



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

EDSON ARAUJO SOARES

**ORGANIZAÇÃO ESTRUTURAL E GEOMÉTRICA DE REDES: GRAFOS ACÍCLICOS E ÁRVORES
CRÍTICAS**

FORTALEZA

2025

EDSON ARAUJO SOARES

ORGANIZAÇÃO ESTRUTURAL E GEOMÉTRICA DE REDES: GRAFOS ACÍCLICOS E ÁRVORES
CRÍTICAS

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Física da Universidade Federal do Ceará, como requisito parcial para a obtenção do Título de Doutor em Física. Área de Concentração: Física da Matéria Condensada.

Orientador: Prof. Dr. André Auto Moreira.

FORTALEZA
2025

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária

- S185r Soares, Edson Araújo .
 Organização Estrutural e Geométrica de Redes: Grafos Acíclicos e Árvores Críticas /
 Edson Araujo Soares. - 2025.
 86 f. : il. color.
- Tese (doutorado) - Universidade Federal do Ceará, Centro de Ciências, Programa de
 Pós-Graduação em Física, Fortaleza, 2025.
 Orientação: Prof. Dr. André Auto Moreira.
1. árvores; 2. percolação; 3. algoritmo de monte carlo. 4. I. Título.

CDD 530

EDSON ARAUJO SOARES

ORGANIZAÇÃO ESTRUTURAL E GEOMÉTRICA DE REDES: GRAFOS ACÍCLICOS E ÁRVORES
CRÍTICAS

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Física da Universidade Federal do Ceará, como requisito parcial para a obtenção do Título de Doutor em Física. Área de Concentração: Física da Matéria Condensada.

Aprovada em 16 / 07 / 2025.

BANCA EXAMINADORA

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

Prof. Dr. Cesar Ivan Nunes Sampaio Filho
Universidade Federal do Ceará (UFC)

Prof. Dr. Saulo Davi Soares e Reis
Universidade Federal do Ceará (UFC)

Prof. Dr. Diego Araújo Frota
Instituto Federal do Ceará
(IFCE)

Prof. Dr. Rilder de Sousa Pires
Universidade de Fortaleza (UNIFOR)

Dedico, com um enorme sentimento de gratidão, esta realização de um sonho aos meus pais Viola e Dodô, que me trouxeram até aqui contra todas as probabilidades.

AGRADECIMENTOS

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

*“Se eu vi mais longe, foi por estar de pé sobre
ombros de gigantes.”* (Isaac Newton)

RESUMO

Nesta tese, estudamos as propriedades estruturais e geométricas em modelos de grafos aleatórios através de duas investigações interligadas. Primeiramente, estudamos a transição de fase estrutural em grafos acíclicos aleatórios, fornecendo insights fundamentais sobre propriedades topológicas de redes. Explorando a enumeração combinatória exata possível para estruturas acíclicas, derivamos uma expressão para a entropia e determinamos pontos críticos de transição e expoentes críticos. Nossos resultados, validados por simulações de Monte Carlo, estabelecem conexões com a teoria de percolação de campo médio, oferecendo novas perspectivas sobre o regime subcrítico de grafos acíclico aleatórios. Na sequência, investigamos propriedades de escala em aglomerados críticos de percolação através de diversas construções de *spanning trees* (MPT, DFS, RFS, MST) imbuídas nestes aglomerados. Revelamos leis de escala universais que governam a relação entre distâncias químicas e euclidianas, caracterizadas por expoentes críticos que dependem da topologia da árvore. Nossa análise mostra que MPTs refletem dimensões fractais de caminho mínimo, enquanto MSTs correspondem a dimensões de caminho ótimo. Para árvores de DFS e RFS, identificamos comportamentos dependentes da dimensão, incluindo um novo expoente crítico ($\theta_1 = 1.575$) para DFS em $d = 3$. Estabelecemos relações fundamentais entre os expoentes de escala θ_1 , θ_2 , a dimensão fractal d_f e a dimensão de espalhamento d_ℓ , demonstrando que $\theta_1 = 1/\theta_2$ deve valer para consistência de escala. Juntos, estes resultados fornecem insights abrangentes tanto sobre transições estruturais em ensembles de grafos quanto sobre a organização geométrica de aglomerados críticos, com implicações para ciência de redes e transporte em sistemas desordenados.

Palavras-chave: árvores; percolação; algoritmo de monte carlo.

ABSTRACT

In this thesis, we study the structural and geometric properties of random graph models through two interconnected investigations. First, we analyze the structural phase transition in random acyclic graphs, providing fundamental insights into the topological properties of networks. By exploiting the exact combinatorial enumeration possible for acyclic structures, we derive an expression for the entropy and determine critical transition points and critical exponents. Our results, validated by Monte Carlo simulations, establish connections with mean-field percolation theory, offering new perspectives on the subcritical regime of random acyclic graphs. Subsequently, we investigate scaling properties in critical percolation clusters through various constructions of *spanning trees* (MPT, DFS, RFS, MST) embedded in these clusters. We reveal universal scaling laws governing the relationship between chemical and Euclidean distances, characterized by critical exponents dependent on the tree topology. Our analysis shows that MPTs reflect fractal dimensions of the shortest path, while MSTs correspond to dimensions of the optimal path. For DFS and RFS trees, we identify dimension-dependent behaviors, including a new critical exponent ($\theta_1 = 1.575$) for DFS in $d = 3$. We establish fundamental relations among the scaling exponents θ_1 , θ_2 , the fractal dimension d_f , and the spreading dimension d_ℓ , demonstrating that $\theta_1 = 1/\theta_2$ must hold for scaling consistency. Together, these results provide comprehensive insights both into structural transitions in graph ensembles and the geometric organization of critical clusters, with implications for network science and transport in disordered systems.

Keywords: trees; percolation; monte carlo algorithm.

LISTA DE TABELAS

Tabela 1 – Estimativa numérica da dimensão fractal das ramificações da árvore de spanning (θ_1) em 2 e 3 dimensões.	56
--	----

LISTA DE FIGURAS

Figura 1 – Exemplo de Grafo	18
Figura 2 – Grafos Direcionados e Não Direcionados	18
Figura 3 – Distribuição de Conectividade	20
Figura 4 – Rede Regular de ordenação 4.	21
Figura 5 – Redes Regulares de especial interesse	22
Figura 6 – Rede Aleatória	22
Figura 7 – Distribuição típica de conectividade de Redes Aleatórias	23
Figura 8 – Distribuição típica de conectividade de Redes Livre de Escala	25
Figura 9 – Esquema simplificado do Modelo de Configuração	26
Figura 10 – Fragmentação a partir da remoção de vértices	28
Figura 11 – Percolação de sítios em uma rede quadrada	30
Figura 12 – Probabilidade da existência de um cluster infinito em função de $\langle k \rangle$	31
Figura 13 – Conectividade do vértice vizinho	31
Figura 14 – Probabilidade da existência de um cluster infinito	33
Figura 15 – Tamanho médio dos clusters finitos	33
Figura 16 – Probabilidade da existência da componente gigante em função da fração de vértices removidos.	34
Figura 17 – Esquema de um sistema composto por dois grafos.	36
Figura 18 – A fórmula de Cayley.	37
Figura 19 – O processo de Monte Carlo.	43
Figura 20 – Sistema composto pela componente gigante e as componentes restantes . . .	44
Figura 21 – Comparação da Distancia Química e Distância Euclidiana em 2d.	49
Figura 22 – Algoritmo de Leath em uma rede quadrada.	50
Figura 23 – Ilustração de um aglomerado percolante em 2d.	53
Figura 24 – Ilustração prática de um cluster crítico em 2d.	53

LISTA DE GRÁFICOS

Gráfico 1 – Entropia para um sistema composto de duas árvores.	38
Gráfico 2 – Distribuição do tamanho das componentes m_n para diferentes valores de grau médio μ	42
Gráfico 3 – A fração ocupada pela maior componente P_∞ em função do grau médio μ . . .	46
Gráfico 4 – Relações na MPT	52
Gráfico 5 – Número médio de sítios na ℓ -ésima camada	55
Gráfico 6 – Probabilidade condicional $P(r \ell)$ em 2 e 3 dimensões para todas as topologias estudadas (MPT, MST, DFS, FRS).	57
Gráfico 7 – probabilidade condicional $P(r \ell)$ reescalada para todas as topologias estudadas, em 2 e 3 d.	58

SUMÁRIO

1	INTRODUÇÃO	14
2	TEORIA DE REDES COMPLEXAS	17
2.1	Teoria de Grafos	17
2.1.1	<i>Definição formal</i>	17
2.1.2	<i>Definições elementares</i>	17
2.1.3	<i>Conectividade</i>	19
2.2	Modelos de Redes	21
2.2.1	<i>Redes Regulares</i>	21
2.2.2	<i>Redes Aleatórias</i>	22
2.2.3	<i>Redes Livres de Escala</i>	24
2.2.3.1	<i>Modelo de Barabási–Albert</i>	25
2.2.4	<i>Modelo de Configuração</i>	26
3	TEORIA DA PERCOLAÇÃO	28
3.1	Percolação em Redes	29
3.1.1	<i>Componente Gigante e o Critério de Molloy-Reed</i>	30
3.1.2	<i>Expoentes Críticos</i>	32
3.1.3	<i>Percolação Inversa e Robustez</i>	33
4	TRANSIÇÃO DE FASE ESTRUTURAL NO CONJUNTO DE GRAFOS ACÍCLICOS ALEATÓRIOS	35
4.1	Entropia para duas árvores que trocam vértices	35
4.1.1	<i>Entropia e maximização de Configurações</i>	36
4.1.2	<i>Uma nota sobre o paradoxo de Gibbs</i>	38
4.2	Entropia para grafos aleatórios acíclicos	39
4.3	Transição de fase no ensemble dos grafos aleatórios acíclicos	41
4.3.1	<i>A fração ocupada pela componente gigante</i>	44
4.4	Comparação com outros modelos de grafos acíclicos	46
5	RELAÇÕES DE ESCALA E MÉTRICAS DE DISTÂNCIA EM CLUSTERS CRÍTICOS E SPANNING TREES	48
5.1	Métricas de distância em clusters críticos	48
5.2	Crescimento de clusters críticos	49
5.3	Relações de escala em métricas de distância de spanning trees de clusters críticos	50
5.4	As topologias de spanning tree utilizadas: MPT, MST, DFS e RFS.	51
5.5	Relações de escala em métricas de distância de spanning trees de clusters críticos para todas as topologias utilizadas: MPT, MST, DFS e RFS.	52
5.6	Resultados numéricos	55
6	CONCLUSÕES	59
	APÊNDICE A – TEOREMA DA INVERSÃO DE LAGRANGE E A SÉRIE DE TAYLOR DA FUNÇÃO $g(x)e^{-g(x)} = x$	61

REFERÊNCIAS	63
ANEXO A – ARTIGO: <i>ENTROPY OF THE ENSEMBLE OF ACYCLIC GRAPHS</i>	69
ANEXO B – ARTIGO: <i>SCALING LAWS OF CHEMICAL AND EUCLIDEAN DISTANCES IN CRITICAL PERCOLATION TREES</i>	80

1 INTRODUÇÃO

A Ciência de Redes é um campo relativamente recente do conhecimento científico, que se consolidou ao longo das últimas duas décadas. O interesse nessa área de pesquisa é evidenciado pelo volume expressivo de publicações e estudos de diferentes níveis de impacto, o que se justifica por diversos fatores. Dentre eles, destaca-se o caráter interdisciplinar do tema, que atrai pesquisadores de áreas tão diversas quanto física, matemática, biologia e sociologia, entre outras. Cada um desses domínios pode apresentar variações na concepção do que se entende por Ciência de Redes. No entanto, de forma geral, podemos defini-la como o estudo sistemático das redes por meio do método científico [1]. As redes, por sua vez, são sistemas que podem ser representados abstratamente por um grafo, estrutura matemática composta por vértices e arestas. Esses elementos correspondem, respectivamente, às entidades que compõem o sistema e às relações entre elas [2]. Ao longo deste trabalho, aprofundaremos essas definições com maior rigor. A presença das redes é verificada em diversos aspectos da vida cotidiana. Entre os exemplos mais notáveis, podemos citar as redes de comunicação, formadas por celulares e telefones; as malhas de distribuição de energia elétrica; as redes de computadores; as redes sociais, em suas múltiplas dimensões; e as redes de disseminação de doenças infecto-contagiosas [3]. Para esses e muitos outros sistemas relatados na literatura, a teoria de redes oferece um poderoso arcabouço teórico, que permite estudar o comportamento e a evolução dessas estruturas [2, 4].

As ferramentas utilizadas na Ciência de Redes vêm sendo desenvolvidas ao longo dos anos com abordagens analíticas, computacionais e estatísticas [5]. Muitas vezes, essas ferramentas são criadas para representar redes de forma abstrata, focando na definição de métricas e conceitos genéricos, independentemente da natureza específica dos sistemas analisados. Assim, embora a Ciência de Redes tenha ampla aplicabilidade prática, é comum encontrar estudos dedicados majoritariamente aos aspectos estatísticos e métricos das mais variadas estruturas de redes.

Uma ferramenta fundamental para a compreensão do comportamento de redes complexas é a teoria da percolação. Originalmente desenvolvida no contexto da física estatística para descrever fenômenos como a propagação de fluidos em meios porosos e sistemas desordenados [6, 7], essa teoria foi posteriormente estendida para o estudo de redes e sistemas fora do equilíbrio, permitindo a análise da conectividade e da robustez estrutural de diversos sistemas [8]. A percolação em redes pode ser entendida como um processo no qual vértices ou arestas são removidos aleatoriamente, afetando a conectividade global do sistema. A partir de um certo limiar crítico, conhecido como limiar de percolação, a rede sofre uma transição de fase geométrica, deixando de ser globalmente conectada e fragmentando-se em componentes menores [9]. Esse fenômeno é essencial para a compreensão de diversos processos, como a resiliência de infraestruturas, a disseminação de epidemias e a propagação de informações em redes sociais [10, 11].

Para evidenciar a importância e aplicabilidade da teoria da percolação no contexto de redes, podemos citar as redes tecnológicas, como redes de computadores ou de distribuição de energia. A teoria da percolação permite avaliar a robustez frente a falhas ou ataques direcio-

nados. Se um número suficiente de nós ou conexões for removido, a infraestrutura pode colapsar, comprometendo sua funcionalidade. Esse tipo de análise é crucial para o desenvolvimento de estratégias de proteção e otimização desses sistemas [12]. No contexto das redes biológicas, a percolação desempenha um papel central na modelagem da disseminação de doenças. Modelos epidemiológicos baseados em redes utilizam conceitos de percolação para prever surtos e identificar estratégias eficazes de contenção, como a vacinação direcionada para indivíduos altamente conectados [13]. Já em redes sociais, a teoria da percolação ajuda a descrever a propagação de informações, rumores e inovações. A existência de um limiar de percolação significa que, abaixo de um certo nível de conectividade ou engajamento, a informação se espalha de maneira limitada. Acima desse limiar, porém, ocorre uma difusão ampla, alcançando grande parte da rede [14].

A teoria da percolação desempenha também um papel crucial na compreensão de fenômenos de transporte em meios porosos, especialmente quando modelados por redes regulares, como as redes quadradas (2D) e cúbicas (3D). Nesses modelos, os vértices representam pontos do meio, enquanto as arestas simbolizam possíveis caminhos para o fluxo de fluidos. A análise da percolação nesses sistemas permite determinar o limiar crítico no qual uma substância consegue atravessar o material de maneira contínua [8, 15]. Uma aplicação prática significativa da teoria da percolação em redes regulares é a análise da permeabilidade de rochas em reservatórios de petróleo. Nesses casos, as rochas podem ser modeladas como redes cúbicas tridimensionais, onde os poros e canais internos são representados por arestas conectando os vértices. A percolação é utilizada para prever a porosidade crítica necessária para que o óleo ou gás se mova através da rocha de forma eficiente. Estudos indicam que, abaixo desse limiar de percolação, o fluido não consegue formar um caminho contínuo, impossibilitando a extração eficaz dos recursos [16].

Através de dois estudos inter-relacionados, esta tese investiga aspectos estruturais e estatísticos de redes complexas sob a ótica da teoria da percolação. Embora abordem sistemas distintos — grafos aleatórios acíclicos e redes regulares —, ambos os estudos compartilham a mesma motivação central: compreender como a estrutura topológica influencia fenômenos críticos de conectividade. O primeiro estudo se concentra na enumeração exata e na caracterização estatística de grafos acíclicos, com o objetivo de definir uma função de entropia associada à fragmentação da rede. O segundo estudo explora relações de escala emergentes em árvores críticas geradas em clusters percolantes, investigando propriedades universais relacionadas à geometria fractal desses sistemas. Os dois enfoques se complementam ao revelar, sob diferentes perspectivas, a natureza das transições de fase em redes, constituindo um corpo coerente de investigação.

Retomando a discussão sobre redes, visando introduzir o primeiro estudo realizado, lembramos que redes são representadas de modo abstrato por uma estrutura matemática denominada *Grafo*, definida como um conjunto finito e não vazio $G(N, E)$ de N vértices e E arestas [2, 17, 18]. Dentre os modelos de redes, o modelo de grafos aleatórios é um dos mais simples e vem sendo amplamente estudado desde sua introdução por Erdős–Rényi [19–21] com especial interesse em descrever as propriedades estatísticas relacionadas ao fenômeno da percolação nestes grafos, como o surgimento de componentes gigantes e a distribuição de tamanho de componentes. Na primeira seção de resultados deste trabalho (capítulo IV), estamos interessados

no ensemble de grafos aleatórios acíclicos, ou seja, grafos nos quais não há ocorrência de loops (um vértice conectado a si mesmo) ou ciclos (um conjunto de arestas que estabeleçam um caminho entre um vértice e ele mesmo). Isso implica dizer que esse tipo de grafo é caracterizado pela existência de no máximo um único caminho possível entre qualquer par de nós pertencentes ao mesmo grafo. As propriedades estatísticas de grafos aleatórios, mencionadas anteriormente, já foram acessadas com sucesso utilizando métodos como o critério de Molloy-Reed [22] e o método das Funções Geradoras [23]. Métodos nos quais se assume que, antes da condição crítica na qual ocorre a formação de uma componente gigante, não existam ciclos ou loops na rede. Todos estes conceitos serão apresentados no decorrer deste texto. Nessa seção de resultados (capítulo IV), demonstramos que a restrição a grafos acíclicos nos proporciona uma maneira de encontrar uma enumeração exata para o ensemble, e utilizamos esses resultados para mostrar como definir exatamente uma entropia para este ensemble. Nós definimos nossa entropia como uma função da distribuição do tamanho das componentes, a partir da enumeração exata do ensemble. Para isso, manusearemos a ideia de grafos acíclicos aleatórios para elucidar os métodos da Mecânica Estatística, de modo didaticamente análogo ao tratamento de um sistema termodinâmico, através dos métodos estatísticos de contagem de estados acessíveis, distribuição de probabilidades, equilíbrio termodinâmico via maximização da entropia e transição de fase, este último associado à teoria de percolação.

Na segunda seção de resultados (capítulo V), trataremos de redes regulares. Para este caso, o modelo de percolação consiste em escolher aleatoriamente vértices (ou arestas) para estarem ativos ou inativos com uma probabilidade P [24, 25]. Para uma probabilidade crítica P_c , ocorre uma transição de fase topológica de segunda ordem, na qual surge um cluster percolante [24, 26]. Na criticidade, um conjunto de leis de escala e expoentes críticos caracteriza os clusters de percolação e sua geometria fractal [24, 27–29]. Uma medida topológica de particular interesse é o menor caminho ℓ entre dois nós no cluster crítico. Nesse capítulo consideraremos a probabilidade condicional $P(r|\ell)$, que expressa a probabilidade de um vértice com uma determinada distância química ℓ estar a uma distância euclidiana r . Determinaremos o comportamento de escala, em 2 e 3 dimensões, dessa função de distribuição para árvores percolantes críticas com topologias especiais. Nosso principal objetivo é explorar como a topologia dessas estruturas influencia seu comportamento de escala e determinar as propriedades universais que emergem no limiar crítico. Utilizamos uma combinação de técnicas analíticas e simulações numéricas para caracterizar esse comportamento de escala.

O restante deste texto está organizado de modo que faremos uma breve revisão teórica acerca dos aspectos e conceitos mais importantes da teoria de redes, bem como uma introdução aos principais modelos de redes, ambos no capítulo II; no capítulo III introduzimos a teoria da percolação, que um aspecto fundamental para um sólido estudo da evolução dinâmica de redes; já no capítulo IV, apresentamos as metodologias e resultados da seção de resultados referente à entropia e transição de fase estrutural no ensemble de grafos acíclicos; no capítulo V apresentamos as metodologias e resultados da segunda seção de resultados, que trata das relações de escala em árvores críticas em redes regulares para diferentes topologias e, no capítulo VI, finalmente elencamos as conclusões, os avanços e perspectivas proporcionadas por este trabalho.

2 TEORIA DE REDES COMPLEXAS

Uma Rede Complexa consiste, essencialmente, em um conjunto de vértices e arestas que reproduzem a topologia e a dinâmica de sistemas das mais diversas naturezas [5,30–32], onde os vértices e arestas representam respectivamente os entes que compõem o sistema e a interação aos pares de cada um desses entes [33]. A Ciência de Redes, em acordo com o que define Lewis (2009), se ocupa do estudo da estrutura, comportamento dinâmico e aplicações de redes em vários ramos do conhecimento [1], por isso, ao introduzir a ideia de Redes Complexas, traremos neste capítulo os conceitos, propriedades e notações mais importantes para o entendimento e familiarização com esta teoria e, principalmente, para o desenvolvimento deste trabalho, incluindo uma revisão sobre Teoria de Grafos, métricas fundamentais envolvendo a caracterização da conectividade, além de alguns dos principais modelos de redes.

2.1 Teoria de Grafos

A Teoria de Redes tem sua fundação matemática baseada na Teoria dos Grafos, o estudo de estruturas matemáticas, chamadas *grafos*, compostas por *vértices* que se conectam das mais diversas maneiras por meio de *arestas*.

2.1.1 Definição formal

Um Grafo G é definido como um conjunto finito, não vazio de V vértices e E arestas [2, 17, 18]. O que se tem na literatura, é que a ideia de usar estruturas desse tipo para resolver problemas matemáticos foi utilizada pela primeira vez em 1736, quando Leonhardt Euler (1707–1783) resolveu com estes métodos o famoso problema das sete pontes de Königsberg [34]. Assim como supostamente teria sido feito por Euler, um grafo é utilizado para representar a topologia de certo objetos, como já dito, por meio de vértices (ou nós e sítios)¹ e arestas (ou links e conexões)². A Fig. (1) caracteriza um grafo genérico a título de conhecimento.

2.1.2 Definições elementares

É amplo o campo de pesquisa que se desenvolve na Teoria dos Grafos e, portanto, são numerosas as definições e teoremas relacionados a ela. Vamos listar alguns dos mais relevantes para uma boa introdução.

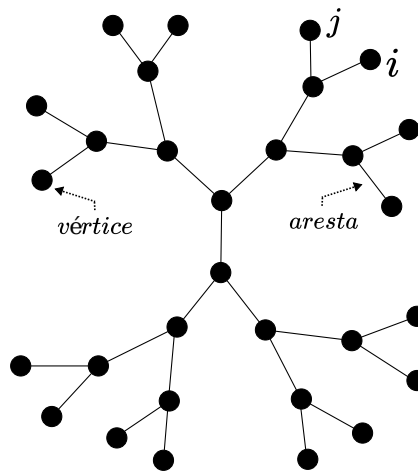
- **Ordem e Tamanho:** A ordem de um grafo é dada pelo somatório do número de vértices $N = |V|$, enquanto o tamanho S é definido como o número de arestas presentes na estrutura³ ($S = |E|$).
- **Adjacência:** Dois vértices i, j que pertencem ao grafo são adjacentes ou vizinhos se existe uma aresta que conecta diretamente i e j .

¹Utilizaremos estes termos como sinônimos ao longo deste trabalho.

²Utilizaremos estes termos como sinônimos ao longo deste trabalho.

³Embora essa definição seja utilizada em Teoria dos Grafos, na Teoria de Redes o tamanho se refere ao somatório do número de vértices presentes na rede.

