). Podemos medir a assortividade entre quantidades escalares x_i e x_j para os nós i, j através da covariância de x_i e x_j , como definida pela equação (2.35), onde $\mu = \frac{1}{2m} \sum_u k_u x_u$ representa a média aritmética dos valores que caracterizam os vértices do grafo.

$$\begin{aligned} \text{cov}(x_i, x_j) &= \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) x_i x_j \\ &= \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) x_i x_j \\ &= \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) x_i x_j \end{aligned} \quad (2.35)$$

A covariância será positiva se, em média, os valores x_i, x_j e tendem a ser ambos grandes ou ambos pequenos, e negativa se eles tendem a variar em direções opostas. Em outras

palavras, a covariância será positiva quando tivermos assortividade e negativa para dissortividade. A medida normalizada, chamada coeficiente de assortividade, é a razão dos dois e é definida pela equação (2.36). Este coeficiente de correlação varia entre um máximo de 1 para uma rede perfeitamente assortativa e um mínimo de -1 para uma perfeitamente dissortativa. Em um grafo, podem existir correlações não-lineares e ainda assim teríamos $r = 0$. O coeficiente de correlação detecta apenas correlações lineares. Portanto, é preciso cautela ao verificar os valores de r para investigar assortividade.

$$r = \frac{\sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) x_i x_j}{\sum_{ij} \left(k_i \delta_{ij} - \frac{k_i k_j}{2m} \right) x_i x_j} \quad (2.36)$$

Assortividade dos graus de vértices

A assortividade em relação aos graus dos vértices, é um caso especial da assortividade por características com relação de ordem. Em um grafo que apresenta essa característica, nós de alto grau tendem a se conectar preferencialmente com outros nós de alto grau e os de baixo grau se conectam com nós de baixo grau. Por exemplo, em uma rede social, há assortividade por grau se pessoas gregárias são amigas de outras pessoas gregárias e os reclusos se conectam com outros reclusos. Em contraste, pode existir dissortividade por grau, o que significaria que as pessoas gregárias interagem com os reclusos e vice-versa (Newman, 2018). O grau, diferentemente de outras propriedades como idade ou renda, é uma propriedade da estrutura da rede. Quando uma propriedade estrutural (os graus) dita outra (as posições das arestas), surgem características interessantes nas redes. Por exemplo, em uma rede assortativa, espera-se um agrupamento ou núcleo de nós de alto grau cercados por uma periferia menos densa de nós de grau inferior. Essa estrutura de núcleo-periferia é uma característica comum de muitas redes, principalmente redes sociais. Por outro lado, em redes dissortativas, nós de alto grau tendem a se conectar com os de baixo grau, criando características estelares na rede. Nessas redes, a estrutura de núcleo-periferia não é comum, sendo a rede mais uniforme. A assortividade por grau pode ser medida da mesma forma que a assortividade por qualquer outra quantidade escalar.

2.2 Percolação

A teoria da percolação, que tem suas raízes em trabalhos de Broadbent e Hammersley, é fundamental no estudo de fractais aleatórios na natureza. Neste contexto, a terminologia padrão

da teoria da percolação difere da teoria dos grafos: os vértices e arestas são chamados de sítios e ligações, respectivamente, e os componentes são chamados de aglomerados ou *clusters* (Bollobás e Riordan, 2006). Quando nosso subgrafo aleatório é obtido selecionando vértices, falamos de percolação de sítios; quando selecionamos arestas, falamos de percolação de ligações. Em ambos os casos, os sítios ou ligações selecionados são chamados de abertos, e aqueles não selecionados são chamados de fechados. No caso da percolação de sítios, o subgrafo aberto é induzido pelos sítios abertos; na percolação de ligações, o subgrafo aberto é formado pelas arestas abertas e todos os vértices. A essência da percolação pode ser visualizada considerando uma rede ou grade, onde cada ligação ou sítio está presente com uma probabilidade p e ausente com $1 - p$. Para um p suficientemente grande, emerge um aglomerado que se estende por toda a rede. Em uma rede infinita, essa formação ocorre ao cruzar um limiar crítico, p_c (Ben-Avraham e Havlin, 2000). Um aspecto intrigante da percolação é a transição de fase que ocorre na probabilidade P_∞ de uma ligação pertencer ao aglomerado infinito. Esta probabilidade é zero para $p < p_c$ e aumenta continuamente para $p > p_c$. Perto e acima do ponto de transição, segue uma lei de potência (Stauffer e Aharony, 2018), como definido na equação (2.37).

$$P_\infty \propto (p - p_c)^\beta \quad (2.37)$$

Este fenômeno é conhecido como a transição de percolação. O nome vem da possível interpretação de ligações como canais abertos para o fluxo de um fluido em um meio poroso (ligações ausentes representam canais bloqueados). No ponto de transição, o fluido pode percolar através do meio pela primeira vez. A taxa de fluxo passa por uma transição de fase semelhante à de P_∞ . De fato, a transição é semelhante a todas as outras transições de fase contínuas (de segunda ordem) em sistemas físicos. P_∞ desempenha o papel de um parâmetro de ordem, análogo à magnetização em um ferromagnetismo, e β é o expoente crítico do parâmetro de ordem (Stauffer e Aharony, 2018). Expoentes críticos são quantidades que descrevem como várias propriedades do sistema mudam quando nos aproximamos do limiar de percolação, p_c . Eles nos dão *insights* sobre como a rede se comporta perto de transições críticas, o que é fundamental em muitos campos da física. Existe uma grande variedade de modelos de percolação. Por exemplo, o modelo pode ser definido em diferentes tipos de reticulados. Na percolação de sítios, os elementos de percolação são sítios da rede, em vez de ligações, e consideramos sítios vizinhos mais próximos como pertencentes ao mesmo aglomerado. Outras regras de conectividade podem ser empregadas, como na percolação de inicialização, onde um subconjunto do aglomerado está

conectado se estiver ligado por pelo menos dois sítios ou ligações. Há também a percolação contínua, que é definida sem a necessidade de uma rede específica, onde os elementos podem ser, por exemplo, formas que se sobrepõem parcialmente. A percolação também pode ser considerada em diferentes dimensões espaciais. Embora o limiar de percolação p_c possa variar dependendo do modelo escolhido, os expoentes críticos, como β , são determinados apenas pela dimensão espacial. Essa característica, chamada de universalidade, mostra que os expoentes críticos são fundamentais para entender a essência dos modelos de percolação. Eles são usados para classificar transições de fase críticas em classes de universalidade. (Ben-Avraham e Havlin, 2000)

O comprimento típico dos *clusters* finitos é caracterizado pelo comprimento de correlação ξ . Ele diverge à medida que p se aproxima de p_c como definido na equação (2.38) com o mesmo expoente crítico ν abaixo e acima da transição. A massa média (o número de sites na percolação de site ou o número de ligações na percolação de ligação) dos *clusters* finitos, S , é análoga à susceptibilidade magnética nas transições de fase ferromagnéticas. Ela diverge em torno de p_c como definido na equação (2.39) (Ben-Avraham e Havlin, 2000).

$$\xi \propto |p - p_c|^{-\nu} \quad (2.38)$$

$$S \propto |p - p_c|^{-\gamma} \quad (2.39)$$

A teoria da percolação tem vastas aplicações práticas. Considere, por exemplo, a falha de roteadores na Internet. A remoção de um roteador afeta não apenas sua capacidade de receber dados, mas também a de outros roteadores conectados a ele. A percolação ajuda a entender os efeitos cascata dessas falhas em toda a rede (Newman, 2018).

Percolação por ataque aleatório

Ataque aleatório é uma metodologia fundamental no estudo da robustez e resiliência de redes (Newman, 2018). Esse tipo de ataque é caracterizado pela remoção sequencial e aleatória de vértices de um grafo, sem qualquer critério preferencial. Matematicamente, dado um grafo $G(V, E)$, onde V é o conjunto de vértices e E é o conjunto de arestas, um ataque aleatório seleciona um vértice $v \in V$ com probabilidade uniforme e o remove, juntamente com todas as

suas arestas incidentes. O processo é repetido até que um determinado critério de parada seja atingido, como a remoção de uma fração específica de vértices ou até que o grafo se desintegre completamente. O principal parâmetro associado ao ataque aleatório é a probabilidade p , que denota a fração de vértices que permanecem na rede após o ataque. No contexto da teoria da percolação, p é frequentemente referido como a probabilidade de ocupação. Quando $p = 1$, todos os vértices estão presentes, e quando $p = 0$, todos os vértices foram removidos (Newman, 2018).

Um aspecto crítico da percolação aleatória é a transição de percolação. Para valores elevados de p , a rede geralmente contém um componente gigante que engloba uma grande fração dos vértices. No entanto, à medida que p diminui, chega-se a um ponto crítico, o limiar de percolação, onde o componente gigante se desintegra, deixando apenas pequenos aglomerados ou componentes isolados. Este limiar é de fundamental importância, pois indica a vulnerabilidade da rede a ataques aleatórios. A natureza dessa transição e o valor exato do limiar de percolação dependem da topologia da rede subjacente. Em grafos aleatórios, por exemplo, a transição ocorre de maneira abrupta em um valor crítico específico de p . No entanto, em redes com distribuições de grau de cauda pesada, a rede pode ser mais resiliente a ataques aleatórios, com um limiar de percolação mais baixo (Newman, 2018).

Percolação por ataque malicioso

Na teoria de redes complexas, um **ataque malicioso** refere-se à remoção sistemática de vértices baseada em certos critérios, em contraste com a remoção aleatória. O critério mais comum para um ataque malicioso é a remoção de vértices em ordem decrescente de grau, dada a importância dos nós de alto grau (hubs) para a manutenção da conectividade em muitas redes (Newman, 2018). Seja $G(V, E)$ um grafo não orientado, onde V e E representam o conjunto de vértices e arestas, respectivamente. Para um ataque malicioso baseado em grau, inicialmente, calcula-se o grau k_i de todos os vértices i no conjunto de vértices V . Em seguida, identifica-se o vértice j que possui o maior grau, de modo que $k_j \geq k_i$ para todos os $i \in V$. Esse vértice j , juntamente com todas as suas arestas associadas, é então removido de V e do conjunto de arestas E . Os passos de cálculo do grau, identificação do vértice de maior grau e sua remoção são repetidos iterativamente até que o conjunto de vértices V esteja vazio ou até que um critério de parada pré-definido seja alcançado. O procedimento pode ser resumido pelo (Algoritmo 1).

Algoritmo 1: Ataque malicioso

Data: Grafo $G(V, E)$
Result: Grafo G após remoção dos vértices seguindo o critério especificado

while $V \neq \emptyset$ e critério de parada não alcançado **do**

 Calcule o grau k_i para todos os vértices $i \in V$;

 Identifique o vértice j tal que $k_j \geq k_i$ para todos $i \in V$;

 Remova o vértice j de V e todas as suas arestas associadas de E ;

end

Neste contexto, o **ataque malicioso** refere-se ao estudo das propriedades de conectividade do grafo G à medida que os vértices são removidos de acordo com o protocolo acima. Especificamente, estamos interessados em como a maior componente conectada de G evolui durante o ataque. Quando comparado com a remoção aleatória, onde vértices são removidos sem considerar suas propriedades, o ataque malicioso é muito mais destrutivo. Em redes com distribuição de grau livre de escala, por exemplo, a remoção de uma pequena fração dos hubs pode fragmentar a rede em muitas componentes desconexas, eliminando a componente gigante. Matematicamente, o impacto de um ataque malicioso pode ser quantificado pela fração de vértices removidos p para reduzir a proporção da maior componente conectada para abaixo de um limiar especificado. Em muitas redes reais, essa fração é surpreendentemente pequena, ilustrando a vulnerabilidade das redes a ataques direcionados. No contexto de redes biológicas ou sociais, o ataque malicioso pode ser interpretado como uma estratégia de intervenção, como a vacinação direcionada em redes de contatos para prevenir a propagação de doenças. A eficácia desta estratégia, no entanto, depende do conhecimento preciso da estrutura da rede e da capacidade de intervir nos *hubs* (Newman, 2018).

Além da centralidade de grau, ataques maliciosos em redes complexas podem ser realizados com base em outras medidas de centralidade, como autovetor, Katz, proximidade e intermediação. Cada uma dessas medidas pode impactar a rede de maneiras distintas, afetando sua estrutura e funcionamento. Por exemplo, a remoção de nós com alta centralidade de autovetor pode desintegrar partes críticas da rede, enquanto a eliminação de nós com alta centralidade de Katz pode prejudicar a eficiência da comunicação. Da mesma forma, atacar nós com alta centralidade de proximidade pode aumentar as distâncias médias na rede, e a remoção de nós com alta centralidade de intermediação pode fragmentar a rede e eliminar rotas críticas. Estas estratégias podem ser adaptadas para diferentes contextos, como interrupção de redes de

comunicação em cenários de segurança ou otimização de estratégias de vacinação em redes de contágio. O impacto e a escolha da medida de centralidade para o ataque dependem das características específicas e da natureza da rede.

Decomposição k-core

A decomposição em k-core é uma técnica utilizada para organizar redes em subgrupos hierárquicos e ordenados. Este método consiste na remoção sequencial de nós que possuem um grau menor que o índice específico da camada atual, continuando até que nenhuma remoção adicional seja viável. Na aplicação prática, a rede é segmentada em várias camadas concêntricas, que vão da mais externa até o núcleo central, ou cerne (do inglês *core*). Durante este processo, são realizadas podas iterativas baseadas no grau k de cada nó, resultando na alocação de cada nó em uma camada apropriada, conhecida como k-core, conforme o seu grau final. Este procedimento fornece informações valiosas sobre a configuração da rede, incluindo aspectos como a ocupação de cada camada, a posição do core em relação ao valor de k , e o ponto crítico de colapso K_γ da rede. O algoritmo de decomposição k-core é utilizado para analisar redes complexas.

O processo de decomposição k-core, conforme descrito por (Burleson-Lesser *et al.*, 2020), começa com a identificação e remoção de todos os nós na rede que têm um grau de conexão de $k = 1$, eliminando-os junto com suas conexões. Após essa remoção inicial, é possível que novos nós com grau 1 surjam, levando à repetição do processo de identificação e remoção desses nós até que não restem mais nós com grau 1, formando assim o k-core para $k = 1$. Este processo é então incrementado e repetido para valores maiores de k , classificando todos os nós da rede em diferentes k-cores, até que cada nó seja atribuído a um k-core específico. O k-core mais interno, ou seja, aquele com o maior valor de k (denominado k_{\max}), é considerado o "core" principal da rede. Os membros da rede são agrupados com base nos k-cores a que pertencem, proporcionando uma visão estratificada da rede. O procedimento é resumido no (Algoritmo 2).

Algoritmo 2: Decomposição k -core

Entrada: Grafo $G(V, E)$
Inicialização: Seja k_v o grau do vértice $v \forall v \in V$.

while existir vértices em G **do**

 Encontre um vértice v com o menor grau k_v no grafo G ;

if $k_v < k$ **then**

 Remova o vértice v e todas as suas arestas incidentes de G ;

 Atualize $k_u = k_u - 1$ para cada vértice u que era vizinho de v ;

else
break;

end

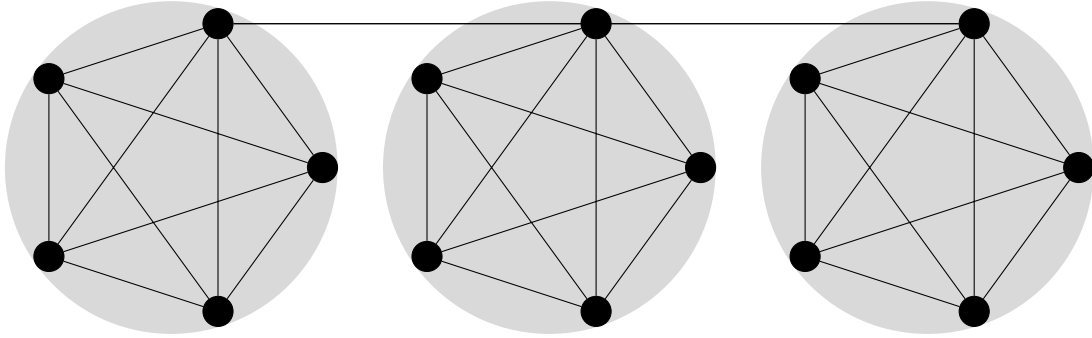
 Incremente k e continue para o próximo nível de k -core se nenhum vértice foi removido;

end

Existe uma relação entre os conceitos de k -core e k -componentes em redes complexas; contudo, eles representam ideias distintas e independentes. Isso implica que é viável a construção de um grafo que possua um k -core específico e, simultaneamente, contenha múltiplos k -componentes. A título de exemplo, a Figura 5 ilustra um grafo com um 4-core que, por sua vez, inclui três 4-componentes distintos.

Em redes complexas, define-se um **clique** como um subconjunto de n vértices onde cada vértice está conectado com os demais $n - 1$ vértices do grupo. Essa configuração corresponde a um $(n - 1)$ -core, ilustrando que todo clique é, por natureza, um k -core. Entretanto, a recíproca não é verdadeira: nem todo k -core configura um clique. Enquanto um clique representa um caso de máxima densidade de conexões entre vértices, um k -core reflete um conceito menos restrito, relacionado à coesão mais entre um conjunto de vértices em um grafo. Na análise de redes complexas, os cliques são frequentemente buscados como indicadores de grupos fortemente interconectados. Por outro lado, k -cores são empregados para identificar áreas de coesão na rede, assegurando um nível mínimo de conectividade sem, no entanto, alcançar a interconexão completa característica de um clique (Newman, 2018).

Figura 5 – Esta figura ilustra uma pequena rede que exemplifica a presença de um 4-core e, simultaneamente, três 4-componentes distintos. A rede, como um todo, representa um único 4-core, pois cada um de seus vértices mantém conexão com pelo menos outros quatro vértices. Contudo, dentro desta mesma rede, identificam-se três 4-componentes independentes, o que é evidenciado pelos círculos sombreados em cinza. Este exemplo demonstra claramente que os conceitos de k -cores e k -componentes são distintos. É importante observar também que cada um dos 4-componentes constitui um clique dentro do grafo, reforçando a ideia de que um clique representa um caso especial de densidade de conexão entre os vértices.



Coefficiente de Molloy-Reed

Para que um grafo com distribuição de grau p_k possua um componente gigante, um nó que é alcançado seguindo uma ligação do componente gigante deve ter em média pelo menos mais uma ligação para permitir a existência do componente. Para que isso ocorra, o grau médio de um nó alcançado (κ) deve ser pelo menos 2 (uma ligação de entrada e uma de saída) (Cohen e Havlin, 2010). Em uma formulação para o modelo configuracional de redes, a probabilidade de um nó ter um vizinho de grau k é calculada com base na estrutura geral da rede. Considerando que a rede consiste em n vértices e m arestas, a soma total dos graus de todos os vértices é $\sum_{i=1}^n k_i = 2m$. Neste contexto, a probabilidade de se conectar a um vértice de grau k é determinada pela probabilidade de se conectar a uma das pontas ou extremidades de aresta desse vértice. Como cada vértice de grau k contribui com k pontas para o total na rede, a probabilidade de escolher um dessas pontas, quando $2m \gg 1$, é $\frac{k}{2m}$ (Newman, 2018). Multiplicando esta probabilidade pelo número de vértices com grau k (ou seja, np_k), obtemos uma expressão para a probabilidade de se conectar a um vértice de grau k através de uma aresta aleatória. Portanto, a probabilidade de um nó escolhido aleatoriamente ter um vizinho de grau k pode ser expressa pela seguinte equação (2.40)

$$\mathbb{P}(\text{nó vizinho tem grau } k) = \frac{k}{2m} np_k. \quad (2.40)$$

Para determinar o grau médio κ de um nó alcançado ao seguir uma ligação em um

componente gigante, aplicamos o cálculo da média sobre a probabilidade dada pela equação (2.40), conforme demonstrado na equação (2.41). Nesta equação, κ é a média ponderada dos graus dos nós, levando em consideração a probabilidade de alcançar cada nó de grau k . $\langle k \rangle$ e $\langle k^2 \rangle$ são, respectivamente, o grau médio e a média dos quadrados dos graus dos nós na rede. A equação (2.42) estabelece o limiar crítico para a formação de um componente gigante. Quando a razão entre a média dos quadrados dos graus e o grau médio é igual a 2, a rede atinge um ponto crítico que favorece a existência de um componente gigante (Molloy e Reed, 1995).

$$\kappa = \sum_{i=1}^n k \frac{k}{2m} n p_k = \frac{n}{2m} \sum_{i=1}^n k^2 p_k = \frac{1}{\langle k \rangle} \langle k^2 \rangle = \frac{\langle k^2 \rangle}{\langle k \rangle} \quad (2.41)$$

$$\frac{\langle k^2 \rangle}{\langle k \rangle} = 2 \quad (2.42)$$

2.3 Redes Complexas

Embora os termos redes e grafos sejam frequentemente empregados como sinônimos em diversos contextos, tais como na teoria dos grafos e na ciência das redes, existem diferenças significativas entre eles. Os grafos são entidades abstratas empregadas na matemática e na ciência da computação para analisar as relações e conexões entre objetos. Por outro lado, as redes representam a aplicação prática desses conceitos abstratos para resolver problemas reais, frequentemente exigindo uma abordagem interdisciplinar.

Redes complexas são estruturas semelhantes a teias que representam uma grande variedade de sistemas importantes do ponto de vista tecnológico e intelectual. Exemplos incluem as células vivas, descritas como uma rede complexa de substâncias químicas conectadas por reações químicas; a Internet, que é uma rede de roteadores e computadores ligados por diferentes tipos de conexões físicas ou sem fio; modas e ideias que se propagam em redes sociais, cujos nós são indivíduos e as arestas representam diversos tipos de relações sociais; e a propagação epidemiológica, na qual um agente infeccioso é transmitido com ou sem um vetor intermediário entre os indivíduos infectados. Estes sistemas são apenas alguns dos inúmeros exemplos que motivaram a comunidade científica a investigar os mecanismos que definem a topologia das redes complexas. Essa investigação e o desejo de compreender tais sistemas interconectados apresentam desafios significativos. A física, por exemplo, que se beneficiou muito do reducionismo, desenvolveu ferramentas para prever o comportamento de sistemas inteiros a partir das

propriedades de seus elementos individuais. No entanto, essas ferramentas são menos eficazes para descrever sistemas nos quais a distância física é irrelevante ou existe ambiguidade na interação entre componentes. Recentemente, identificou-se que a mecânica estatística oferece um quadro adequado para descrever tais sistemas interligados. Tradicionalmente, o estudo de redes complexas foi dominado pela teoria dos grafos, inicialmente com um foco em grafos regulares. Mais adiante, redes de grande escala sem princípios de design aparentes começaram a ser descritas dentro dessa teoria. Isso influenciou o entendimento sobre redes complexas por muitos anos. No entanto, com o aumento do interesse em sistemas complexos, cientistas começaram a questionar se as redes reais por trás de sistemas tão diversos como células ou a Internet são fundamentalmente aleatórias. Percebeu-se que esses sistemas complexos devem apresentar princípios organizacionais, que são codificados em sua topologia. Se a topologia dessas redes realmente difere de um grafo aleatório, é necessário desenvolver novas ferramentas e medidas para quantificar esses princípios organizacionais. Nos últimos anos, houve avanços notáveis nessa direção, impulsionados pelo aumento da capacidade computacional, que permite analisar redes com milhões de vértices; pela quebra das barreiras entre disciplinas, que oferece acesso a diversos bancos de dados; e pela crescente necessidade de superar abordagens reducionistas para entender o comportamento do sistema como um todo. Compreender a topologia das interações entre os componentes, isto é, as redes, tornou-se um passo essencial nesse processo (Albert e Barabási, 2002).

Na área de estudo das redes complexas, a aplicação de modelos de grafos aleatórios é fundamental para compreender a estrutura e o comportamento de redes que ocorrem no mundo real. Esses modelos constituem uma ferramenta vital na geração de redes artificiais com características e parâmetros bem definidos, facilitando o entendimento de aspectos cruciais tanto da estrutura dessas redes quanto dos processos que nelas ocorrem (Newman, 2018). A relevância dos grafos aleatórios na modelagem de redes complexas não se limita apenas à sua simplicidade conceitual; eles também oferecem um campo fértil para testar teorias e hipóteses sobre a dinâmica de redes. Por exemplo, em epidemiologia, os grafos aleatórios podem ser usados para modelar a propagação de doenças, permitindo aos pesquisadores estudar o impacto de diferentes estratégias de vacinação ou de isolamento social. Da mesma forma, em ecologia, esses modelos ajudam a compreender as complexas redes de interações entre espécies, fornecendo noções sobre a estabilidade e resiliência de ecossistemas. Outro aspecto importante dos grafos aleatórios é sua aplicabilidade no estudo de fenômenos de redes sociais. Eles podem ser usados

para analisar padrões de conexão e formação de comunidades, assim como para entender como informações ou influências se propagam através de redes sociais. Além disso, em sistemas de comunicação e redes de computadores, os grafos aleatórios auxiliam no entendimento de questões de robustez e eficiência de rede, fornecendo informações valiosas para o design e otimização desses sistemas. A análise matemática e computacional de grafos aleatórios também desempenha um papel crítico no desenvolvimento de novos algoritmos e técnicas para a análise de grandes conjuntos de dados. À medida que a quantidade de dados disponíveis cresce exponencialmente, métodos eficientes para analisar e interpretar esses dados em termos de redes são cada vez mais necessários. Grafos aleatórios oferecem um meio para testar e refinar esses métodos, garantindo que sejam robustos e confiáveis. Finalmente, o estudo de grafos aleatórios contribui para o avanço da teoria dos grafos como um todo. Ao explorar as propriedades e comportamentos emergentes desses grafos, pesquisadores podem desenvolver uma compreensão mais profunda dos princípios fundamentais que regem a teoria dos grafos e, por extensão, de redes complexas em geral. Este conhecimento, por sua vez, tem implicações de longo alcance, influenciando áreas tão diversas quanto neurociência, economia e engenharia. Assim, a incorporação de modelos de grafos aleatórios no estudo de redes complexas oferece uma abordagem poderosa para desvendar os mistérios das redes naturais e artificiais, ampliando significativamente nossa capacidade de entender e manipular esses sistemas complexos.

Grafos aleatórios de Poisson (modelo de Erdős-Rényi)

Dentro do espectro dos modelos de grafos aleatórios de Poisson, os mais prevalentes são conhecidos como $G(n, m)$ e $G(n, p)$. O modelo $G(n, m)$ é desenvolvido estabelecendo-se um número fixo de nós n e de arestas m , onde as conexões entre os pares de nós são estabelecidas de maneira aleatória e uniforme, conforme o (Algoritmo 3)

Algoritmo 3: Modelo modelo de Erdős-Rényi

Data: Número de vértices n , número de arestas m

Result: Grafo aleatório $G(n, m)$

Inicialize um conjunto vazio de arestas E ;

Calcule o número total de pares de vértices $N = \binom{n}{2}$;

for $i = 1$ **to** m **do**

Selecione um par de vértices (v, w) aleatoriamente sem reposição de todos os pares possíveis;

if $(v, w) \notin E$ **then**

Adicione a aresta (v, w) ao conjunto de arestas E ;

else

Repita a seleção de um novo par (v, w) até encontrar um que não esteja em E ;

end

end

Crie o grafo $G(n, m)$ usando o conjunto de vértices V e o conjunto de arestas E ;

Este processo pode ser visualizado como uma seleção aleatória dentro de um conjunto (*ensemble*) de grafos, cada um com n vértices e m arestas. Ao considerar um grafo com n vértices, o total de arestas possíveis (N) corresponde ao número de combinações de pares distintos de vértices, expresso como a combinação de n elementos tomados dois a dois, ou seja $N = \binom{n}{2}$. Portanto, ao restringir a análise para grafos com exatamente m arestas, a cardinalidade do conjunto de possíveis grafos é dada pela combinação do número total de arestas possíveis (N), escolhidas m a m , representada matematicamente por $\binom{N}{m}$. Assumindo que todos os grafos neste conjunto são equiprováveis, a equação (2.43) define uma distribuição de probabilidade sobre todos os grafos possíveis neste conjunto. Este modelo revela-se particularmente útil para investigar propriedades médias em redes, como o diâmetro médio, fornecendo valiosas compreensões acerca do comportamento típico em tais grafos, especialmente no limite termodinâmico, onde o número de vértices é muito grande e as propriedades médias podem ser calculadas com precisão (Newman, 2018).

$$\mathbb{P}(G) = \frac{1}{\binom{N}{m}} \quad (2.43)$$

No modelo $G(n, p)$, a estrutura do grafo é definida estabelecendo-se uma probabilidade fixa p para a formação de arestas entre cada par de vértices em um conjunto de n vértices.

Nesse modelo, o número total de arestas varia, podendo ir de zero até o máximo teórico de N , que representa todas as combinações possíveis de pares de vértices. Devido à equivalência entre os vértices, a probabilidade de formação de um grafo específico G com m arestas pode ser calculada considerando-se que, com a probabilidade p , m arestas são formadas com sucesso. Ao mesmo tempo, com base na propriedade de complementaridade (A.1), arestas não são formadas com probabilidade $1 - p$. Como a formação de cada aresta é um evento independente (ou seja, com covariância nula), as probabilidades associadas a cada evento de formação de aresta podem ser multiplicadas para determinar a probabilidade do grafo $G(n, m)$ sob este modelo, como expresso na equação simplificada (2.44). Este modelo, frequentemente associado aos trabalhos de Erdős e Rényi, é crucial para entender a estrutura e características fundamentais das redes aleatórias, incluindo a distribuição de graus e a configuração de seus componentes conectados (Barabási, 2013).

$$\mathbb{P}(G(n, m)) = p^m (1 - p)^{N-m} \quad (2.44)$$

Considerando que o número de arestas (m) é uma variável aleatória discreta, podemos calcular seu valor médio para entender melhor o comportamento típico desses grafos aleatórios. A equação (2.44) fornece a probabilidade de obtermos um grafo específico dentro do *ensemble* de grafos $\Omega(G(n, m))$, que possuem n vértices e m arestas. Para calcular a probabilidade de observar um grafo qualquer com n vértices e m arestas dentro do espaço amostral de todos os grafos $G(n, p)$, devemos considerar todos os grafos possíveis com n vértices e m arestas. Essa probabilidade é expressa pela equação (2.45), já que o número total de grafos com n vértices e m arestas é dado por $\binom{N}{m}$. Essa é uma distribuição de probabilidades binomial, e portanto, o valor esperado de m , denotado por $\langle m \rangle$, é Np .

$$\mathbb{P}(E = m) = \binom{N}{m} p^m (1 - p)^{N-m} \quad (2.45)$$

O resultado acima nos permite calcular o grau médio ($\langle k \rangle$) dos vértices de um grafo $G(n, p)$. Considerando que o grau médio de um grafo com exatamente n vértices e m arestas é dado por $2m/n$, podemos desenvolver a expressão para $\langle k \rangle$. Este cálculo é mostrado em (2.46).

$$\langle k \rangle = \left\langle \frac{2m}{n} \right\rangle = \frac{2\langle m \rangle}{n} = \frac{2}{n} Np = \frac{2}{n} \frac{n!}{(n-2)!2!} p = (n-1)p \quad (2.46)$$

Com raciocínio semelhante, podemos desenvolver a função mássica de probabilidade - do inglês *Probability Mass Function* (PMF) - dos graus dos vértices desse grafo aleatório. Considerando um vértice específico na rede, ele possui a probabilidade p de se conectar a um dos outros $n - 1$ vértices do grafo $G(n, p)$. Assim, a probabilidade desse vértice estar conectado a somente k outros vértices do total, é $p^k(1 - p)^{n-1-k}$. Podemos escolher esses k vizinhos de $\binom{n-1}{k}$ maneiras do total de $n - 1$ vértices. Portanto, a probabilidade total de um vértice estar conectado exatamente a k outros vértices é dada pela equação (2.47) e portanto temos novamente uma distribuição binomial.

$$p_k = \binom{n-1}{k} p^k (1-p)^{n-1-k} \quad (2.47)$$

Geralmente, nosso interesse está nas propriedades dos sistemas no limite termodinâmico, ou seja, quando consideramos valores grandes de n . Além disso, como discutido previamente, a maioria das redes do mundo real é esparsa. Isso significa que apenas uma pequena fração das $\binom{n}{2}$ arestas possíveis está de fato presente, e o grau médio $\langle k \rangle$ aumenta mais lentamente do que n à medida que n se torna grande. Neste cenário, podemos considerar que p é suficientemente pequeno, permitindo a aproximação da distribuição Binomial por uma distribuição de Poisson. Para isso, definimos $\lambda = (n - 1)p$ e desenvolvemos p_k conforme a equação (2.48):

$$\begin{aligned} p_k &= \binom{n-1}{k} p^k (1-p)^{n-1-k} = \frac{(n-1)!}{(n-1-k)!k!} \left[\frac{\lambda}{(n-1)} \right]^k \left[1 - \frac{\lambda}{(n-1)} \right]^{n-1-k} \\ &= \left[\frac{(n-1)(n-2)\cdots(n-1-k)}{k!} \right] \left[\frac{\lambda^k}{(n-1)^k} \right] \frac{[1 - \lambda/(n-1)]^{n-1}}{[1 - \lambda/(n-1)]^k} \\ &= \left[\frac{(n-1)(n-2)\cdots(n-1-k)}{(n-1)^k} \right] \frac{[1 - \lambda/(n-1)]^{n-1} \lambda^k}{[1 - \lambda/(n-1)]^k k!} \end{aligned} \quad (2.48)$$

Para calcular o valor de $\lim_{n \rightarrow \infty} p_k$, analisamos o limite de cada componente individualmente e, em seguida, combinamos os resultados. Portanto, com base nas igualdades (2.49), (2.50) e (2.51), concluímos que os graus dos vértices seguem uma distribuição de Poisson. Assim, este é o fundamento para a denominação de "grafo aleatório de Poisson" para esse modelo, utilizado ocasionalmente para diferenciá-lo de outros modelos de grafos aleatórios que serão apresentados e que, em geral, não possuem distribuições de graus de Poisson.

$$\lim_{n \rightarrow \infty} \frac{(n-1)(n-2)\cdots(n-1-k)}{(n-1)^k} = \lim_{n \rightarrow \infty} \frac{n^k (1-1/n)(1-2/n)\cdots(1-1/n-k/n)}{n^k} = 1 \quad (2.49)$$

$$\lim_{n \rightarrow \infty} \left[1 - \frac{\lambda}{(n-1)} \right]^{n-1} = e^{-\lambda} \quad (2.50)$$

$$\lim_{n \rightarrow \infty} \left[1 - \frac{\lambda}{(n-1)} \right]^k = 1 \quad (2.51)$$

Grafos aleatórios em lei de potência e a propriedade livre de escala

Uma distribuição em lei de potência, aplicável tanto a fenômenos discretos quanto contínuos, é caracterizada pela propriedade de que a probabilidade de um evento decresce como uma função de potência do tamanho desse evento. Esta relação é descrita pela equação (2.53). Essa relação é válida para valores de k maiores ou iguais a um mínimo k_{\min} . As distribuições em lei de potência são notáveis por suas caudas longas e pesadas, uma característica que se traduz em um decaimento mais lento do que o observado em distribuições exponenciais e de Poisson, conforme descrito em (Barabási, 2013). Isso indica uma probabilidade relativamente alta de ocorrências extremas, em comparação com essas distribuições mais comuns. Em termos de momentos estatísticos, uma distribuição com cauda pesada pode apresentar momentos de ordem superior, como variância e curtose, muito grandes ou até infinitos, o que é discutido em (Casella e Berger, 2021). Outra característica fundamental dessas distribuições é a invariância de escala, na qual a multiplicação do tamanho do evento por uma constante resulta apenas na multiplicação da probabilidade por outra constante, mantendo inalterada a forma funcional da distribuição. Esta propriedade de invariância de escala é semelhante à observada em *fractais*, como vimos anteriormente. Distribuições em lei de potência são encontradas em uma ampla variedade de campos, incluindo física, biologia, economia, geologia, demografia e ciências sociais, conforme relatado em (Barabási, 2013). Elas são particularmente úteis para modelar fenômenos que envolvem mecanismos multiplicativos ou de retroalimentação positiva. Tais distribuições são especialmente apropriadas para fenômenos nos quais eventos extremos são mais prováveis ou mais impactantes do que aqueles previstos por modelos baseados em distribuições normais ou exponenciais.

$$P(k) \propto k^{-\gamma} \quad (2.52)$$

Demonstrar que todos os momentos divergem em uma distribuição em lei de potência com $\gamma < 2$ é um processo direto, particularmente ao considerar o caso contínuo. Para ilustrar isso,

vamos calcular o n -ésimo momento e analisar o comportamento da função resultante. O caso discreto segue uma lógica análoga. Em uma distribuição de probabilidade em lei de potência, utilizando a função densidade de probabilidade dada por (2.52), expressamos a função como $f(k) = Ck^{-\gamma}$, onde C é uma constante de normalização. Concentrando-nos no momento de ordem n dessa distribuição e substituindo a função de densidade na expressão do momento, chegamos à equação (2.53). Observa-se que para a convergência de $E[K^n]$, é necessário que $n - \gamma + 1 < 0$. Isso implica que $n + 1 < \gamma$, ou seja, para a existência do primeiro momento (e, por extensão, dos momentos subsequentes), deve-se ter $2 < \gamma$. Perceba que quando $2 < \gamma < 3$, a distribuição em lei de potência possui o primeiro momento, mas não os demais.

$$E[K^n] = \int_{k_{\min}}^{+\infty} k^n Ck^{-\gamma} dk = C \int_{k_{\min}}^{+\infty} k^{n-\gamma} dk = \lim_{k \rightarrow \infty} C \frac{k^{n-\gamma+1}}{n-\gamma+1} - C \frac{k_{\min}^{n-\gamma+1}}{n-\gamma+1} \quad (2.53)$$

Redes cuja distribuição de graus adere à lei de potência são categorizadas como redes **livres de escala**. Esta característica peculiar destaca-se nos graus de entrada e saída em redes direcionadas. A análise matemática dessas redes pode ser efetuada empregando tanto o formalismo discreto quanto o contínuo, sendo a natureza livre de escala independente da abordagem metodológica utilizada, conforme explicado por (Barabási, 2013). Em contraste, redes aleatórias com uma distribuição de graus de Poisson exibem uma variância $\sigma_k = \langle k \rangle^{1/2}$, consistentemente menor que a média $\langle k \rangle$. Isso implica que a maioria dos graus dos nós se concentra em torno de $\langle k \rangle$, com uma probabilidade mais alta dentro do intervalo $k = \langle k \rangle \pm \langle k \rangle^{1/2}$. Portanto, os nós em uma rede aleatória tendem a ter graus comparáveis, fazendo com que o grau médio $\langle k \rangle$ funcione como uma "escala" para a rede. Por outro lado, nas redes que seguem uma distribuição de graus em lei de potência com $2 < \gamma < 3$, observa-se que, embora o primeiro momento (média) seja finito, o segundo momento (variância) tende ao infinito. Esse fenômeno, destacado anteriormente, sugere que as flutuações em torno da média podem ser extremamente grandes à medida que o tamanho da rede N cresce. Consequentemente, a seleção aleatória de um nó pode resultar em um grau extremamente baixo ou surpreendentemente alto. Dessa forma, redes com $2 < \gamma < 3$ não apresentam uma escala interna definida, classificando-se, portanto, como livres de escala (Barabási, 2013). A variação do valor de γ impacta significativamente nas propriedades da rede, conforme descrito a seguir:

- **Regime Anômalo** ($\gamma = 2$): Neste regime, o grau do maior *hub* cresce linearmente com o tamanho do sistema (n vértices), isto é, $k_{\max} \sim n$. Isso força a rede a uma configuração na qual todos os nós estão próximos uns dos outros, pois todos se conectam ao mesmo *hub*

central. Neste regime, o comprimento médio do caminho não depende de n , além disso, a rede é robusta a falhas aleatórias, mas vulnerável a ataques direcionados aos *hubs* (Cohen e Havlin, 2010).

- **Mundo Ultra-Pequeno** ($2 < \gamma < 3$): Neste regime, a distância média aumenta como $\ln \ln n$, um crescimento significativamente mais lento do que o $\ln n$ derivado para redes aleatórias. Redes neste regime são chamadas de ultra-pequenas, pois os *hubs* reduzem radicalmente o comprimento do caminho ao se ligarem a um grande número de nós de pequeno grau, criando distâncias curtas entre eles. A rede mantém robustez substancial a falhas aleatórias e vulnerabilidade a ataques direcionados (Cohen e Havlin, 2010).
- **Ponto Crítico** ($\gamma = 3$): Este valor é de interesse teórico particular, pois o segundo momento da distribuição de grau não diverge mais. Portanto, $\gamma = 3$ é chamado de ponto crítico. Neste ponto, a dependência $\ln N$ encontrada em redes aleatórias retorna, mas com uma correção logarítmica dupla $\ln \ln N$, que encurta as distâncias em comparação com uma rede aleatória de tamanho similar. As características de robustez a falhas aleatórias e vulnerabilidade a ataques direcionados são mantidas, mas em menor grau (Cohen e Havlin, 2010).
- **Mundo Pequeno** ($\gamma > 3$): Neste regime, $\langle k^2 \rangle$ é finito e a distância média segue o resultado do mundo pequeno derivado para redes aleatórias. Embora os *hubs* continuem presentes, para $\gamma > 3$, eles não são suficientemente grandes ou numerosos para impactar significativamente na distância entre os nós. A robustez a falhas aleatórias e a vulnerabilidade a ataques direcionados são mais balanceadas, comparadas aos valores menores de γ (Cohen e Havlin, 2010). Neste regime, as propriedades são difíceis de distinguir das propriedades de uma rede aleatória de Poisson com tamanho similar (Barabási, 2013).

É viável construir uma rede de maneira determinística que manifeste as propriedades típicas de uma rede livre de escala, exemplificada pela família de redes denominadas *Mandala*. Neste modelo específico, a construção da rede é realizada pela adição sucessiva de vértices a uma nova camada externa, assemelhando-se à formação de novas cascas sobre as anteriores. Os vértices recém-adicionados formam conexões tanto dentro de sua própria camada (intra-camada) quanto com camadas mais internas (inter-camada), seguindo uma lógica hierárquica de descendência. Essa abordagem promove um aumento contínuo tanto na complexidade quanto na conectividade da rede a cada nova etapa. Notavelmente, os vértices de cada camada, especialmente aqueles mais conectados, conhecidos como *hubs*, estão estrategicamente localizados nas

camadas mais internas, consolidando o caráter hierárquico e integrado da rede (Sampaio *et al.*, 2015).