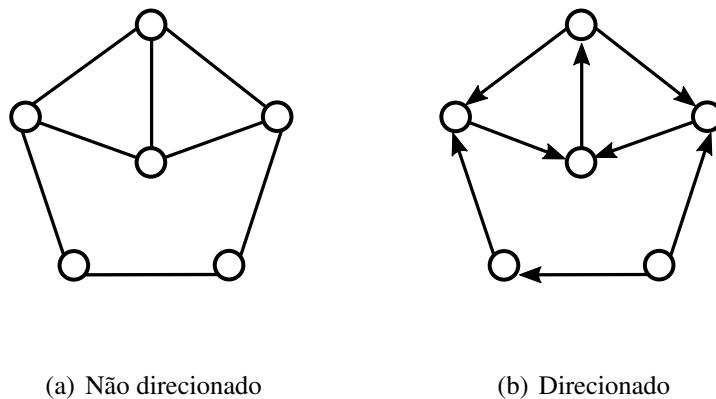
Figura 1: Exemplo de Grafo



Fonte: Produzido pelo autor. Representação de um grafo. Os vértices são ligados através de aresta, por exemplo uma aresta $\{i, j\}$ liga um par i e j de vértices.

- **Grafos não Direcionados:** Um grafo $G(V, E)$ é dito não direcionado quando o conjunto de arestas não possui direção, ou seja, não há informação acerca do fato de qual vértice é a fonte e qual é o alvo das arestas. Note que, desta forma, não há distinção entre uma aresta $\{i, j\}$ e outra $\{j, i\}$.
- **Grafos Direcionados:** Similar ao caso anterior, com a clara exceção do fato das arestas serem direcionadas, ou seja, a aresta $\{i, j\}$ parte do vértice i rumo ao vértice j e é distinguível da aresta $\{j, i\}$. Veja uma comparação entre os dois casos na Fig. (2).

Figura 2: Grafos Direcionados e Não Direcionados



Fonte: Produzido pelo autor. (a) mostra o exemplo de um grafo não direcionado, em detrimento a Fig. (b) que exemplifica um grafo direcionado, com a direção das ligações indicadas pela seta.

- **Grau de conectividade k_i :** é a propriedade que carrega a informação acerca de quantas arestas incidem sobre o vértice i , ou seja, a quantos vértices este estabelece conexão. Em grafos direcionados, esta propriedade se ramifica em grau de saída (k_{out}) e grau de entrada (k_{in}), sendo (k_{out}) o somatório de todas as arestas que tem origem no referido vértice, e (k_{in}) o somatório de todas as arestas que tem este vértice como alvo.

- **Subgrafo:** Um subgrafo de $G(V, E)$, é um grafo $G'(V', E')$ no qual $V' \subset V$ e $E' \subset E$.
- **Caminho:** Um caminho é uma sequência de arestas do tipo $\{\{i_0, i_1\}, \{i_1, i_2\}, \dots, \{i_{l-1}, i_l\}\}$ que conecta respectivamente os vértices $\{(i_0, i_1), (i_1, i_2), \dots, (i_{l-1}, i_l)\}$. Diz-se que o caminho vai do vértice i_0 até i_l e vice versa em grafos não direcionados e, em grafo direcionados (tendo todas as arestas a mesma direção), vai somente do vértice i_0 até i_l .
- **Ciclo:** Trata-se de um caminho fechado no grafo, ou seja, uma sequência de arestas $\{\{i_0, i_1\}, \{i_1, i_2\}, \dots, \{i_{l-1}, i_l\}\}$, na qual $i_0 = i_l$.
- **Loop:** Caracterizado no caso onde temos um vértice ligado diretamente à ele mesmo, ou seja, a existência de uma aresta $\{i, i\}$.
- **Aresta Múltipla:** A ocorrência de mais de uma aresta conectando diretamente o mesmo par de vértices.
- **Grafo Conectado:** Um grafo não direcionado é dito conectado quando existe um caminho entre qualquer par de vértices i e j . Similarmente, um grafo direcionado é dito fortemente conectado quando para qualquer par de vértices i e j , existe um caminho direcionado entre i e j bem como entre j e i .
- **Árvores:** Uma classe de grafos de especial relevância na teoria de grafos são as chamadas *Árvores*, que podemos definir com o sendo um Grafo $G(V, E)$ caracterizado pelas seguintes propriedades [35]:
 1. Não há ciclos ou loops no grafo, ou seja, trata-se de um grafo acíclico,
 2. É um grafo conectado,
 3. Existe um e apenas um caminho que conecta quaisquer pares de vértices no grafo,
 4. O número arestas do grafo é dado por $E = N - 1$, onde E é número de arestas e N é o número de vértices.

As árvores apresentam mais uma característica importante para a estrutura, que as classificam em *árvores enraizadas* e *não enraizadas*. O primeiro caso significa que a árvore é direcionada, e que exatamente um e somente um vértice tem grau de conectividade de saída igual a zero. Este vértice que não é origem de nenhuma aresta é chamado de *raiz* [36]. Os vértices com grau de entrada igual a zero são chamados de *folhas*, e cada vértice, exceto a própria raiz, tem grau de saída exatamente igual a 1. Discutiremos mais a respeito das árvores no decorrer do trabalho.

 5. **Florestas:** Uma floresta é um grafo composto por um conjunto de árvores, ou seja, cada componente (ou cluster, ou agregado) é uma árvore. Caso as árvores sejam enraizadas, a floresta também é dita enraizada.

2.1.3 Conectividade

O grau de conectividade k é uma característica de cada vértice da rede. Uma visão mais ampla acerca dessa informação pode ser dada por propriedades mais gerais, como por

exemplo o grau de conectividade média $\langle k \rangle$, dada pela razão entre o somatório de graus de conectividade de cada vértice pelo número N de vértices da rede, conforme a Eq. (2.1),

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i = \frac{2L}{N}. \quad (2.1)$$

Onde L é o número de arestas da rede⁴.

Para redes direcionadas faz-se necessário a distinção entre conectividade média de entrada e de saída, e a conectividade total da rede é dada pela Eq. (2.2),

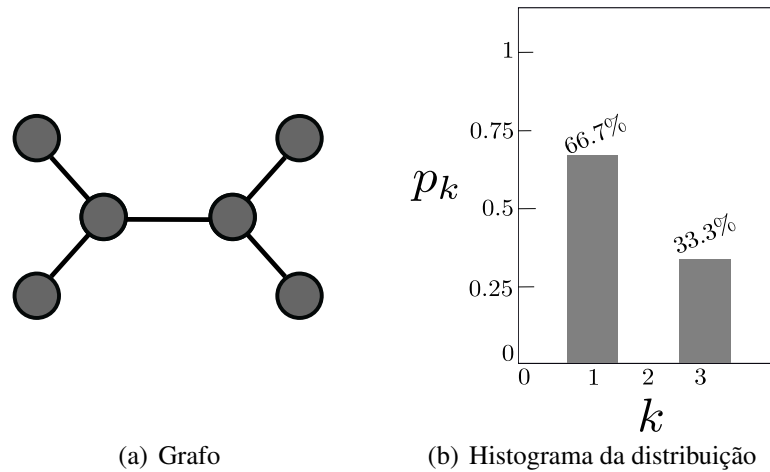
$$L = \sum_{i=1}^N K_i^{in} + \sum_{i=1}^N K_i^{out}. \quad (2.2)$$

Outra informação geral a respeito de conectividade da rede é a *distribuição de conectividade*, comumente denotada por $P(k)$. Formalmente, essa distribuição $P(k)$ é definida como a probabilidade de que um vértice pertencente ao conjunto de vértices da rede, aleatoriamente escolhido, tenha exatamente grau k de conectividade [2, 5, 30, 37]. Matematicamente podemos expressar essa informação na forma da Eq. (2.3),

$$P_k = \frac{N_k}{N}. \quad (2.3)$$

Onde N_k é o número de vértices com conectividade k , e é satisfeita a condição de normalização. Veja na Fig. (3) um exemplo prático da captura dessa propriedade.

Figura 3: Distribuição de Conectividade



Fonte: Produzido pelo autor. O grafo mostrado em (a) possui 6 vértices, dos quais 4 têm grau de conectividade igual a 1 e 2 tem grau de conectividade 3. Em (b) essas informações são utilizadas para caracterizar a respectiva distribuição de conectividade, sendo a probabilidade de aleatoriamente escolher, dentre os 6, um vértice de conectividade 1 é de $4/6 = 66,7\%$, e a de escolher um vértice de conectividade 3 é de $2/6 = 33,3\%$. todos os outros valores de conectividade possuem probabilidade nula.

⁴Observe que o fator 2 surge devido ao fato de uma aresta aumentar em uma unidade o grau de conectividade de um par de vértices

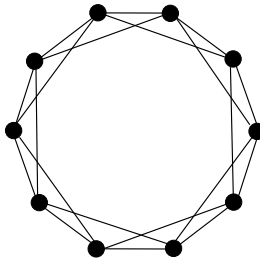
2.2 Modelos de Redes

As mais variadas formas de distribuição de conectividade dão origem a propriedades distintas às redes, e caracterizam os vários modelos existentes, cada um com suas particularidades, que se adéquam aos problemas, sendo uma mais representativa para certos problemas do que outros. Vamos apresentar adiante alguns dos principais modelos existentes, bem como os aspectos gerais das suas propriedades e aplicabilidades. Em tempo, mostramos também uma alternativa de como construir redes de cada uma dessas classes.

2.2.1 Redes Regulares

Como sugere a própria nomenclatura, o modelo de Redes Regulares é caracterizado por uma distribuição de conectividade que evidencia um padrão repetitivo [1], ou seja, um padrão regular no grau de conectividade de cada vértice, tendo cada um exatamente o mesmo valor de k vértices adjacentes ou exatamente o mesmo grau de conectividade. Veja na Fig. (4) um exemplar de uma rede regular onde cada vértice tem grau de conectividade $k = 4$.

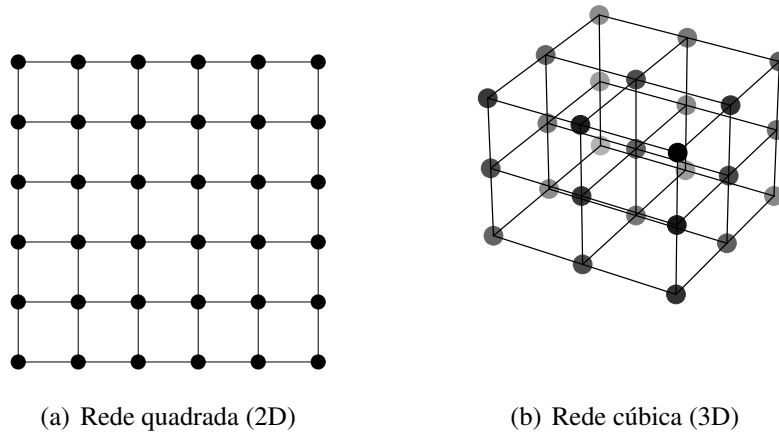
Figura 4: Rede Regular de ordenação 4.



Fonte: Produzido pelo autor. A estrutura possui dez vértices e cada um deles apresenta o mesmo padrão de conectividade $k = 4$.

São vários os exemplos de redes regulares que podemos apresentar em adição ao da Fig. (4). Veja na Fig. (5) exemplos ilustrativos de duas redes com estruturas regulares de especial interesse pela sua aplicação em modelos e problemas práticos: a rede quadrada (2D) na Fig. (5).a e a rede cúbica (3D) na Fig. (5).b.

Figura 5: Redes Regulares de especial interesse



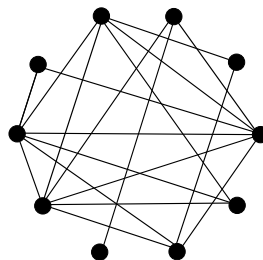
Fonte: Produzido pelo autor.

2.2.2 Redes Aleatórias

Em completa oposição às redes regulares, as Redes Aleatórias são caracterizadas por uma distribuição de conectividade aleatória [38], traduzindo o fato de que cada par de vértice pode ser conectado com uma probabilidade p [30]. Esta classe de redes foi introduzida inicialmente por Pál Erdős (1913-1996) e Alfréd Rényi (1921-1970) [20], aos quais é atribuído o nome do modelo (*Modelo Erdős–Rényi (ER)*). O modelo de **ER** produz um ensemble de redes aleatórias, tendo alguns parâmetros fixos, como veremos adiante nas duas variantes do modelo:

1. O modelo $G(N, E)$, que consiste em um conjunto fixo de N vértices e E arestas, no qual as arestas são aleatoriamente dispostas entre pares de vértices [30].
2. Já o modelo $G(N, p)$ consiste inicialmente apenas de um conjunto de N vértices, cada par de vértices será conectado com uma probabilidade p . Esta variante possui o atributo positivo de não fixar o número de arestas na estrutura, o que o torna mais factível no que se refere à representatividade de redes reais, uma vez que, em redes reais, raramente o número de arestas é fixo [30]. Veja na Fig. (6) uma representação de uma rede aleatória, e compare com o da Fig. (4).

Figura 6: Rede Aleatória



Fonte: Produzido pelo autor. O aspecto desordenado da disposição das arestas remete visualmente ao aspecto da aleatoriedade inerente ao modelo aleatório, isso em completa oposição ao grafo da Fig. (4), onde a regularidade na disposição das arestas é predominante. Este aspecto torna visível a diferença entre os dois modelos.

A distribuição de conectividade p_k em redes aleatórias é capturada por uma distribuição binomial, na forma da equação 2.4,

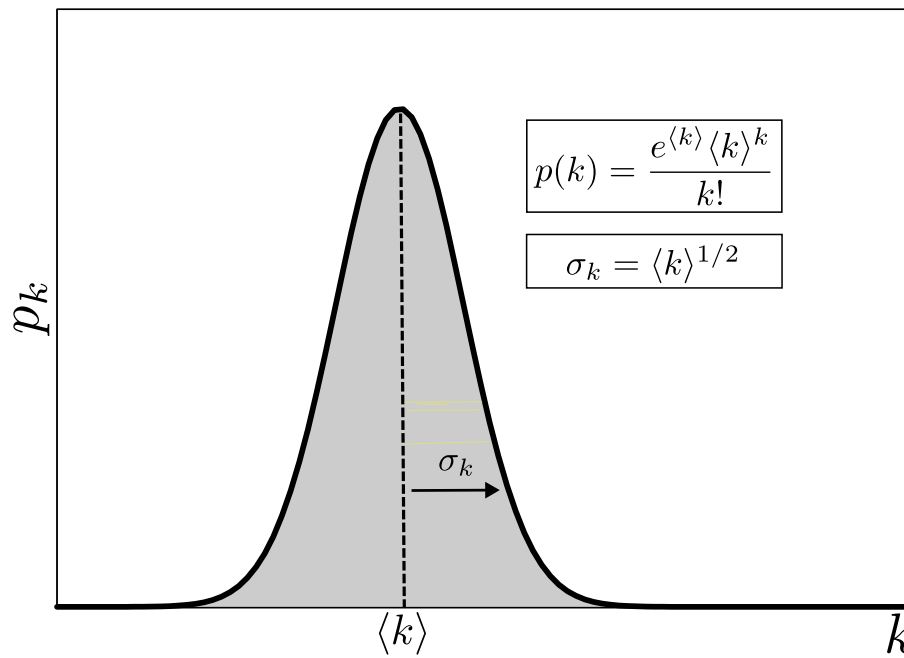
$$p_k = \binom{N-1}{k} p^k (1-p)^{N-1-k}. \quad (2.4)$$

Que depende do tamanho da rede (N). No entanto, desde que N seja suficientemente grande⁵, a distribuição é bem aproximada por uma *Distribuição de Poisson* [23], como na equação (2.5),

$$p(k) = \frac{e^{-\langle k \rangle} \langle k \rangle^k}{k!}. \quad (2.5)$$

Que depende diretamente apenas do grau médio de conectividade $\langle k \rangle$. A distribuição de conectividade escrita na forma da Eq. (2.5) traz consigo uma série de simplificações analíticas, que nos permitem acessar com maior facilidade informações e propriedades estatísticas acerca da rede. Por isso, sempre que necessário no decorrer deste trabalho nos referiremos a segunda forma para p_k como sendo a distribuição de conectividade das redes aleatórias. A Fig. (7) exemplifica o comportamento da distribuição de conectividade para redes aleatórias e sumariza suas propriedades mais importantes.

Figura 7: Distribuição típica de conectividade de Redes Aleatórias



Fonte: Produzido pelo autor. A figura mostra o formato típico da curva para a distribuição conforme a Eq. (2.5). Podemos observar que para Redes Aleatórias a probabilidade de encontramos vértices com graus de conectividade distantes da conectividade média $\langle k \rangle$ vai a zero rapidamente, revelando o fato de que nesse tipo de distribuição a maioria dos graus de conectividades tem valores próximos ao valor médio e, portanto, não prevê a existência de vértices altamente conectados. Formalmente isso é capturado pela variância da distribuição ($\sigma_k = \langle k \rangle^{1/2}$), que carrega a informação acerca de o quão dispersos são valores da distribuição em relação ao valor médio, ou seja, os valores devem ser distribuídos ao longo do intervalo $\langle k \rangle \pm \sigma_k$.

Embora o modelo aleatório seja amplamente explorado no campo da Teoria de Re-

⁵Disse-se: desde que N esteja no limite termodinâmico. Na maioria das aplicações de interesse essa condição é válida.

des, este modelo de fato não consegue reproduzir a complexidade da maioria dos problemas reais. Ao contrário do que acontece nas redes aleatórias, os modelos que melhor capturam a distribuição de conectividade de dados reais apresentam a característica de possuírem vértices com grau de conectividade muito distantes do valor médio $\langle k \rangle$, ou seja, uma dispersão maior do que de fato exhibe o modelo aleatório, isso nos leva ao nosso terceiro modelo.

2.2.3 Redes Livres de Escala

Como mencionado na seção anterior, apesar da sua ampla exploração no campo da Teoria de Redes, as redes aleatórias em geral não conseguem reproduzir as propriedades de redes reais. Isto foi reportado pela primeira vez em 1999 por Hawoong, em colaboração com Réka Albert e Albert-László Barabási, no artigo intitulado “*The Diameter of the World Wide Web*” (O Diâmetro da World Wide Web) [39]⁶. As Redes Livres de Escala, surgem justamente a partir dessa ineficácia do modelo aleatório no que se refere à capacidade de representar bem as configurações de redes reais. São numerosos os sistemas reais cujas estruturas vem sendo modeladas por redes do tipo Livre de Escala, como por exemplo, a colaboração entre atores em filmes, redes de malhas energéticas, redes de citações em artigos científicos, etc [40].

As Redes Livres de escala seguem distribuições de conectividade do tipo *Lei de Potência* [30,41], na forma da Eq. (2.6),

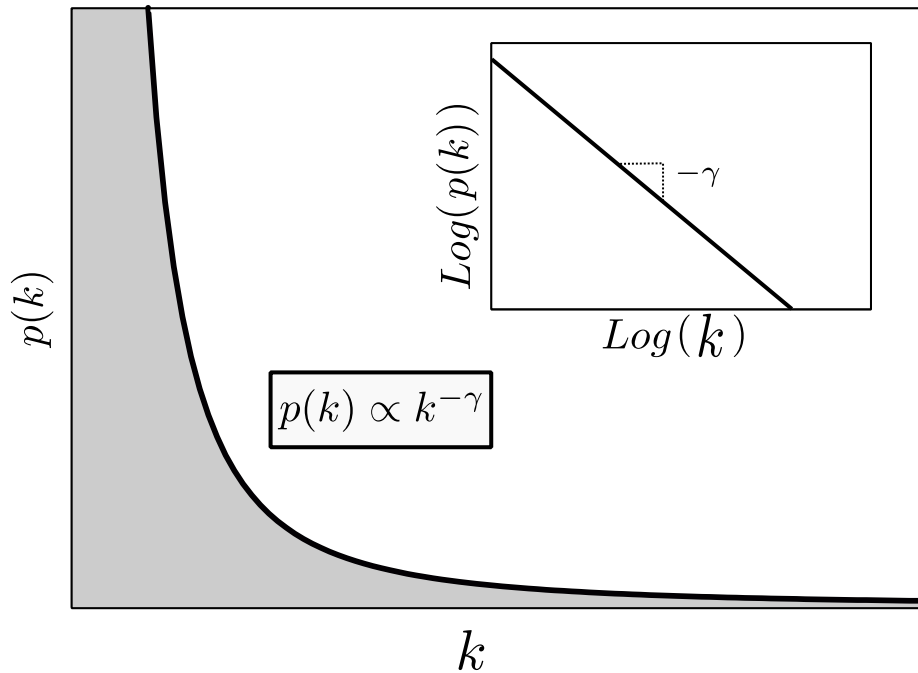
$$p_k = Ck^{-\gamma}. \quad (2.6)$$

Onde C é uma constante a ser determinada pela condição de normalização e γ é o expoente característico da distribuição chamado *parâmetro de escala*, que carrega a informação acerca da inclinação da reta típica da distribuição de conectividade, quando exibida em escala logarítmica, veja este comportamento na Fig. (8). Este fator traduz, de acordo com seu valor, características estatísticas e estruturais fundamentais de tais redes [30,42]. Para um regime específicos ($2 < \gamma < 3$) do expoente γ , as flutuações em torno do valor da conectividade média divergem⁷ [42], o que de fato confere a propriedade livre de escala, que dá nome ao modelo. Essa grande flutuação em torno da conectividade média $\langle k \rangle$ é o que leva a uma das principais características de redes livres de escala, a existência de vértices altamente conectados, em relação ao valor médio e aos vértices com menor grau de conectividade, ou seja, verifica-se nesse tipo de rede a existência de uma pequena quantidade de vértices com um número muito grande de conexões, em compensação, muitos vértices apresentam graus de conectividade muito baixos. Estes vértices altamente conectados são chamados de *Hubs* e desempenham um papel fundamental na estrutura da rede. A existência dos Hubs confere às redes livres de escala uma alta resistência à falhas aleatórias, em contrapartida, estas são vulneráveis à remoções direcionadas aos vértices mais conectados [41,43]. Apresentamos a seguir um dos mais difundidos modelos para gerar ensembles de redes livre de escala.

⁶A WWW (World Wide Web) é uma rede real onde os vértices são documentos e os links são URLs, que permitem acessar um documento a partir de outro através de um click [30]. Para uma rede dessa natureza, Hawoong Jeong mapeou o domínio nd.edu [30], tal domínio constava de 325.729 documentos e 1.469.680 links. Este trabalho reproduziu a topologia dessa rede e, em resumo, percebeu-se que existe uma falha no ajuste da distribuição de conectividade da rede referente à uma distribuição normal, enquanto uma curva em lei de potência conseguia reproduzir bem o comportamento da distribuição.

⁷Considerando o limite $N \rightarrow \infty$. A variância diverge ($\sigma_k \rightarrow \infty$).

Figura 8: Distribuição típica de conectividade de Redes Livre de Escala



Fonte: Produzido pelo autor. Comportamento típico da distribuição de conectividade de Redes Livre de Escala ($p_k \propto k^{-\gamma}$), ainda que de modo ilustrativo, podemos verificar que esta distribuição prevê a existência de muitos vértices com baixo grau de conectividade, e também uma pequena quantidade de vértices altamente conectados (ambos em relação ao valor médio $\langle k \rangle$), os Hubs. Em Escala logarítmica, o expoente γ fornece a inclinação da reta que representa a distribuição.

2.2.3.1 Modelo de Barabási–Albert

O Modelo de Barabási–Albert (Modelo BA ou Modelo Livre de Escala) é um algoritmo capaz de gerar redes livres de escala. Este método se baseia em duas características:

- **Crescimento:** Ao contrário das variantes do modelo aleatório, no modelo BA o número de vértices não é fixo, mas sim aumenta gradativamente durante a execução do algoritmo com a criação de novos vértices.
- **Ligação Preferencial:** Em mais uma distinção em relação aos modelos que tem como produto uma rede aleatória, a distribuição de arestas não é aleatória, sendo mais provável que um vértice mais conectado receba uma nova aresta, ou seja, os novos vértices adicionado à rede se ligam preferencialmente à vértices com maior grau de conectividade.

Essas características são intuitivamente relacionados ao crescimento de uma rede real, como podemos verificar em vários exemplos (veja o capítulo 5 da referência ([30])). O algoritmo do modelo BA segue a seguintes etapas[30, 40]:

1. A rede é iniciada com m_0 vértices que são arbitrariamente conectados entre si, cada vértice tem grau de conectividade não nula.
2. A cada passo i é criado um novo vértice v_i com grau de conectividade k_{v_i} menor ou igual ao número de vértices presentes no passo anterior.