Grafos aleatórios com associação preferencial (preferential attachment)

Os modelos de associação preferencial, particularmente o modelo de Price e o modelo de Barabási-Albert, representam um avanço fundamental no estudo de redes complexas, especialmente aquelas que exibem distribuições de graus em lei de potência. O modelo de Price, que se concentra na análise de grafos aleatórios direcionados, foi um dos primeiros a introduzir o conceito de associação preferencial. O modelo de Barabási-Albert (BA), o qual desprenderemos mais atenção, é um caso particular do modelo de Price, com sua análise focada em grafos não direcionados. Esses modelos são essenciais para entender como as redes crescem e evoluem, enfatizando o papel crítico da associação preferencial na formação de redes sem escala (Newman, 2018).

O modelo BA parte de um pequeno número inicial de nós, m_0 , interconectados de maneira que cada vértice tenha ao menos uma ligação. A cada passo temporal, um novo nó é adicionado à rede, ligando-se ao i -ésimo vértice preexistente com probabilidade $\Pi(k_i)$ definida na equação (2.54), produzindo $m(\leq m_0)$ novas ligações. Após t passos temporais, o modelo BA resulta em uma rede com $N = t + m_0$ nós e $m_0 + mt$ ligações, conforme o (Algoritmo 4).

Algoritmo 4: Modelo de Associação Preferencial de Barabási-Albert

Data: Número inicial de vértices m_0 , número final de vértices N , número de arestas a adicionar por novo vértice $m \leq m_0$

Result: Grafo G seguindo o modelo de Barabási-Albert

Inicialize o grafo G com m_0 vértices conectados de forma arbitrária, garantindo que não haja vértices isolados;

for $n = m_0 + 1$ **to** N **do**

 Adicione um novo vértice v_n ao grafo G ;

 Calcule $\sum_j k_j$, a soma dos graus de todos os vértices existentes em G ;

for $i = 1$ **to** m_0 **do**

 Escolha um vértice v_i existente em G com probabilidade $\Pi(k_i) = \frac{k_i}{\sum_j k_j}$;

 Conecte o novo vértice v_n ao vértice v_i com uma aresta;

end

end

Simulações mostram que distribuição de graus resultantes desse processo é aproximadamente uma lei de potência com um expoente de grau $\gamma = 3$ (Barabási, 2013). Assim, a probabilidade de um vértice adquirir novas ligações é proporcional ao seu grau, resultando em uma dinâmica em que os nós com grau mais elevado tendem a gerar novas ligações mais rapidamente do que os menos conectados, ao que damos o nome de **associação preferencial** (preferential attachment) e que levam à formação de *hubs*. Para entender como a propriedade livre de escala emerge, precisamos focar na evolução temporal do modelo de Barabási-Albert.

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} \quad (2.54)$$

Neste modelo, um vértice existente tem a oportunidade de aumentar seu grau a cada vez que um novo vértice é adicionado ao grafo. Este novo vértice estabelece conexões com m dos $n(t)$ vértices já existentes, onde $n(t)$ representa a dependência temporal do tamanho do grafo. Para simplificar a análise, o grau k_i de um vértice é tratado como uma variável real contínua (Cohen e Havlin, 2010). Supondo que cada novo vértice adiciona em média $m\Pi(k_i)$ arestas ao i -ésimo vértice existente, a taxa de aquisição de novas ligações por esse vértice é descrita pela equação (2.55). A soma no denominador dessa equação abrange todos os vértices preexistentes. Considerando que a cada unidade de tempo t , são adicionadas m novas ligações ao grafo, e que cada ligação aumenta o grau de dois vértices, temos um total de $2mt$ graus adicionados. No entanto, ao considerar somente os vértices preexistentes, é necessário subtrair os m graus atribuídos ao novo vértice, resultando na equação (2.56). Substituindo esta equação em (2.57) e aproximando para grandes valores de t (i.e., $t \gg 1$), obtemos uma equação diferencial ordinária. A condição inicial é estabelecida pelo fato de que o vértice indexado por i foi adicionado na i -ésima iteração, portanto, $k_i(i) = m$. Utilizando esta condição na equação (2.57), determinamos que $C = mi^{-1/2}$. Consequentemente, a dinâmica de incremento de graus é descrita pela equação (2.58), onde definimos $\beta = 1/2$ como o expoente dinâmico do processo.

$$\frac{dk_i}{dt} = m\Pi(k_i) = \frac{mk_i}{\sum_{j=1}^{N-1} k_j} \quad (2.55)$$

$$\sum_{j=1}^{N-1} k_j = 2mt - m \quad (2.56)$$

$$\frac{dk_i}{dt} = \frac{k_i}{2t-1} \approx \frac{1}{2} \frac{k_i}{t} \Rightarrow \frac{dk_i}{k_i} = \frac{1}{2} \frac{dt}{t} \Rightarrow k_i(t) = Ct^{1/2} \quad (2.57)$$

$$k_i(t) = m \left(\frac{t}{i} \right)^{1/2} = m \left(\frac{t}{i} \right)^\beta \quad (2.58)$$

A equação (2.58) revela que o grau de cada nó na rede aumenta segundo uma lei de potência com um expoente dinâmico constante, $\beta = 1/2$. Esta constância implica que todos os nós seguem a mesma regra de crescimento, caracterizando um padrão uniforme em toda a rede. Interessantemente, o aumento dos graus é sublinear, isto é, $\beta < 1$, uma característica resultante da natureza evolutiva do modelo: cada novo nó tem um leque mais amplo de nós para se conectar do que os adicionados anteriormente. Além disso, o momento em que um nó é adicionado à rede influencia significativamente o seu grau. Nós adicionados mais cedo, representados por um menor valor de i , tendem a ter graus maiores, demonstrando a vantagem do pioneiro, onde os *hubs* ou nós centrais são maiores devido ao seu tempo prolongado na rede. A taxa na qual um nó adquire novas arestas, conforme descrita pela equação (2.59), reforça essa observação. Os nós mais antigos adquirem mais ligações a cada etapa temporal, ao mesmo tempo que frequência de aquisição de novas arestas diminui com o tempo, seguindo uma taxa proporcional a $t^{-1/2}$.

$$\frac{dk_i(t)}{dt} = \frac{m}{2\sqrt{it}} \quad (2.59)$$

Para deduzir a distribuição de graus do grafo no modelo de Barabási-Albert, começamos invertendo a equação (2.58) em relação a i , resultando na equação (2.60). Considerando que a cada unidade de tempo transcorrido (e assumindo $t \gg 1$), a probabilidade p_i de um novo vértice escolher o i -ésimo vértice previamente adicionado é aproximadamente $1/t$, podemos realizar uma mudança de variável conforme definido pela equação (2.61). Aplicando a transformação (2.61) e considerando a relação entre i e k obtida na equação (2.59), nos leva ao resultado expresso na equação (2.62), indicando que o modelo de Barabási-Albert produz uma rede com uma distribuição de graus que segue uma lei de potência com $\gamma = 3$.

$$i = m^2 \frac{t}{k^2} \quad (2.60)$$

$$p_k = p_i \left| \frac{di}{dk} \right| \quad (2.61)$$

$$p_k = \frac{1}{t} \left| \frac{d(-m^2tk^{-2})}{dk} \right| = 2m^2k^{-3} = 2m^2k^{-\gamma} \quad (2.62)$$

Embora (Barabási, 2013) sugira que o crescimento da rede e a associação preferencial são requisitos indispensáveis para que uma rede exiba a propriedade Livre de Escala, é importante notar que o crescimento isoladamente não é um fator crucial. (Tsallis, 2008) mostrou que, mesmo em redes com um número fixo de nós e ligações, características de distribuição de grau semelhantes às observadas em redes Livres de Escala podem emergir. Esse fenômeno ocorre em redes cuja distribuição de grau se assemelha a uma q-exponencial no limite de grandes números, indicando que a invariância de escala pode ser alcançada mesmo sem o crescimento contínuo da rede.

Esta sequência de equações matemáticas e transformações fornece um caminho claro para entender como o modelo de Barabási-Albert gera uma rede cuja distribuição de graus segue uma lei de potência, enfatizando a relação intrínseca entre o tempo de adição de um vértice e a probabilidade de seleção desse vértice em processos subsequentes (Barabási, 2013).

Modelo configuracional

O **modelo configuracional**, frequentemente referido como a "construção de Bollobás" e em alguns contextos como a "construção de Molloy-Reed", é uma abordagem notável na criação de grafos aleatórios. Este modelo se distingue de outras abordagens por desconsiderar a evolução temporal da rede, focando exclusivamente na distribuição de graus dos nós e desconsiderando quaisquer correlações entre eles (Cohen e Havlin, 2010). A característica fundamental deste modelo é a independência na formação das conexões: a probabilidade de um nó estar conectado a outro é determinada unicamente pelos graus de ambos os nós, independente de qualquer propriedade intrínseca ou histórico de conexões prévias dos nós envolvidos. Assim, ao seguir uma aresta de um nó, o nó de destino é escolhido de maneira aleatória, sem relação com o nó de origem (Newman, 2018). A implementação do modelo configuracional ocorre em duas etapas fundamentais. Na primeira etapa, denominada Atribuição de Graus, cada nó no grafo é designado um grau específico, baseado em uma distribuição de graus pré-estabelecida,

que define o número de conexões (arestas) que o nó formará. A segunda etapa, conhecida como Formação de Arestas, começa após a atribuição dos graus, onde cada nó possui terminações, ou "stubs", correspondentes ao seu grau. Estes stubs são emparelhados de maneira aleatória para formar arestas completas entre os nós, assegurando que cada terminação de um nó se conecte com uma terminação de outro nó, completando assim a estrutura de arestas do grafo (Barabási, 2013). Como resultado, obtemos um grafo aleatório com uma sequência de graus pré-definida. Além do grau de cada nó k_i , a estrutura do grafo é aleatória. Repetindo este método para a mesma sequência de graus, diferentes redes podem ser geradas, todas com a mesma distribuição de graus p_k . Neste modelo, a probabilidade de formação de uma ligação entre dois nós de graus k_i e k_j é dada por $p_{ij} = \frac{k_i k_j}{2E(t)-1}$, onde $E(t)$ representa o número total de arestas após a t -ésima iteração do processo. Esta probabilidade surge do fato de que cada terminação de um nó tem a possibilidade de se conectar a várias outras terminações, e essa probabilidade é proporcional ao número de terminações do nó de destino. Um aspecto importante do modelo configuracional tradicional é que a rede gerada pode incluir laços e arestas repetidas, pois o algoritmo não impede explicitamente que um nó se conecte a si mesmo ou crie múltiplas ligações com outro nó. Contudo, é possível rejeitar essas ocorrências sem comprometer a integridade da rede e a probabilidade de outras combinações, especialmente sob condições específicas, como a exigência de um grau mínimo maior ou igual a dois (Sampaio *et al.*, 2023).

No Modelo Configuracional, a construção de uma rede aleatória inicia-se pela seleção aleatória de pares de vértices, denotados por v e w . Estes vértices são escolhidos com base na condição de que ainda não atingiram o número de ligações estipulado pelos seus graus predeterminados, k_v e k_w , respectivamente. Uma vez selecionados, estabelece-se uma aresta entre v e w , adicionando-a ao conjunto de arestas da rede. Este processo é repetidamente executado até que todos os vértices da rede tenham alcançado o número de conexões correspondente aos seus graus especificados. O procedimento detalhado para a geração da rede no Modelo Configuracional está resumido no (Algoritmo 5). Este algoritmo fornece um guia passo a passo para a construção eficiente da rede, garantindo que a estrutura final reflita com precisão a distribuição de grau desejada.

Algoritmo 5: Modelo Configuracional

Entrada: Grafo $G(V, E = \emptyset)$ e mapeamento $k_i \rightarrow i | i \in V$

```

while  $k_i > 0 \forall i \in [1, N] \cap \mathbb{N}$  do
  | 1. Selecionamos aleatoriamente dois vértices  $v$  e  $w \in V$ ;
  | 2. if  $k_v > 0$  e  $k_w > 0$  then
  |   | 2.1 reduzimos uma unidade de  $k_v$  e  $k_w$ ;
  |   | 2.2 adicionamos a aresta  $v - w$  ao conjunto  $E$ ;
  | else
  |   | Realizamos novamente o passo 5.;
  | end
end

```

2.4 Mecânica estatística de redes complexas

Estatística não extensiva

A teoria de redes se revelou uma ferramenta essencial no estudo de sistemas complexos, abrangendo desde processos biológicos até questões socioeconômicas. Sua eficácia deriva da capacidade de modelar e analisar estatisticamente estruturas em grande escala, permitindo examinar a interconexão e interação entre diferentes elementos. Esta análise revela padrões universais em redes, categorizando-as como aleatórias, sem escala ou de mundo pequeno. Compreender estes padrões é fundamental para decifrar a estrutura, robustez e função das redes, que frequentemente emergem de regras microscópicas variando de determinísticas a estatísticas. Estendendo essa compreensão para a mecânica estatística, encontramos uma analogia poderosa no estudo das redes. As propriedades microscópicas dos elementos das redes, como as regras de formação de ligações e restrições de recursos, influenciam significativamente suas propriedades macroscópicas. Assim, a abordagem de mecânica estatística não se limita apenas à estrutura atual das redes, mas também aborda como elas podem evoluir ao longo do tempo, explorando uma gama de possíveis configurações (Dehmer e Emmert-Streib, 2009).

O caos em sistemas dinâmicos é um fenômeno que ocorre quando pequenas variações nas condições iniciais levam a grandes diferenças no comportamento do sistema ao longo do tempo, caracterizado por ser imprevisível e aparentemente aleatório, mesmo que o sistema seja determinístico. Este comportamento é frequentemente associado a atratores estranhos,

que são tipos de atratores nos sistemas dinâmicos. Atratores são estados ou conjuntos de estados (pontos no espaço de fase) para os quais o sistema tende a evoluir a partir de uma ampla gama de condições iniciais. Enquanto atratores regulares, como pontos fixos e ciclos limites, são geométrica e matematicamente simples, os atratores estranhos são complexos, com estruturas que muitas vezes exibem propriedades fractais, significando que eles têm detalhes em todas as escalas e são auto-similares (Vulpiani *et al.*, 2009). A análise de sistemas dinâmicos complexos é enriquecida pelo estudo dos expoentes de Lyapunov. Estes quantificam a taxa de separação de trajetórias infinitesimalmente próximas, fornecendo uma medida de previsibilidade e caos. Um expoente positivo de Lyapunov indica um comportamento caótico, refletindo uma sensibilidade acentuada às condições iniciais. Essa sensibilidade é evidente na divergência exponencial das trajetórias no espaço de fases, uma característica central dos sistemas dinâmicos. Avançando nessa análise, até mesmo sistemas determinísticos com poucas variáveis podem exibir complexidades comparáveis a sistemas de muitas partículas, especialmente se sua dinâmica for caótica. Esses sistemas, embora pequenos em termos de graus de liberdade, apresentam uma interação complexa entre propriedades dinâmicas microscópicas e comportamento estatístico macroscópico. Isso se torna particularmente notável em cenários fora do equilíbrio, como aqueles expostos a gradientes ou campos externos (Vulpiani *et al.*, 2009).

Na dinâmica de sistemas complexos, a sensibilidade às condições iniciais é um aspecto crucial para entender o comportamento caótico. Essa sensibilidade é quantificada pela função $\xi(t) \equiv \lim_{\Delta x(0) \rightarrow 0} [\Delta x(t) / \Delta x(0)]$, que mede como pequenas variações nas condições iniciais podem resultar em grandes diferenças nas trajetórias do sistema ao longo do tempo. A expressão $\xi(t) = e^{\lambda_1 t}$ captura essa sensibilidade, onde λ_1 representa o expoente de Lyapunov. O expoente de Lyapunov, λ_1 , é uma medida vital que indica se um sistema é caótico ($\lambda_1 > 0$) ou não ($\lambda_1 < 0$). Um λ_1 positivo significa que pequenas diferenças nas condições iniciais aumentam exponencialmente com o tempo, caracterizando um comportamento caótico intenso. Por outro lado, valores negativos de λ_1 indicam que as trajetórias tendem a convergir, refletindo um sistema mais estável e previsível. Em cenários de caos fraco, onde $\lambda_1 = 0$, a função $\xi(t)$ não diverge exponencialmente, mas sim de maneira sub-exponencial, geralmente como uma lei de potência. Nessas situações, um novo expoente, $\lambda_q > 0$, é introduzido na expressão $\xi = e_q^{\lambda_q t}$, onde $e_q^z = e_q(z)$, definido pela equação (2.63). Aqui, q é um parâmetro que reflete a natureza do caos fraco no sistema. Para o caos forte, fazemos $q \rightarrow 1$, tal que a forma exponencial clássica de ξ é preservada, conforme o desenvolvimento (2.64). No entanto, para o caos fraco, q

assumindo valores maiores que 1 e λ_q positivo, $\xi(t)$ cresce mais lentamente do que uma função exponencial pura. Isso implica que $\xi(t)$ é proporcional a $t^{1/(1-q)}$ para tempos longos, refletindo uma sensibilidade às condições iniciais que evolui de forma mais controlada do que em sistemas com caos forte (Tsallis, 2023).

$$e_q(x) = \begin{cases} [1 + (1-q)x]^{1/(1-q)} & \text{se } [1 + (1-q)x] > 0 \\ 0 & \text{outros casos} \end{cases} . \quad (2.63)$$

$$\begin{aligned} \lim_{q \rightarrow 1} e_q(z) &= \lim_{q \rightarrow 1} [1 - (q-1)z]^{1/(q-1)} \\ &= \lim_{t \rightarrow 0} \left\{ \exp \left[\frac{1}{t} \ln(1-zt) \right] \right\} \quad (t = q-1) \\ &= \exp \left[\lim_{t \rightarrow 0} \frac{\ln(1-zt)}{t} \right] \\ &= \exp \left[\lim_{t \rightarrow 0} \frac{\frac{-z}{1-zt}}{1} \right] \quad \text{aplicando L'Hopital} \\ &= e^{-z} \end{aligned} \quad (2.64)$$

A mecânica estatística padrão, ou seja, a teoria de Boltzmann-Gibbs (BG), fundamenta-se na entropia como funcional entrópico para variáveis discretas, representada pela equação (2.65), onde k é uma constante positiva e W representa o número de configurações microscópicas possíveis (Tsallis, 1988). Esta entropia, é aditiva em sistemas independentes A e B , no sentido da relação (2.66). Este formalismo se aplica em amplas classes de sistemas naturais considerados simples, porém, falha ao ser aplicado a sistemas complexos. A compreensão detalhada dos sistemas caóticos leva a uma reavaliação da mecânica estatística de Boltzmann-Gibbs, especialmente em contextos de caos fraco, onde o expoente de Lyapunov máximo é nulo ou próximo de zero. Nestes cenários, a abordagem tradicional baseada na função entropia aditiva revela-se inadequada, levando à proposta da mecânica estatística não extensiva. Esta generalização, utilizando uma função entropia não aditiva, se adapta melhor à complexidade dos sistemas com caos fraco, criando uma ponte mais ampla entre as ciências naturais e sociais (Tsallis, 2023). Redes, mesmo em suas formas mais simples, resultam em entropias que são não aditivas ou não extensivas. Isso é intuitivamente compreensível: se imaginarmos a separação de uma rede dada em duas sub-redes A e B , todas as conexões entre A e B serão perdidas. Isso pode representar uma parte substancial da rede e claramente induz à não aditividade (Dehmer e Emmert-Streib, 2009).

$$S_{BG} = -k \sum_{i=1}^W p_i \ln p_i, \quad (2.65)$$

$$S_{BG}(A+B) = S_{BG}(A) + S_{BG}(B). \quad (2.66)$$

A generalização à entropia de Boltzmann-Gibbs proposta em (Tsallis, 1988) é definida pela equação (2.67). Neste modelo, a função S_q é introduzida como uma medida alternativa de entropia, caracterizada por um parâmetro livre q e um número total de estados W . Esta nova formulação converge para a expressão clássica de Boltzmann para a entropia quando q se aproxima de 1, detalhado no desenvolvimento (2.68). Ao maximizar a função de entropia generalizada S_q , sujeita às restrições $\sum_{i=1}^W p_i = 1$ e $\sum_{i=1}^W \varepsilon_i p_i = U_q$, onde $\{\varepsilon_i\}$ e U_q são números reais conhecidos (o mesmo valor ε_i pode estar associado a mais de uma configuração possível), e são denominados espectro generalizado e energia interna generalizada, obtemos a função de partição generalizada do Ensemble Canônico (2.69). Essa abordagem como um todo apresenta uma maneira inovadora de descrever sistemas complexos e caóticos.

$$S_q = k \frac{1 - \sum_{i=1}^W p_i^q}{q - 1}, \quad (2.67)$$

$$\begin{aligned} \lim_{q \rightarrow 1} S_q &= \lim_{q \rightarrow 1} k \frac{1 - \sum_{i=1}^W p_i^q}{q - 1} \\ &= k \lim_{q \rightarrow 1} \frac{\sum_{j=1}^W p_j - \sum_{i=1}^W p_i^q}{q - 1} \\ &= k \lim_{q \rightarrow 1} \frac{\sum_{i=1}^W p_i (1 - p_i^{q-1})}{q - 1} \\ &= -k \sum_{i=1}^W p_i \lim_{t \rightarrow 0} \frac{(p_i^t - 1)}{t} \quad (t = q - 1) \\ &= -k \sum_{i=1}^W p_i \ln p_i \\ &= S_{BG} \end{aligned} \quad (2.68)$$

$$Z_q = \sum_{i=1}^W [1 - \beta(q-1)\varepsilon_i]^{1/(q-1)} \quad (2.69)$$

O desenvolvimento (2.70) mostra que a função de partição (2.69) é uma generalização da função de partição usual do Ensemble Canônico, oriundo da entropia de Boltzmann (Z_1). A consideração da não aditividade das entropias em sistemas independentes resulta na expressão (2.71), que indica que S_q é não aditiva, exceto quando $\frac{1-q}{k}$ se anula (Tsallis, 2023).