3. O novo vértice v_i é conectado à k_{v_i} vértices pré-existent na estrutura de modo preferencial aos vértices mais conectados da rede, ou seja, a probabilidade do novo vértice i estabelecer uma conexão com um dado vértice pré-existente j é proporcional ao grau de conectividade k_j de j , na forma

$$\Pi(k_j) = \frac{k_j}{\sum_q k_q}. \quad (2.7)$$

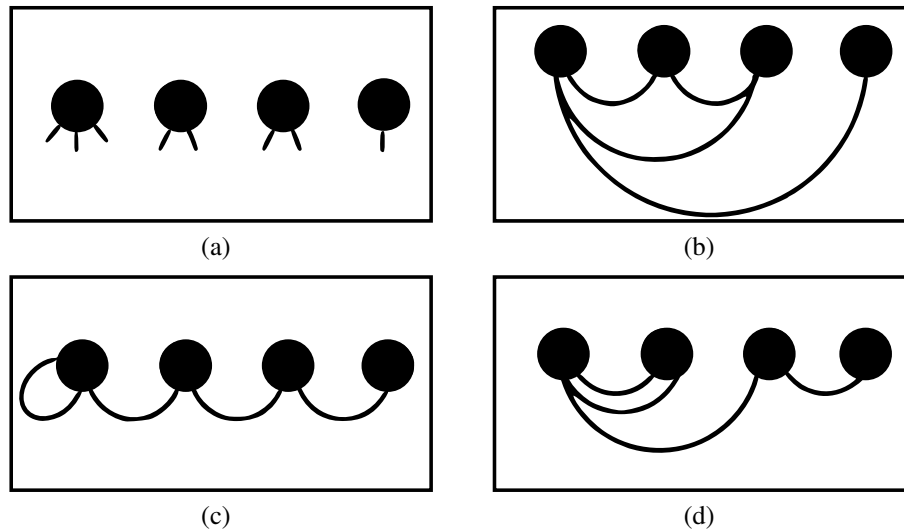
Onde o denominador é soma sobre todas os graus de conectividade.

Redes de todos os tipos aqui descritas podem ser geradas também a partir de uma distribuição de conectividade pré-determinada, ou seja, a partir de uma sequência que caracteriza a distribuição de conectividade. Este modelo onde cada vértice carrega seu grau de conectividade previamente é chamado de **Modelo de Configuração** e é apresentado na seção seguinte.

2.2.4 Modelo de Configuração

O Modelo de Configuração [5] consiste em uma metodologia de criar redes aleatoriamente a partir de uma lista genérica com informações sobre a distribuição de conectividade, independente da forma da distribuição. Na prática, através de uma distribuição de conectividade arbitrária $p(k)$, é criada uma sequência de informações sobre o grau de conectividade k_i de cada vértice i e cada um deles permanece com conectividade fixa durante o processo de estabelecimento das ligações entre os pares de vértices como exemplificado na Fig. (9). O

Figura 9: Esquema simplificado do Modelo de Configuração



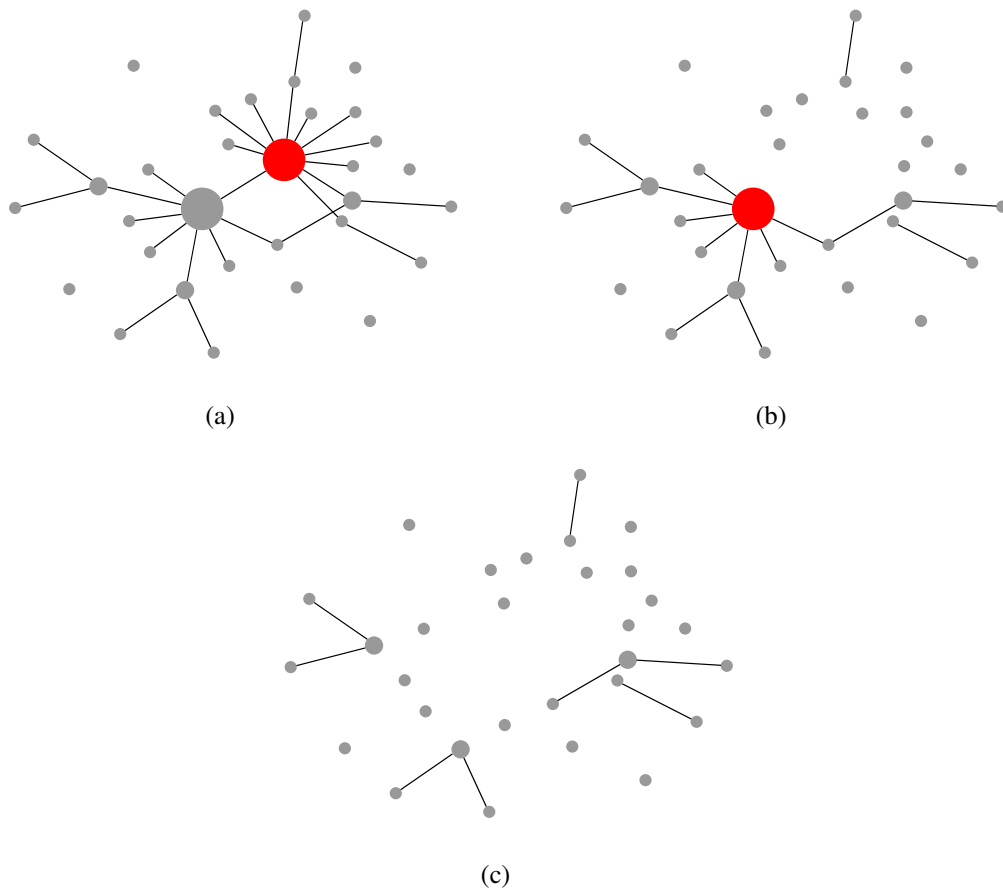
Fonte: Adaptado de Barabási: Network Science. (a) Um conjunto de quatro vértices com grau de conectividade pré-determinado, com $k_1 = 3, k_2 = 2, k_3 = 2$ e $k_4 = 1$. O índice de k faz referência ao rótulo do vértice por ordem. Outra maneira de armazenar essa informações é criar uma lista onde o vértice i entra k_i vezes, para este caso teríamos $L = [1, 1, 1, 2, 2, 3, 3, 4]$. O modelo de Configuração cria redes de modo aleatório conectando aos pares estes vértices até que todas as conexões tenham sido usadas. O resultado final da estrutura pode ser todas as naturezas já conhecidas, como exemplo (b) redes com ciclos, (c) com loops, (d) arestas múltiplas entre outras, tudo para uma mesma sequência de graus de conectividade.

modelo de Configuração é um importante recurso no âmbito computacional, no sentido de viabilizar a criação de redes com as mais diversas propriedades, a garantia de que as propriedades particulares de cada topologia seja observada [44].

3 TEORIA DA PERCOLAÇÃO

Um dos atributos de maior relevância de uma rede é a sua *Robustez* [45–48], ou seja, a maneira como o funcionamento do sistema representado é sensível à remoções sistêmicas ou falhas de vértices e arestas. Este aspecto da Ciência de Redes é amplamente explorado na literatura [10, 12, 41, 45–47, 49]. As falhas as quais os sistemas possam estar suscetíveis podem ser de natureza aleatória [12], ou seja, atingir aleatoriamente um vértice qualquer da rede, ou ataques maliciosos (ou direcionados), atrelados a algum atributo do vértice [50]. Em geral esse atributo trata-se do grau de conectividade do vértice, sendo os ataques maliciosos direcionados aos vértices mais conectados da estrutura. Veja na Fig. (10) um exemplo prático de fragmentação de uma rede a partir do ataque aos seus vértices mais conectados.

Figura 10: Fragmentação a partir da remoção de vértices



Fonte: Produzido pelo autor. (a) O vértice destacado em vermelho na rede é o mais conectado e portanto, de acordo com a forma de ataque direcionado aos hubs, será o primeiro removido da rede. (b) Após a remoção do mais conectado, bem como todas as sua arestas, o ataque direcionado atinge e deleta o vértice mais conectado naquele momento, destacado em vermelho. (c) Apenas devido à remoção de dois vértices da rede, esta se torna um conjunto praticamente fragmentado de vértices.

A maneira como as redes que seguem as mais difundidas distribuições de conectividade, especialmente as expostas no capítulo 2, as Redes Aleatórias e Livres de Escala, são bem características e revelam importantes informações acerca dos sistemas que representam. Pode-

mos sumarizar os principais resultados existentes para esse dois tipos de falhas, especialmente para Redes Livres de Escala, da seguinte forma [10, 41, 49]:

- Redes Livre de Escala são altamente resistentes à falhas aleatórias[30], o que pode ser explicado pela existência do hubs, que são responsáveis por manter a rede conectada. Como pela própria essência do modelo, existem bem mais vértices com conectividade baixa (e que portanto desempenham um papel menos importante na estrutura) do que os altamente conectados, o que faz com que a escolha aleatória de vértices leve majoritariamente a vértices pouco conectados.
- Já em relação a ataques maliciosos, direcionados aos hubs, Redes Livre de escala são expressivamente suscetíveis e se desintegram para uma baixa fração de vértices removidos [30]. Mais uma vez isso pode ser explicado pela relevância já destacada do papel dos hubs na estrutura.

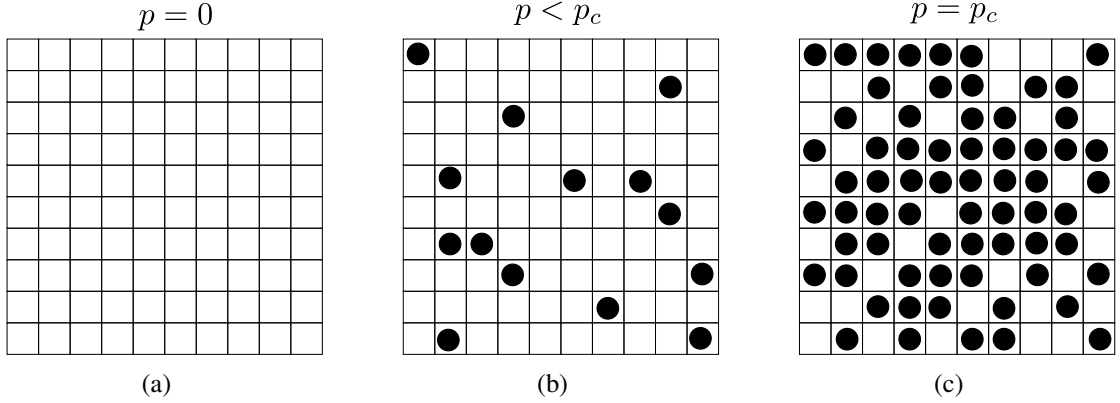
Como veremos adiante, este problema pode ser visto de duas formas distintas. A primeira foi e que foi apresentada até aqui, consiste em responder o quão resistente é uma rede, em relação à remoção categórica de vértices e/ou arestas, ou seja, de que forma a estrutura se comporta sob esse tipo de ação até se desintegrar em pequenos fragmentos e suas implicações para cada sistema [12, 41, 49, 51, 52]. O problema inverso, que é a segunda forma, é de similar relevância e consiste em responder a seguinte questão: quando um conjunto fragmentado de vértices torna-se uma rede praticamente totalmente conectada? A *Teoria da Percolação* nos possibilita, por meio do seu ferramental matemático, abordar esse tipo de problema. Esse campo de estudo foi fundado por *Broadbent* e *Flantnansley* [6, 53] em um artigo publicado no ano de 1957 e intitulado “*Percolation processes: I. Crystals and mazes*” (Processos de percolação: I. Cristais e desordens), e nos aprofundaremos um pouco sobre ela a partir de agora.

3.1 Percolação em Redes

Um exemplo de grande valor didático para um primeiro entendimento acerca da Teoria da Percolação consiste em tomarmos um estrutura quadrada, tal como na Fig. (11) [8, 30, 54], onde temos um total de $N = 100$ sítios. Cada sítio pode ser ocupado aleatoriamente com probabilidade p ou permanecer vazio com probabilidade $(1 - p)$, e dois ou mais sítios vizinhos formam um cluster ou uma componente. Para baixos valores de p , existem apenas alguns poucos clusters pequenos e isolados e, logicamente, é de se esperar que à medida que p cresça, cresça também a fração da estrutura ocupada pelo maior cluster. A previsão mais relevante da teoria da Percolação é que, ao contrário do que possa parecer intuitivo, o tamanho dos clusters não cresce gradualmente com p , na verdade existe um limiar de probabilidade p_c que separa uma fase onde existem apenas numerosos clusters isolados e uma fase onde emerge um cluster gigante, comumente denotado por P_∞ . Dessa forma, podemos dizer que nosso sistema percola, quando passa da fase de vários clusters isolados para a fase onde surge o *cluster gigante* ou *componente gigante*. Este fenômeno é entendido como uma transição de fase geométrica, onde o ponto crítico é a probabilidade p_c [54].

Um clássico exemplo físico para percolação como ocorre na Fig. (11) consiste em imaginarmos que cada sítio ocupado é um condutor elétrico e um sítio não ocupado é um iso-

Figura 11: Percolação de sítios em uma rede quadrada



Fonte: Produzido pelo autor. (a) Uma estrutura quadrada com probabilidade $p = 0$ de ter algum sítio ocupado, naturalmente, com esse valor de p , nenhum dos sítios são ocupados. (b) Para $p < p_c$, ou seja, para uma probabilidade abaixo da probabilidade crítica, apenas alguns clusters isolados existem. (c) Para $p > p_c$ o sistema experimenta uma transição de fase geométrica, onde passa a existir uma única componente que ocupa uma fração massiva dos sítios da estrutura, ou seja, o sistema percola a partir do surgimento da componente gigante (p_∞).

lante. Para baixos valores de p , existem apenas alguns poucos condutores ou grupos de condutores vizinhos, isso implica dizer que a corrente não consegue transitar de um lado a outro da estrutura e portanto o material é um isolante. No momento em que é atingido o limiar crítico p_c , o material experimenta uma transição de fase e se torna um condutor. Dizemos que a corrente percola a partir desse ponto crítico [54]. Além da percolação de sítios, da qual tratamos durante esta seção, existe também a percolação de arestas. Neste caso, os vértices são dispostos na estrutura e as arestas é que são distribuídas aleatoriamente e, equivalentemente ao caso anterior, uma probabilidade crítica separa uma fase de clusters de tamanho finito e outra com um cluster de tamanho infinito [54].

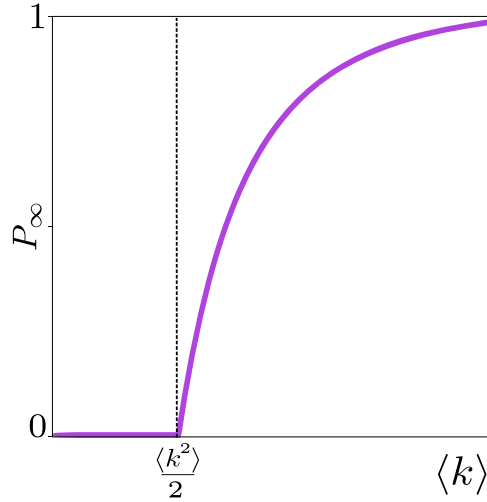
3.1.1 Componente Gigante e o Critério de Molloy-Reed

A emergência da componente gigante é uma questão chave na descrição da dinâmica de crescimento da rede, entre outras coisas, essa característica pode garantir o funcionamento do sistema em questão. Por essa e outras razões, é importante mapear em que situação e como isso ocorre. Uma das formas de determinar esse momento é analisando as propriedades estatísticas da distribuição de conectividade da rede, através do critério de Molloy-Reed [22], que pode ser escrito na forma da Eq. (3.1),

$$\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle} > 2. \quad (3.1)$$

Onde $\langle k \rangle$ e $\langle k^2 \rangle$ são respectivamente o primeiro e o segundo momento da distribuição de conectividade $p(k)$. O critério de Molloy-Reed estabelece que, independentemente da forma da distribuição $p(k)$, a condição necessária para que exista uma componente gigante é que, em média, cada vértice da rede esteja conectado a mais de outros dois vértices, veja na Fig. (12) a interpretação gráfica dessa informação. Podemos justificar este critério da seguinte maneira [55]: seguindo uma das arestas de um vértice aleatório, encontramos outro vértice com grau k_{viz} , no qual teremos três possibilidades, na primeira encontramos um vértice de grau $k_{viz} = 1$,

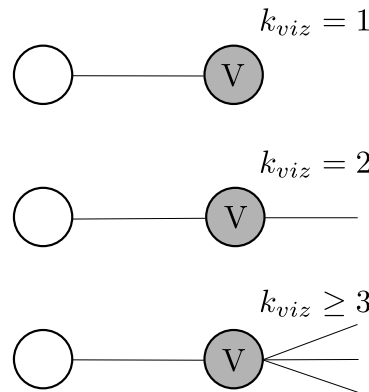
Figura 12: Probabilidade da existência de um cluster infinito em função de $\langle k \rangle$.



Fonte: Produzido pelo autor. De acordo com o critério de Molloy-Reed (Eq. 3.1) a componente gigante surge quando $\langle k \rangle = \frac{\langle k^2 \rangle}{2}$, sendo essa a condição necessária, independente da forma da distribuição de conectividade $p(k)$.

nesse caso o ramo dessa conexão é um terminal. A segunda possibilidade é encontramos um vértice de grau $k_{viz} \geq 3$, nesse caso o ramo se bifurca. Finalmente, na terceira possibilidade, encontramos um vértice de grau $k_{viz} = 2$ e o ramo segue sem bifurcar (ver Fig. 13). Ou seja, se $\langle k_{viz} \rangle > 2$ o número de vértices encontrados deve crescer rapidamente a cada passo, atingindo uma fração massiva da rede. Caso contrário, se $\langle k_{viz} \rangle < 2$ esse número provavelmente vai diminuir e vai se atingir apenas uma pequena fração de vértices. Portanto, $\langle k_{viz} \rangle = 2$ é o ponto crítico¹.

Figura 13: Conectividade do vértice vizinho



Fonte: Produzido pelo autor. Os vértices destacados à direita exemplificam as situações onde um ramo chega um terminal ($k_{viz} = 1$), segue sem se bifurcar ($k_{viz} = 2$) ou se ramifica ($k_{viz} \geq 3$).

¹Note que $\langle k_{viz} \rangle$ difere do grau médio $\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i$.

3.1.2 Expoentes Críticos

Para um melhor entendimento acerca do papel dos expoentes críticos em Percolação de Redes, nos deixe recuperar um clássico e de grande valor didático problema físico: a transição de fase magnética [54,56]. Este fenômeno ocorre quando um material, que exibe uma magnetização espontânea² ($m > 0$) para baixas temperaturas, experimenta um determinado acréscimo de temperatura e a magnetização diminui continuamente até desaparecer para uma temperatura crítica T_c . Acima da temperatura limiar, o material transita para a fase paramagnética. Próximo à temperatura crítica, a magnetização pode ser descrita conforme a Eq. (3.2) [54],

$$m(T) \propto (T_c - T)^\tau. \quad (3.2)$$

A grandeza m é chamada de *parâmetro de ordem* para este problema³ e τ é o expoente crítico que determina o comportamento do sistema⁴ perto de T_c . Transições de fase como essa, onde o parâmetro de ordem varia continuamente até divergir no ponto crítico⁵, apresentam uma propriedade de especial relevância: a universalidade. Isso significa que os expoentes críticos não dependem dos detalhes da estrutura, mas somente da dimensão do sistema em questão [58]. No caso da magnetização, por exemplo, todo material tridimensional é descrito pelo mesmo expoente⁶ τ . Por esse motivo, os expoentes críticos são mais relevantes que os próprios limiares de criticidade, uma vez que, estes sim são bastante sensíveis aos demais detalhes de estrutura e interação [58].

De modo análogo à transição de fase magnética, podemos caracterizar as mais relevantes grandezas que envolvem a percolação em redes, a saber:

- A probabilidade P_∞ de um vértice aleatório pertencer à componente gigante

$$P_\infty \propto (p - p_c)^\beta, \quad (3.3)$$

sendo P_∞ é o parâmetro de ordem. Assim como antes, o expoente β depende somente da dimensão da rede. Veja na Fig. (14) um gráfico ilustrativo do comportamento da Eq. (3.3).

- O tamanho médio das componentes (excluindo a componente gigante, quando ela existe)

$$\langle S \rangle \propto |p - p_c|^{-\gamma}. \quad (3.4)$$

O que mostra que as componentes finitas crescem à medida que p se aproxima de p_c e divergem quando $p = p_c$. Veja agora, na Fig. (15), uma representação ilustrativa do comportamento da Eq. (3.4).

- E por fim, a distancia média entre todos os possíveis pares de vértices no mesmo cluster

²Sem a presença de um campo magnético externo.

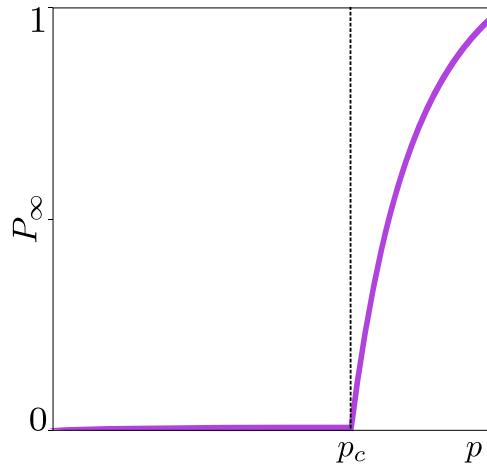
³ m carrega a informação acerca da orientação dos spins, para altas temperaturas os spins apontam aleatoriamente em todas as possíveis direções, já abaixo de T_c todos os átomos orientam seus spins na mesma direção. Estas duas situações separam uma fase ordenada de uma fase desordenada, e daí a ideia de parâmetro de ordem.

⁴Para maiores detalhes sobre esta transição de fase, consulte o capítulo XI da referência [57].

⁵Chamadas de transições de fase de segunda ordem.

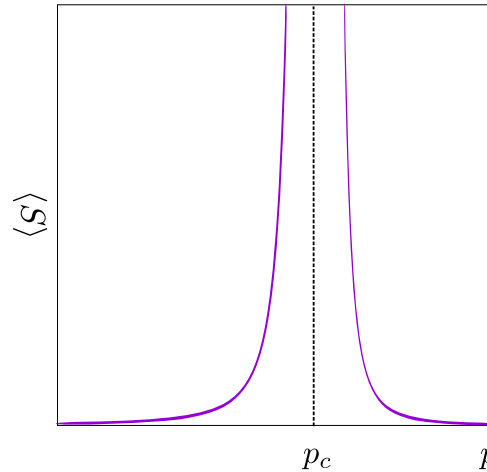
⁶Desde que as interações entre os spins sejam de curto alcance.

Figura 14: Probabilidade da existência de um cluster infinito



Fonte: Produzido pelo autor. Para $p < p_c$ a probabilidade de escolher um vértice que pertença à componente gigante p_∞ é nula, ou seja, a componente gigante não existe, mas emerge para $p > p_c$.

Figura 15: Tamanho médio dos clusters finitos



Fonte: Produzido pelo autor. Para baixos valores de p , existem apenas numerosos e pequenos clusters isolados, à medida que a probabilidade de ligar pares de vértices aumenta e diverge no ponto crítico, onde muitos clusters de grande tamanho coexistem.

finito, ou comprimento de correlação ξ ,

$$\xi \propto |p - p_c|^{-\nu}. \quad (3.5)$$

Os expoentes β , γ e ν apresentam a propriedade de universalidade, e podemos encontrar vários trabalhos que se ocupam em determinar os valores de cada um deles para as dimensões de interesse ⁷. Existem outros expoentes de interesse em percolação, no entanto todos eles são relacionados e podemos determiná-los a partir dos demais.

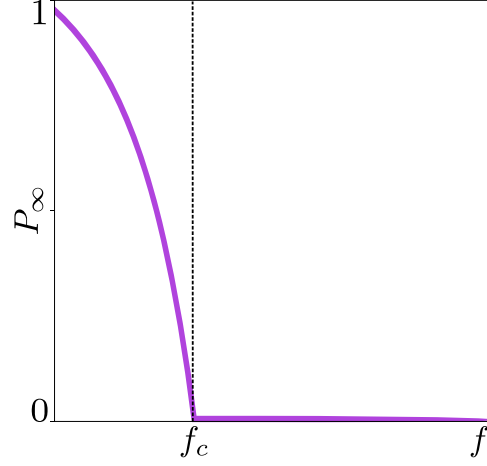
3.1.3 Percolação Inversa e Robustez

Podemos retomar agora nossa discussão a respeito da robustez de redes, que pode ser a partir deste momento enxergada sob a ótica da teoria de Percolação bem a como a apresentamos nas ultimas seções. Podemos, portanto, tratar a fragmentação de uma rede a partir de

⁷Veja a página 65 da referência [54] para valores com $d = 2, 3, \geq 6$

remoção categórica de vértices ou arestas, como um problema de percolação inversa, ou equivalentemente o desaparecimento da componente gigante, veja na Fig. (16) o comportamento genérico da probabilidade P_∞ como função da fração de vértices removidos.

Figura 16: Probabilidade da existência da componente gigante em função da fração de vértices removidos.