$$\lim_{q \rightarrow 1} Z_q = \lim_{q \rightarrow 1} \sum_{i=1}^W e_q(\beta \varepsilon_i) = \sum_{i=1}^W \lim_{q \rightarrow 1} [e_q(\beta \varepsilon_i)] = \sum_{i=1}^W \exp(-\beta \varepsilon_i) Z_{(1)} \quad (2.70)$$

$$S_q(A+B) = S_q(A) + S_q(B) + \frac{1-q}{k} S_q(A) S_q(B) \quad (2.71)$$

As relações matemáticas da mecânica estatística não extensiva baseada na entropia não aditiva S_q , conduzem naturalmente a uma distribuição de graus na forma q-exponencial. Esta distribuição surge de forma ubíqua em redes que apresentam leis de potência assintóticas, tanto em contextos naturais, artificiais quanto sociais. Por essa razão, essas redes são frequentemente denominadas redes sem escala. Entretanto, a descrição assintoticamente sem escala seria mais precisa. A distribuição de graus mais comum nestas redes é descrita pela equação (2.72), onde $k_0 > 0$ e $\gamma > 0$, e k representa o número de ligações ou o grau de um vértice. Uma mudança de notação leva ao uso da função q-exponencial (e_q^x). Esta abordagem mostra que as redes sem escala são caracterizadas por uma distribuição de graus que decresce não de maneira exponencial, mas conforme uma lei de potência modificada pela função q-exponencial, como indicado na equação (2.73) (Tsallis, 2008).

$$p(k) \propto \frac{1}{(k_0 + k)^\gamma} \quad \left(\gamma = \frac{1}{q-1}, \quad k_0 = \frac{\kappa}{q-1} \right) \quad (2.72)$$

$$p(k) \propto e_q^{-k/\kappa} = e_q^{-\lambda k} \quad \left(\lambda = \frac{1}{\kappa} \right) \quad (2.73)$$

Distribuição de probabilidade q-exponencial

A **distribuição de probabilidade q-exponencial**, uma generalização da distribuição de probabilidade exponencial tradicional, incorpora um parâmetro adicional q juntamente com a taxa λ , expandindo seu escopo e aplicabilidade. Diferentemente da distribuição exponencial padrão, a q-exponencial modifica esta taxa de decaimento introduzindo o parâmetro q . Quando

$q \rightarrow 1$, a distribuição reverte para a forma exponencial clássica, mas desvia-se dela à medida que q se afasta de 1, tornando as caudas da distribuição mais pesadas ou mais leves. Esta flexibilidade faz da distribuição q -exponencial uma ferramenta mais adaptável para uma variedade de cenários, especialmente onde a suposição de decaimento exponencial puro não é adequada. A interação entre q e λ na q -exponencial oferece uma modelagem mais abrangente. Por exemplo, com q maior que 1, quando λ tende a zero, a q -exponencial aproxima-se de uma distribuição uniforme no intervalo $[0, \infty)$, enquanto que para $\lambda \rightarrow \infty$ faz a distribuição assemelhar-se a uma distribuição em lei de potência, na qual podemos fazer a associação $\gamma = 1/(q - 1)$, que é conhecida por suas caudas longas e pesadas.

Assim, podemos generalizar a distribuição de probabilidade exponencial, cuja função de densidade de probabilidade é definida por (2.74), quando fazemos q se aproximar de 1 e a distribuição de probabilidade em lei de potência, cuja função de densidade de probabilidade é definida por (2.75), quando $\lambda \rightarrow \infty$, através da família de distribuições de probabilidade q -exponencial, cujo funcional de Densidade de Probabilidade é definido por (2.76), onde o prefator $(2 - q)$ aparece para manter a normalização da distribuição. A figura 6 torna explícito a diferença entre o comportamento dessas distribuições de probabilidade.

$$f(x, \lambda) = \begin{cases} \lambda e^{-\lambda x} & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases}. \quad (2.74)$$

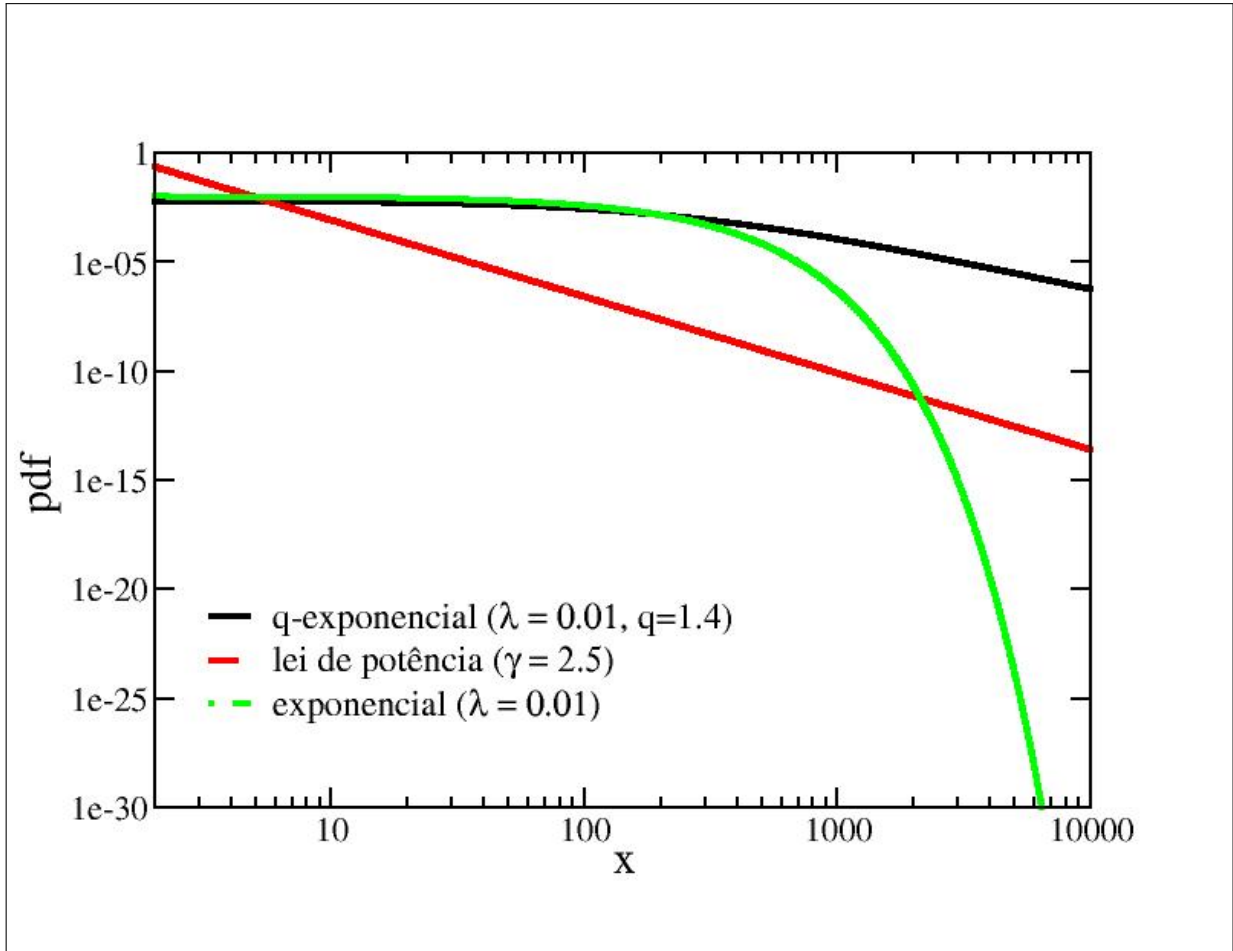
$$f(x, \gamma) = \begin{cases} Cx^{-\gamma} & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases}. \quad (2.75)$$

$$f(x, q, \lambda) = \begin{cases} (2 - q)\lambda e_q(-\lambda x) & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases}, \quad (2.76)$$

Considerando a $T(x)$ definido pela equação (2.77), λ e q fixos, podemos definir a Função de Distribuição Acumulada - *Cumulative Distribution Function* (CDF) - da distribuição de probabilidade q -exponencial através da equação (2.78). Para os casos em que desejarmos truncar a distribuição num intervalo arbitrário $[x_{min}, x_{max}]$, devemos calcular a constante de normalização apropriada, conforme o desenvolvimento (2.79).

$$T(x) = [1 - (1 - q)\lambda x]. \quad (2.77)$$

Figura 6 – Esta figura compara as funções de densidade de probabilidade de três distribuições: a distribuição q -exponencial com parâmetros $\lambda = 0.01$ e $q = 1.4$, a lei de potência com parâmetro $\gamma = 2.5$, e a exponencial com parâmetro $\lambda = 0.01$. É interessante notar que a distribuição q -exponencial é versátil: ao fazer q tender a 1, ela assume um perfil semelhante à distribuição exponencial com $\lambda = 0.01$. Alternativamente, quando λ se aproxima do infinito, a q -exponencial se assemelha mais à lei de potência com parâmetro $\gamma = 2.5$.



Fonte: (Sampaio *et al.*, 2023)

$$\int_0^x (2-q)e_q(-\lambda y) dy = -T(y)^{\frac{2-q}{1-q}} \Big|_0^x = 1 - T(x)^{\frac{2-q}{1-q}}. \quad (2.78)$$

$$1 = \int_{x_{min}}^{x_{max}} Cf(y, q, \lambda) dy = C \left(-T(y)^{\frac{2-q}{1-q}} \right) \Big|_{x_{min}}^{x_{max}} = C(T(x_{min})^{\frac{2-q}{1-q}} - T(x_{max})^{\frac{2-q}{1-q}}) \quad (2.79)$$

$$\implies C = \frac{1}{T(x_{min})^{\frac{2-q}{1-q}} - T(x_{max})^{\frac{2-q}{1-q}}}$$

No contexto da teoria de probabilidades e estatísticas, a análise da distribuição q -exponencial requer a avaliação rigorosa de seus momentos estatísticos. Os momentos, particularmente o primeiro (média) e o segundo (variância), são cruciais na caracterização da

natureza da distribuição, fornecendo informações sobre a centralidade e a dispersão dos dados. O primeiro momento, ou expectância matemática, da distribuição q-exponencial é obtido através da integração ponderada da variável aleatória sobre o intervalo de interesse. Esta integração, matematicamente realizada como demonstrado na equação (2.80), é fundamental para representar a localização central da massa de probabilidade da distribuição.

$$\begin{aligned}
\langle x \rangle &= \int_{x_{\min}}^{x_{\max}} Cyf(y, q, \lambda) dy \\
&= C \left[y \left(-T(y)^{\frac{2-q}{1-q}} + C_1 \right) \Big|_{x_{\min}}^{x_{\max}} + \int_{x_{\min}}^{x_{\max}} T(y)^{\frac{2-q}{1-q}} - C_1 dx \right] \\
&= C \left[-y T(y)^{\frac{2-q}{1-q}} \Big|_{x_{\min}}^{x_{\max}} - \frac{T(y)^{\frac{3-2q}{1-q}}}{(3-2q)\lambda} \Big|_{x_{\min}}^{x_{\max}} \right] \\
&= CT(y)^{\frac{2-q}{1-q}} \left(-y - \frac{T(y)}{(3-2q)\lambda} \right) \Big|_{x_{\min}}^{x_{\max}}
\end{aligned} \tag{2.80}$$

Por outro lado, o segundo momento, conhecido como variância, é uma medida estatística que quantifica a dispersão dos valores da variável aleatória em torno da média. Esta medida é crucial para avaliar o grau de incerteza ou variabilidade inerente à distribuição. A variância da distribuição q-exponencial é derivada conforme estabelecido na equação (2.81).

$$\begin{aligned}
\langle x^2 \rangle &= \int_{x_{\min}}^{x_{\max}} Cy^2 f(y, q, \lambda) dy \\
&= C \left[y^2 \left(-T(y)^{\frac{2-q}{1-q}} + C_1 \right) \Big|_{x_{\min}}^{x_{\max}} + \int_{x_{\min}}^{x_{\max}} y(T(y)^{\frac{2-q}{1-q}} - C_1) dy \right] \\
&= C \left[-y^2 T(y)^{\frac{2-q}{1-q}} \Big|_{x_{\min}}^{x_{\max}} - y \frac{T(y)^{\frac{3-2q}{1-q}}}{(3-2q)\lambda} \Big|_{x_{\min}}^{x_{\max}} + \int_{x_{\min}}^{x_{\max}} \frac{T(y)^{\frac{3-2q}{1-q}}}{(3-2q)\lambda} dy \right] \\
&= C \left[-y^2 T(y)^{\frac{2-q}{1-q}} \Big|_{x_{\min}}^{x_{\max}} - y \frac{T(y)^{\frac{3-2q}{1-q}}}{(3-2q)\lambda} \Big|_{x_{\min}}^{x_{\max}} - \frac{T(y)^{\frac{4-3q}{1-q}}}{(3-2q)(4-3q)\lambda^2} \Big|_{x_{\min}}^{x_{\max}} \right] \\
&= CT(y)^{\frac{2-q}{1-q}} \left(-y^2 - y \frac{T(y)}{(3-2q)\lambda} - \frac{T(y)^2}{(3-2q)(4-3q)\lambda^2} \right) \Big|_{x_{\min}}^{x_{\max}}
\end{aligned} \tag{2.81}$$

O cálculo da média e da variância da distribuição q-exponencial vai além de um mero exercício teórico, constituindo um passo essencial para uma compreensão abrangente do comportamento desta distribuição. A análise destes momentos proporciona uma descrição analítica detalhada das propriedades fundamentais da distribuição, permitindo sua aplicação em diversos campos, como física estatística, teoria das filas e economia.

3 METODOLOGIA

3.1 Geração dos grafos

No contexto da simulação de grafos aleatórios, adotamos o modelo configuracional para a geração das redes com a restrição de que não existiriam loops e arestas múltiplas. Essa escolha foi motivada pela capacidade do modelo de reproduzir características importantes de redes complexas, como a distribuição de graus. Particularmente, definimos o grau mínimo (k_{\min}) como 2, o que assegura a integridade estatística da rede. Esta definição garante que todos os nós na rede possuem pelo menos dois vizinhos, contribuindo para a conectividade da rede e evitando a trivialidade de nós isolados.

No nosso algoritmo de simulação de randomizada em grafos, precisamos de uma representação que fosse possível selecionar as pontas dos vértices uniformemente em todo o grafo, para isso, empregamos uma técnica que cria uma nova lista onde cada identificador de vértice é repetido um número de vezes igual ao seu grau. Inicialmente, dispomos de uma lista de vértices com seus respectivos graus. Cada entrada nesta lista consiste em um par de valores: o identificador do vértice e seu grau. O identificador é um rótulo ou número único que identifica cada vértice no grafo, enquanto o grau é uma contagem de quantas arestas incidem sobre esse vértice. Para criar a nova lista, chamada aqui de "lista de repetição de graus", seguimos um processo iterativo. Percorremos a lista original de vértices e, para cada vértice, repetimos seu identificador na nova lista tantas vezes quanto for o seu grau. Por exemplo, se um vértice com identificador v tem grau 3, o identificador v aparecerá três vezes na lista de repetição de graus. O resultado é uma lista onde vértices com graus mais altos têm uma presença proporcionalmente maior. Esta lista de repetição de graus permite uma série de operações úteis, especialmente em contextos de simulação ou algoritmos que requerem amostragem proporcional ao grau. Por exemplo, ao selecionar um identificador aleatoriamente desta lista, a probabilidade de escolher um vértice específico é diretamente proporcional ao seu grau no grafo original. Isso é particularmente útil em algoritmos que dependem da ideia de "preferência pela conectividade", onde vértices com mais conexões são mais propensos a serem escolhidos. Após gerar a lista de repetições de graus, selecionamos e removemos pares de elementos de maneira aleatória e uniforme. Se não houver ligação prévia entre os elementos selecionados e se eles forem distintos, estabelece-se uma nova ligação entre eles. Este processo é repetido até que a lista não contenha mais elementos disponíveis para extração.

Para a estrutura de dados subjacente aos grafos, optamos por uma variação da lista de adjacências. Tradicionalmente, listas de adjacências são implementadas usando listas encadeadas ou arrays. No entanto, para melhorar a eficiência na manipulação dos dados, especialmente em operações de busca e atualização, substituímos as listas padrão por árvores rubro-negras. As árvores rubro-negras são uma forma de árvore binária de busca autoequilibrada. Elas mantêm suas propriedades essenciais através de regras de coloração (cada nó é vermelho ou preto) e de balanceamento (incluindo rotações e recolorações após inserções e exclusões) (Cormen *et al.*, 2022). As características distintivas de uma árvore rubro-negra incluem:

- **Altura Balanceada:** Garante que a árvore permaneça aproximadamente equilibrada, com a altura logarítmica em relação ao número de nós (Cormen *et al.*, 2022).
- **Eficiência de Operações:** As operações de busca, inserção e exclusão em uma árvore rubro-negra têm complexidade de tempo $O(\log n)$, onde n é o número de nós na árvore. Isso é particularmente vantajoso em cenários onde ocorrem frequentes alterações na estrutura do grafo, como é o caso de simulações dinâmicas (Cormen *et al.*, 2022).

Além disso, para gerenciar eficientemente a percolação maliciosa, empregamos uma estrutura de dados do tipo heap para manter os graus dos vértices ordenados. A estrutura heap é ideal para este propósito, pois permite a inserção e a remoção de elementos mantendo a propriedade de ordenação com uma complexidade de tempo $O(\log n)$. Isso possibilita o acesso rápido ao nó de maior grau em cada etapa da percolação maliciosa, uma operação crítica para simular ataques direcionados à rede (Cormen *et al.*, 2022).

3.2 Implementação

Os métodos implementados em C++ e disponíveis no nosso repositório do GitHub abrangem uma gama diversificada de funcionalidades além da manipulação de grafos. Além dos algoritmos já mencionados, o repositório inclui implementações de várias métricas analíticas e distribuições probabilísticas, como a lei de potência, que são essenciais para a análise de dados em sistemas complexos. Estas ferramentas adicionais permitem uma análise mais profunda e variada, abrindo portas para uma compreensão mais rica de padrões e comportamentos em dados. Além disso, reconhecendo a importância do desempenho em processamento de dados e simulações, incorporamos mecanismos de processamento multithread no código. Essa abordagem multithread aproveita ao máximo os recursos de hardware modernos, permitindo que as operações sejam executadas de maneira mais eficiente e em paralelo. Isso é particularmente vantajoso para tarefas

computacionalmente intensivas, como simulações em larga escala e análise de grandes conjuntos de dados. Portanto, o repositório no GitHub oferece uma suíte abrangente de ferramentas para análise e simulação, todas implementadas em C++ para eficiência e eficácia. O repositório com a implementação pode ser acessado em <https://github.com/marciomacielbastos/RandomGraph.git>.

4 REDES ALEATÓRIAS COM DISTRIBUIÇÃO q -EXPONENCIAL

Em uma grande variedade de campos, encontram-se experimentos, resultados numéricos e modelos teóricos que concordam bastante bem com q -exponenciais. Isso inclui aplicações em turbulência completamente desenvolvida (Beck *et al.*, 2001), difusão anômala em plasmas (Liu e Goree, 2008), estatísticas de raios cósmicos (Beck, 2004), econometria (Souza *et al.*, 2006) e (Borges, 2004), biofísica (Upadhyaya *et al.*, 2001)], e muitos outros. Em particular, muitas redes complexas empíricas foram encontradas seguindo distribuições de grau q -exponenciais (Carro *et al.*, 2016) e (Wedemann *et al.*, 2009). O mesmo comportamento também foi detectado em várias redes modelo propostas (Turner *et al.*, 2007), a maioria delas geradas através de modelos de crescimento baseados no princípio de **associação preferencial** (Soares *et al.*, 2005) e (Ochiai e Nacher, 2009). Esses modelos produzem distribuições q -exponenciais empiricamente, que em alguns casos podem até ser confirmadas analiticamente em (Turner e Tsallis, 2005) e (Ochiai e Nacher, 2009). Aqui, usamos o modelo de **configuração** (Newman, 2010) para gerar e estudar sistematicamente as propriedades de redes aleatórias com distribuição de grau q -exponencial arbitrariamente escolhida.