Fonte: Produzido pelo autor. A probabilidade p_∞ de uma componente gigante que existe quando $f = 0$, ou seja, a rede mantida em sua integridade. À medida que a fração de vértices removidos cresce, eventualmente a rede se fragmenta em clusters finitos para uma fração crítica f_c , o valor da fração crítica varia de acordo com as propriedades topológicas da rede e da natureza da remoção dos vértices.

Assim como na situação inversa ao representado na Fig. (16), o processo de fragmentação da rede ocorre de modo não gradual, mas sim abruptamente para um valor característico f_c , ou seja, a fração crítica de vértices removidos. Como dito no início do capítulo, cada classe de estrutura possui tolerância específica para as duas formas de remoção de vértices: aleatória ou direcionada. Mas algumas informações que independem dos detalhes da estrutura ainda podem ser capturadas, vejamos:

- O expoente crítico que caracteriza a transição de regimes é o mesmo (β) que foi apresentado para o problema inverso, na Eq. (3.3), conforme segue na Eq. (3.6)

$$P_\infty \propto |f - f_c|^\beta. \quad (3.6)$$

E podemos verificar facilmente essa condição apenas atentando para o fato de que na Eq. (3.6) mapeia a Eq. (3.3) com a transposição $f = 1 - p$.

- O limiar de criticidade f_c também pode ser determinado pelo critério Molloy-Reed (equação 3.1) para uma distribuição de conectividade genérica sob a ocorrência de remoções aleatórias de vértices, conforme a Eq. (3.7)[10],

$$f_c = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1}. \quad (3.7)$$

Novamente determinado apenas pelo primeiro e segundo momento de $p(k)$.

Na sequência deste trabalho, exploraremos os conceitos aqui discutidos para fundamentar uma série de resultados numéricos e fenomenológicos associados à teoria da percolação no ensemble de grafos acíclicos aleatórios.

4 TRANSIÇÃO DE FASE ESTRUTURAL NO CONJUNTO DE GRAFOS ACÍCLICOS ALEATÓRIOS

O comportamento crítico de grafos aleatórios, particularmente no modelo clássico de Erdős–Rényi, é um problema já amplamente resolvido na literatura, com resultados estabelecidos há décadas sobre a emergência de componentes gigantes e os expoentes associados às transições de fase [19–21]. De forma análoga, o problema de árvores acíclicas rotuladas também possui uma solução combinatória conhecida, com destaque para a clássica fórmula de Cayley [59], que determina exatamente o número de tais estruturas. Em ambos os casos, os expoentes críticos obtidos correspondem ao regime de campo médio, como é característico de sistemas sem correlação espacial. O presente trabalho, portanto, não visa apresentar novos resultados sobre esses modelos, mas sim reformular o problema sob uma perspectiva inspirada nos métodos da Mecânica Estatística, oferecendo uma abordagem alternativa e potencialmente mais acessível para físicos. Ao invés da linguagem tradicional da Combinatória e da Matemática Discreta, adotamos aqui conceitos como entropia, maximização de estados acessíveis e equilíbrio termodinâmico para rederivar resultados conhecidos e, ao mesmo tempo, propor um enquadramento conceitual que pode ser facilmente generalizado. Essa estratégia abre caminho para futuras extensões, como a modelagem de transições abruptas em percolação explosiva, e aproxima o estudo estrutural de redes dos métodos consagrados em sistemas estatísticos. Nossa abordagem aproveita o fato de que grafos acíclicos permitem uma enumeração combinatória exata de todo o conjunto, o que leva à definição de uma entropia para este sistema. Com essa definição de entropia, podemos determinar o início da transição crítica, bem como os expoentes críticos associados à transição. Nossos resultados são ilustrados com resultados obtidos a partir do método de Monte Carlo e são discutidos no contexto de grafos aleatórios em geral, bem como em comparação com outro modelo de grafos acíclicos.

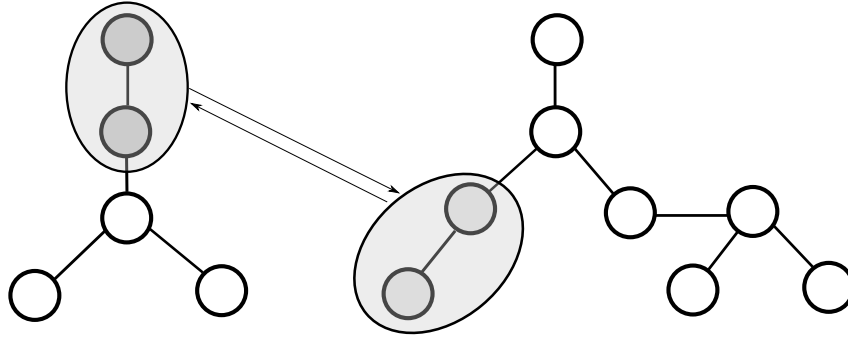
4.1 Entropia para duas árvores que trocam vértices

Vamos explorar inicialmente a ideia de grafos para perscrutar os métodos da Mecânica Estatística. Consideremos dois grafos acíclicos de tamanhos n_1 e n_2 , respectivamente, compostos por vértices distinguíveis. Podemos modelar um sistema composto em que ocorrem trocas mútuas de vértices entre os dois grafos, conforme ilustrado na Fig. (17), onde $N = n_1 + n_2$ é número total de vértices do sistema composto.

A troca de vértices entre os grafos gera diferentes configurações, que podemos associar a estados distintos do sistema. Este fator de permutação de vértices se traduz na forma de tomar n_1 elementos para a árvore 1, dentre N vértices disponíveis no sistema composto, dado pelo coeficiente multinomial $\binom{N}{n_1}$. Este fator vem da dependência da entropia de Boltzmann no número de partículas, o qual requer o cálculo da probabilidade do número de partículas distinguíveis em cada subsistema de um sistema composto e está diretamente relacionado ao chamado “*Paradoxo de Gibbs*”, que surge em função de uma definição imprecisa da entropia, como discutiremos adiante.

Além disso, devemos considerar as possíveis configurações internas de cada árvore. O número de grafos acíclicos conectados com N vértices rotulados é dado pela *Fórmula de*

Figura 17: Esquema de um sistema composto por dois grafos.



Fonte: Produzido pelo autor. Ilustração esquemática de um sistema composto, no qual os dois grafos podem trocar conjuntos de vértices entre si, análogo à troca de partículas. Na figura, supondo que os vértices sejam distinguíveis, a permuta dos ramos destacados entre os dois grafos leva à uma configuração distinta em relação à anterior, configurando um novo estado para o sistema composto.

Cayley [59,60], que estabelece que existem N^{N-2} diferentes grafos acíclicos conectados com n vértices rotulados. Este resultado pode ser provado de vários modos [59], como por exemplo o Teorema da Matriz-Árvore de Kirchhoff [61], o método de dupla contagem implementado por Jim Pitman [59,60], e uma maneira especialmente engenhosa, que se dá pelo uso de *sequências de Prüfer*. Uma sequência de Prüfer [62] é uma sequência numérica única de tamanho $n - 2$, associada à uma árvore de n vértices e rotulada de 1 a n . A sequência de Prüfer é obtida a partir de uma árvore por meio de um processo onde são removidos iterativamente os vértices da árvore até que apenas dois vértices permaneçam. O processo começa identificando a folha com o menor rótulo na árvore, ou seja, o vértice que possui apenas uma aresta conectada. O rótulo do único vizinho dessa folha é então adicionado à sequência de Prüfer. Em seguida, a folha é removida da árvore junto com sua aresta, e o processo se repete com a nova folha de menor rótulo, de modo que, no passo i , é removido a folha com o menor rótulo e definido o i -ésimo elemento da sequência de Prüfer como o rótulo do vizinho dessa folha. Existe uma correspondência biunívoca entre árvores rotuladas com n vértices e sequências de Prüfer de tamanho $n - 2$, o que significa que a sequência pode ser usada tanto para descrever uma árvore quanto para reconstruí-la. Além disso, a fórmula de Cayley, n^{n-2} , coincide com o número de sequências possíveis de $n - 2$ elementos, onde cada elemento pode assumir valores entre 1 e n .

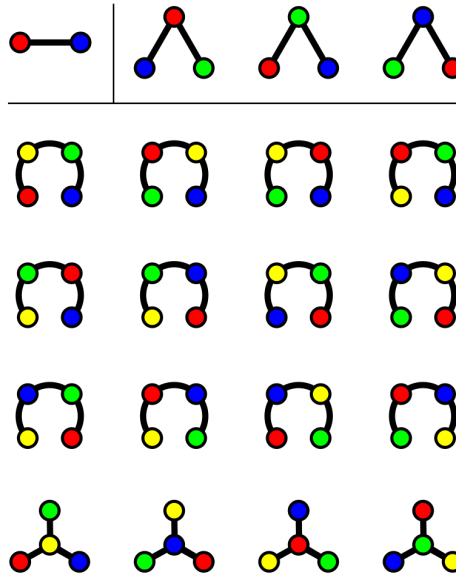
A partir destes dois termos, o coeficiente multinomial e a fórmula de Cayley, podemos determinar o número $\Omega_0(n_1)$ de configurações acessíveis à um sistema composto como o descrito, da seguinte forma

$$\Omega_0(n_1) = n_1^{n_1-2} (N - n_1)^{(N-n_1-2)} \left(\frac{N!}{n_1! (N - n_1)!} \right). \quad (4.1)$$

4.1.1 Entropia e maximização de Configurações

De acordo com os métodos da Mecânica Estatística, a entropia do sistema é proporcional ao logaritmo natural do número de estados acessíveis, então, tomando o logaritmo da Eq. (4.1), escrevemos a entropia para este sistema na forma

Figura 18: A fórmula de Cayley.



Fonte: wikipedia.org. Todas as árvores possíveis para: $n = 2$ vértices (1 árvore possível); $n = 3$ vértices (3 árvores possíveis) e; $n = 4$ vértices (16 árvores possíveis), em acordo à fórmula de Cayley, $T_n = n^{n-2}$.

$$S \propto \ln(\Omega_0(n_1)) = \ln \left(n_1^{n_1-2} (N - n_1)^{(N-n_1-2)} \left(\frac{N!}{n_1! (N - n_1)!} \right) \right). \quad (4.2)$$

A maximização da entropia, no caso de um sistema termodinâmico, está relacionada à maximização da probabilidade [63]. Tomamos então a diferenciação da entropia em relação n_1 , que nos permite encontrar o ponto crítico para o qual a função $\ln \Omega_0(n_1)$ experimenta um valor extremo, alcançado quando $d \ln \Omega_0 / dn_1 = 0$. Tomando n_1 e $N \gg 1$ e fazendo uso da *aproximação de Stirling* ($\ln x! \sim x(\ln x - 1)$), chegamos a

$$\frac{\partial \ln \Omega_0(n_1)}{\partial n_1} = \left(\frac{n_1 - 2}{n_1} \right) - \left(\frac{N - n_1 - 2}{N - n_1} \right). \quad (4.3)$$

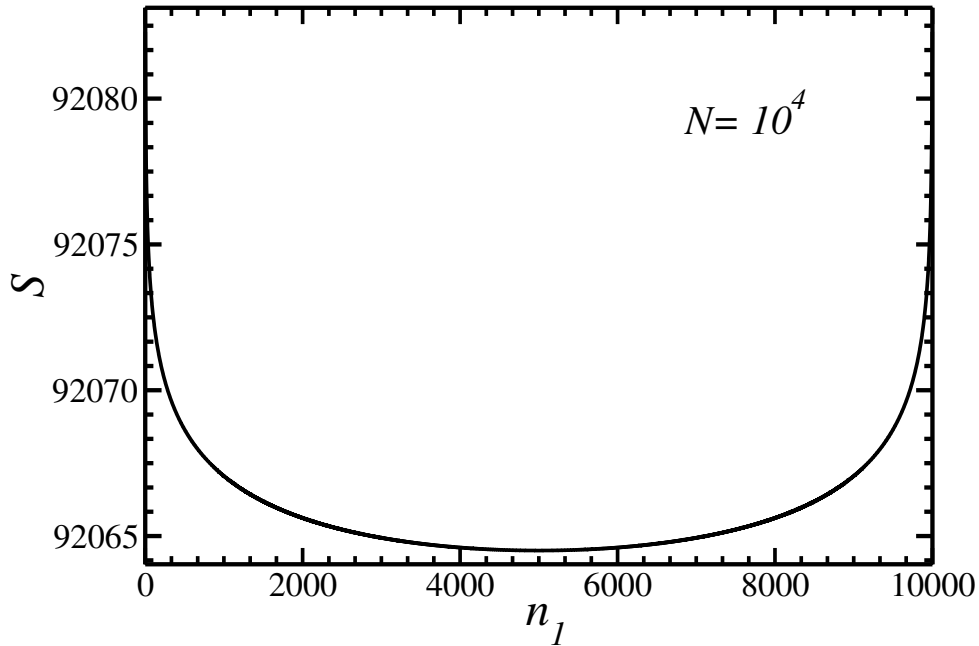
Igualando a Eq. (4.3) a zero e resolvendo para n_1 , encontramos que o ponto crítico da função se dá quando $n_1 = N/2$, ou seja, quando as duas componentes possuem o mesmo tamanho. E mais, a avaliação da derivada segunda de $\ln \Omega_0$ aplicada no ponto $n_1 = N/2$ é maior do que zero, o que significa que o extremo ocorre em um ponto de mínima entropia. Com isso, constatamos que na verdade a maior probabilidade é de que o sistema encontre-se com quase todos os vértices em uma só árvore, veja por exemplo o Gráfico (1).

Uma vez que a quantidade de configurações possíveis do sistema é dado por Ω_0 , a probabilidade de encontrarmos tal sistema em um estado no qual uma das árvores tenha um tamanho n_1 deve ser proporcional à Ω_0 , na forma

$$P(n_1) \propto n_1^{n_1-2} (N - n_1)^{(N-n_1-2)} \left(\frac{N!}{n_1! (N - n_1)!} \right). \quad (4.4)$$

Se tomarmos $N \gg n_1$ (o que de fato é o mais provável que aconteça, como vimos), podemos

Gráfico 1: Entropia para um sistema composto de duas árvores.



Fonte: Produzido pelo autor. Conforme a Eq. (4.2), a entropia não tem um máximo definido mas sim um mínimo que ocorre quando $n_1 = N/2$, sendo os maiores valores e portanto os estados mais prováveis correspondem à quando uma das árvore é bem maior que a outra.

usar a Fórmula de Stirling ($x! \sim \sqrt{2\pi x} \left(\frac{x}{e}\right)^x$ desde que $x \gg 1$) para escrever

$$P(n_1) \propto \frac{n_1^{n_1-2} (N-n_1)^{(N-n_1-2)} N!}{n_1! (2\pi)^{1/2} (N-n_1)^{1/2} N^{(N-n_1)} e^{-N} e^{n_1}}. \quad (4.5)$$

Os termos que não dependem de n_1 podem ser englobados no fator de proporcionalidade, com isso temos,

$$P(n_1) \propto \frac{n_1^{n_1-2} e^{-n_1} (N-n_1)^{(-5/2)}}{n_1!}. \quad (4.6)$$

o último termo no numerador da Eq. (4.6) depende muito fracamente de n_1 , portanto, podemos finalmente estabelecer que a probabilidade $P(n_1)$ é

$$P(n_1) \propto \frac{n_1^{n_1-2} e^{-n_1}}{n_1!}. \quad (4.7)$$

4.1.2 Uma nota sobre o paradoxo de Gibbs

Devido à introdução do termo binomial, os fatores $1/N!$ aparecem naturalmente na definição da entropia dos vértices distinguíveis, $S = \ln(\Omega/N!)$. Sem esses fatores não seria possível garantir a aditividade das entropias dos subsistemas, $S = S_1 + S_2$. Gibbs percebeu que dividir a enumeração de estados acessíveis Ω por $N!$ seria necessário para garantir a aditividade e extensividade da entropia. No entanto, ele não atribuiu a presença desse termo ao fator binomial, essencial para obter uma enumeração correta dos estados do sistema composto. De fato, Gibbs considerou a necessidade desse fator como paradoxal, e justificou com a ideia de que a permutação de partículas idênticas não deve levar a estados distintos. Muitos criticaram essa

explicação desse termo, já que no modelo clássico, que estava em voga no século XIX quando Gibbs fez seus estudos, as partículas idênticas são distinguíveis por sua trajetória. Dezenas de anos depois dos trabalhos de Gibbs, com o advento da Mecânica Quântica, mostrou-se que de fato que a permutação de partículas idênticas não leva a estados distintos, confirmando a intuição física de Gibbs. No entanto, esse fato não faz com que essa explicação, que se mostrou finalmente correta, seja a única explicação possível. Como vimos aqui, quando tratamos de elementos distinguíveis, onde permutações levam a estados distintos na enumeração, o fator $1/N!$ aparece naturalmente da lógica combinatorial. Em 1907 Paul e Tatjana Erenfest [64] publicaram um artigo onde o termo binomial já aparecia na enumeração do sistema composto de partículas clássicas. Apesar disso, a explicação combinatória para a introdução desse termo não foi difundida na comunidade de físicos e, é comum ler-se em livros de Mecânica Estatística que não é possível definir-se uma entropia para sistemas de partículas distinguíveis e que somente na Mecânica Quântica seria possível construir uma Física Estatística consistente. Essas afirmações já foram fortemente rechaçadas [65]. De fato, distinguível e indistinguível não são propriedades que podem ser atribuídas a uma única partícula, mas a grupos de partículas, e mesmo na Mecânica Quântica é possível ter um gás de múltiplas partículas mutualmente distinguíveis, e é possível definir uma entropia nesse caso [66]. A forma da entropia para sistemas de partículas distinguíveis foi discutida por Jaynes [67] que tenta entender o que levou Gibbs a não perceber a necessidade do termo $1/N!$. A discussão da introdução desse termo também é tratada de forma simples por Swendsen [68].

4.2 Entropia para grafos aleatórios acíclicos

Vamos agora determinar a entropia de um conjunto de grafos acíclicos dada uma distribuição particular de tamanhos de componentes m_n . Como consideramos que os N vértices do grafo são rotulados ou distinguíveis, a entropia é dada por $S = \ln(\Omega[m_n]/N!)$, onde $\Omega[m_n]$ é a enumeração do conjunto de grafos acíclicos, dado a distribuição de tamanhos de componentes m_n . O fator $N!$ é novamente introduzido para tornar a entropia dos vértices distinguíveis aditiva [65, 67, 68].

O primeiro passo para obter a enumeração Ω é distribuir os N vértices pelas várias componentes, o que introduz um fator na forma de coeficiente multinomial $\frac{N!}{\prod_{i=1} (n_i!)}$, onde os n_i são tamanhos das componentes e temos a restrição de que $\sum_i n_i = N$. Como estamos interessados na distribuição de tamanhos das componentes m_n , é útil escrever esse fator em termos dessas variáveis como $\frac{N!}{\prod_{n=1} (n!)^{m_n}}$, e a restrição torna-se $\sum_{n=1} n m_n = N$. Uma vez que enumeramos as maneiras de dividir N vértices entre os várias componentes, precisamos incluir alguns outros fatores. O primeiro carrega a informação acerca da enumeração das distintas formas de construir os grafos, novamente dado pela fórmula de Cayley, que introduz um fator da forma $\prod_{n=1} n^{m_n(n-2)}$ na enumeração. Além disso, como as componentes não são distinguíveis a priori, ainda estamos contando demais nossa enumeração. Por exemplo, se tivermos m_1 componentes com apenas um vértice, permutar os vértices dessas componentes não leva a uma nova partição dos vértices. Da mesma forma, se pegarmos dois componentes de tamanho $n = 2$ e permutarmos o par de vértices de uma componente com os dois vértices do outra componente, essa permutação também não levará a uma nova partição. O mesmo acontecerá se permutarmos

qualquer grupo de vértices que estejam em componentes do mesmo tamanho. Para levar em conta essa múltipla contagem, precisamos dividir nossa enumeração por $\prod_{n=1} m_n!$, obtendo, a partir da junção de todas essas contribuições, finalmente

$$\Omega[m_n] = N! \prod_{n=1} \left[\left(\frac{n^{n-2}}{n!} \right)^{m_n} \frac{1}{m_n!} \right], \quad (4.8)$$

o que leva à entropia na forma

$$S = \ln \left(\frac{\Omega}{N!} \right) = \sum_{n=1} m_n \left((n-2) \ln n - \ln n! - \ln m_n + 1 \right), \quad (4.9)$$

onde utilizamos mais uma vez a aproximação de Stirling, $\ln(m_n!) \approx m_n(\ln(m_n) - 1)$.

Para obter a distribuição mais provável do tamanho das componentes m_n , precisamos maximizar a entropia levando em conta as restrições das variáveis m_n . Como mencionado, uma dessas restrições dá o número total de vértices N . No entanto, considerando que o grau médio dos vértices do grafo é especificado, temos outra restrição. Isso acontece porque em grafos acíclicos, cada conexão entre um par de vértices reduz necessariamente o número total de componentes M em uma unidade. Como consequência, temos que o número total de conexões é dado por $E = N - M$, e o grau médio é dado por $\mu = 2E/N = 2(1 - M/N)$. Portanto, em grafos acíclicos, fixar o grau médio também fixa o número total de componentes. Temos então as seguintes restrições

$$\sum_{n=1} m_n = M \quad \text{e} \quad \sum_{n=1} n m_n = N. \quad (4.10)$$

Como temos essas restrições para as entradas da Eq. (4.9), devemos utilizar o método dos multiplicadores de Lagrange, que consiste em uma técnica de otimização que permite encontrar os valores extremos de uma função sujeita a uma ou mais restrições [69], a partir da qual podemos obter a entropia máxima resolvendo

$$\frac{\partial}{\partial m_n} [S + \lambda_1 (M - \sum m_n) + \lambda_2 (N - \sum n m_n)]_{m_n=m_n^*} = 0, \quad (4.11)$$

onde λ_1 e λ_2 são os multiplicadores de Lagrange. Fazendo a derivada chegamos a

$$\sum \ln \frac{n^{n-2}}{n!} - \ln m_n - \lambda_1 n - \lambda_2 = 0. \quad (4.12)$$

Com a introdução dos multiplicadores de Lagrange podemos considerar as variáveis m_n independentes, o que permite tomar nulo cada termo do somatório em (4.12), chegando a

$$m_n^* = \frac{n^{n-2} e^{-\lambda_1} e^{-n\lambda_2}}{n!}. \quad (4.13)$$

Os multiplicadores de Lagrange devem ser determinados a partir das restrições (4.10). Definindo $x = e^{-\lambda_2}$ e $y = e^{-\lambda_1}$, temos

$$N = y \sum_n \frac{n^{n-1}}{n!} x^n \quad (4.14)$$

e

$$M = y \int_0^x \sum_n \frac{n^{n-1}}{n!} x'^{n-1} dx'. \quad (4.15)$$

A soma em (4.14) e (4.15) pode ser relacionada à expansão de Taylor de uma função $g(x) = \sum_{n=1}^{\infty} \frac{n^{n-1}}{n!} x^n$. Usando o teorema de inversão de Lagrange [70], é possível mostrar que a função implicitamente definida pela equação transcendental $g(x)e^{-g(x)} = x$ terá uma série de Taylor precisamente nessa forma (veja a demonstração no apêndice (A)). De fato, nossa função $g(x)$ pode ser relacionada ao ramo principal da função W de Lambert [71] na forma $g(x) = -W_0(-x)$. A partir disso, temos

$$N = yg(x) \quad \text{e} \quad M = y \int_0^x \frac{g(x')}{x'} dx'. \quad (4.16)$$

Diferenciando ambos os lados da definição implícita $g(x)e^{-g(x)} = x$, obtém-se que $\frac{dg(x)}{dx} = \frac{e^{g(x)}}{1-g(x)}$, a partir de onde pode-se verificar que $\frac{d}{dx}(g(x)(1-g(x)/2)) = \frac{g(x)}{x}$ (veja as demonstrações no apêndice (A)), ou seja, a integral que aparece na segunda restrição (4.16), pode ser substituída por $g(x)(1-g(x)/2)$. O que nos permite escrever para uma das restrições $M = yg(x) \left[1 - \frac{g(x)}{2}\right]$ e para o grau médio, a partir das restrições (4.16) e da definição de grau médio em grafos acíclicos, $\mu = 2(1 - M/N) = g(x)$.

Isso nos permite determinar os parâmetros $x = \mu e^{-\mu}$ e $y = N/\mu$, chegando à forma final da distribuição de tamanho de componentes [72],

$$m_n^* = N \frac{n^{n-2}(\mu e^{-\mu})^n}{\mu n!}. \quad (4.17)$$

Se a distribuição de tamanho de componentes, Eq. (4.17), for usada na fórmula para a entropia, Eq. (4.9), é possível determinar a entropia sob a condição de equilíbrio

$$S(N, \mu \leq 1) = N \left[\left(1 - \frac{\mu}{2}\right) \left(1 - \ln \frac{N}{\mu}\right) + \mu - \ln \mu \right]. \quad (4.18)$$

4.3 Transição de fase no ensemble dos grafos aleatórios acíclicos

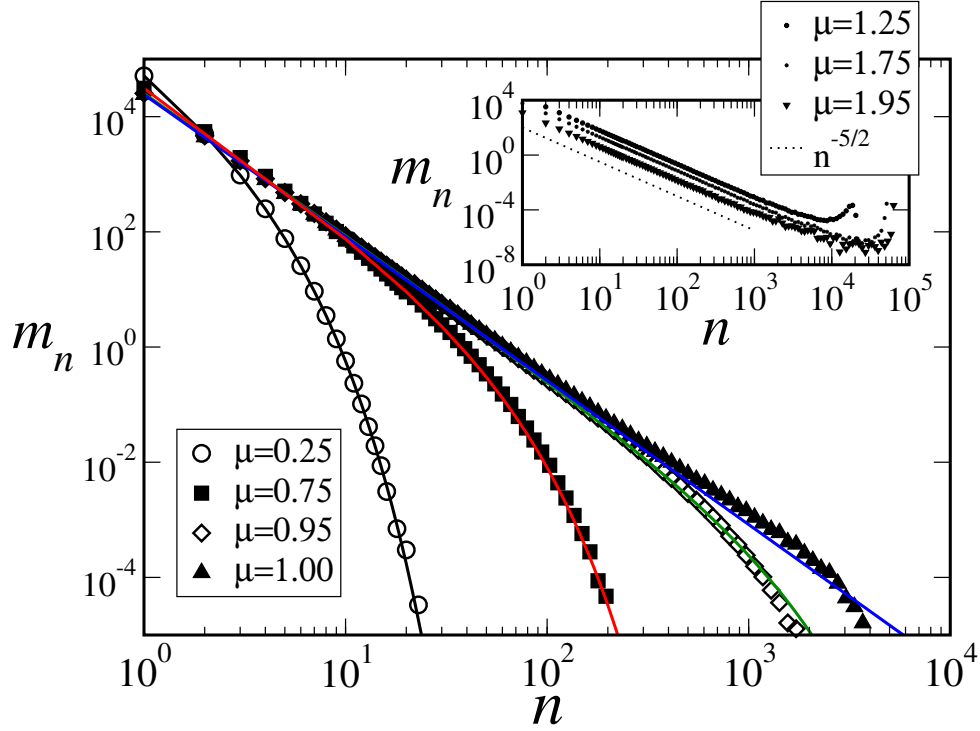
É comum definir a probabilidade $P_n = nm_n/N$ de um vértice dado ser encontrado em uma componente de tamanho n . Para obter o comportamento assintótico dessa probabilidade, precisamos empregar a aproximação de Stirling $n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$, obtendo

$$P_n \approx n^{-3/2} \exp(-(\mu - 1 - \ln \mu)n) / \sqrt{2\pi\mu^2}. \quad (4.19)$$

O comportamento assintótico descrito em (4.19) tem uma contribuição de um decaimento de lei de potência, $n^{-3/2}$, e um decaimento exponencial, $e^{-[\mu-1-\ln\mu]n}$. Para $\mu = 1$, os termos no decaimento exponencial se cancelam, e apenas a lei de potência $n^{-3/2}$ permanece. Esse ponto em que o decaimento exponencial se anula define o ponto crítico $\mu_c = 1$. Observe

que a condição crítica $\mu = 1$ ocorre quando o número de componentes é metade do número de vértices, $M = N/2$. O expoente que controla o decaimento de lei de potência da distribuição do tamanho das componentes é chamado de expoente de Fisher, e $\tau = 3/2$ é o valor esperado no campo médio [73], e é o que é obtido para o modelo de grafo aleatório [72]. Para $\mu \lesssim 1$, temos $\mu - 1 - \ln \mu \approx (1 - \mu)^2/2$, e o decaimento exponencial dominará quando $n > 2/(1 - \mu)^2$. Uma divergência da escala de tamanho com $(\mu - \mu_c)^{-2}$ também é consistente com o esperado no limite de campo médio [73].

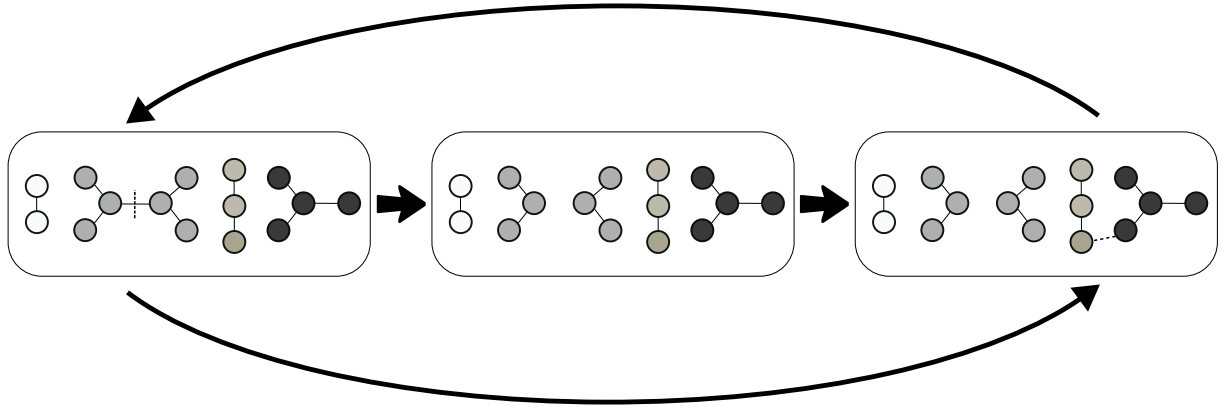
Gráfico 2: Distribuição do tamanho das componentes m_n para diferentes valores de grau médio μ .



Fonte: Produzido pelo autor. No painel principal, mostramos os resultados para $\mu \leq 1$. Os símbolos são resultados de simulações de Monte Carlo para um sistema com $N = 2^{16}$ vértices, enquanto as linhas contínuas são as previsões da Eq. (4.17). Há uma boa concordância entre a teoria e a simulação, com exceção do caso $\mu = 1$, onde se pode observar um pequeno desvio na cauda da distribuição. A razão desse desvio é o tamanho finito do sistema. À medida que $\mu \rightarrow 1$, a escala da transição exponencial diverge, no entanto, as maiores componentes são limitadas pelo tamanho finito do sistema, daí o motivo do desvio. No painel secundário, apresentamos resultados para $\mu > 1$. Neste regime, a distribuição do tamanho das componentes segue a forma descrita pela Eq. (4.25), que decai como uma lei de potência com expoente $-5/2$. Os picos observados na cauda da distribuição quando $\mu > 1$ indicam a presença de uma componente gigante neste regime.

Para ilustrar nossos resultados, procedemos a realizar simulações de Monte-Carlo do nosso ensemble de grafos acíclicos. Para fazer isso, criamos um grafo acíclico com um dado grau médio μ e, em cada etapa do algoritmo, removemos uma conexão aleatória do grafo e incluímos uma nova conexão aleatória que não cria um ciclo, veja a ilustração desse processo na Fig. (19).

Figura 19: O processo de Monte Carlo.



Fonte: Produzido pelo autor. O processo consiste na criação de um conjunto de grafos com grau médio de conectividade fixo μ , a partir do qual é iniciado o processo de monte Carlo, removendo uma ligação aleatória do conjunto e adicionando outra que não forma um ciclo ou um loop na rede. Na configuração inicial da figura, a aresta destacada com a linha tracejada foi removida, gerando o estado intermediário da parte mais central da ilustração, que por sua vez leva ao estado final a partir a inserção da nova ligação destacada em traços na figura mais à direita. Observe que este processo não altera a conectividade média da rede e, portanto, não altera o número M de componentes. Este processo obedece ao balanço detalhado, ou seja, a probabilidade de transitar do estado inicial da figura para o estado final é mesma de fazer o caminho inverso. Isso permite gerar amostras do ensemble de grafos acíclicos aleatórios para um dado conjunto de N vértices e grau médio de conectividade μ .

É possível ver que a probabilidade de transitar de um grafo **A** para outro grafo **B** é a mesma que a de transitar de volta de **B** para **A**. O balanço detalhado [74] garante que, nessas condições, cada grafo possível no conjunto tem a mesma chance de aparecer no processo de Monte-Carlo. Nossos resultados de Monte-Carlo são apresentados na Gráfico (2). No painel principal deste gráfico, mostramos a distribuição do tamanho das componentes de sistemas com grau médio $\mu \leq 1$. Pode-se ver que nossos resultados numéricos seguem de perto as previsões da Eq. (4.17). No painel secundário da Gráfico (2), apresentamos resultados para $\mu > 1$. A forma da distribuição nessa região será discutida a seguir.

A distribuição apresentada em (4.17) será válida desde que o grau médio esteja abaixo da condição crítica $\mu < 1$. Acima da condição crítica, $\mu > 1$, m_n como definido em (4.17) não pode ser normalizado, pois nesse regime a exponencial se tornaria crescente com o tamanho da componente n . O que acontece acima da condição crítica, na verdade, é o surgimento de uma componente gigante que ocupa uma fração não nula P_∞ do grafo.

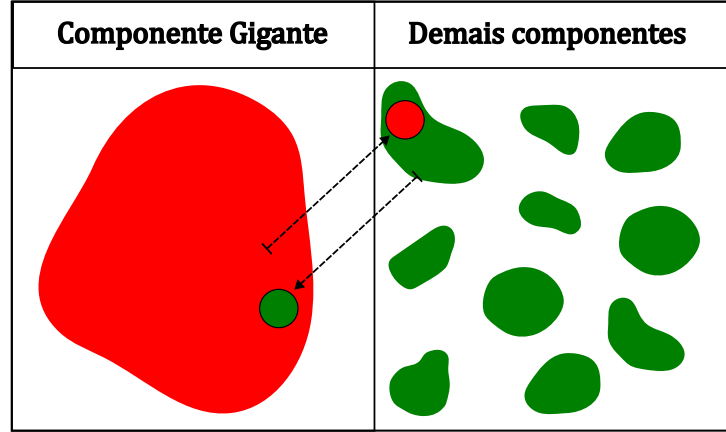
Para obter a fração P_∞ quando $\mu > 1$, precisamos investigar a entropia de um sistema composto, composto por um subsistema sendo a componente gigante e outro subsistema que inclui as componentes restantes, este processo é ilustrado na Fig. (20). A enumeração Ω do número de grafos acíclicos possíveis para este sistema composto é dada por

$$\Omega(N_\infty, N) = \Omega_R(N_R) \Omega_\infty(N_\infty) \frac{N!}{N_\infty! (N_R)!}, \quad (4.20)$$

onde Ω_R é a enumeração das possíveis configurações para os N_R vértices fora da componente gigante. Da mesma forma, Ω_∞ é a enumeração das possíveis configurações para a componente gigante e $N_\infty = N - N_R$ é o número de vértices dentro da componente gigante. O fator binomial precisa ser introduzido para levar em conta as possíveis maneiras pelas quais os N vértices distinguíveis podem ser divididos entre a componente gigante e as componentes restantes [68].

Precisamos garantir, assim como antes, que a entropia seja aditiva, ou seja, a entropia do sistema composto deve ser a soma das entropias da componente gigante e da entropia do resto do grafo, $S = S_\infty + S_R$. Isso leva a $S = \ln[\Omega/N!]$, $S_\infty = \ln[\Omega_\infty/N_\infty!]$ e $S_R = \ln[\Omega_R/N_R!]$.

Figura 20: Sistema composto pela componente gigante e as componentes restantes



Fonte: Produzido pelo autor. De modo similar ao da Fig. (17), definimos um sistema composto pela componente gigante e um subsistema que contém todas as componentes restantes. Os vértices podem ser permutados entre estes dois subsistemas, de modo análogo à troca de partículas.

4.3.1 A fração ocupada pela componente gigante

A enumeração das possíveis configurações para a componente gigante pode ser facilmente obtida a partir da fórmula de Calley [59], levando a:

$$S_\infty = \ln \left(\frac{N_\infty^{N_\infty-2}}{N_\infty!} \right) \approx N_\infty - \frac{5}{2} \ln(N_\infty), \quad (4.21)$$

onde a aproximação pode ser deduzida a partir da fórmula de Stirling para o fatorial. A partir disso, temos $dS_\infty/dN_\infty \approx 1 - 5/(2N_\infty) \approx 1$.

Também precisamos obter a variação da entropia das componentes restantes. Note que os vértices podem se mover dentro e fora da componente gigante, mas o número de componentes deve permanecer constante desde que μ permaneça constante. Para obter a entropia S_R do subsistema de vértices fora da componente gigante, aproveitamos a fórmula (4.18), definindo:

$$S_R(N_R, M-1) = S \left(N_R, \mu_R = 2 \left[1 - \frac{M-1}{N_R} \right] \right), \quad (4.22)$$

Onde N_R e μ_R são o número e o grau médio dos vértices fora da componente gigante, respectivamente. O número de componentes restantes é $M-1$, já que uma das componentes é a componente gigante. Ao diferenciar com relação a N_R , enquanto o número de componentes é mantido constante, pode-se mostrar que

$$\left(\frac{\partial S_R}{\partial N_R} \right)_M = \mu_R - \ln \mu_R, \quad (4.23)$$

onde se deve lembrar que $\mu_R = 2 \left(1 - \frac{M-1}{N_R}\right)$. Finalmente, temos

$$\left(\frac{\partial S}{\partial N_\infty}\right)_\mu = \frac{dS_\infty}{dN_\infty} - \left(\frac{\partial S_R}{\partial N_R}\right)_M \approx 1 - \mu_R + \ln \mu_R \leq 0. \quad (4.24)$$

Observe que a variação da entropia é nula quando $\mu_R = 1$, isto é, $[(\partial S/\partial N_\infty)]_{\mu_R=1} = 0$. No entanto, a segunda derivada também é nula neste ponto, $\left[(\partial^2 S/\partial N_\infty^2)_\mu\right]_{\mu_R=1} = 0$. Isso sugere que $\mu_R = 1$ não é um ponto de entropia máxima, mas sim um ponto de inflexão. No entanto, deve-se lembrar que $\mu_R \leq 1$, caso contrário a distribuição do tamanho das componentes não será normalizável. Como a entropia S sempre aumentará à medida que N_∞ diminui, a entropia máxima ocorrerá quando N_∞ for mínimo, ou quando N_R for máximo. Como mencionado, o número de vértices fora da componente gigante é limitado a ser menor do que duas vezes o número de componentes, $N_R \leq 2(M-1)$. Para maximizar a entropia S , portanto, N_R deve atingir seu valor máximo possível, levando a $\mu_R = 1$. Concluimos que, semelhante ao que acontece na condensação de Bose-Einstein [75, 76], onde as partículas fora do condensado apresentam uma distribuição crítica com um potencial químico que tende a zero, a condição de equilíbrio do conjunto de grafos acíclicos ocorrerá quando a distribuição do tamanho da componente fora da componente gigante for crítica, $\mu_R = 1$. Portanto, acima do ponto crítico $\mu > 1$, temos

$$m_n = (1 - P_\infty)N \frac{n^{n-2}e^{-n}}{n!}, \quad (4.25)$$

com os restantes $P_\infty N$ vértices colocados na componente gigante. Essa previsão é consistente com os resultados do algoritmo de Monte Carlo mostrados no painel secundário da Gráfico (2).

A dependência de P_∞ com o grau médio μ pode ser obtida de forma simples. Isso é feito observando que o grau médio é a média ponderada do grau esperado dos vértices dentro e fora da componente gigante,

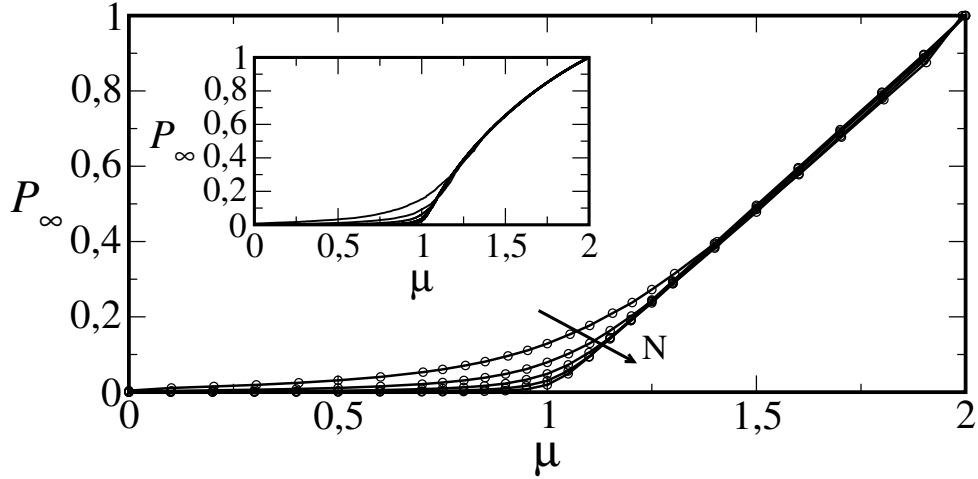
$$\mu = P_\infty \mu_\infty + (1 - P_\infty) \mu_R, \quad (4.26)$$

onde μ_∞ é o grau médio dos vértices dentro da componente gigante, e mencionamos que μ_R é o grau médio dos vértices restantes. Para componentes acíclicos, o grau médio de uma componente de tamanho n é dado por $2(1 - 1/n)$, e para uma componente gigante, com o tamanho indo para infinito, temos $\mu_\infty = 2$. Como os vértices restantes exibirão uma distribuição crítica, temos $\mu_R = 1$. A partir disso, concluimos que

$$P_\infty = \mu - 1. \quad (4.27)$$

O parâmetro de ordem acima do ponto crítico segue $P_\infty = (\mu - \mu_c)^\beta$, e temos o expoente crítico $\beta = 1$. Mais uma vez, esse valor do expoente crítico é o esperado no limite de campo médio [73], e concorda com o observado no modelo de grafo aleatório [72]. A Eq. (4.27) deve ser válida para qualquer grau médio na faixa $1 < \mu < 2$. Em $\mu = 2$ temos $P_\infty = 1$, esse limite indica uma única componente acíclico que compreende todos os vértices do grafo. Essas previsões são ilustradas pelos resultados do algoritmo de Monte Carlo apresentados na Gráfico (3), onde mostramos que à medida que o tamanho do sistema cresce, a fração ocupada pelo maior componente P_∞ se aproxima do comportamento previsto pela Eq. (4.27).

Gráfico 3: A fração ocupada pela maior componente P_∞ em função do grau médio μ .



Fonte: Produzido pelo autor. Resultados do algoritmo de Monte Carlo para grafos acíclicos de tamanhos $N = 2^8, 2^{10}, 2^{12}, 2^{14}$ e 2^{16} . À medida que os tamanhos dos sistemas aumentam, o comportamento se aproxima do esperado pela Eq. (4.27), com $P_\infty \rightarrow 0$ para $\mu < 1$ e $P_\infty = \mu - 1$ para $\mu > 1$. No painel secundário, mostramos resultados para um modelo em que os grafos acíclicos são criados pela adição sucessiva de conexões que não criam ciclos. Este modelo de grafos acíclicos difere do conjunto não tendencioso de grafos acíclicos [77].

Finalmente, podemos derivar a entropia do conjunto de grafos acíclicos acima da condição crítica, $\mu > 1$. Como mencionado, isso pode ser feito somando a entropia da componente gigante e a entropia das componentes restantes. A entropia da componente gigante deve ser dada pela Eq. (4.21), com $N_\infty = (\mu - 1)N$, enquanto a entropia dos demais pode ser obtida a partir da Eq. (4.18) com um número de vértices $N_R = (2 - \mu)N$ e um grau médio crítico para os vértices restantes $\mu_R = 1$. Com isso, pode-se derivar

$$S(N, \mu \geq 1) = N \left\{ 2 - \frac{\mu}{2} - \left(1 - \frac{\mu}{2} \right) \ln[N(2 - \mu)] \right\}, \quad (4.28)$$

onde negligenciamos os termos que são sublineares em N . Note que as Eqs. (4.18) e (4.28) são idênticas no limite $\mu = 1$, onde $S(N, \mu = 1) = N[3/2 - \ln(N)]$.

4.4 Comparação com outros modelos de grafos acíclicos

Abaixo da condição crítica, impor a ausência de ciclos não deve afetar as propriedades estatísticas dos grafos, e as propriedades do ensemble de grafos acíclicos aleatórios devem ser equivalentes às propriedades do ensemble de grafos aleatórios gerais, pelo menos até o ponto crítico. Acima da condição crítica, onde emerge uma componente gigante, esses dois conjuntos devem diferir, mas a classe de universalidade da transição deve seguir as previsões de campo médio em ambos os casos. Portanto, podemos comparar nossos resultados em relação ao modelo geral de grafos aleatórios, bem como em relação a diferentes modelos de grafos acíclicos. No entanto, isso deve ser feito observando que, acima da condição crítica, grafos aleatórios devem apresentar ciclos [72], e nossa abordagem, que é exata para o conjunto de grafos acíclicos aleatórios, pode não ser adequada para o modelo geral de grafos aleatórios. Na verdade, o modelo de grafo aleatório foi mostrado ser similar a um modelo diferente de grafos acíclicos (veja [77] e referências lá contidas), em que um grafo acíclico com N vértices e E conexões é obtido adicionando-se uma conexão que não cria um ciclo a um grafo acíclico com

N vértices e $E - 1$ conexões. vértices nos referiremos a este modelo modificado como modelo de crescimento. O modelo de crescimento de grafos acíclicos não apresenta os possíveis grafos com a mesma probabilidade e, portanto, produz médias diferentes daquelas do conjunto não tendencioso de grafos acíclicos. Isso pode ser visto no painel secundário da Gráfico (3), onde apresentamos resultados numéricos para o modelo de crescimento. Curiosamente, o modelo de crescimento tem uma relação mais próxima com o modelo de grafos aleatórios [77]. Vamos definir $\xi = 2(1 - M/N)$, lembrando que N e M são o número de vértices e componentes, respectivamente. Em grafos acíclicos, ξ pode ser relacionado ao grau médio, $\xi = \mu$. Para grafos aleatórios, no entanto, devido à possível presença de ciclos, a mesma relação não é verdadeira e $\xi < \mu$. Essa distinção é particularmente relevante acima da condição crítica, uma vez que abaixo da criticidade, grafos aleatórios não devem apresentar ciclos. A relação entre o modelo de crescimento e o modelo de grafos aleatórios é que o comportamento observado no painel secundário da Gráfico (3) é exatamente o que seria observado para grafos aleatórios em um gráfico de P_∞ vs ξ [77].

5 RELAÇÕES DE ESCALA E MÉTRICAS DE DISTÂNCIA EM CLUSTERS CRÍTICOS E SPANNING TREES

Neste capítulo investigamos as métricas de distância em clusters críticos, a distância química e a distância euclidiana. Primeiramente, apresentamos as relações entre essas métricas e sua relação com dimensão fractal do menor caminho. Em seguida, discutimos a construção de clusters críticos utilizando o algoritmo de Leath e a extração de spanning trees a partir desses clusters. São exploradas diferentes topologias de spanning trees em duas e três dimensões, incluindo a Árvore de Caminho Mínimo (MPT), Minimal Spanning Tree (MST), e variações da Busca em Profundidade (DFS e RFS). Por fim, estabelecemos relações de escala entre as métricas de distância para essas diferentes abordagens, permitindo uma análise comparativa de suas propriedades topológicas e determinação da dimensão fractal do menor caminho em cada uma delas.

5.1 Métricas de distância em clusters críticos

Na Seção 2.1.2, introduzimos o conceito de caminho em redes, o qual também se aplica a estruturas regulares, árvores e clusters críticos. Dentre todos os caminhos possíveis entre um par de vértices V_1 e V_2 , há um interesse especial no menor caminho entre eles, denominado **caminho mínimo** ou **distância química** ℓ . Ao longo deste trabalho, adotaremos o termo **distância química**. Esta pode ser definida como o menor caminho que conecta dois vértices, isto é, o número mínimo de arestas que compõem a trajetória entre V_1 e V_2 [78].

Em estruturas regulares, também podemos definir uma métrica de distância euclidiana r em relação a um vértice localizado na origem da estrutura, dada por

$$r = \sqrt{\sum_{i=1}^d r_i^2}, \quad (5.1)$$

onde d é a dimensão da estrutura na qual o objeto está imbuído, e r_i corresponde à distância em cada uma das d coordenadas entre o vértice de interesse e o vértice situado na origem da estrutura. A Fig. (21) traz um exemplo visual da caracterização destas duas métricas em um cluster crítico em duas dimensões.