Redes são chamadas de livre de escala se a distribuição de graus segue uma lei de potência com expoente γ , $P(k) \sim k^{-\gamma}$, onde k é o grau de um nó, definido como o número de conexões que ele tem com outros nós. A distribuição q -exponencial dada por

$$P_q(k) = (2 - q)\lambda [1 - (1 - q)\lambda k]^{-1/(q-1)}, \quad (4.1)$$

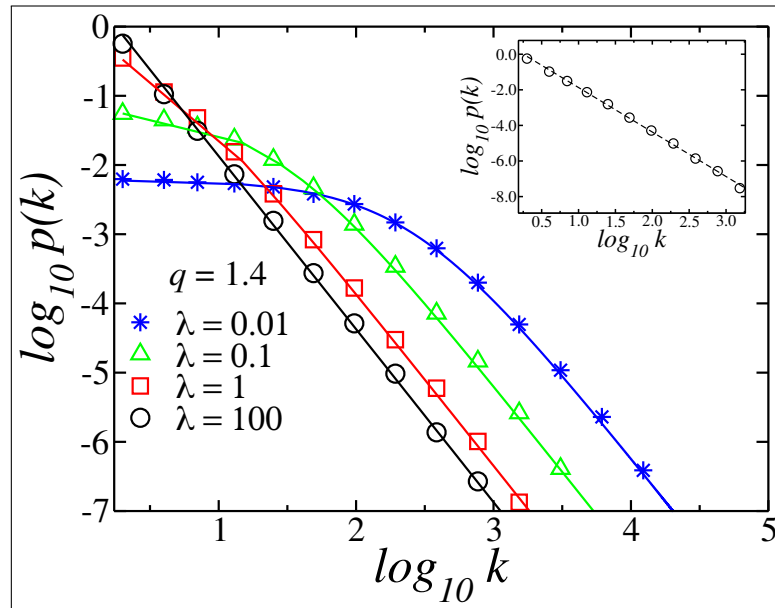
tem dois parâmetros, $q \geq 1$ e $\lambda \geq 0$. É uma generalização de uma distribuição de lei de potência, já que para grandes valores de k ($k \gg \lambda^{-1}$) ela decai como $k^{-\gamma}$ com $\gamma = 1/(q - 1)$. Para graus pequenos $k \ll \lambda^{-1}$, no entanto, tende a uma distribuição de patamar de altura $(2 - q)\lambda$. O parâmetro λ^{-1} determina portanto a transição entre esses dois regimes. A distribuição Eq. (1) representa um ingrediente fundamental no formalismo matemático da termostatística generalizada e suas aplicações (Tsallis, 2009) e (Adib *et al.*, 2003).

Embora a maioria das propriedades observadas em redes complexas com distribuições de cauda pesada seja determinada pela forma da cauda, desvios da lei de potência em graus menores alteram a ocorrência dos nós menos conectados, resultando em mudanças estruturais que podem afetar as principais propriedades e processos que ocorrem nessas redes. No que segue, daremos uma olhada mais de perto nesses efeitos.

Diferente de trabalhos anteriores baseados em modelos de fixação preferencial de

crescimento para obter redes com distribuições de grau imitando o comportamento q-exponencial, aqui, como já mencionado, empregamos o modelo de configuração (Newman, 2010) para construir nossas redes aleatórias. Dessa forma, garantimos que as propriedades topológicas particulares observadas, incluindo correlações intrínsecas de nó (Catanzaro *et al.*, 2005), não são induzidas pelo processo de crescimento, mas são sim uma consequência direta da forma q-exponencial da distribuição de grau.

Figura 7 – Distribuição de grau de redes q-exponenciais (símbolos) comparada com a distribuição esperada da equação (4.1) (linhas contínuas) para $q = 1,4$ e para $\lambda = 0.01$ (estrelas azuis), $\lambda = 0.1$ (triângulos verdes), $\lambda = 1$ (quadrados vermelhos) e $\lambda = 100$ (círculos pretos). O detalhe mostra uma comparação entre as distribuições de grau de uma q-exponencial com $\lambda = 100$ e uma distribuição puramente livre de escala (linha tracejada preta) com $q = 1.4$ ($\lambda = 2.5$). Esses resultados são obtidos para redes com tamanho $N = 500000$ pela média de 100 amostras. Como pode ser visto, para pequenos valores de λ , as distribuições atingem um platô em graus pequenos e o regime de lei de potência se torna maior com o aumento de λ .

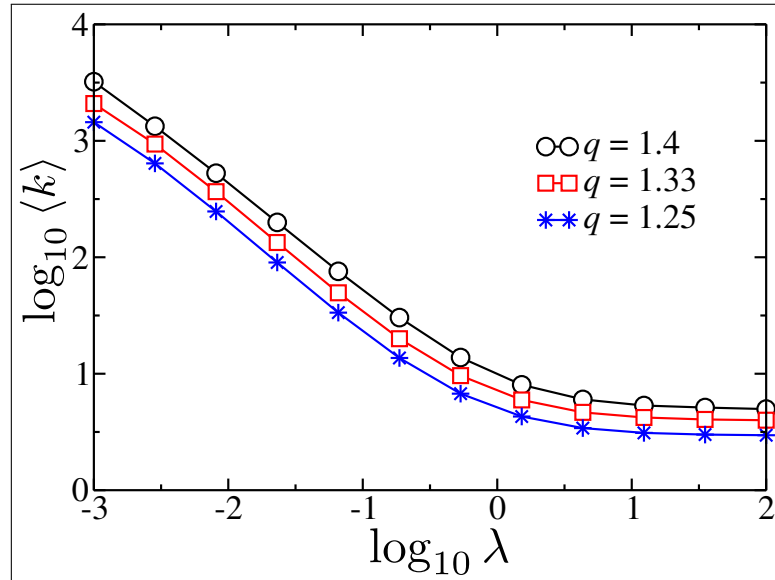


Fonte: (Sampaio *et al.*, 2023)

4.1 Resultados e discussão

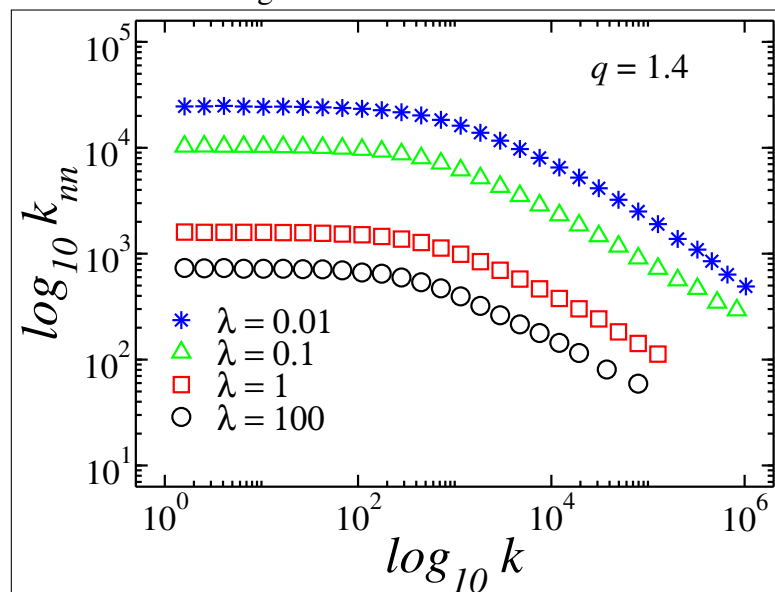
Como mostrado na Figura 7, as distribuições de grau obtidas por este método seguem de muito perto a forma q-exponencial esperada da equação (4.1). Para $\lambda \ll 1$, observa-se um platô pronunciado para pequenos valores de k . Por outro lado, quando $\lambda \gg 1$, efetivamente, nossa distribuição de grau será idêntica a uma distribuição de grau livre de escala com o mesmo k_{min} .

Figura 8 – Dependência do grau médio $\langle k \rangle$ no parâmetro λ para diferentes valores de q . Estas curvas correspondem a $q = 1.4$ (círculos pretos), $q = 1.33$ (quadrados vermelhos) e $q = 1.25$ (estrelas azuis). Para valores de $\lambda < 1$, o grau médio segue $\langle k \rangle \sim \lambda^{-1}$, conforme esperado para distribuições q -exponenciais. No limite de lei de potência pura ($\lambda \gg 1$), o grau médio satura em um valor independente de λ .



Fonte: (Sampaio *et al.*, 2023)

Figura 9 – Dependência do grau médio do vizinho mais próximo k_{nn} dos vértices no grau k para redes com tamanho $N = 10^7$ nós, $q = 1.4$, e $\lambda = 100$ (círculo preto), $\lambda = 1$ (quadrado vermelho), $\lambda = 0.1$ (triângulo verde) e $\lambda = 0.01$ (estrela azul). Os patamares de cada curva correspondem ao valor $\langle k^2 \rangle / \langle k \rangle$ (Catanzaro *et al.*, 2005). Todas as curvas apontam para um comportamento intrinsecamente dissortativo para valores suficientemente grandes de k .



Fonte: (Sampaio *et al.*, 2023)

Portanto, ao diminuir o parâmetro λ , podemos nos afastar continuamente de uma distribuição de grau livre de escala e ampliar o platô. Assim, variar λ nos permitirá identificar o efeito dos desvios da livre de escala pura que são particulares às distribuições q-exponenciais.

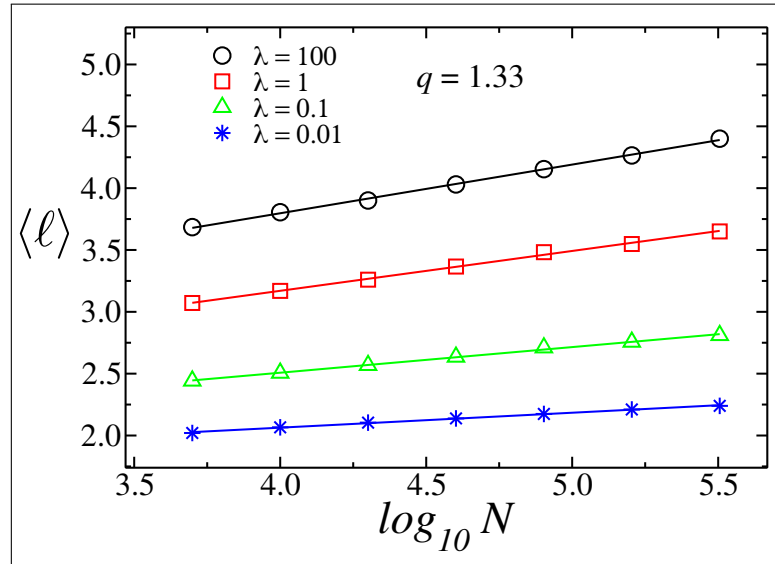
Na Figura 8, mostramos como o grau médio $\langle k \rangle$ depende de λ para diferentes valores do parâmetro q . Como esperado para uma rede q-exponencial, para pequenos valores de λ , o grau médio é proporcional a λ^{-1} . No entanto, para valores suficientemente grandes de λ , a distribuição de grau se transforma em uma lei de potência. Neste último regime, o grau médio torna-se independente de λ . Portanto, por um lado, redes livres de escala têm muitos mais nós menos conectados do que redes com distribuição de grau q-exponencial com pequenos λ . Quanto menor o λ , mais densas as redes se tornam, aumentando o número e o grau de seus *hubs*. Diferenças topológicas como essas podem levar a mudanças substanciais nas propriedades estruturais de redes complexas, bem como no comportamento estático e dinâmico de modelos quando implementados nesses substratos.

Foi mostrado por (Catanzaro *et al.*, 2005), (Newman, 2003) e (Newman, 2002) que redes aleatórias com distribuição de grau de cauda pesada frequentemente exibem assortatividade intrínseca. Este também é o caso para redes q-exponenciais. Na Figura 9, mostramos o grau médio esperado $k_{nn}(k)$ dos vizinhos mais próximos de um dado nó com grau k . Como pode ser visto, os nós mais conectados têm um k_{nn} menor, apontando para uma assortatividade intrínseca negativa. A razão para esse comportamento dissortativo em redes aleatórias é que os nós mais conectados nem podem se conectar a si mesmos nem ter múltiplas conexões entre eles (Catanzaro *et al.*, 2005). Também vemos na Figura 3 que k_{nn} se torna maior quanto mais a distribuição se desvia de uma pura livre de escala.

Redes q-exponenciais aleatórias também exibem comportamento de pequeno mundo, o que significa que seu caminho mais curto médio aumenta logaritmicamente com o tamanho da rede, $\langle l \rangle = \alpha \log_{10} N$, conforme mostrado na Figura 10. Claramente, o caminho mais curto torna-se consideravelmente mais curto à medida que as distribuições de grau se desviam sistematicamente da lei de potência pura, ou seja, para valores suficientemente grandes de λ . A Figura 11 mostra a variação do pré-fator α para diferentes valores de q na faixa $1.25 \leq q \leq 1.4$, e $\lambda = 0.1, 1$ e 100 . Para todos os fins práticos, essa faixa de q , que corresponde a $2.5 \leq \gamma \leq 4.0$, cobre todos os decaimentos interessantes da lei de potência. Como já mencionado, o caso $\lambda = 100$ é equivalente a uma rede livre de escala com $P(k) \sim k^{-\gamma}$ e $k_{min} = 2$.

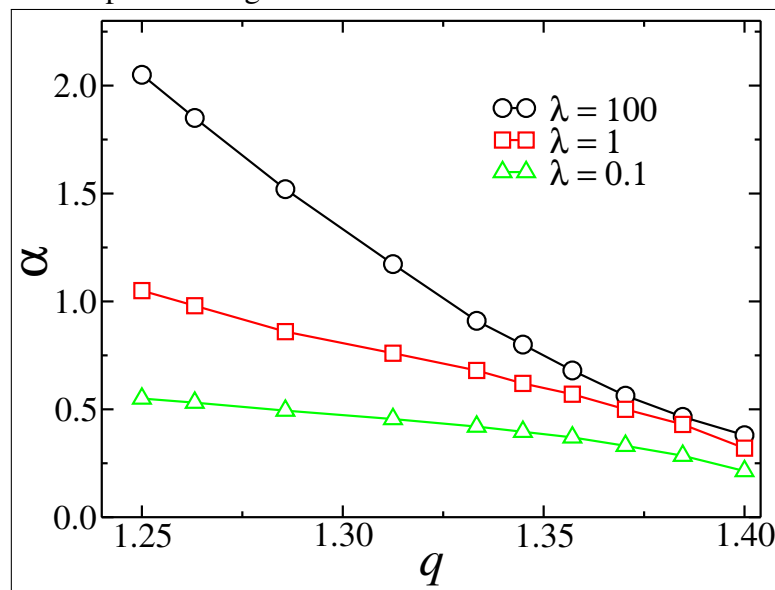
Uma propriedade particularmente importante para fins práticos é a robustez das redes

Figura 10 – Dependência do tamanho no comprimento médio do menor caminho $\langle \ell \rangle$ para $q = 1.33$ e para $\lambda = 0.01$ (estrelas azuis), $\lambda = 0.1$ (triângulos verdes), $\lambda = 1$ (quadrados vermelhos) e $\lambda = 1$ (círculos pretos). Os resultados são obtidos através da média de 105, 104, 104, 104, 104, 103 e 103 amostras de tamanho $N = 5000, 10000, 20000, 40000, 80000, 160000$ e 320000 , respectivamente.



Fonte: (Sampaio *et al.*, 2023)

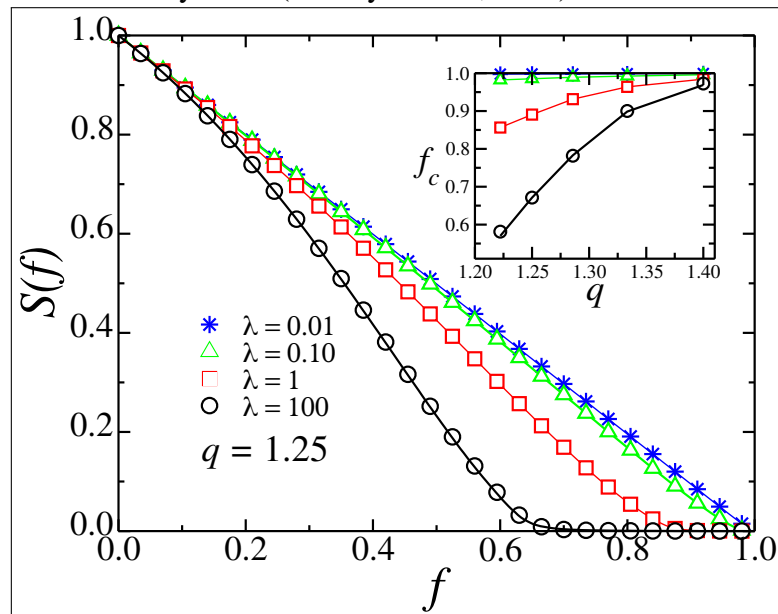
Figura 11 – O pré-fator α como função de q para $\lambda = 100$ (círculos pretos), $\lambda = 1$ (quadrados vermelhos) e $\lambda = 0.1$ (triângulos verdes). Cada ponto é obtido de um ajuste de mínimos quadrados aos dados, $\alpha \sim \log_{10} N$, com barras de erro menores que os símbolos. As linhas contínuas representam guias visuais.



Fonte: (Sampaio *et al.*, 2023)

contra falhas aleatórias. Na Figura 12, traçamos a densidade de nós no maior *cluster* $S(f)$ em função da fração f de nós removidos aleatoriamente, para $N = 320000$ $q = 1.25$ e vários valores de λ . Para grandes valores de λ ($\lambda \gg 1$), essa distribuição de grau se torna uma lei de potência com $\gamma = 4$. Considerando que $k_{min} = 2$, essa rede deveria exibir um ponto crítico $f_c < 1$ (Burda *et al.*, 2001). Isso é exatamente o que se vê na Figura 6 para $\lambda = 100$. Para valores menores de λ , a rede se torna mais robusta, com $f_c \rightarrow 1$ à medida que $\lambda \rightarrow 0$. No detalhe da Figura 12, vemos como f_c depende de q para diferentes valores de λ . Claramente, quanto mais uma rede q -exponencial se assemelha a uma livre de escala, ou seja, para grandes valores de λ , mais ela é frágil contra falhas aleatórias. Além disso, esse efeito é dramaticamente ampliado com a diminuição do parâmetro q .

Figura 12 – A densidade de nós no maior aglomerado $S(f)$ como função da fração f de nós removidos para falhas aleatórias em redes q -exponenciais com $q = 1.25$, $\lambda = 0.01$ (estrelas azuis), $\lambda = 0.1$ (triângulos verdes), $\lambda = 1$ (quadrados vermelhos) e $\lambda = 100$ (círculos pretos). Estes resultados correspondem a redes de tamanho $N = 320000$ por média de 200 amostras. As linhas contínuas para $\lambda = 0.01, 0.1, 1$ representam guias visuais, enquanto para $\lambda = 100$, a linha sólida preta corresponde ao resultado para ataques aleatórios em uma rede puramente livre de escala. O detalhe mostra a fração crítica f_c determinada pelos critérios de Molloy-Reed (Molloy e Reed, 1995).



Fonte: (Sampaio *et al.*, 2023)

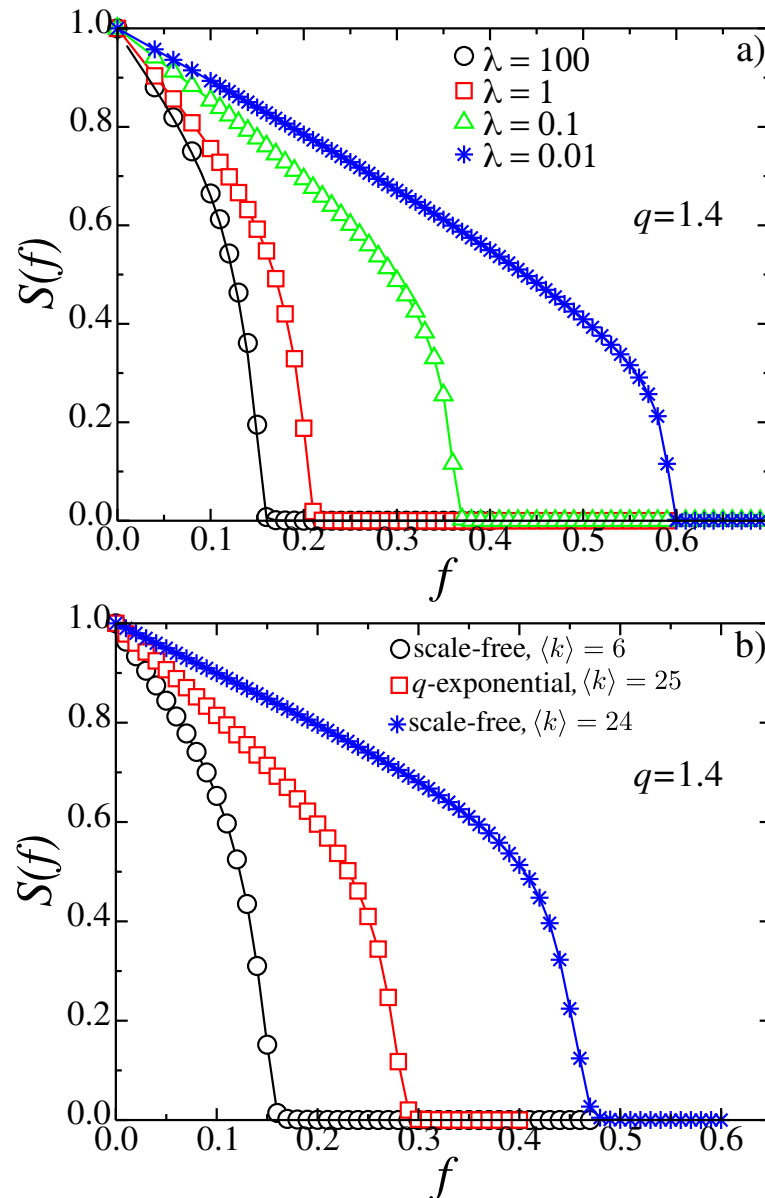
No caso de ataque malicioso, mostramos na Figura 13(a) a variação da densidade do maior *cluster* S em função da fração f de nós removidos direcionados pelo maior grau, para $q = 1.4$ ($\gamma = 2.5$) e diferentes valores de λ . Aqui, a distribuição de grau é atualizada após cada

remoção de nó, conforme descrito nas em (Schneider *et al.*, 2011) e (Wu e Holme, 2011). Como retratado, as redes q-exponenciais se tornam cada vez menos resilientes à medida que o *crossover* λ aumenta, já que uma fração maior de nós precisa ser removida antes que o ponto crítico seja alcançado. Esse comportamento persiste até um ponto em que o valor de λ é suficientemente grande, de modo que o comportamento livre de escala da distribuição de grau domina. Como resultado, a curva q-exponencial S versus f para $\lambda = 100$ e a correspondente gerada a partir de redes com distribuição puramente livre de escala são perfeitamente coincidentes.