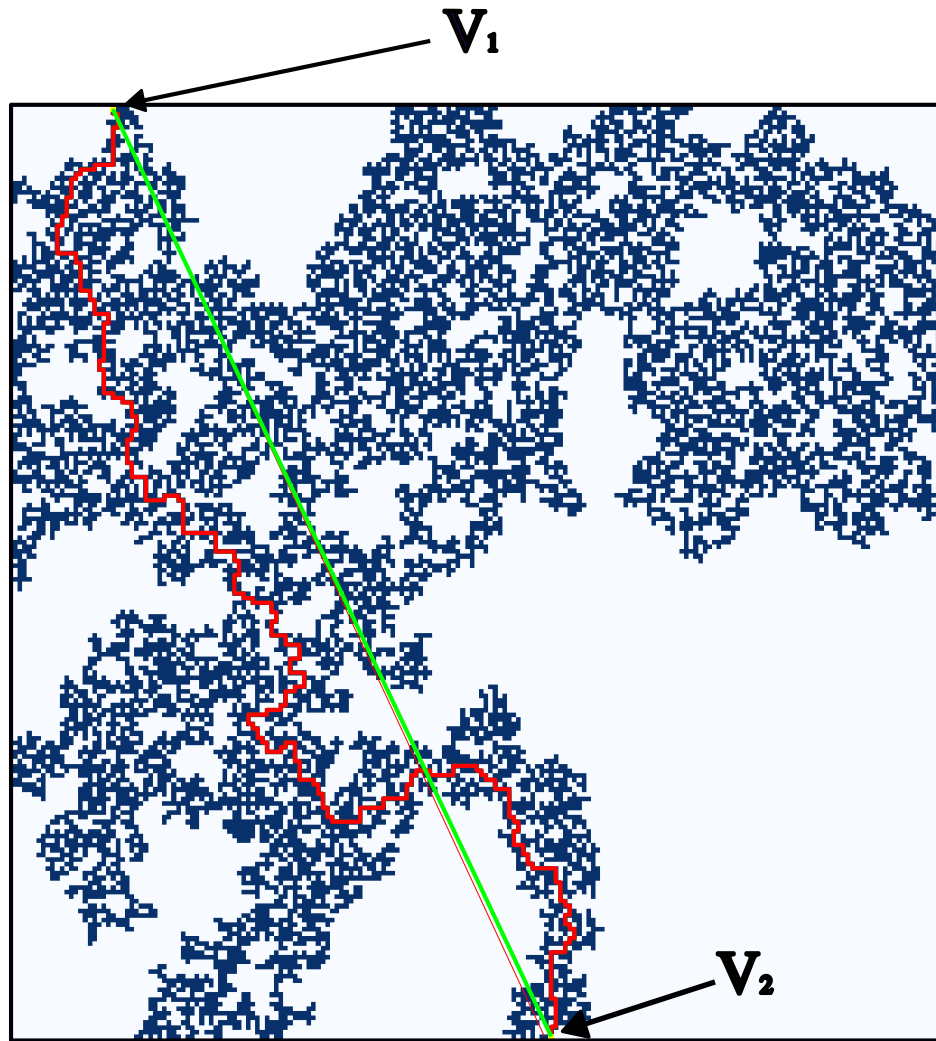
A distância química média está relacionada com a distância Euclidiana através da relação

$$\langle \ell \rangle \sim r^{d_{min}}. \quad (5.2)$$

Onde d_{min} é a dimensão fractal que caracteriza a distância química média, chamado de dimensão fractal do menor caminho [7, 79–88].

Além da dimensão fractal do menor caminho d_{min} , outras quantidades fractais desempenham um papel importante na caracterização geométrica de clusters críticos e suas árvores geradoras. Uma delas é a **dimensão fractal do caminho ótimo**, d_{opt} , que descreve o comprimento médio do caminho de menor custo em uma minimal spanning tree (MST), quando pesos aleatórios são atribuídos às arestas do cluster [89]. Esse caminho representa uma trajetória globalmente otimizada e, diferentemente do menor caminho topológico (distância química), in-

Figura 21: Comparação da Distância Química e Distância Euclidiana em 2d.



Fonte: Produzido pelo autor. No cluster 2d representado, escolhemos dois vértices sinalizados como V_1 e V_2 para exemplificar a ideia de Distância Química e Distância Euclidiana. A Distância Química é o menor caminho que liga estes dois vértices e é destacado pela linha vermelha, já a Distância Euclidiana entre o mesmo par de vértices é dada pelo comprimento da reta verde que os une, dada em unidades apropriadas.

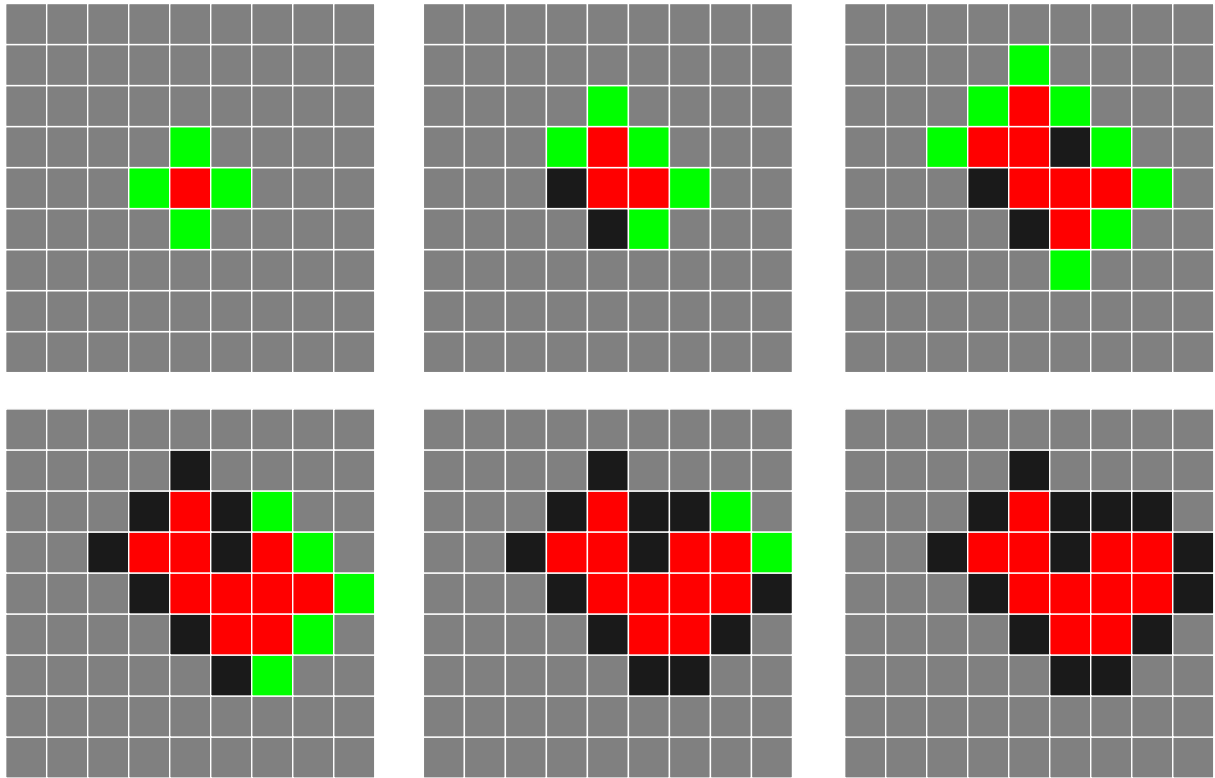
corpora a heterogeneidade induzida pelos pesos. Outra grandeza relevante é o **Percolation Hull Exponent**, d_{hull} , que caracteriza a dimensão fractal do contorno externo de um cluster percolante crítico. Essa borda é composta pelos sítios periféricos acessíveis ao longo de uma caminhada autoevitante que margeia o cluster, e seu valor é conhecido exatamente em duas dimensões, $d_{\text{hull}} = 4/3$, com base na correspondência entre percolação crítica e teorias conformes [90].

5.2 Crescimento de clusters críticos

Para gerar as amostras de clusters críticos utilizadas na investigação numérica deste trabalho, utilizamos o algoritmo de Leath [79,91]. Este é um método de crescimento estocástico

utilizado para gerar clusters críticos em sistemas de percolação, sendo aplicável em duas ou mais dimensões. Seu funcionamento baseia-se em uma exploração aleatória a partir de um sítio inicial, onde cada vizinho é adicionado ao cluster com uma determinada probabilidade p . O crescimento funciona do seguinte modo: começamos com um sítio inicial ocupado na origem de uma rede. Com uma probabilidade P , cada vizinho do sítio inicial é ativado e adicionado ao cluster em crescimento, enquanto com uma probabilidade $1 - P$, o vizinho é desativado e excluído do crescimento futuro. Uma vez que todos os vizinhos dos sítios ativos atuais são processados, o crescimento continua a partir dos sítios recém-ativados até que não seja possível adicionar mais sítios, ou seja, o processo de crescimento do cluster termina quando o conjunto de sítios ativos é confinado por um contorno externo de sítios inativos, como mostrado na Fig. (22).

Figura 22: Algoritmo de Leath em uma rede quadrada.



Fonte: Produzido pelo autor. A sequência de figuras exemplifica o crescimento de um cluster por meio do algoritmo de Leath. A cor cinza representa um sítio com estado indefinido, a cor vermelha indica um sítio ativo, a cor verde indica sítios indefinidos, mas que podem ser visitados na etapa seguinte, e marcados como ativos com probabilidade P , ou inativos com probabilidade $1 - P$ e, por fim, a cor preta indica sítios inativos. Observe que o processo se encerra quando um aglomerado de sítios ativos é confinado por sítios inativos.

5.3 Relações de escala em métricas de distância de spanning trees de clusters críticos

O algoritmo de Leath é uma implementação de um Breadth-First Search (BFS), que em português significa “busca em largura”. Podemos visualizar esse algoritmo como uma forma de construir uma spanning tree, considerando que, durante o processo de busca, incluímos uma aresta entre o sítio atualmente ativo e o vizinho que está sendo ativado. Chamaremos isso de Árvore de Caminho Mínimo (MPT - *Minimal Patch Tree*). Nesta árvore, cada sítio está conectado à origem da rede pelo caminho mínimo no cluster crítico original.

Uma quantidade chave e de interesse é a probabilidade condicional $P(r|\ell)$, que descreve a probabilidade de um nó estar a distância euclidiana r a partir da origem dado que este esteja a uma distância química ℓ desta mesma origem [92]. Como a definição da MPT não introduz nenhuma escala característica, a invariância de escala do cluster crítico deve ser preservada na MPT. Portanto, esperamos que $P(r|\ell)$ obedeça a uma lei de escala onde a distância química ℓ é a única escala relevante [93],

$$P(r|\ell) = \ell^{-\theta} g_x \left(\frac{r}{\ell^\theta} \right), \quad (5.3)$$

onde $g(x)$ é uma função de escala universal. Essa relação de escala é suportada pelos resultados numéricos mostrados na Fig. (4). Também medimos o número de sítios ativados $N(\ell)$ como função de ℓ , e no regime crítico esperamos que a média estatística $\langle N(\ell) \rangle$ siga a forma de escala

$$N(\ell) = \ell^{d_\ell - 1}, \quad (5.4)$$

onde d_ℓ é a dimensão de propagação [94], e deve estar relacionada à dimensão fractal d_f do cluster e à dimensão fractal do caminho mínimo por $d_\ell = d_f/d_{min}$ [7]. Isso significa que a inclinação $\langle N_\ell \rangle$ vs. ℓ deve ser dada por $d_f/d_{min} - 1$. Essa dependência em lei de potência é mostrada no inset mais à direita do Gráfico (4).

O expoente θ na Eq. (5.3) deve ser identificado com $1/d_{min}$ e pode ser determinado a partir da inclinação α do gráfico $\langle N_\ell \rangle$ vs. ℓ , presente no inset inferior do Gráfico (4), através da relação $\theta = (d_f/\alpha + 1)$. De fato, é com esta metodologia que encontramos, para todas as topologias estudadas neste trabalho, o expoente que governa a lei de escala para $P(r|\ell)$. Os resultados do Gráfico. (4) são consistentes com essas relações. Em [24, 26], foi sugerido que a relação $d_\ell = d_f/d_{min}$ não deveria se manter. De fato, como mostramos a seguir, essa relação pode ser deduzida a partir da hipótese de uma relação de escala universal, na forma da Eq. (5.3).

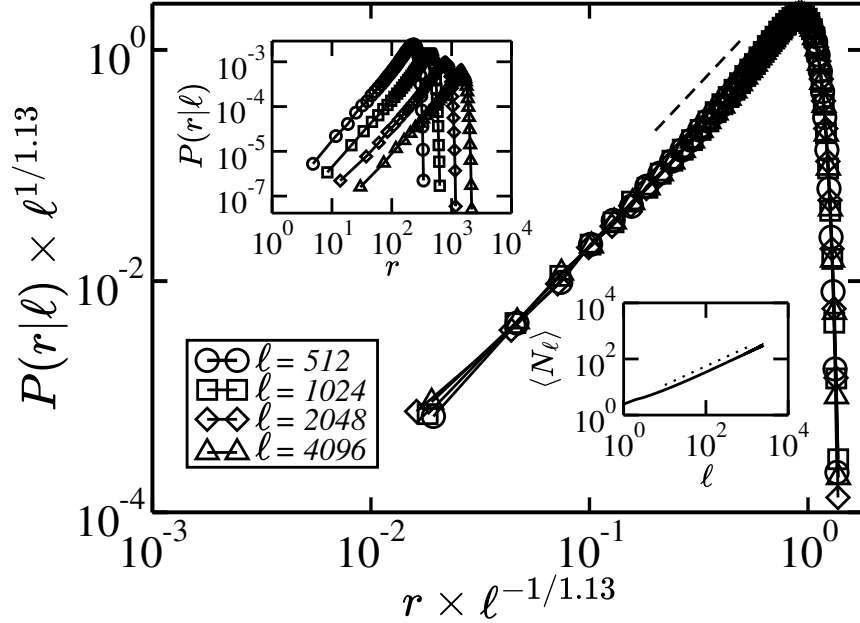
5.4 As topologias de spanning tree utilizadas: MPT, MST, DFS e RFS.

Além da estratégia de busca em largura, outras estratégias podem ser utilizadas para construir spanning trees do cluster de percolação crítica. Evidentemente, a adoção de diferentes estratégias deve levar a diferentes conjuntos de distâncias químicas ℓ na spanning tree. Por exemplo, podemos construir uma spanning tree por meio da Busca em Profundidade (DFS - *Depth-First Search*), que visita todos os sítios do cluster a partir da origem explorando cada ramo o máximo possível antes de retroceder até a última ramificação.

Realizamos a busca em profundidade de duas maneiras distintas. Quando nossa busca atinge uma bifurcação, no DFS escolhemos aleatoriamente uma direção (1ª maneira) ou sempre escolhemos o caminho inexplorado mais à direita (2ª maneira). Chamamos essa última estratégia de Busca Prioritária à Direita (RFS - *Right-First Search*). Cada uma dessas estratégias leva à construção de uma árvore com características topológicas típicas. Um exemplo das árvores MPT, DFS e RFS em um pequeno cluster de percolação é ilustrado na Fig. (23), e de modo mais prático na Fig. (24).

Outra possível maneira de construir uma árvore geradora para o cluster de percolação crítica é identificar a chamada Minimal spanning tree (MST) [96, 97]. Nesse procedimento, pe-

Gráfico 4: Relações na MPT



Fonte: Produzido pelo autor. Para $d = 2$, o painel principal mostra a probabilidade reescalada $P(r|\ell), \ell^{1/\theta}$ plotada em função da distância reescalada $r \times \ell^{-1/\theta}$ no aglomerado percolante na criticalidade ($p_c = 0.592$) [95]. O aglomerado foi gerado usando o método de Leath, iniciando a partir de um sítio semente ativo no centro da rede, e a MPT foi extraída. A dimensão fractal é $d_f = 1.89$ e o expoente de escala obtido é $\theta = 1.130$. Os símbolos correspondem a diferentes distâncias químicas nas árvores geradoras: 512 (círculo), 1024 (quadrado), 2048 (losango) e 4096 (triângulo). A linha tracejada possui inclinação 2, correspondendo à fase crescente da probabilidade reescalada. O inset à esquerda mostra a probabilidade original não reescalada $P(r|\ell)$. O inset à direita exibe o número médio de sítios a uma distância química ℓ , com uma inclinação $\alpha = 0.672 \pm 0.005$ (inclinação da linha pontilhada). Esse valor foi utilizado para determinar o expoente de escala θ por meio de $\theta = (d_f/\alpha + 1) = 1.130 \pm 0.003$. Esses resultados são baseados em 50000 configurações de aglomerados em uma rede quadrada com lado $L = 10^4$.

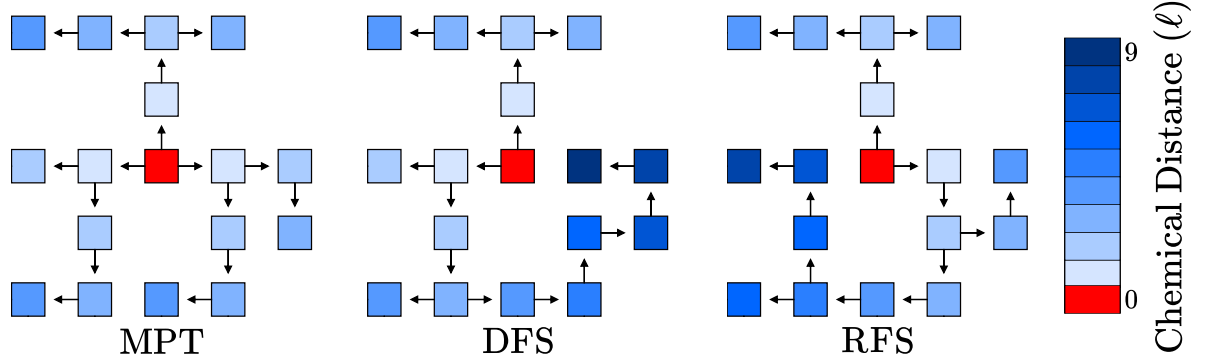
sos aleatórios são atribuídos a cada aresta do cluster percolante, e a árvore geradora é construída a partir do sítio-semente localizado no centro da rede. O crescimento prossegue iterativamente incorporando sítios vizinhos, selecionando a aresta com o menor peso a cada etapa, garantindo que a árvore se expanda ao longo do caminho de menor peso total. Esse processo continua até que todos os sítios do cluster formem uma árvore.

5.5 Relações de escala em métricas de distância de spanning trees de clusters críticos para todas as topologias utilizadas: MPT, MST, DFS e RFS.

Nenhuma das diferentes estratégias mencionadas acima introduz escalas características. Assim, para todos os casos, a distribuição das distâncias euclidianas r deve apresentar uma relação de escala semelhante à Eq. (5.3). De fato, o mesmo deve valer para a distribuição $P(\ell|r)$, das distâncias químicas através da spanning tree, dada uma distância euclidiana. Portanto, devemos ter as seguintes relações de escala:

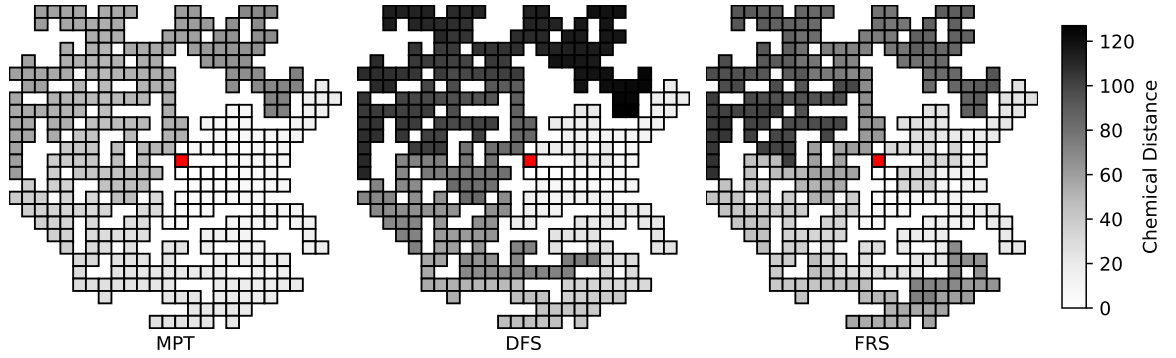
$$P(\ell|r) = r^{-\theta_1} g_\ell \left(\frac{\ell}{r^{\theta_1}} \right) \quad \text{e} \quad P(r|\ell) = \ell^{-\theta_2} g_r \left(\frac{r}{\ell^{\theta_2}} \right). \quad (5.5)$$

Figura 23: Ilustração de um aglomerado percolante em 2d.



Fonte: Produzido pelo autor. Ilustração de um aglomerado percolante pictórico em uma rede quadrada com lado $L = 5$. Para cada caso, o mesmo aglomerado foi percorrido a partir do sítio semente (sítio vermelho) utilizando uma das três estratégias, representadas pelas direções das setas. Cada estratégia leva a uma árvore diferente e a um conjunto distinto de distâncias químicas ℓ , de acordo com a barra de cores. O primeiro caso é a Busca em Largura (Breadth-First Search, BFS), uma estratégia que visita todos os sítios na profundidade atual antes de avançar para os sítios no próximo nível de profundidade. Usando essa estratégia, obtemos a Árvore de Caminho Mínimo (Minimal Path Tree, MPT). No segundo caso, o aglomerado foi percorrido utilizando a Busca em Profundidade (Depth-First Search, DFS), que visita todos os sítios explorando cada ramo o máximo possível antes de retroceder até o último ponto de bifurcação. Na DFS, quando há mais de uma opção de caminho, a escolha da ramificação a seguir primeiro é feita aleatoriamente. Por fim, o terceiro caso é a Busca Prioritária à Direita (Right First Search, RFS), uma DFS enviesada onde, diferentemente da DFS normal, ao se deparar com mais de uma opção de caminho, a decisão é sempre tomar a ramificação à direita. Na RFS, a única escolha arbitrária ocorre ao determinar de qual vizinho não visitado do sítio semente iniciar um novo ramo.

Figura 24: Ilustração prática de um cluster crítico em 2d.



Fonte: Produzido pelo autor. Comportamento típico da distancia química em MPT, DFS e RFS em um pequeno cluster crítico de lado $L = 26$. A barra de cores caracteriza a escala de distancia permite verificar a diferença na distribuição de distancia em um caso prático (efetivamente gerado a partir do algoritmo de Leath). O vértice destacado em vermelho é a origem da estrutura e corresponde à distancia química $\ell = 0$. Observe que, assim como na Fig. (23), os três clusters são exatamente o mesmo cluster crítico, apenas percorridos por cada uma das diferentes estratégias.

A partir dessas relações, temos que a distância química média segue $\bar{\ell}(r) = c_\ell r^{\theta_1}$, com $c_\ell = \int x g_\ell(x) dx$. Portanto, para o caso da MPT, definida anteriormente, o expoente θ_1 será a dimensão fractal do caminho mínimo d_{min} . Da mesma forma, $\bar{r}(\ell) = c_r \ell^{\theta_2}$, e $c_r = \int x g_r(x) dx$. Um leitor informado pode se surpreender com o fato de que, neste ponto, assumimos que os expoentes θ_1 e θ_2 são independentes. Como demonstraremos a seguir, esses dois expoentes, na verdade, precisam estar relacionados entre si.

Definimos o número $\mathcal{N}(\ell, r)$ de sítios que estão simultaneamente a uma dada distância

química e euclidiana da origem. Temos que

$$\mathcal{N}(\ell, r) = P_\ell(\ell|r)N_r(r) = P_r(r|\ell)N_\ell(\ell), \quad (5.6)$$

onde $N_r(r)$ e $N_\ell(\ell)$ são o número de sítios a uma dada distância euclidiana e química em relação à origem, respectivamente. Essa equação é uma expressão do teorema de Bayes e, como mostraremos, essa identidade, juntamente com as formas de escala da Eq. (5.5), pode ser usada para relacionar os expoentes d_f , d_ℓ , θ_1 e θ_2 .

Consideramos que $N_r(r) = r^{d_f-1}$ e $N_\ell(\ell) = \ell^{d_\ell-1}$. O expoente d_ℓ aqui tem o mesmo papel que na Eq. (5.4), porém, devemos notar que o valor desse expoente depende da spanning tree específica que está sendo considerada. A partir das Eqs. (5.5) e (5.6), obtemos:

$$r^{d_f-\theta_1-1} g_\ell\left(\frac{\ell}{r^{\theta_1}}\right) = \ell^{d_\ell-\theta_2-1} g_r\left(\frac{r}{\ell^{\theta_2}}\right). \quad (5.7)$$

Agora definimos $f_r(x) = g_r(x^{-\theta_2})$, o que leva a $g_r\left(\frac{r}{\ell^{\theta_2}}\right) = f_r\left(\frac{\ell}{r^{1/\theta_2}}\right)$. Note que a função $g_r(x)$ molda a distribuição $P_r(r|\ell)$, que possui um modo bem definido e é não-monotônica. Portanto, podemos confiar que a função $f_r(x)$ também é não-monotônica.

Podemos também definir $\zeta = \frac{\ell}{r^{\theta_1}}$, e reescrever a Eq. (5.7) como

$$\frac{g_\ell(\zeta)}{f_r\left(\zeta r^{\theta_1-\frac{1}{\theta_2}}\right)} = \zeta^{d_\ell-\theta_2-1} r^\gamma, \quad (5.8)$$

onde $\gamma = \theta_1(d_\ell - \theta_2 - 1) - (d_f - \theta_1 - 1)$. Dadas suas definições, as variáveis ζ e r são independentes, e podemos considerar um caso onde r varia enquanto ζ permanece constante. O lado esquerdo sugere uma dependência não-monotônica de r através de $f_r(\zeta r^{\theta_1-1/\theta_2})$, enquanto o lado direito apresenta uma dependência monotônica de potência em r . Consequentemente, para que a Eq. (5.8) seja consistente, precisamos definir os expoentes de r como sendo ambos zero. Verifica-se que isso leva a:

$$\theta_1 = \frac{1}{\theta_2}, \quad d_\ell = \frac{d_f}{\theta_1}. \quad (5.9)$$

Para o caso da MPT, temos $\theta_1 = d_{min}$. Foi alegado em [24, 26] que, para a MPT em dimensão $d = 2$, a Eq. (5.9) não é válida, com $d_\ell < d_f/d_{min}$. Além disso, [24, 26] propõem um intervalo para o expoente d_ℓ e indicam que resultados numéricos para $d = 4$ e $d = 5$ colocam d_{min} fora desse intervalo. Como mostramos, em qualquer dimensão, uma quebra das relações (5.9) seria incompatível com a hipótese de escala para as probabilidades condicionais (5.5).

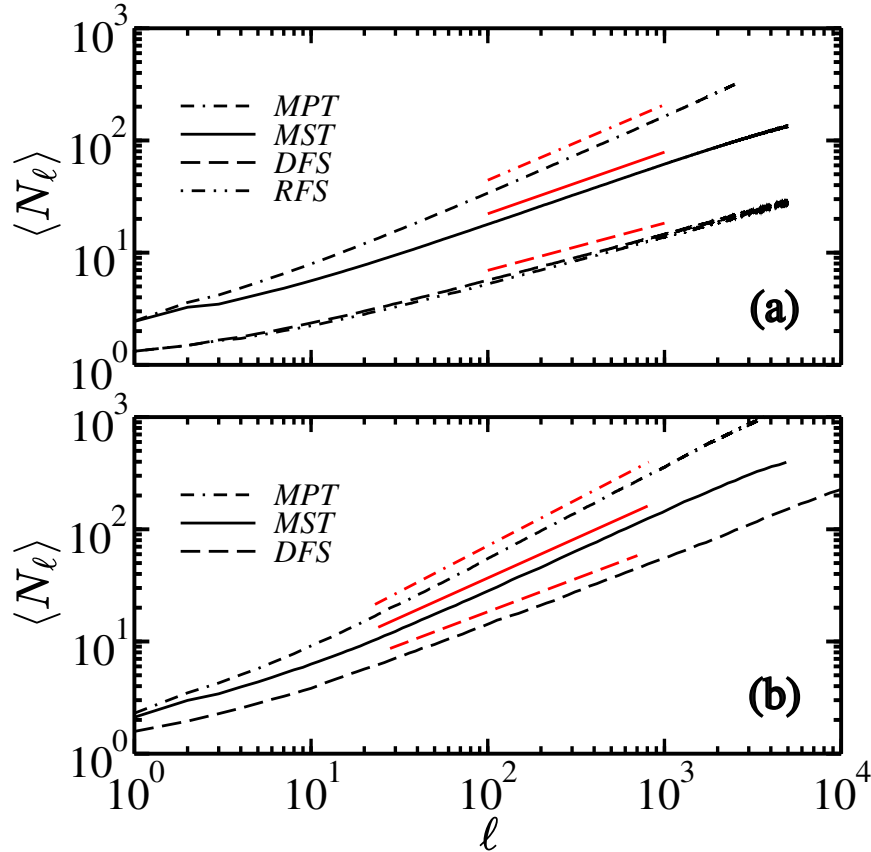
A Eq. (5.9) nos mostra que os expoentes que controlam a escala das distribuições condicionais na Eq. (5.5) estão, de fato, relacionados. Além disso, esses expoentes podem ser obtidos a partir da função $N_\ell(\ell) \sim \ell^{d_\ell-1}$, com $\theta_1 = d_f/d_\ell$. Isso nos fornece uma maneira indireta de obter numericamente a dimensão fractal da distância química média, $\bar{\ell}(r) \sim r^{\theta_1}$, examinando a forma de $N_\ell(\ell)$. A seguir, fazemos isso para cada uma das possíveis estratégias de construção de uma spanning tree.

5.6 Resultados numéricos

Começamos gerando uma amostra de aglomerados críticos usando o algoritmo de Leath [79,91]. Cada amostra cresce até a distância mínima $\ell_{min} = 512$, e amostras que não atinjam esse tamanho são descartadas. O tamanho da rede usada excede 512, portanto, na prática, essa distância química máxima é a única restrição de escala do crescimento do aglomerado. Para cada aglomerado, obtemos as múltiplas árvores geradoras mínimas seguindo cada uma das estratégias descritas anteriormente. As médias foram obtidas a partir de 10^5 amostras.

O Gráfico (5) apresenta nossos resultados para $N_\ell(\ell)$ para cada estratégia, em 2 e 3 dimensões. Os declives de $N_\ell(\ell)$ no gráfico log-log nos fornecem os expoentes d_ℓ , e usando os valores de $d_f = 91/48$, para $d = 2$, e $d_f = 2.52$ para $d = 3$, obtemos os valores de $\theta_1 = d_f/d_\ell$ apresentados na Tabela I.

Gráfico 5: Número médio de sítios na ℓ -ésima camada



Fonte: Produzido pelo autor. Para $d = 2$ (a) e $d = 3$ (b), o número médio de sítios na ℓ -ésima casca $\langle N_\ell \rangle$ é apresentado em função da distância química ℓ . (a) A inclinação α obtida a partir dos dados numéricos em $d = 2$ é 0.672 ± 0.005 para a MPT, 0.540 ± 0.005 para a MST, e 0.419 ± 0.005 para a DFS e para a RFS. Assim, os valores obtidos para $\theta_1 = d_f/(\alpha + 1)$ foram 1.130 ± 0.003 para a MPT, 1.227 ± 0.004 para a MST e 1.331 ± 0.004 para a DFS e a RFS. Os expoentes para a RFS e a DFS são consistentes com o expoente do contorno externo em $d = 2$. (b) Para $d = 3$, a inclinação obtida a partir dos dados numéricos é 0.832 ± 0.005 para a MPT, 0.700 ± 0.005 para a MST e 0.600 ± 0.005 para a DFS. Assim, os valores obtidos para o expoente de escala θ_1 foram 1.375 ± 0.003 para a MPT, 1.482 ± 0.004 para a MST e 1.575 ± 0.008 para a DFS. As médias foram obtidas a partir de 10^5 configurações de aglomerados para a MST e 50000 configurações de aglomerados para os demais casos.

Tabela 1: Estimativa numérica da dimensão fractal das ramificações da árvore de spanning (θ_1) em 2 e 3 dimensões.

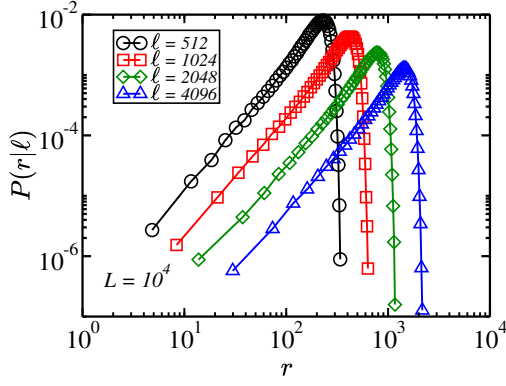
	2 d	3 d
MPT	1.130 ± 0.003	1.375 ± 0.003
MST	1.227 ± 0.004	1.482 ± 0.004
DFS	1.331 ± 0.004	1.575 ± 0.008
RFS	1.331 ± 0.004	

Em todos os casos, o expoente θ_1 fornece a dimensão fractal das ramificações da spanning tree. Para o caso da MPT, tanto em $d = 2$ quanto em $d = 3$, os valores obtidos para o expoente de escala θ_1 concordam com os esperados para a dimensão fractal do caminho mínimo $d_{min} = 1.130$ [7] e 1.375 [7], respectivamente. Da mesma forma, para a MST obtemos valores consistentes com a dimensão do caminho ótimo, $d_{opt} = 1.22$ [89] e 1.482 .

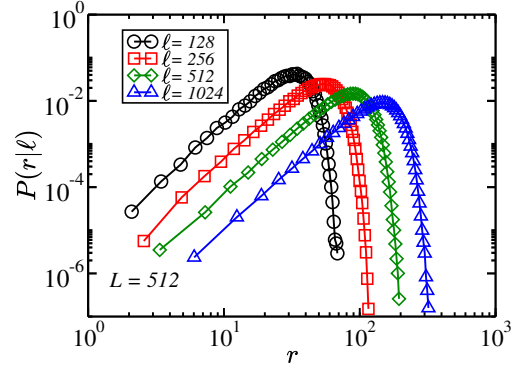
No caso do DFS e RFS, em $d = 2$, obtemos $\theta_1 = 1.331$, um valor consistente com a dimensão do casco externo $d_{hull} = 4/3$ [90]. Em três dimensões, o casco externo tem a mesma dimensão fractal do aglomerado crítico. A Busca Prioritária à Direita não é definida em $d = 3$, mas para o DFS obtemos um expoente $\theta_1 = 1.575$. Esse parece ser um novo expoente, que coincide com d_{hull} apenas em duas dimensões.

Para ilustrar a utilidade de nossa abordagem, no conjunto de gráficos. (6) apresentamos o conjunto de todos os histogramas em 2 e 3d pra a probabilidade condicional $P(r|\ell)$ para todas as topologias estudadas. Já no Gráfico (7), mostramos o resultado de escala da probabilidade $P(r|\ell) \times \ell^{\theta_2}$ vs $r \times \ell^{\theta_2}$ para cada uma dessas topologias, onde utilizamos os valores da Tabela I e o fato de que $\theta_2 = 1/\theta_1$. Essas escalas fornecem um suporte adicional aos expoentes de escala obtidos.

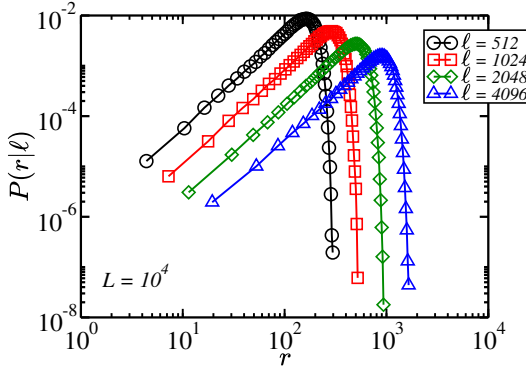
Gráfico 6: Probabilidade condicional $P(r|\ell)$ em 2 e 3 dimensões para todas as topologias estudadas (MPT, MST, DFS, FRS).



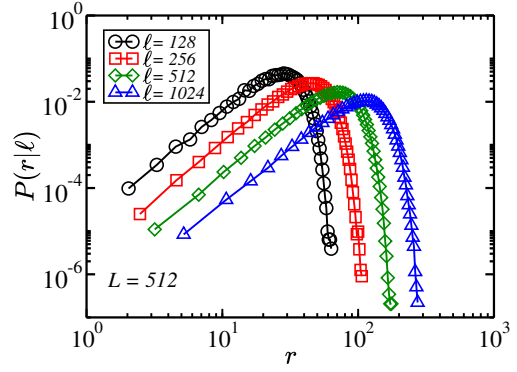
(a) MPT 2d.



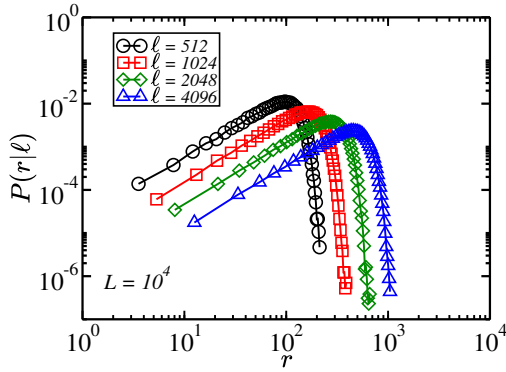
(b) MPT 3d.



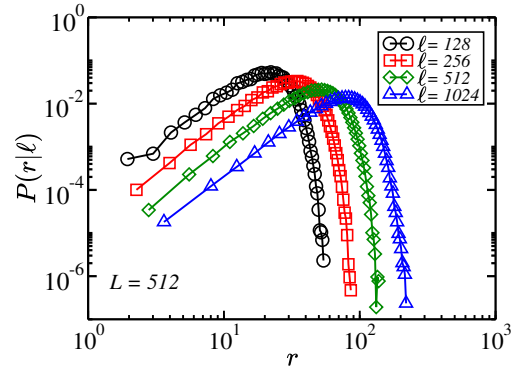
(c) MST 2d.



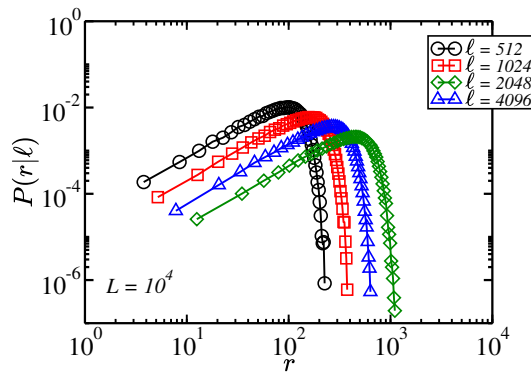
(d) MST 3d.



(e) DFS 2d.

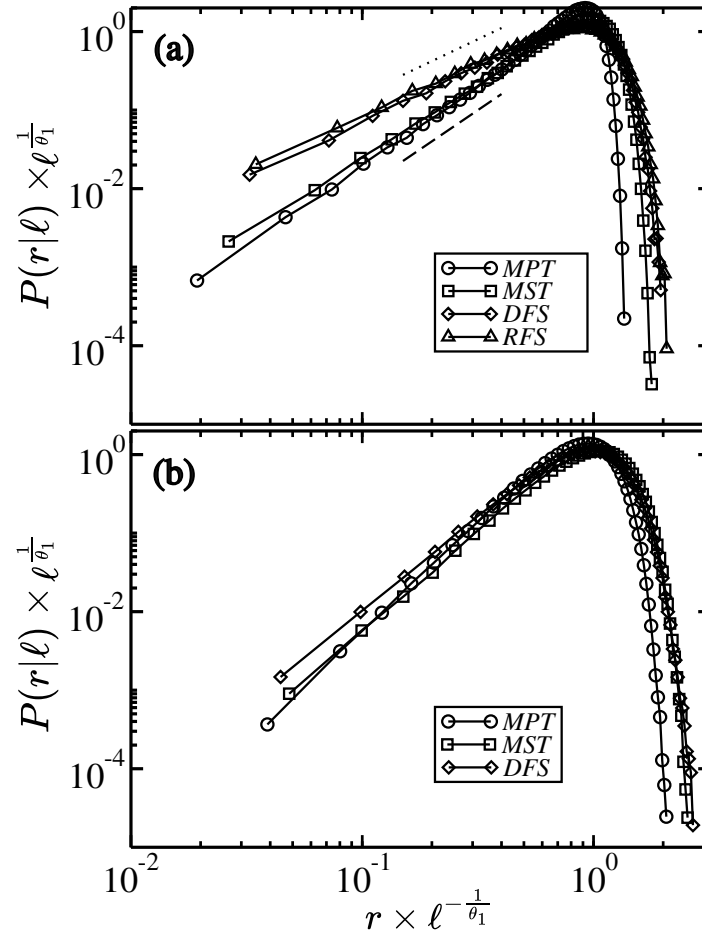


(f) DFS 3d.



(g) FRS 2d.

Gráfico 7: probabilidade condicional $P(r|\ell)$ reescalada para todas as topologias estudadas, em 2 e 3 d.



Fonte: Produzido pelo autor. Gráfico reescalado da probabilidade condicional $P(r|\ell)$ para árvores geradoras do aglomerado percolante crítico. Esses resultados referem-se à Árvore de Caminho Mínimo (MPT) (círculo), Árvore Geradora Mínima (MST) (quadrado), Busca em Largura (DFS) (losango) e busca prioritária à direita, Right First Search (RFS) (triângulo). Os resultados para $d = 2$ são apresentados em (a) e para $d = 3$ em (b). A busca prioritária à direita não pode ser definida em $d = 3$. Os aglomerados críticos amostrados foram obtidos usando o algoritmo de Leath. Em ambos os casos, $d = 2$ e $d = 3$, apenas os aglomerados que crescem até uma distância mínima de 512 são considerados, sem impor tamanho fixo para a rede. As médias foram obtidas a partir de 10^5 aglomerados. Os valores utilizados para o expoente de escala θ_1 em $d = 2$ foram 1.130 para a MPT, 1.227 para a MST e 1.331 para a DFS e a RFS. A fase crescente da probabilidade reescalada para a MPT e a MST possui inclinação aproximadamente igual a 2, representada pela linha pontilhada, enquanto para a DFS e a RFS a inclinação é 1.4, representada pela linha tracejada. Para $d = 3$, os valores utilizados para o expoente de escala θ_1 foram 1.375 para a MPT, 1.482 para a MST e 1.575 para a DFS. Em $d = 3$, a fase crescente da probabilidade reescalada apresenta inclinação aproximadamente igual a 2.7 para a MPT, 2.53 para a MST e 2.31 para a DFS.

Os resultados apresentados neste capítulo fornecem uma caracterização detalhada das relações de escala entre distâncias químicas e euclidianas em árvores geradoras de *clusters* críticos, revelando como diferentes estratégias de construção de *spanning trees* impactam suas propriedades geométricas e críticas. A demonstração da relação recíproca entre os expoentes θ_1 e θ_2 , bem como a consistência dos resultados numéricos com a hipótese de escala, reforça a robustez do formalismo adotado e da estratégia de obter os expoentes de escala a partir do comportamento do número de nós ativos para dadas distâncias químicas.

6 CONCLUSÕES

Neste trabalho, investigamos a estrutura e o comportamento crítico de redes complexas por meio de duas abordagens complementares, ambas fundamentadas na teoria da percolação. Embora tratem de sistemas distintos — grafos aleatórios acíclicos e redes regulares com árvores geradoras críticas —, os dois estudos compartilham a motivação central de compreender como características topológicas influenciam a emergência de fenômenos críticos. A primeira abordagem tem natureza combinatória e analítica, enquanto a segunda é essencialmente geométrica e baseada em simulações numéricas. Em conjunto, esses estudos revelam aspectos estruturais fundamentais de sistemas complexos, permitindo caracterizar tanto a formação de componentes gigantes em grafos acíclicos quanto as leis de escala que governam a organização espacial de árvores críticas. A seguir, detalhamos os principais resultados obtidos em cada uma dessas frentes.

Na primeira parte do estudo, mostramos como obter uma enumeração combinatória exata para o conjunto de grafos acíclicos aleatórios, dada uma distribuição particular de tamanhos de componentes. Esta enumeração levou a uma definição de entropia para este conjunto, que quando maximizada nos dá a distribuição de tamanhos de componentes na condição de equilíbrio. A partir disso obtivemos a entropia de equilíbrio como função do número de nós e grau médio do grafo. Nossos resultados indicam o início de uma transição de fase, onde uma componente gigante emerge. Tratando o grafo após a condição crítica como um sistema composto, fomos capazes de determinar o tamanho da componente gigante procurando a condição de equilíbrio do sistema composto. Nossa abordagem nos permite determinar o início da transição crítica, bem como os expoentes críticos associados. Embora alguns dos resultados que obtivemos sejam conhecidos [19–21, 72], os métodos e a linguagem que usamos devem ser mais familiares aos pesquisadores no campo da Física Estatística, bem como de modo mais amplo aos estudantes dos cursos de Ciências Exatas e da Terra. Esperamos que esta solução para o problema de grafos acíclicos aleatórios possa ser expandida em novas soluções para outros problemas gerais no campo da ciência de redes.

Na segunda parte deste estudo, investigamos a relação entre distâncias químicas e euclidianas em árvores de percolação crítica em duas e três dimensões, explorando diferentes estratégias para construção dessas árvores: Árvore de Caminho Mínimo (MPT), Busca em Profundidade (DFS), Busca Prioritária à Direita (RFS) e Árvore Geradora Mínima (MST). Demonstramos que a distribuição de probabilidade condicional $P(r|\ell)$ obedece a uma lei de escala universal, caracterizada por expoentes críticos que dependem da topologia da árvore considerada. Nossos resultados confirmam que, para a MPT, o expoente de escala θ_1 coincide com a dimensão fractal do caminho mínimo d_{min} , em concordância com resultados estabelecidos na literatura. Para a MST, obtivemos valores consistentes com a dimensão do caminho ótimo, enquanto no caso da DFS e RFS em $d = 2$, o expoente corresponde à dimensão do contorno externo (d_{hull}). Em três dimensões, a DFS apresenta um novo expoente crítico ($\theta_1 = 1.575$), sugerindo um comportamento distinto em dimensões superiores. Além disso, derivamos relações gerais entre os expoentes θ_1 , θ_2 , d_f e d_ℓ , mostrando que a validade da hipótese de escala exige que $\theta_1 = 1/\theta_2$ e que ao contrário do que encontramos em [24, 26] em qualquer dimensão, uma

quebra das relações (5.9) seria incompatível com a hipótese de escala para as probabilidades condicionais (5.5). Enfatizamos, portanto, que a identidade $d_\ell = d_f/d_{\min}$ não é meramente uma observação empírica, mas decorre diretamente das formas de escala e do teorema de Bayes. É uma consequência estrita da própria hipótese de escala. Portanto, qualquer violação dessa identidade implicaria uma inconsistência no arcabouço teórico. Essas relações foram verificadas numericamente, reforçando a consistência do formalismo proposto.

Este trabalho fornece uma compreensão profunda da estrutura geométrica de aglomerados de percolação crítica e das propriedades universais de diferentes árvores geradoras, com implicações para o estudo de transporte em meios desordenados, condutividade em materiais heterogêneos e dinâmica em redes complexas. Estudos futuros poderão explorar a generalização desses resultados para outras geometrias e modelos de percolação.