Neste ponto, argumentamos que a base para comparar a robustez de redes livres de escala e q-exponenciais não é óbvia, no sentido de que a primeira pode ser vista como um caso particular da segunda, se considerarmos os mesmos valores de q e k_{min} e para um valor suficientemente grande do parâmetro λ . A situação se torna bastante diferente para baixos valores de λ , uma vez que essas redes q-exponenciais têm valores substancialmente maiores de $\langle k \rangle$, sendo, portanto, mais resilientes. Para obter redes livres de escala com o mesmo grau médio que essas q-exponenciais, portanto, temos que mudar seu k_{min} . Em todo caso, ainda é interessante comparar a resiliência de redes q-exponenciais e livres de escala com graus médios semelhantes. Como exemplo, se fixarmos $q = 1.4$ e considerarmos q-exponenciais geradas com $k_{min} = 2$ e $\lambda = 10^{-1}$, e redes livres de escala com $k_{min} = 8$, ambas com tamanho $N = 10^6$, seus graus médios resultantes se tornam bastante próximos, ou seja, $\langle k \rangle = 24$ e 25 , respectivamente. No caso de ataques aleatórios, as redes q-exponenciais e as livres de escala respondem de forma semelhante, sendo necessário remover a maioria dos nós para degradar completamente seus maiores *clusters* (não mostrado). Essa forte resiliência deve ser esperada, uma vez que o parâmetro $q = 1.4$ ($\gamma = 2.5$) cai no regime em que até mesmo redes livres de escala são robustas a falhas aleatórias (Burda *et al.*, 2001). Os resultados apresentados na Figura 13(b) mostram que, nessas condições particulares, redes q-exponenciais (com pequeno λ) são menos resilientes para ataques maliciosos do que as livres de escala (ou q-exponenciais com grande λ). Precisamente, a fração do maior *cluster*, $S(f)$, dessas redes q-exponenciais decai mais rápido com f do que a de suas contrapartes livres de escala com grau médio semelhante, com o maior *cluster* desaparecendo após uma fração menor dos nós serem removidos. Para comparação, também mostramos a resposta a ataques maliciosos das redes livres de escala menos robustas geradas com $k_{min} = 2$, levando a $\langle k \rangle = 6$.

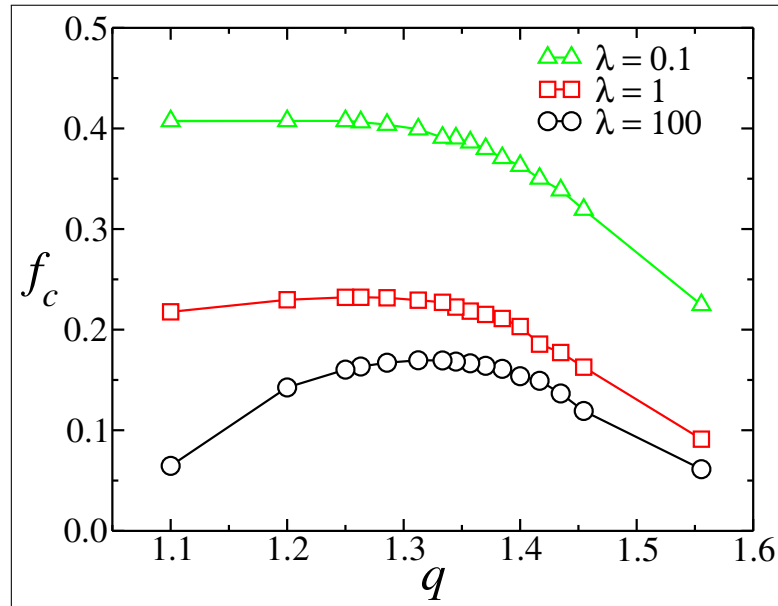
A Figura 14 mostra a dependência da fração crítica f_c no parâmetro q , conforme determinado pelo critério de Molloy-Reed (Molloy e Reed, 1995) e para diferentes valores

Figura 13 – (a) A densidade de nós no maior *cluster* $S(f)$ como função da fração f de nós removidos para ataques maliciosos direcionados pelo maior grau, para $q = 1.4$ e $\lambda = 0.01$ (estrelas azuis), $\lambda = 0.1$ (triângulos verdes), $\lambda = 1$ (quadrados vermelhos) e $\lambda = 100$ (círculos pretos). As linhas contínuas para $\lambda = 0.01, 0.1$ e 1 representam guias para o olho, enquanto para $\lambda = 100$ a linha contínua preta corresponde ao resultado para ataques maliciosos em uma rede puramente livre de escala. Esses resultados correspondem a redes de tamanho $N = 320000$ pela média de 200 amostras. (b) A densidade $S(f)$ como função da fração f de redes geradas com $q = 1.4$ e submetidas a ataques maliciosos direcionados pelo maior grau. Os resultados correspondem a redes livres de escala com $k_{min} = 2$ e $\langle k \rangle = 6$ (círculos pretos), $k_{min} = 8$ e $\langle k \rangle = 24$ (estrelas azuis) e para redes q -exponenciais com $\lambda = 10, k_{min} = 2, \langle k \rangle = 25$ (quadrados vermelhos). Eles foram obtidos com redes de tamanho $N = 10^6$ pela média de 100 amostras.



Fonte: (Sampaio *et al.*, 2023)

Figura 14 – A fração crítica f_c para ataques maliciosos como função de q para $\lambda = 0.1$ (triângulos verdes), $\lambda = 1$ (quadrados vermelhos) e $\lambda = 100$ (círculos pretos). Os resultados são obtidos para redes de tamanho $N = 500000$ pela média de 2000 amostras. A fração a partir da qual o maior *cluster* não obedece ao critério de Molloy-Reed é a fração crítica f_c . As redes q -exponenciais com valores menores de λ são claramente mais robustas, já que uma maior fração de nós precisa ser removida para atingir o ponto crítico. As linhas contínuas representam guias para o olho.



Fonte: (Sampaio *et al.*, 2023)

de q . É interessante notar que o comportamento de f_c muda substancialmente com λ . Por exemplo, considerando $\lambda = 0.1$, vemos um platô até $q \approx 4/3$, seguido por um declínio. Por outro lado, para $\lambda = 100$ observamos um máximo claro em f_c em $q \approx 4/3$. Como mencionado, para $\lambda = 100$, a distribuição de grau se aproxima da forma de uma lei de potência, e esse máximo na condição crítica é consistente com o esperado para redes livres de escala pura (Pósfai e Barabasi, 2016). Notamos que esse máximo é devido a um compromisso entre dois efeitos. Para $q > 4/3$, a distribuição de grau decai assintoticamente como uma lei de potência com expoente controlador $\gamma < 3$, atingindo $\gamma = 2$ à medida que q se aproxima de $3/2$. Neste limite, remover apenas alguns *hubs* resulta em um colapso total da rede. No outro limite, à medida que γ diverge quando $q \rightarrow 1$, a distribuição de grau já não é mais pesada na cauda, diminuindo rapidamente. Além disso, definindo $k_{min} = 2$, a rede gerada já está próxima do estado crítico. Este comportamento pode ser suprimido impondo $k_{min} \geq 3$ no caso de redes livres de escala pura (Pósfai e Barabasi, 2016), ou usando pequenos valores de λ em redes q -exponenciais.

Em seguida, estendemos a análise da topologia das redes q -exponenciais investigando

sua estrutura hierárquica em termos do método de decomposição de k-core (Alvarez-Hamelin *et al.*, 2005) e (Serafino *et al.*, 2022). O k-core de um gráfico G é o maior subgrafo conectado em que todos os seus nós têm um grau maior ou igual a k . Para obter o k-core, removemos todos os nós com um grau menor que k . Em seguida, verificamos se alguns nós ainda têm um grau corrente menor que k e os removemos. Repetimos essa verificação até que nenhuma remoção adicional seja possível. A partir dessa decomposição, podemos definir para cada nó um *rank* na rede, tal que um nó será mais periférico quanto menor for o seu k . Subgrafos de k-core são resilientes contra falhas, pois preservam sua convexidade após $(k - 1)$ reorganizações aleatórias de arestas ou nós. Esse tipo de robustez tende a aumentar para os nós mais internos. Aqui analisamos a dependência de tamanho finito do k-core mais alto, k_h , e sua massa, M_h , para redes q-exponenciais com $q = 1.4$ e diferentes valores do parâmetro λ . A partir da referência (Dorogovtsev *et al.*, 2006), esperamos obter escalonamento de tamanho finito como

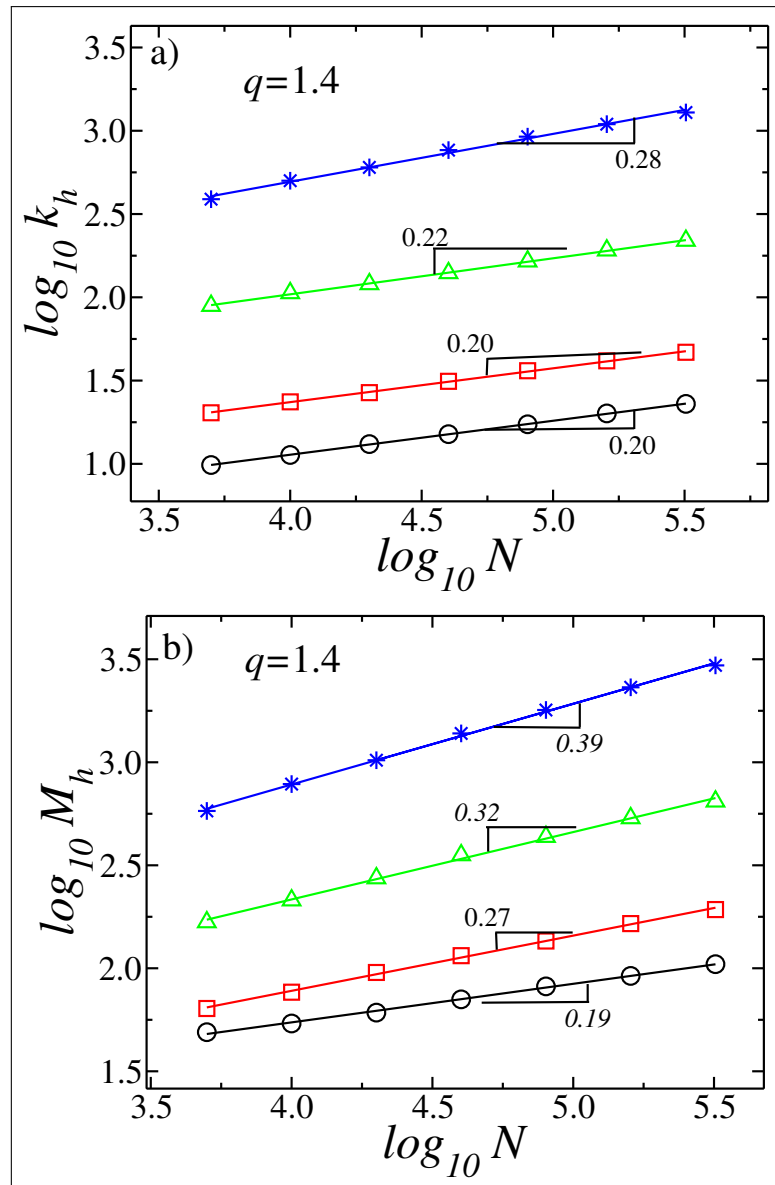
$$k_h \sim N^\delta \tag{4.2}$$

para o k-core mais alto, e

$$M_h \sim N^\Delta \tag{4.3}$$

para a massa do k-core mais alto. Descobrimos que os expoentes δ e Δ são funções dos parâmetros de distribuição q e λ . A Figura 15(a) mostra a dependência no tamanho da rede N de k_h em um gráfico log-log para $q = 1.4$ e diferentes valores de λ . Em comparação com a dependência moderada do expoente δ em λ , o fator de escala da relação da equação (4.2), no entanto, aumenta consideravelmente com λ^{-1} . Por outro lado, como mostrado na Figura 15(b), a dependência de M_h em N indica que tanto o expoente Δ quanto o fator de escala da equação (4.3) diminuem substancialmente com λ . Isso mostra que redes com um platô maior (menor λ) têm um k-core mais alto que é maior e tem um k maior. Isso está de acordo com nossas descobertas anteriores de que as redes q-exponenciais se tornam mais robustas quanto menor for o λ .

Figura 15 – (a) k-core mais alto e (b) massa do k-core mais alto versus N para $q = 1.4$ e $\lambda = 0.01$ (estrelas azuis), $\lambda = 0.1$ (triângulos verdes), $\lambda = 1$ (quadrados vermelhos) e $\lambda = 100$ (círculos pretos). Os resultados são obtidos pela média de 105, 104, 104, 104, 104, 103 e 103 amostras de tamanho $N = 5000, 10000, 20000, 40000, 80000, 160000$ e 320000 ci, respectivamente.



Fonte: (Sampaio *et al.*, 2023)

5 CONCLUSÕES

Em conclusão, nossa pesquisa resultou na geração de redes complexas não enviesadas, exibindo distribuições de grau q -exponenciais com uma diversidade de valores paramétricos. Essas redes representam uma generalização das redes livres de escala, pois permitem a obtenção de distribuições de grau de lei de potência pura ajustando-se o parâmetro de transição λ . Este parâmetro é fundamental, pois controla a extensão do platô q -exponencial em graus de nó mais baixos, oferecendo uma nova camada de flexibilidade nas características das redes.

Os resultados alcançados demonstram que este grau adicional de liberdade nas redes q -exponenciais proporciona uma flexibilidade substancial em suas propriedades topológicas e de transporte. Observamos em nossas análises uma notável assortatividade, comportamento de pequeno mundo, e resiliência superior contra falhas aleatórias e ataques maliciosos, ultrapassando muitas das características das redes livres de escala tradicionais. Além disso, a análise do escalonamento de tamanho finito nos elementos da decomposição de k -core fornece uma visão mais aprofundada da estrutura e robustez dessas redes.

Considerando a riqueza dos resultados obtidos e a relevância prática das redes q -exponenciais, sugere-se para futuros estudos a investigação de suas propriedades dinâmicas. Áreas como sincronização, propagação de epidemias e formação de opiniões emergem como campos promissores para esses estudos. Tais investigações podem revelar insights valiosos sobre o comportamento das redes q -exponenciais sob variados processos dinâmicos e condições, expandindo ainda mais a aplicabilidade dessas redes em sistemas complexos de múltiplas disciplinas, desde a física até as ciências sociais e biológicas.

REFERÊNCIAS

- ADIB, A. B.; MOREIRA, A. A.; JR, J. S. A.; ALMEIDA, M. P. Tsallis thermostatics for finite systems: a hamiltonian approach. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 322, p. 276–284, 2003.
- AKIBA, T.; NAKAMURA, K.; TAKAGUCHI, T. Fractality of massive graphs: scalable analysis with sketch-based box-covering algorithm. In: **2016 IEEE 16th International Conference on Data Mining**. [S. l.]: ICDM, 2016. p. 769–774.
- ALBERT, R.; BARABÁSI, A.-L. Statistical mechanics of complex networks. **Reviews of modern physics**, APS, v. 74, n. 1, p. 47, 2002.
- ALVAREZ-HAMELIN, J.; DALL’ASTA, L.; BARRAT, A.; VESPIGNANI, A. Large scale networks fingerprinting and visualization using the k-core decomposition. **Advances in neural information processing systems**, v. 18, 2005.
- ANDERSON, I. Graph theory (2nd edn.), by reinhard diestel. pp. 312.£ 24. 2000. isbn 0 387 98976 5 (springer-verlag). **The Mathematical Gazette**, Cambridge University Press, v. 85, n. 502, p. 176–177, 2001.
- ASH, R. B.; DOLÉANS-DADE, C. A. **Probability and measure theory**. [S. l.]: Academic press, 2000.
- AXLER, S. **Measure, integration & real analysis**. [S. l.]: Springer Nature, 2020.
- BARABÁSI, A.-L. **Network science**. [S. l.]: The Royal Society Publishing, 2013. v. 371. 20120375 p.
- BECK, C. Generalized statistical mechanics of cosmic rays. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 331, n. 1-2, p. 173–181, 2004.
- BECK, C.; LEWIS, G. S.; SWINNEY, H. L. Measuring nonextensivity parameters in a turbulent couette-taylor flow. **Physical Review E**, APS, v. 63, n. 3, p. 035303, 2001.
- BEN-AVRAHAM, D.; HAVLIN, S. **Diffusion and reactions in fractals and disordered systems**. [S. l.]: Cambridge university press, 2000.
- BOLLOBÁS, B.; RIORDAN, O. **Percolation**. [S. l.]: Cambridge University Press, 2006.
- BORGES, E. P. Empirical nonextensive laws for the county distribution of total personal income and gross domestic product. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 334, n. 1-2, p. 255–266, 2004.
- BRITO, S.; NUNES, T. C.; SILVA, L. R. da; TSALLIS, C. Scaling properties of d-dimensional complex networks. **Physical Review E**, APS, v. 99, n. 1, p. 012305, 2019.
- BRITO, S.; SILVA, L. D.; TSALLIS, C. Role of dimensionality in complex networks. **Scientific reports**, Nature Publishing Group UK London, v. 6, n. 1, p. 27992, 2016.
- BURDA, Z.; CORREIA, J.; KRZYWICKI, A. Statistical ensemble of scale-free random graphs. **Physical Review E**, APS, v. 64, n. 4, p. 046118, 2001.

- BURLESON-LESSER, K.; MORONE, F.; TOMASSONE, M. S.; MAKSE, H. A. K-core robustness in ecological and financial networks. **Scientific reports**, Nature Publishing Group UK London, v. 10, n. 1, p. 3357, 2020.
- CARRO, A.; TORAL, R.; MIGUEL, M. S. Coupled dynamics of node and link states in complex networks: a model for language competition. **New Journal of Physics**, IOP Publishing, v. 18, n. 11, p. 113056, 2016.
- CASELLA, G.; BERGER, R. L. **Statistical inference**. [S. l.]: Cengage Learning, 2021.
- CATANZARO, M.; BOGUNÁ, M.; PASTOR-SATORRAS, R. Generation of uncorrelated random scale-free networks. **Physical review e**, APS, v. 71, n. 2, p. 027103, 2005.
- COHEN, R.; HAVLIN, S. **Complex networks: structure, robustness and function**. [S. l.]: Cambridge university press, 2010.
- CORMEN, T. H.; LEISERSON, C. E.; RIVEST, R. L.; STEIN, C. **Introduction to algorithms**. [S. l.]: MIT press, 2022.
- CPPREFERENCE.COM. **C++ Random Number Generation**. 2023. Disponível em: <https://en.cppreference.com/w/cpp/numeric/random>. Acesso em: 21 nov. 2023.
- DEHMER, M.; EMMERT-STREIB, F. **Analysis of complex networks: from biology to linguistics**. [S. l.]: John Wiley & Sons, 2009.
- DOROGOVTSEV, S. N.; GOLTSEV, A. V.; MENDES, J. F. F. K-core organization of complex networks. **Physical review letters**, APS, v. 96, n. 4, p. 040601, 2006.
- FRALEIGH, J. B. **A first course in abstract algebra**. [S. l.]: Pearson Education India, 2003.
- JAMES, B. R. **Probabilidade: Um curso em nível intermediário**, 2a. edição. [S. l.]: IMPA, Rio de Janeiro, 2002.
- KATZ, L. A new status index derived from sociometric analysis. **Psychometrika**, Springer, v. 18, n. 1, p. 39–43, 1953.
- KURATOWSKI, K.; WHYBURN, G. T. Sur les éléments cycliques et leurs applications. **Fundamenta Mathematicae**, Polska Akademia Nauk. Instytut Matematyczny PAN, v. 16, n. 1, p. 305–331, 1930.
- LIMA, E. L. **Curso de análise**. [S. l.]: IMPA, 1999. v. 1.
- LIMA, E. L. **Curso de análise**. [S. l.]: IMPA, 2008. v. 2.
- LIU, B.; GOREE, J. Superdiffusion and non-gaussian statistics in a driven-dissipative 2d dusty plasma. **Physical review letters**, APS, v. 100, n. 5, p. 055003, 2008.
- MADER, W. Existenz n-fach zusammenhängender teilgraphen in graphen genügend grosser kantendichte. In: SPRINGER. **Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg**. [S. l.], 1972. v. 37, p. 86–97.
- MANDELBROT, B. B. Les objets fractals: forme, hasard et dimension. **(No Title)**, 1975.
- MENGER, K. Zur allgemeinen kurventheorie. **Fundamenta Mathematicae**, Polska Akademia Nauk. Instytut Matematyczny PAN, v. 10, n. 1, p. 96–115, 1927.

- MOLLOY, M.; REED, B. A critical point for random graphs with a given degree sequence. **Random structures & algorithms**, Wiley Online Library, v. 6, n. 2-3, p. 161–180, 1995.
- NEWMAN, M. **Networks: an introduction**. [S. l.]: Oxford University Press, 2010.
- NEWMAN, M. **Networks**. [S. l.]: Oxford university press, 2018.
- NEWMAN, M. E. Assortative mixing in networks. **Physical review letters**, APS, v. 89, n. 20, p. 208701, 2002.
- NEWMAN, M. E. Mixing patterns in networks. **Physical review E**, APS, v. 67, n. 2, p. 026126, 2003.
- OCHIAI, T.; NACHER, J. On the construction of complex networks with optimal tsallis entropy. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 388, n. 23, p. 4887–4892, 2009.
- PÓSFAL, M.; BARABASI, A.-L. **Network Science**. [S. l.]: Citeseer, 2016.
- ROBERTSON, N.; SEYMOUR, P. D. Graph minors. i. excluding a forest. **Journal of Combinatorial Theory, Series B**, Elsevier, v. 35, n. 1, p. 39–61, 1983.
- ROSS, S. M. **Introduction to probability models**. [S. l.]: Academic press, 2014.
- ROSS, S. M. **Simulation**. [S. l.]: Academic Press, 2022.
- SAMPAIO, C. I.; BASTOS, M. M.; HERRMANN, H. J.; MOREIRA, A. A.; JR, J. S. A. Random networks with q-exponential degree distribution. **Physical Review Research**, APS, v. 5, n. 3, p. 033088, 2023.
- SAMPAIO, C. I.; MOREIRA, A. A.; ANDRADE, R. F.; HERRMANN, H. J.; JR, J. S. A. Mandala networks: ultra-small-world and highly sparse graphs. **Scientific reports**, Nature Publishing Group UK London, v. 5, n. 1, p. 9082, 2015.
- SCHNEIDER, C. M.; KESSELRING, T. A.; JR, J. S. A.; HERRMANN, H. J. Box-covering algorithm for fractal dimension of complex networks. **Physical Review E**, APS, v. 86, n. 1, p. 016707, 2012.
- SCHNEIDER, C. M.; MOREIRA, A. A.; JR, J. S. A.; HAVLIN, S.; HERRMANN, H. J. Mitigation of malicious attacks on networks. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 108, n. 10, p. 3838–3841, 2011.
- SERAFINO, M.; MONTEIRO, H. S.; LUO, S.; REIS, S. D.; IGUAL, C.; NETO, A. S. L.; TRAVIZANO, M.; JR, J. S. A.; MAKSE, H. A. Digital contact tracing and network theory to stop the spread of covid-19 using big-data on human mobility geolocalization. **PLOS Computational Biology**, Public Library of Science San Francisco, CA USA, v. 18, n. 4, p. e1009865, 2022.
- SOARES, D. J.; TSALLIS, C.; MARIZ, A. M.; SILVA, L. R. da. Preferential attachment growth model and nonextensive statistical mechanics. **Europhysics Letters**, IOP Publishing, v. 70, n. 1, p. 70, 2005.
- SONG, C.; HAVLIN, S.; MAKSE, H. A. Self-similarity of complex networks. **Nature**, Nature Publishing Group UK London, v. 433, n. 7024, p. 392–395, 2005.

SOUZA, J. de; MOYANO, L. G.; QUEIRÓS, S. D. On statistical properties of traded volume in financial markets. **The European Physical Journal B-Condensed Matter and Complex Systems**, Springer, v. 50, p. 165–168, 2006.

STAUFFER, D.; AHARONY, A. **Introduction to percolation theory**. [S. l.]: CRC press, 2018.

THURNER, S.; KYRIAKOPOULOS, F.; TSALLIS, C. Unified model for network dynamics exhibiting nonextensive statistics. **Physical Review E**, APS, v. 76, n. 3, p. 036111, 2007.

THURNER, S.; TSALLIS, C. Nonextensive aspects of self-organized scale-free gas-like networks. **Europhysics Letters**, IOP Publishing, v. 72, n. 2, p. 197, 2005.

TSALLIS, C. Possible generalization of boltzmann-gibbs statistics. **Journal of statistical physics**, Springer, v. 52, n. 1, p. 479–487, 1988.

TSALLIS, C. Connection between scale-free networks and nonextensive statistical mechanics. **The European Physical Journal Special Topics**, Springer, v. 161, n. 1, p. 175–180, 2008.

TSALLIS, C. **Introduction to nonextensive statistical mechanics: approaching a complex world**. [S. l.]: Springer, 2009. v. 1.

TSALLIS, C. Non-additive entropies and statistical mechanics at the edge of chaos: a bridge between natural and social sciences. **Philosophical Transactions of the Royal Society A**, The Royal Society, v. 381, n. 2256, p. 20220293, 2023.

UPADHYAYA, A.; RIEU, J.-P.; GLAZIER, J. A.; SAWADA, Y. Anomalous diffusion and non-gaussian velocity distribution of hydra cells in cellular aggregates. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 293, n. 3-4, p. 549–558, 2001.

VULPIANI, A.; CECCONI, F.; CENCINI, M. **Chaos: from simple models to complex systems**. [S. l.]: World Scientific, 2009. v. 17.

WEDEMANN, R. S.; DONANGELO, R.; CARVALHO, L. A. de. Nonextensivity in a memory network access mechanism. **Brazilian Journal of Physics**, SciELO Brasil, v. 39, p. 495–499, 2009.

WU, Z.-X.; HOLME, P. Onion structure and network robustness. **Physical Review E**, APS, v. 84, n. 2, p. 026106, 2011.

APÊNDICE A – PROBABILIDADE E VARIÁVEIS ALEATÓRIAS

A teoria da medida fornece uma base matemática rigorosa para definição de probabilidade, permitindo-nos quantificar e estudar a medida de conjuntos em um espaço amostral. Para entender essa teoria, precisamos nos familiarizar com vários conceitos-chave, como álgebras, σ -álgebras, pré-medidas e extensões. Uma álgebra é uma coleção \mathcal{A} de subconjuntos de um conjunto Ω que satisfaz (ASH; DOLÉANS-DADE, 2000):

1. Ω está em \mathcal{A} .
2. Se A está em \mathcal{A} , então o complemento de A também está em \mathcal{A} .
3. Se A e B estão em \mathcal{A} , então a união $A \cup B$ também está em \mathcal{A} .

Para exemplificar, considere $\Omega = \{1, 2, 3, 4\}$. Uma álgebra \mathcal{A} pode ser: $\mathcal{A} = \{\emptyset, \{1, 2\}, \{3, 4\}, \Omega\}$. Uma σ -álgebra é uma extensão da ideia de álgebra. Uma coleção \mathcal{F} de subconjuntos de um conjunto Ω é chamada de σ -álgebra se (ASH; DOLÉANS-DADE, 2000):

1. Ω está em \mathcal{F} .
2. Se A está em \mathcal{F} , então o complemento de A também está em \mathcal{F} .
3. Se A_1, A_2, \dots são conjuntos em \mathcal{F} , então a união (possivelmente infinita) $\bigcup_{i=1}^{\infty} A_i$ também está em \mathcal{F} .

Assim, considerando o mesmo conjunto $\Omega = \{1, 2, 3, 4\}$, a potência de Ω , ou seja, o conjunto de todos os subconjuntos possíveis de Ω é um exemplo de σ -álgebra. Uma função de pré-medida estabelece uma noção inicial de tamanho para certos subconjuntos de um conjunto universal. É um passo preliminar para a definição de uma medida. Uma função $\mu_0 : \mathcal{A} \rightarrow [0, \infty]$, onde \mathcal{F} uma coleção de subconjuntos de um conjunto Ω , é chamada de pré-medida se:

1. $\mu_0(\emptyset) = 0$.
2. É σ -aditiva em sua coleção de domínio.

Em outras palavras, a pré-medida do conjunto vazio é zero e se A_n é uma sequência de conjuntos mutuamente disjuntos em \mathcal{F} e a união (possivelmente infinita) $\bigcup_{n=1}^{\infty} A_n$ também está em \mathcal{F} , então $\mu_0\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu_0(A_n)$. Exemplificando, seja $\Omega = \{1, 2, 3\}$ e considere a coleção $\mathcal{A} = \{\emptyset, \{1\}, \{2, 3\}, \Omega\}$. Uma pré-medida μ_0 pode ser definida como $\mu_0(\emptyset) = 0$, $\mu_0(\{1\}) = 1$, $\mu_0(\{2, 3\}) = 2$ e $\mu_0(\Omega) = 3$.

Extensão é o processo pelo qual uma função de pré-medida é estendida para se tornar uma medida em uma σ -álgebra completa. O Teorema da Extensão de Carathéodory nos diz que, se uma função μ_0 definida em uma álgebra \mathcal{A} é σ -aditiva, então existe uma única extensão de μ_0 para uma medida μ definida em uma σ -álgebra gerada por \mathcal{A} (ASH; DOLÉANS-DADE, 2000).

Uma vez que temos uma pré-medida e uma σ -álgebra, podemos definir uma medida. Uma medida é uma função que atribui um tamanho a cada conjunto em uma σ -álgebra, estendendo a ideia da pré-medida. A ideia é que, uma vez que tenhamos uma pré-medida em uma coleção inicial de conjuntos, podemos usar o Teorema da Extensão de Carathéodory para estender essa pré-medida para uma medida completa definida em uma σ -álgebra que contém \mathcal{A} .

Uma probabilidade \mathbb{P} é uma função definida sobre a σ -álgebra \mathcal{F} que associa a cada evento A (um conjunto em \mathcal{F}) uma medida pertencente ao conjunto dos Números Reais, de acordo com as seguintes propriedades:

Propriedade A.1 (Complementaridade).

$$\mathbb{P}(A^c) = 1 - \mathbb{P}(A) \quad (A^c \text{ é o conjunto complementar de } A) \quad (\text{A.1})$$

Propriedade A.2 (Limitação).

$$0 \leq \mathbb{P}(A) \leq 1 \quad (\text{A.2})$$

Propriedade A.3 (Monotonicidade).

$$A_1 \subset A_2 \implies \mathbb{P}(A_1) \leq \mathbb{P}(A_2) \quad (\text{A.3})$$

Propriedade A.4 (Não-negatividade).

$$\forall A \in \mathcal{F} \implies \mathbb{P}(A) \geq 0. \quad (\text{A.4})$$

Propriedade A.5 (Totalidade).

$$\mathbb{P}(\Omega) = 1 \quad (\text{A.5})$$

Propriedade A.6 (Aditividade contável).

$$\forall (A_n)_{n=1}^{\infty} \in \mathcal{F} / \bigcap_{n=1}^{\infty} A_n = \emptyset \implies \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n) \quad (\text{A.6})$$

Essas propriedades definem as características fundamentais de uma probabilidade e nos permitem entender e manipular eventos e seus respectivos espaços amostrais de uma maneira matematicamente rigorosa (JAMES, 2002).

Uma **variável aleatória** é uma função que associa a cada resultado de um experimento aleatório (ou evento) um número real. Para ser rigorosamente definida, a variável aleatória deve satisfazer certas propriedades relacionadas à teoria da medida. Vamos detalhar essa definição:

Considere um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$, onde:

- Ω é o espaço amostral, que contém todos os possíveis resultados de um experimento aleatório.
- \mathcal{F} é uma σ -álgebra, que é uma coleção de subconjuntos de Ω , contendo eventos aos quais podemos atribuir probabilidades.
- \mathbb{P} é uma função de probabilidade que associa a cada evento em \mathcal{F} um número real entre 0 e 1.

A variável aleatória X é definida como uma função $X : \Omega \rightarrow \mathbb{R}$, que mapeia cada resultado $\omega \in \Omega$ a um número real $X(\omega) \in \mathbb{R}$, de tal forma que, para todo número real x , $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$. Em outras palavras, o conjunto de resultados no espaço amostral Ω para os quais a variável aleatória X é menor ou igual a um valor real x deve ser um evento em \mathcal{F} . Quando essa condição é satisfeita, dizemos que X é **mensurável** em relação à σ -álgebra \mathcal{F} . A importância dessa definição é que ela garante que podemos atribuir probabilidades aos valores (ou intervalos de valores) que a variável aleatória pode assumir. Por exemplo, a probabilidade de que X assumira valores menores ou iguais a x é dada por $\mathbb{P}(X \leq x)$. Essa probabilidade é bem definida porque o conjunto $\{\omega \in \Omega : X(\omega) \leq x\}$ é um evento em \mathcal{F} devido à mensurabilidade de X . Em resumo, uma variável aleatória é uma maneira de quantificar numericamente os resultados de um experimento aleatório de forma que possamos atribuir probabilidades aos seus valores (JAMES, 2002).

Um conjunto \mathcal{X} é *enumerável* se ele é finito ou se pode ser colocado em correspondência bijetiva com os números naturais \mathbb{N} . Se tal bijeção $f : \mathbb{N} \rightarrow \mathcal{X}$ existe, então \mathcal{X} é denominado *infinito enumerável*. Para essa bijeção, definindo $x_n = f(n)$ para cada $n \in \mathbb{N}$, podemos representar $\mathcal{X} = \{x_1, x_2, \dots\}$. Esta bijeção f é referida como uma enumeração de \mathcal{X} (LIMA, 1999). Uma **variável aleatória discreta** X é definida como uma função que mapeia eventos de um espaço amostral Ω para valores numéricos. Se os possíveis valores de X pertencem a um conjunto enumerável \mathcal{X} , então X é classificada como discreta. Para tal X , há um conjunto $\mathcal{X} = \{x_1, x_2, \dots\} \subset \mathbb{R}$, de modo que para qualquer $\omega \in \Omega$, o valor $X(\omega)$ é igual a x_i para algum $i \in \mathbb{N}$. A função definida por $p(x_i) = \mathbb{P}(X = x_i)$, onde $i = 1, 2, \dots$, é chamada de **função de massa de probabilidade (fmp)** de X , que também é referida por outros nomes como **função de probabilidade discreta** e **função de frequência** (JAMES, 2002).

Alguns exemplos de variáveis aleatórias discretas, junto com suas respectivas distribuições de probabilidade, média, variância e aplicações, são (ROSS, 2014):

1. Distribuição de Bernoulli:

- **PMF:** $P(X = k) = p^k(1 - p)^{1-k}$ para $k = 0, 1$.
- **Média (μ):** p .
- **Variância (σ^2):** $p(1 - p)$.

Exemplo: O resultado de lançar uma moeda uma vez (cara ou coroa).

2. Distribuição Binomial:

- **PMF:** $P(X = k) = \binom{N}{k} p^k(1 - p)^{N-k}$ para $k = 0, 1, 2, \dots, N$.
- **Média (μ):** Np .
- **Variância (σ^2):** $Np(1 - p)$.

Exemplo: O número de caras resultantes ao lançar uma moeda três vezes.

3. Distribuição de Poisson:

- **PMF:** $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$ para $k = 0, 1, 2, \dots$
- **Média (μ):** λ .
- **Variância (σ^2):** λ .

Exemplo: O número de carros que passam por um ponto específico em um determinado intervalo de tempo, como uma hora.

Antes de definirmos as **variáveis aleatórias contínuas**, precisamos definir alguns conceitos de Teoria da Medida e Análise Real.

Uma função $f : \mathbb{R} \rightarrow \mathbb{R}$ é dita *absolutamente contínua* em um intervalo $[a, b]$ se para cada $\varepsilon > 0$, existe um $\delta > 0$ tal que para qualquer coleção finita de subintervalos disjuntos (x_i, y_i) de $[a, b]$ com $\sum (y_i - x_i) < \delta$, temos que $\sum |f(y_i) - f(x_i)| < \varepsilon$. Uma função absolutamente contínua é mais forte que uma função de variação limitada e implica que a função é quase em toda parte diferenciável em $[a, b]$ (LIMA, 1999).

A σ -álgebra de Borel, representada por $\mathcal{B}(\mathcal{X})$, é gerada pelos conjuntos abertos de um espaço topológico \mathcal{X} . No caso da reta real \mathbb{R} , ela inclui todos os intervalos abertos e também formas mais complexas, como intervalos fechados e semiabertos. Especificamente no intervalo $[0, 1]$, a σ -álgebra de Borel é essencial para a probabilidade. Ela fornece a estrutura adequada para definir probabilidades em subconjuntos desse intervalo, que é fundamental para a teoria da probabilidade e processos estocásticos. Os elementos dessa σ -álgebra são chamados de Borelianos (JAMES, 2002).

A *medida de Lebesgue*, denotada por m , é uma extensão da noção usual de comprimento para um subconjunto mais amplo de \mathbb{R} . Ela é definida em uma σ -álgebra \mathcal{L} sobre \mathbb{R} , que contém todos os intervalos abertos de \mathbb{R} e, portanto, todos os conjuntos Borelianos. A medida de

Lebesgue de um intervalo aberto (a, b) é definida pela equação (A.7). Assim, considerando que uma cobertura de um conjunto é uma coleção de subconjuntos cuja união contém o conjunto original, para conjuntos mais gerais em \mathcal{L} , a medida de Lebesgue é definida considerando coberturas por intervalos abertos. Especificamente, a medida de Lebesgue de um conjunto é o *ínfimo* das somas dos comprimentos dos intervalos abertos que fornecem uma cobertura para esse conjunto. Formalmente, para qualquer conjunto $E \subseteq \mathbb{R}$, definimos a medida de Lebesgue pela equação (A.8), onde o ínfimo é tomado sobre todas as coleções de intervalos abertos que cobrem E (AXLER, 2020).

$$m((a, b)) = b - a \quad (\text{A.7})$$

$$m(E) = \inf \left\{ \sum_{i=1}^{\infty} (b_i - a_i) : E \subseteq \bigcup_{i=1}^{\infty} (a_i, b_i) \right\} \quad (\text{A.8})$$

Um conjunto tem medida zero se ele pode ser coberto por intervalos abertos cujos comprimentos totais são arbitrariamente pequenos. Intuitivamente, esses conjuntos são tão pequenos que não têm comprimento no sentido da medida de Lebesgue (JAMES, 2002). Formalmente, dizemos que o conjunto $E \subseteq \mathbb{R}$ é dito ter *medida zero* se, para qualquer $\varepsilon > 0$, existem intervalos abertos $\{(a_i, b_i)\}_{i=1}^{\infty}$ tal que:

1. $E \subseteq \bigcup_{i=1}^{\infty} (a_i, b_i)$
2. $\sum_{i=1}^{\infty} (b_i - a_i) < \varepsilon$

A *integral de Lebesgue* é uma generalização da integral de Riemann, desenvolvida para lidar com certas limitações da integral de Riemann. Enquanto a integral de Riemann se baseia em particionar o domínio de uma função, a integral de Lebesgue se concentra em particionar o conjunto imagem. Dado um espaço mensurável $(\Omega, \mathcal{F}, \mathbb{P})$, onde Ω é o espaço amostral, \mathcal{F} é uma σ -álgebra de subconjuntos de Ω , e \mathbb{P} é uma medida definida em \mathcal{F} , a integral de Lebesgue de uma função mensurável $f : \Omega \rightarrow \mathbb{R}$ é dada pela equação (A.9). Uma das características notáveis da integral de Lebesgue é sua capacidade de integrar funções que possuem descontinuidades em muitos de pontos do domínio, diferentemente da integral de Riemann (AXLER, 2020).

$$\int_{\Omega} f d\mathbb{P} = \int_{\Omega} f(\omega) \mathbb{P}(d\omega) \quad (\text{A.9})$$

Assim, uma **variável aleatória contínua** X é uma variável aleatória que toma valores em um conjunto não enumerável e tem uma função de distribuição $F_X(x)$ que é absolutamente contínua em relação à medida de Lebesgue. Isso significa que existe uma função não negativa $f_X(x)$, chamada **função densidade de probabilidade (fdp)** de X , tal que para todo $x \in \mathbb{R}$ temos F_X definida pela equação (A.10). Se X é absolutamente contínua, então F_X , sendo uma integral de Lebesgue indefinida em f_X , é contínua. Nesse contexto, X tem densidade se, e somente se, F_X é absolutamente contínua. Ou seja, F_X é a integral de sua derivada. Neste caso, $f_X(x) = F_X'(x)$ em todo ponto, exceto num conjunto de medida de Lebesgue nula (diz-se que $f_X = F_X'$ em quase toda parte). A função $f_X(x)$ tem a propriedade de que, para quaisquer dois números a e b com $a < b$, é válida a relação dada pela equação (A.11). Por fim, $f_X(x) \geq 0$ é fdp de X se, e somente se, $\lim_{x \rightarrow \infty} F_X(x) = 1$ (JAMES, 2002).

$$F_X(x) = \int_{-\infty}^x f_X(t) dt \quad (\text{A.10})$$

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(t) dt \quad (\text{A.11})$$

Alguns exemplos de variáveis aleatórias contínuas, junto com suas respectivas distribuições de probabilidade, média, variância e aplicações, são (ROSS, 2014):

1. Distribuição Uniforme Contínua:

- **PDF:** $f(x) = \frac{1}{b-a}$ para $a \leq x \leq b$.
- **Média** (μ): $\frac{a+b}{2}$.
- **Variância** (σ^2): $\frac{(b-a)^2}{12}$.

2. Distribuição Normal (Gaussiana):

- **PDF:** $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$.
- **Média** (μ): μ (parâmetro da distribuição).
- **Variância** (σ^2): σ^2 (parâmetro da distribuição).

3. Distribuição Exponencial:

- **PDF:** $f(x) = \lambda e^{-\lambda x}$ para $x \geq 0$.
- **Média** (μ): $\frac{1}{\lambda}$.
- **Variância** (σ^2): $\frac{1}{\lambda^2}$.

APÊNDICE B – MOMENTOS DE UMA DISTRIBUIÇÃO DE PROBABILIDADE

Os momentos de uma distribuição de probabilidade são medidas estatísticas essenciais que capturam diversas características dessa distribuição. Eles desempenham um papel importante na descrição da forma e das propriedades de uma distribuição, sendo considerados elementos fundamentais no contexto estatístico (CASELLA; BERGER, 2021). O n -ésimo momento de uma distribuição de probabilidade de uma variável aleatória é definido diferentemente para variáveis discretas e contínuas. Para uma variável aleatória discreta X , onde x_i representa os valores possíveis que X pode assumir e $\mathbb{P}(X = x_i)$ são as respectivas probabilidades, o momento é dado pela equação (B.1). Para o caso de uma variável aleatória contínua, onde $f(x)$ é a função densidade de probabilidade, o momento é expresso pela equação (B.2). Vale ressaltar que algumas distribuições de probabilidade podem não possuir todos os momentos, já que estes podem divergir (CASELLA; BERGER, 2021).

$$\mathbb{E}[X^n] = \sum_i x_i^n \mathbb{P}(X = x_i) \quad (\text{B.1})$$

$$\mathbb{E}[X^n] = \int_{-\infty}^{\infty} x^n f(x) dx \quad (\text{B.2})$$

Dentre os momentos, alguns possuem interpretações diretas e importantes:

- **Primeiro Momento (Média ou Esperança Matemática):** Este é o momento de ordem um e é a média de uma distribuição. Matematicamente, é a soma dos produtos de cada valor e sua probabilidade associada. Para uma variável aleatória contínua X com função densidade de probabilidade $f(x)$, a média é dada por $\mu = E[X] = \int_{-\infty}^{\infty} x f(x) dx$. Para uma variável aleatória discreta, a média é $\mu = E[X] = \sum x P(X = x)$, onde $P(X = x)$ é a probabilidade de X assumir o valor x .
- **Segundo Momento (Variância):** Este é o momento de ordem dois e mede a dispersão ou a variação dos valores em relação à média. A variância é o valor esperado do quadrado do desvio de X em relação à sua média, ou seja, $\text{Var}(X) = E[(X - \mu)^2]$.
- **Terceiro Momento (Assimetria):** Este é o momento de ordem três e está relacionado com a assimetria da distribuição. Uma distribuição simétrica tem uma assimetria (ou skewness) de zero. Matematicamente, é definido como $E[(X - \mu)^3]$ e normalmente é padronizado dividindo-se pelo desvio padrão ao cubo.

- **Quarto Momento (Curtose):** Este é o momento de ordem quatro e está relacionado com a "cauda" da distribuição, fornecendo uma medida de "peso" dos outliers. A curtose é dada por $E[(X - \mu)^4]$ e, quando padronizada, compara a cauda da distribuição com a de uma distribuição normal.

APÊNDICE C – FRACTAIS

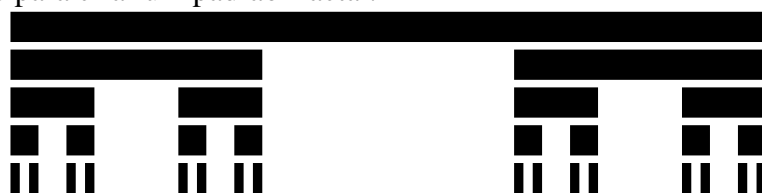
O termo *fractal*, conforme definido em (MANDELBROT, 1975), origina-se do termo latino *fractus*, implicando uma estrutura fraturada. Estas entidades matemáticas manifestam-se tanto em fenômenos naturais quanto em construções matemáticas deliberadas, ou seja, construídos através de processos iterativos ou recursivos. Uma característica distintiva dos fractais é a autossimilaridade, implicando que sua estrutura intrincada se mantém consistentemente independente da escala de observação. Esta propriedade leva à descrição informal de fractais como possuindo complexidade infinita, em contraste com as estruturas geométricas simples. A dimensão de um fractal, frequentemente não-inteira, serve como um indicador quantitativo de sua complexidade e rugosidade, uma divergência notável em relação aos objetos euclidianos padrão que possuem dimensões inteiras claramente definidas. A geometria fractal e a geometria euclidiana distinguem-se não apenas em termos de dimensão, mas também na natureza fundamental dos objetos descritos. A geometria euclidiana, originária dos trabalhos de Euclides articula-se em torno de formas e estruturas regularmente definidas, como linhas retas, polígonos e sólidos. Em contrapartida, a geometria fractal centra-se em estruturas autossimilares com complexidades que desafiam a simplicidade euclidiana. Enquanto os métodos da geometria euclidiana se baseiam em conceitos matemáticos estabelecidos, a abordagem fractal frequentemente incorpora processos iterativos e recursivos. A geometria fractal tem sido instrumental em áreas que examinam fenômenos complexos, abrangendo desde padrões naturais até a teoria do caos e áreas emergentes de análise financeira.

Em (MANDELBROT, 1975) propõe-se que um conjunto F no espaço Euclidiano \mathbb{R}^n é um fractal se apresentar *autossimilaridade*, *dimensão anômala* e *condição de conjunto aberto*. A autossimilaridade é uma das propriedades fundamentais dos fractais.

Matematicamente, a **autossimilaridade** refere-se à propriedade de um objeto ou conjunto ser invariante sob uma determinada transformação de escala. Isso significa que partes do objeto são semelhantes ao objeto inteiro. Para ser matematicamente preciso, seja F um conjunto e S uma transformação de escala. Dizemos que F é autossimilar se, para cada transformação s em S , existe uma parte F_S de F tal que $s(F_S) = F$. Aqui, $s(F_S)$ refere-se à aplicação da transformação de escala s à parte F_S de F . Em outras palavras, ao aplicarmos a transformação s a F_S , obtemos o conjunto F inteiro. É importante notar que a autossimilaridade pode ser exata ou estatística (MANDELBROT, 1975). A autossimilaridade exata refere-se a objetos que são idênticos em todas as escalas. A autossimilaridade estatística refere-se a objetos

que são semelhantes em termos de propriedades estatísticas em diferentes escalas, mas não necessariamente idênticos em forma. Muitos fractais naturais, como paisagens montanhosas ou linhas costeiras, exibem autossimilaridade estatística. Historicamente, a visão predominante era de que as redes complexas não apresentavam autossimilaridade quando submetidas a mudanças de escala. Esta percepção era amplamente baseada na propriedade conhecida como "mundo pequeno" (*small world*), comum a muitas dessas redes. Nessa propriedade, à medida que o diâmetro da rede cresce, o número de nós aumenta em uma taxa exponencial. Essa expansão exponencial parecia contrastar com o crescimento mais previsível de lei de potência visto em estruturas claramente autossimilares. Contudo, (SONG *et al.*, 2005) mostrou que esse era uma hipótese errada. Utilizando-se de uma técnica de renormalização (semelhante à contagem de caixas que veremos adiante), identificou-se que estas redes complexas possuem, de fato, padrões consistentes em todas as escalas. Esse padrão sugere uma autossimilaridade intrínseca à rede. De forma mais específica, a relação entre o número de caixas usadas e o tamanho dessas caixas segue uma lei de potência, resultando na determinação de um expoente específico de autossimilaridade.

Figura 16 – Representação do Conjunto de Cantor em quatro iterações. Cada linha representa uma iteração no processo de construção do conjunto de Cantor, com intervalos centrais removidos progressivamente para criar um padrão fractal.



A **dimensão anômala** é um conceito que surge ao tentar quantificar a complexidade de fractais, que não se encaixam nas definições tradicionais de dimensão. Diferentemente da dimensão euclidiana clássica, que estamos familiarizados, na qual as dimensões são inteiras e são intuitivamente claras, como um ponto tem dimensão 0, uma linha que tem dimensão 1, uma superfície que tem dimensão 2 ou um volume que tem dimensão 3, fractais complicam essa noção porque eles se situam entre essas dimensões inteiras. A dimensão anômala, frequentemente chamada de dimensão fractal, tenta quantificar essa complexidade (BEN-AVRAHAM; HAVLIN, 2000). Uma abordagem comum para definir a dimensão fractal (d) de objetos é o algoritmo da

cobertura ou contagem de caixas. O conceito básico por trás deste método é ver como o número de caixas (coberturas) necessárias para cobrir um objeto muda à medida que o tamanho da caixa é alterado. O algoritmo segue os seguintes passos (BEN-AVRAHAM; HAVLIN, 2000):

1. Escolha uma Escala: Comece escolhendo um tamanho de caixa ε . Estas caixas podem ser quadrados (para objetos 2D) ou cubos (para objetos 3D), ou hiper-cubos para dimensões mais altas;
2. Cubra o Objeto: Use caixas de tamanho ε para cobrir todo o objeto ou conjunto fractal. Isso significa que você coloca as caixas de forma que o objeto esteja completamente contido dentro das caixas, sem se preocupar se algumas caixas contêm apenas uma pequena parte do objeto;
3. Conte as Caixas: Conte o número de caixas $N(\varepsilon)$ necessárias para cobrir o objeto. Note que nem todas as caixas conterão partes do objeto, então só conte as caixas que realmente contêm uma parte do objeto
4. Repita para Diferentes Escalas: Reduza o tamanho da caixa ε e repita os passos 2 e 3. Faça isso para muitos tamanhos de caixa diferentes, gerando um conjunto de pares $(\varepsilon, N(\varepsilon))$;
5. Análise Log-Log: Plote o logaritmo de $N(\varepsilon)$ (eixo vertical) contra o logaritmo de $1/\varepsilon$ (eixo horizontal) em um gráfico. Se o objeto é fractal (ao menos em certa escala), os pontos devem se alinhar aproximadamente em uma linha reta;
6. Calcule a Dimensão Fractal: A dimensão fractal d é dada pelo declive negativo da linha no gráfico log-log. Matematicamente, se $N(\varepsilon) \propto \varepsilon^{-d}$, então d é a dimensão fractal, mais precisamente definido pela equação C.1:

$$d = -\lim_{\varepsilon \rightarrow 0} \frac{\log(N(\varepsilon))}{\log(\varepsilon)} \quad (\text{C.1})$$

A cobertura das caixas em geral é um problema NP-completo e portanto é difícil encontrar com precisão a melhor cobertura para que a contagem de caixas seja ótima em relação a minimização de caixas para um determinado objeto. Existem diversas variações do algoritmo de contagem de caixas, entre eles, (SCHNEIDER *et al.*, 2012) apresenta uma solução determinística próxima da ótima e baseada no princípio da coloração gulosa para a determinação da dimensão fractal em grafos. Além disso, é possível obter uma solução aproximada através de algoritmos estocásticos, como o apresentado em (AKIBA *et al.*, 2016), de forma que torna viável a estimativa da dimensão fractal em grafos densos e grandes.

Para um conjunto F no espaço Euclidiano \mathbb{R}^n , a **condição de conjunto aberto** estabelece que, para algum valor positivo de r , toda bola de raio r com centro em algum ponto de F intersecta o complemento de F em um conjunto de dimensão positiva (LIMA, 2008). Intuitivamente, dado um objeto (que pode ser um fractal) no espaço, escolha qualquer ponto desse objeto e desenhe uma pequena esfera (ou "bola" em terminologia matemática) de raio r em torno desse ponto. A Condição de Conjunto Aberto afirma que, se o objeto é um fractal, então essa esfera irá intersectar partes do espaço que não são parte do fractal. Em outras palavras, a esfera não pode ser totalmente contida dentro do fractal. Além disso, a parte da esfera que está fora do fractal não pode ser apenas um conjunto de pontos insignificantes; ela deve ter uma dimensão positiva, implicando uma substância real ou significativa. Essa condição é importante porque garante que os fractais não são localmente densos. Se você pensar em um fractal como uma estrutura irregular ou porosa, a condição de conjunto aberto é uma maneira de garantir que essa irregularidade ou porosidade seja preservada em todas as escalas, até as mais pequenas.

O **Conjunto de Cantor** é um exemplo clássico de um fractal, definido inicialmente por Georg Cantor em 1883. É construído através de um processo iterativo começando com o segmento de reta unitário $[0, 1]$ e removendo repetidamente partes centrais desse segmento. Após um número infinito de etapas, o que resta é o Conjunto de Cantor, a Figura 16 ilustra o processo construtivo do conjunto de Cantor após quatro iterações de remoção do terço central de cada intervalo presente na iteração anterior. É um conjunto não contável de pontos que tem medida zero, mesmo que não seja um conjunto vazio (LIMA, 1999). O Conjunto de Cantor, apresenta as três propriedades acima mencionadas:

- **Autossimilaridade:** O Conjunto de Cantor é autossimilar porque é invariante sob uma transformação de escala. Cada segmento pode ser transformado no conjunto inteiro através de uma transformação de escala e translação. Isso é evidente no processo iterativo de construção: cada etapa produz segmentos que são uma versão reduzida do conjunto original.
- **Dimensão Anômala:** O Conjunto de Cantor desafia a noção convencional de dimensão. Enquanto um segmento de reta tem uma dimensão de 1 em um espaço euclidiano, o Conjunto de Cantor, apesar de ser um subconjunto dessa reta, tem uma dimensão fractal que é aproximadamente 0,63 (como pode ser mostrado no teorema C.0.1).
- **Condição de Conjunto Aberto:** O Conjunto de Cantor satisfaz a condição de conjunto aberto. Para qualquer ponto no Conjunto de Cantor, se desenharmos uma pequena esfera

(ou intervalo, neste caso) em torno desse ponto, essa esfera intersectará partes do espaço que não fazem parte do Conjunto de Cantor. Isso é evidente pelo fato de que estamos constantemente removendo segmentos (os terços médios) durante a construção.

Teorema C.0.1. *O Conjunto de Cantor apresenta dimensão fractal 0,63*

Demonstração. A dimensão fractal do Conjunto de Cantor é determinada pela relação $N = r^d$. Onde N é o número de segmentos auto-similares, r é a razão pela qual o segmento original é dividido e d é a dimensão fractal (anômala). No Conjunto de Cantor, a cada iteração, o segmento é dividido em 3 partes, e 2 dessas partes permanecem. Portanto, $N = 2$ e $r = 3$. Resolvendo para d , temos que $2 = 3^d$, assim:

$$\log(2) = d \cdot \log(3)$$

$$d = \frac{\log(2)}{\log(3)} \tag{C.2}$$

$$d \approx 0,6309$$

□

Até agora, consideramos apenas fractais isotrópicos, que exibem a mesma propriedade auto-similar em todas as direções. De forma mais geral, um objeto pode possuir uma simetria de dilatação anisotrópica. Nesse caso, falamos em auto-afinidade. Considere um fractal cujo iniciador seja um quadrado unitário. Este é o nosso ponto de partida, chamado de "iniciador". O gerador consiste em subdividir esse quadrado em várias células menores, como se estivéssemos cortando-o em pedaços retangulares. O número desses retângulos é determinado pelas medidas b_1 na horizontal e b_2 na vertical. Em um exemplo específico, dividimos o quadrado em 3 células horizontalmente e 2 verticalmente, como se estivéssemos formando uma grade de 3x2. No entanto, nem todas essas células são mantidas. Algumas são descartadas, e mantemos apenas n delas. Em nosso caso, mantemos 3 desses retângulos e descartamos os outros. Esse formato que criamos não é igual em todas as direções. Se olharmos na direção horizontal (ou direção x), e quisermos ver o padrão original (o quadrado) novamente, precisamos ampliar essas células 3 vezes. Por outro lado, na direção vertical (ou direção y), precisamos ampliar apenas 2 vezes. Esse processo cria uma imagem que parece se repetir, mas de uma forma que muda dependendo da direção em que você olha. É isso que chamamos de objeto auto-afim (BEN-AVRAHAM; HAVLIN, 2000).

Para calcular a dimensão fractal de um objeto auto-afim, temos que considerar as diferentes escalas de dilatação ao longo de cada direção. A equação (C.3) descreve a dimensão

fractal local, $d_{f,local}$, representando a dimensão fractal aparente quando observamos o objeto em escalas de comprimento cada vez menores. Especificamente, ela se baseia na relação entre o número $N(\varepsilon)$ de caixas de tamanho ε necessárias para cobrir o fractal e a relação inversa do tamanho da caixa $1/\varepsilon$ (BEN-AVRAHAM; HAVLIN, 2000).

$$d_{f,local} = \lim_{\varepsilon \rightarrow 0} \frac{\ln N(\varepsilon)}{\ln N(1/\varepsilon)} = \frac{\ln(nb_1/b_2)}{\ln b_1} \quad (C.3)$$

Já a equação (C.4) define a dimensão fractal global, $d_{f,global}$, que é relevante para caixas de dimensão crescente e caracteriza o objeto nas maiores escalas de comprimento. Para fractais auto-similares isotrópicos, ambas as dimensões, local e global, coincidem. No entanto, no contexto de auto-afinidade, essa igualdade não é garantida (BEN-AVRAHAM; HAVLIN, 2000).

$$d_{f,global} = \lim_{\varepsilon \rightarrow \infty} \frac{\ln N(\varepsilon)}{\ln N(1/\varepsilon)} = \frac{\ln(nb_1/b_2)}{\ln b_2} \quad (C.4)$$

Além dessas abordagens, a dimensão fractal de objetos auto-afins também pode ser caracterizada considerando o escalonamento anisotrópico. Em termos simples, o fractal não se expande ou contrai da mesma forma em todas as direções. A equação (C.5) generaliza esse conceito, onde $M(L_x, L_y)$ denota a massa do fractal com dimensões L_x e L_y nas direções x e y , respectivamente. Aqui, as potências $1/d_f^x$ e $1/d_f^y$ descrevem o escalonamento anisotrópico nas direções x e y . No exemplo fornecido, a relação entre as massas em diferentes escalas fornece as dimensões fractais d_f^x e d_f^y em termos dos parâmetros b_1 , b_2 e n (BEN-AVRAHAM; HAVLIN, 2000).

$$M(b^{1/d_f^x} L_x, b^{1/d_f^y} L_y) = b M(L_x, L_y) \quad (C.5)$$

Fractais afins aleatórios também são comuns na natureza. Paisagens montanhosas são um exemplo clássico, mas, de forma mais geral, tendem a aparecer em fenômenos envolvendo superfícies e interfaces. A nova ciência do crescimento superficial se beneficia imensamente desse conceito. Os métodos práticos para medir d_f de fractais isotrópicos aleatórios são facilmente generalizados para fractais afins (BEN-AVRAHAM; HAVLIN, 2000).

APÊNDICE D – GERAÇÃO DE NÚMEROS ALEATÓRIOS SEGUNDO A DISTRIBUIÇÃO Q-EXPONENCIAL

Utilizamos o método da inversa da função de distribuição cumulativa (CDF) para gerar a lista de graus aleatórios dos grafos q-exponencialmente distribuídos. Esse método é uma técnica essencial para a geração de números aleatórios em simulações estatísticas, particularmente quando se deseja amostras de uma distribuição de probabilidade específica. A função de distribuição cumulativa, para uma variável aleatória contínua X , é definida como a probabilidade de X ser menor ou igual a um valor x . Matematicamente, isto é expresso como $F(x) = P(X \leq x)$. Uma propriedade chave da CDF é que ela é uma função não decrescente que varia de 0 a 1 à medida que x varia de $-\infty$ a $+\infty$. O método da inversa da CDF tira proveito do fato de que, se U é uma variável aleatória uniformemente distribuída no intervalo $[0, 1]$, então a variável aleatória $X = F^{-1}(U)$ segue a distribuição com a CDF F . O processo começa com a geração de uma amostra u de uma distribuição uniforme no intervalo $[0, 1]$. Em seguida, aplica-se a inversa da CDF F^{-1} a esta amostra. A função F^{-1} é a inversa da CDF, significando que para cada valor u de U , encontramos um valor x tal que $F(x) = u$ (ROSS, 2022), que para o nosso caso discreto, foi arredondado com a função piso.

Consideremos a distribuição q-exponencial, uma generalização da distribuição exponencial frequentemente usada em sistemas complexos. A CDF da distribuição q-exponencial é dada por $F(x) = [1 - (1 - q)\lambda x]^{1/(1-q)}$ para $x \geq 0$, onde $\lambda > 0$ e q é um parâmetro da distribuição. Para encontrar a inversa desta CDF, resolvemos a equação para x . Dada uma variável aleatória uniforme u no intervalo $[0, 1]$, a inversa da CDF da q-exponencial, $F^{-1}(u)$, é encontrada da seguinte maneira:

1. Primeiro, reorganizamos a equação para isolar x : $[1 - (1 - q)\lambda x]^{1/(1-q)} = u$.
2. Elevamos ambos os lados da equação à potência $(1 - q)$: $1 - (1 - q)\lambda x = u^{1-q}$.
3. Isolamos x : $x = \frac{1 - u^{1-q}}{(1 - q)\lambda}$.

Portanto, a função inversa da CDF da q-exponencial é $F^{-1}(u) = \frac{1 - u^{1-q}}{(1 - q)\lambda}$. Utilizando esta função inversa em uma amostra uniforme u , obtemos uma variável x que segue a distribuição q-exponencial. Para gerar os números uniformemente distribuídos no intervalo $[0, 1]$, foi empregada biblioteca padrão <random>, introduzida no C++11. Esta biblioteca fornece várias funcionalidades para geração de números aleatórios e diferentes tipos de distribuições. Para gerar números uniformemente distribuídos, você pode usar um gerador de números aleatórios (como `std::mt19937`, que é um Motor de Mersenne Twister (CPPREFERENCE.COM, 2023)).