APÊNDICE A – TEOREMA DA INVERSÃO DE LAGRANGE E A SÉRIE DE TAYLOR DA FUNÇÃO

$$g(x)e^{-g(x)} = x$$

O teorema da inversão de Lagrange diz: Dado uma função $f(g)$, que obedece $f(0) = 0$, além de $\frac{df}{dg}|_{g=0} \neq 0$, se

$$x = f(g),$$

então

$$g(x) = \sum_{k=1}^{\infty} \left[\frac{d^{k-1}}{dg^{k-1}} \left(\frac{g'}{f(g')} \right) \right]_{g'=0} \frac{x^k}{k!}.$$

Podemos usar isso para encontrar a série de Taylor de $g(x)$, onde

$$g(x)e^{-g(x)} = x. \quad (\text{A.1})$$

Sabendo que $f(a=0) = 0 = x_0$, escrevemos:

$$g(x) = a + \sum_{k=1}^{\infty} g_k \frac{(x-x_0)^k}{k!} = \sum_{k=1}^{\infty} g_k \frac{(x)^k}{k!}$$

, onde:

$$g_k = \lim_{g \rightarrow 0} \left[\frac{d^{k-1}}{dg^{k-1}} \left(\frac{g-a}{f(g)-f(a)} \right)^k \right] = \lim_{g \rightarrow 0} \left[\frac{d^{k-1}}{dg^{k-1}} \left(\frac{g}{f(g)} \right)^k \right].$$

Dessa forma, temos,

$$\begin{aligned} g(x) &= \sum_{k=1}^{\infty} \lim_{g \rightarrow 0} \left[\frac{d^{k-1}}{dg^{k-1}} \left(\frac{g}{f(g)} \right)^k \right] \frac{(x)^k}{k!} \\ g(x) &= \sum_{k=1}^{\infty} \lim_{g \rightarrow 0} \left[\frac{d^{k-1}}{dg^{k-1}} \left(\frac{g}{ge^{-g}} \right)^k \right] \frac{(x)^k}{k!} \\ g(x) &= \sum_{k=1}^{\infty} \lim_{g \rightarrow 0} \left[\frac{d^{k-1}}{dg^{k-1}} (e^{gk}) \right] \frac{(x)^k}{k!} \\ g(x) &= \sum_{k=1}^{\infty} \lim_{g \rightarrow 0} \left[(k)^{k-1} e^{gk} \right] \frac{(x)^k}{k!} \\ g(x) &= \sum_{k=1}^{\infty} \left[(k)^{k-1} \right] \frac{(x)^k}{k!}. \end{aligned} \quad (\text{A.2})$$

Podemos mostrar ainda que vale a relação $\frac{dg(x)}{dx} = \frac{e^{g(x)}}{1-g(x)}$, fazendo

$$\begin{aligned} g(x)e^{-g(x)} &= x \\ \frac{d}{dx}(g(x)e^{-g(x)}) &= \frac{d}{dx}x \\ g'(x)e^{-g(x)} - g(x)g'(x)e^{-g(x)} &= 1 \\ g'(x)\left(e^{-g(x)} - g(x)e^{-g(x)}\right) &= 1. \end{aligned} \tag{A.3}$$

Resolvendo a equação A.3 para $g'(x)$, temos

$$\begin{aligned} g'(x) &= \frac{1}{e^{-g(x)} - g(x)e^{-g(x)}} \\ g'(x) &= \frac{1}{e^{-g(x)}(1 - g(x))} \\ g'(x) &= \frac{e^{g(x)}}{(1 - g(x))}. \end{aligned} \tag{A.4}$$

O resultado A.4 serve para mostrar que $\frac{d}{dx}\left(g(x)\left(1 - \frac{g(x)}{2}\right)\right) = \frac{g(x)}{x}$, como segue

$$\begin{aligned} \frac{d}{dx}\left(g(x)\left(1 - \frac{g(x)}{2}\right)\right) &= g'(x)\left(1 - \frac{g(x)}{2}\right) + g(x)\left(-\frac{g'(x)}{2}\right) \\ \frac{d}{dx}\left(g(x)\left(1 - \frac{g(x)}{2}\right)\right) &= g'(x) - g'(x)\frac{g(x)}{2} - g(x)\frac{g'(x)}{2} \\ \frac{d}{dx}\left(g(x)\left(1 - \frac{g(x)}{2}\right)\right) &= g'(x) - g(x)g'(x) \\ \frac{d}{dx}\left(g(x)\left(1 - \frac{g(x)}{2}\right)\right) &= g'(x)(1 - g(x)). \end{aligned} \tag{A.5}$$

Usando o resultado de A.4 em A.5, temos

$$\begin{aligned} \frac{d}{dx}\left(g(x)\left(1 - \frac{g(x)}{2}\right)\right) &= \frac{e^{g(x)}}{(1 - g(x))}(1 - g(x)) \\ \frac{d}{dx}\left(g(x)\left(1 - \frac{g(x)}{2}\right)\right) &= e^{g(x)}. \end{aligned} \tag{A.6}$$

Da definição da função $g(x)$ (A.1) decorre diretamente o fato de que $e^{-g(x)} = x/g(x)$. Substituindo essa informação na equação A.5, temos finalmente que

$$\frac{d}{dx}\left(g(x)\left(1 - \frac{g(x)}{2}\right)\right) = \frac{g(x)}{x}. \tag{A.7}$$

REFERÊNCIAS

- [1] LEWIS, T. G. *Network science*. [S.l.]: New Jersey: Wiley and Sons, 2009.
- [2] BARRAT, A.; BARTHELEMY, M.; VESPIGNANI, A. *Dynamical processes on complex networks*. [S.l.]: Cambridge university press, 2008.
- [3] COHEN, R.; HAVLIN, S. *Complex networks: structure, robustness and function*. [S.l.]: Cambridge university press, 2010.
- [4] ALBERT, R.; BARABÁSI, A.-L. Statistical mechanics of complex networks. *Reviews of modern physics*, APS, v. 74, n. 1, p. 47, 2002.
- [5] NEWMAN, M. *Networks: An introduction*. [S.l.]: Oxford University Press, 2010.
- [6] BROADBENT, S. R.; HAMMERSLEY, J. M. Percolation processes: I. crystals and mazes. In: CAMBRIDGE UNIVERSITY PRESS. *Mathematical Proceedings of the Cambridge Philosophical Society*. [S.l.], 1957. v. 53, n. 3, p. 629–641.
- [7] ZHOU, Z. *et al.* Crossover from isotropic to directed percolation. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, APS, v. 86, n. 2, p. 021102, 2012.
- [8] STAUFFER, D.; AHARONY, A. *Introduction to percolation theory*. [S.l.]: CRC press, 2018.
- [9] DOROGOVTSSEV, S. N.; GOLTSEV, A. V.; MENDES, J. F. Critical phenomena in complex networks. *Reviews of Modern Physics*, APS, v. 80, n. 4, p. 1275–1335, 2008.
- [10] COHEN, R. *et al.* Resilience of the internet to random breakdowns. *Physical review letters*, APS, v. 85, n. 21, p. 4626, 2000.
- [11] PASTOR-SATORRAS, R.; VESPIGNANI, A. Epidemic spreading in scale-free networks. *Physical review letters*, APS, v. 86, n. 14, p. 3200, 2001.
- [12] CALLAWAY, D. S. *et al.* Network robustness and fragility: Percolation on random graphs. *Physical review letters*, APS, v. 85, n. 25, p. 5468, 2000.
- [13] PASTOR-SATORRAS, R.; VESPIGNANI, A. Immunization of complex networks. *Physical review E*, APS, v. 65, n. 3, p. 036104, 2002.
- [14] WATTS, D. J. A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, The National Academy of Sciences, v. 99, n. 9, p. 5766–5771, 2002.
- [15] SAHIMI, M. *Applications of percolation theory*. [S.l.]: CRC Press, 1994.
- [16] SAHIMI, M. Flow phenomena in rocks: from continuum models to fractals, percolation, cellular automata, and simulated annealing. *Reviews of modern physics*, APS, v. 65, n. 4, p. 1393, 1993.
- [17] HARARY, F. *Graph theory*. [S.l.]: Addison-Wesley, 1969.

- [18] NETTO BOAVENTURA, Paulo O. *Grafos: teorias, modelos e algoritmos*. 4. ed.. ed. São Paulo: Blucher, 2006.
- [19] ERDŐS, P.; RÉNYI, A. On the strength of connectedness of a random graph. *Acta Mathematica Hungarica*, v. 12, n. 1, p. 261–267, 1961.
- [20] ERDŐS, P.; RÉNYI, A. *et al.* On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, v. 5, n. 1, p. 17–60, 1960.
- [21] ERDŐS, P.; RÉNYI, A. On random graphs. *Publicationes Mathematicae*, v. 6, p. 290–297, 1959.
- [22] 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.
- [23] NEWMAN, M. E.; STROGATZ, S. H.; WATTS, D. J. Random graphs with arbitrary degree distributions and their applications. *Physical review E, APS*, v. 64, n. 2, p. 026118, 2001.
- [24] BALANKIN, A. S. A survey of fractal features of bernoulli percolation. *Chaos, Solitons & Fractals*, Elsevier, v. 184, p. 115044, 2024.
- [25] PAUL, G.; ZIFF, R. M.; STANLEY, H. E. Percolation threshold, fisher exponent, and shortest path exponent for four and five dimensions. *Physical Review E, APS*, v. 64, n. 2, p. 026115, 2001.
- [26] BALANKIN, A. S.; MENA, B.; CRUZ, M. M. Topological hausdorff dimension and geodesic metric of critical percolation cluster in two dimensions. *Physics Letters A*, Elsevier, v. 381, n. 33, p. 2665–2672, 2017.
- [27] STANLEY, H. E. Cluster shapes at the percolation threshold: and effective cluster dimensionality and its connection with critical-point exponents. *Journal of Physics A: Mathematical and General*, IOP Publishing, v. 10, n. 11, p. L211, 1977.
- [28] STAUFFER, D. Scaling theory of percolation clusters. *Physics reports*, Elsevier, v. 54, n. 1, p. 1–74, 1979.
- [29] ISICHENKO, M. B. Percolation, statistical topography, and transport in random media. *Reviews of modern physics*, APS, v. 64, n. 4, p. 961, 1992.
- [30] BARABÁSI, A.-L. *et al.* *Network science*. [S.l.]: Cambridge university press, 2016.
- [31] WATTS, D. J. *Six degrees: The science of a connected age*. [S.l.]: WW Norton & Company, 2004.
- [32] ALBERT, R.; BARABÁSI, A.-L. Statistical mechanics of complex networks. *Reviews of modern physics*, APS, v. 74, n. 1, p. 47, 2002.
- [33] GOH, K.-I.; KAHNG, B.; KIM, D. Universal behavior of load distribution in scale-free networks. *Physical review letters*, APS, v. 87, n. 27, p. 278701, 2001.
- [34] HARTMANN, A. K.; WEIGT, M. *Phase transitions in combinatorial optimization problems: basics, algorithms and statistical mechanics*. [S.l.]: John Wiley & Sons, 2006.
- [35] BENDER, E. A.; WILLIAMSON, S. G. *Lists, decisions and graphs*. [S.l.]: S. Gill Williamson, 2010.
- [36] LIU, C. L. *Elements of discrete mathematics*. [S.l.]: McGraw-Hill Computer Science Series, 1986. v. 2.

- [37] DOROGOVTSSEV, S. N.; MENDES, J. F. Scaling properties of scale-free evolving networks: Continuous approach. *Physical Review E*, APS, v. 63, n. 5, p. 056125, 2001.
- [38] WANG, X.; CHEN, G. Complex networks: Small-world, scale-free and beyond. *Circuits and Systems Magazine, IEEE*, v. 3, p. 6 – 20, 02 2003.
- [39] ALBERT, R.; JEONG, H.; BARABÁSI, A.-L. Diameter of the world-wide web. *nature*, Nature Publishing Group, v. 401, n. 6749, p. 130–131, 1999.
- [40] BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. *science*, American Association for the Advancement of Science, v. 286, n. 5439, p. 509–512, 1999.
- [41] ALBERT, R.; JEONG, H.; BARABÁSI, A.-L. Error and attack tolerance of complex networks. *nature*, Nature Publishing Group, v. 406, n. 6794, p. 378–382, 2000.
- [42] COHEN, R.; HAVLIN, S. Scale-free networks are ultrasmall. *Physical review letters*, APS, v. 90, n. 5, p. 058701, 2003.
- [43] BARABÁSI, A.-L.; BONABEAU, E. Scale-free networks. *Scientific american*, JSTOR, v. 288, n. 5, p. 60–69, 2003.
- [44] FILHO, C. I. S. *et al.* Random networks with q-exponential degree distribution. *Physical Review Research*, APS, v. 5, n. 3, p. 033088, 2023.
- [45] TANAKA, G.; MORINO, K.; AIHARA, K. Dynamical robustness in complex networks: the crucial role of low-degree nodes. *Scientific reports*, Nature Publishing Group, v. 2, p. 232, 2012.
- [46] HOLME, P. *et al.* Attack vulnerability of complex networks. *Physical review E*, APS, v. 65, n. 5, p. 056109, 2002.
- [47] SHAO, S. *et al.* Percolation of localized attack on complex networks. *New Journal of Physics*, IOP Publishing, v. 17, n. 2, p. 023049, 2015.
- [48] SABERI, A. A. Recent advances in percolation theory and its applications. *Physics Reports*, Elsevier, v. 578, p. 1–32, 2015.
- [49] COHEN, R. *et al.* Breakdown of the internet under intentional attack. *Physical review letters*, APS, v. 86, n. 16, p. 3682, 2001.
- [50] ZENG, A.; LIU, W. Enhancing network robustness against malicious attacks. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, APS, v. 85, n. 6, p. 066130, 2012.
- [51] COHEN, R.; BEN-AVRAHAM, D.; HAVLIN, S. Percolation critical exponents in scale-free networks. *Physical Review E*, APS, v. 66, n. 3, p. 036113, 2002.
- [52] SCHNEIDER, C. M. *et al.* Mitigation of malicious attacks on networks. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 108, n. 10, p. 3838–3841, 2011.
- [53] BOLLOBÁS, B. *et al.* *Percolation*. [S.l.]: Cambridge University Press, 2006.

- [54] BUNDE, A.; HAVLIN, S. *Fractals and disordered systems*. [S.l.]: Springer Science & Business Media, 2012.
- [55] SOARES, E. A. *Estudo de redes complexas sob uma perspectiva combinatória*. Dissertação (Mestrado) — Universidade Federal do Ceará, 2021.
- [56] STANLEY, H. E. *Phase transitions and critical phenomena*. [S.l.]: Clarendon Press, Oxford, 1971. v. 7.
- [57] PATHRIA, R. K. *Statistical mechanics*. [S.l.]: Elsevier, 2016.
- [58] YEOMANS, J. M. *Statistical mechanics of phase transitions*. [S.l.]: Clarendon Press, 1992.
- [59] AIGNER, M.; ZIEGLER, G. M. *Proofs from The Book*. [S.l.]: Springer, Berlin, 1999.
- [60] PITMAN, J. Coalescent random forests. *Journal of Combinatorial Theory, Series A*, v. 85, n. 2, p. 165–193, 1999.
- [61] KIRCHHOFF, G. On the solution of the equations obtained from the investigation of the linear distribution of galvanic currents. *IRE transactions on circuit theory*, IEEE, v. 5, n. 1, p. 4–7, 1958.
- [62] PRUFER, H. Neuer beweis eines satzes uber per mutationen. *Archiv der Mathematik und Physik*, v. 27, p. 742–744, 1918.
- [63] SALINAS, S. R. *Introdução à física estatística*. [S.l.]: Edusp, 1999.
- [64] EHRENFEST, P.; EHRENFEST-AFANASSJEWA, T. *Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem*. [S.l.]: Hirzel, 1907.
- [65] KAMPEN, N. G. van. *The Gibbs Paradox*. [S.l.]: Pergamon, New York, 1984.
- [66] PETERS, H. Demonstration and resolution of the gibbs paradox of the first kind. *European Journal of Physics*, IOP Publishing, v. 35, n. 1, p. 015023, 2013.
- [67] JAYNES, E. T. *The Gibbs Paradox*. [S.l.]: Kluwer Academic Publishers, Dordrecht, 1992.
- [68] SWENDSEN, R. H. Gibbs' paradox and the definition of entropy. *Entropy*, v. 10, p. 15, 2008.
- [69] STEWART, J. *Calculus*. 7nd. ed. [S.l.]: Cengage Learning, 2011.
- [70] ABRAMOWITZ, M.; STEGUN, A. I. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. [S.l.]: Dover, New York, 1972.
- [71] CORLESS, R. M. *et al.* On the lambertw function. *Advances in Computational Mathematics*, v. 5, p. 329–359, 1996.
- [72] BOLLOBÁS, B. *Random Graphs*. 2nd. ed. [S.l.]: Academic Press, New York, 2001.
- [73] STAUFFER, D.; AHARONY, A. *Introduction to percolation theory*. [S.l.]: CRC Press, 2018.
- [74] KRAUTH, W. *Statistical mechanics: algorithms and computations*. [S.l.]: OUP Oxford, 2006. v. 13.
- [75] REIF, F. *Fundamentals of statistical and thermal physics*. [S.l.]: Waveland Press, 2009.

- [76] HUANG, K. *Statistical Mechanics*. 2nd. ed. [S.l.]: John Wiley & Sons, New York, 1987.
- [77] ŁUCZAK, T.; PITTEL, B. Components of random forests. *Combinatorics, Probability and Computing*, v. 1, n. 1, p. 35–52, 1992.
- [78] DOKHOLYAN, N. V. *et al.* Distribution of shortest paths in percolation. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 266, n. 1-4, p. 55–61, 1999.
- [79] ALEXANDROWICZ, Z. Critically branched chains and percolation clusters. *Physics Letters A*, Elsevier, v. 80, n. 4, p. 284–286, 1980.
- [80] PIKE, R.; STANLEY, H. E. Order propagation near the percolation threshold. *Journal of Physics A: Mathematical and General*, IOP Publishing, v. 14, n. 5, p. L169, 1981.
- [81] HAVLIN, S.; NOSSAL, R. Topological properties of percolation clusters. *Journal of Physics A: Mathematical and General*, IOP Publishing, v. 17, n. 8, p. L427, 1984.
- [82] RAMMAL, R.; D'AURIAC, J. A.; BENOIT, A. Metric properties of fractal lattices. *Journal of Physics A: Mathematical and General*, IOP Publishing, v. 17, n. 9, p. L491, 1984.
- [83] GRASSBERGER, P. On the spreading of two-dimensional percolation. *Journal of Physics A: Mathematical and General*, IOP Publishing, v. 18, n. 4, p. L215, 1985.
- [84] HERRMANN, H.; STANLEY, H. E. The fractal dimension of the minimum path in two-and three-dimensional percolation. *Journal of Physics A: Mathematical and General*, IOP Publishing, v. 21, n. 17, p. L829, 1988.
- [85] GRASSBERGER, P. Pair connectedness and shortest-path scaling in critical percolation. *Journal of Physics A: Mathematical and General*, IOP Publishing, v. 32, n. 35, p. 6233, 1999.
- [86] DENG, Y. *et al.* Some geometric critical exponents for percolation and the random-cluster model. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, APS, v. 81, n. 2, p. 020102, 2010.
- [87] GRASSBERGER, P. Spreading and backbone dimensions of 2d percolation. *Journal of Physics A: Mathematical and General*, IOP Publishing, v. 25, n. 21, p. 5475, 1992.
- [88] ROMAN, H. E. *et al.* Distributions of polymers in disordered structures. *Physical Review E*, APS, v. 52, n. 6, p. 6303, 1995.
- [89] FEHR, E. *et al.* Corrections to scaling for watersheds, optimal path cracks, and bridge lines. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, APS, v. 86, n. 1, p. 011117, 2012.
- [90] SALEUR, H.; DUPLANTIER, B. Exact determination of the percolation hull exponent in two dimensions. *Physical review letters*, APS, v. 58, n. 22, p. 2325, 1987.
- [91] LEATH, P. Cluster size and boundary distribution near percolation threshold. *Physical review B*, APS, v. 14, n. 11, p. 5046, 1976.
- [92] PORTO, M. *et al.* Probability distribution of the shortest path on the percolation cluster, its backbone, and skeleton. *Physical Review E*, APS, v. 58, n. 5, p. R5205, 1998.

- [93] PORTO, M. *et al.* Structural and dynamical properties of the percolation backbone in two and three dimensions. *Physical Review E*, APS, v. 56, n. 2, p. 1667, 1997.
- [94] VANNIMENUS, J.; NADAL, J.; MARTIN, H. On the spreading dimension of percolation and directed percolation clusters. *Journal of Physics A: Mathematical and General*, IOP Publishing, v. 17, n. 6, p. L351, 1984.
- [95] NEWMAN, M.; ZIFF, R. M. Efficient monte carlo algorithm and high-precision results for percolation. *Physical Review Letters*, APS, v. 85, n. 19, p. 4104, 2000.
- [96] DOBRIN, R.; DUXBURY, P. Minimum spanning trees on random networks. *Physical Review Letters*, APS, v. 86, n. 22, p. 5076, 2001.
- [97] WILKINSON, D.; WILLEMSSEN, J. F. Invasion percolation: a new form of percolation theory. *Journal of physics A: Mathematical and general*, IOP Publishing, v. 16, n. 14, p. 3365, 1983.

ANEXO A – ARTIGO: *ENTROPY OF THE ENSEMBLE OF ACYCLIC GRAPHS*

Entropy of the ensemble of acyclic graphs

Edson A. Soares* and André A. Moreira†

*Departamento de Física
Universidade Federal do Ceará
Caixa Postal 6030, Campus do Pici
60455-760 Fortaleza, Ceará, Brazil*

**edson.soares@fisica.ufc.br*

†auto@fisica.ufc.br

Received 10 May 2023

Accepted 6 June 2023

Published 14 August 2023

The structural phase transition in complex network models is known to yield knowledge relevant for several problems of practical application. Examples include resilience of artificial networks, dynamics of epidemic spreading, among others. The model of random graphs is probably the simplest model of complex networks, and the solution of this model falls into the universality class mean field percolation. In this work, we concentrate on a similar problem, namely, the structural phase transition in the ensemble of random acyclic graphs. It should be noted that several approaches to the problem of random graphs, such as generating functions or the Molloy–Reed criterion, rely on the fact that before the critical point, cycles should not be present in random graphs. In this way, up to the critical point our solution should produce results that are equivalent to these other methods. Our approach takes advantage of the fact that acyclic graphs allow for an exact combinatorial enumeration of the whole ensemble, what leads to an exact expression for the entropy of this system. With this definition of entropy we can determine the onset of the critical transition as well as the critical exponents associated with the transition. Our results are illustrated with Monte-Carlo results and are discussed within the context of general random graphs, as well as in comparison with another model of acyclic graphs.

Keywords: Random graphs, networks, percolation problems, exact results, classical phase transitions

PACS numbers: 89.75.Hc, 05.70.Jk, 05.70.Fh

1. Introduction

Models of networks have been an important topic of study over the past few decades, particularly due to their applicability in Physics problems^{1–3} and various other fields.^{4–7} Among these network models, the random graph model is one of the simplest and has been widely studied since its introduction by Erdos–Rényi,^{8–10}

†Corresponding author.

with a special interest in describing the statistical properties related to the percolation phase transition of these graphs,^{11–13} such as the emergence of giant components and the distribution of component sizes.¹⁴ These statistical properties of random graphs have already been successfully accessed using methods such as the Molloy–Reed criterion^{15,16} and generating function.¹⁷ It is noteworthy that most of these methods assume that, before the critical condition in which the formation of a giant component occurs, there are no cycles in the network.¹⁷ A few works make efforts to overcome this restriction by assuming the possibility of cycles under certain conditions.^{18,19}

There are works that define a measure of entropy for network models.^{20,21} However, these works focus mainly on the entropy of graphs with constraints imposed on the degree distributions or assortative degree correlations. Here, we define our entropy as a function of the component size distribution. We do this by concentrating in the problem of random acyclic graphs,²² where an exact enumeration of the ensemble can be obtained. As mentioned, below the critical condition, imposing the absence of cycles should bear no effect on the statistical properties of the graphs, and the properties of the ensemble of random acyclic graphs should be equivalent to the properties of the ensemble of general random graphs, at least up to the critical point. Above the critical condition, where a giant component emerges, these two ensembles should differ, but the universality class of the transition should follow the mean-field predictions in both cases. We can, therefore, view our results in relation to the general model of random graphs, as well as in relation to different models of acyclic graphs.

2. Entropy of Random Acyclic Graphs

The enumeration $\Omega_T(N, M)$ of the ensemble of acyclic graphs with N nodes and M components has no simple form that would allow for an exact definition of the entropy. To overcome this hurdle we will compute instead the enumeration $\Omega[m_n]$ of the number of acyclic graphs given a particular distribution of component sizes m_n . From that, we can obtain the equilibrium distribution $m_n^*(N, M)$ that maximizes Ω while satisfying the constraints on the number of vertices N and components M , and define $\Omega^*(N, M) = \Omega[m_n^*(N, M)]$. As we will show, $\ln \Omega_T = \ln \Omega^* + (1/2)\sum_n \ln(m_n^*) + \text{const.}$, and the difference between the entropy $S_T = \ln \Omega_T$ and the entropy $S = \ln \Omega^*$ is negligible. Defining an expression for this entropy will allow us to obtain the properties of the critical transition for this ensemble.

Our initial objective, therefore, is to determine the entropy of an ensemble of acyclic graphs given a particular distribution of component sizes m_n . Since we consider that the N nodes of the graph are labeled, or distinguishable, we have that the entropy is given as $S = \ln(\Omega[m_n]/N!)$, where $\Omega[m_n]$ is the enumeration of the ensemble of acyclic graphs for a given distribution of component sizes m_n . The factor $N!$ is introduced to make the entropy of the system of distinguishable nodes additive.^{23–25} The first step to obtain the enumeration Ω is to distribute the N nodes

over the several components, this introduces a factor in the form of the multinomial coefficient $\frac{N!}{\prod_{i=1} (n_i!)}$, where n_i are sizes of the components, and we have the constraint that $\sum_i n_i = N$. Since we are interested in the component size distribution m_n , it is useful to write this factor in terms of these variables as $\frac{N!}{\prod_{n=1} (n!)^{m_n}}$, and the constraint becomes $\sum_{n=1} n m_n = N$. Once we have enumerated the ways N nodes can be divided among the several components, we need to include the correct factors necessary to obtain $\Omega[m_n]$. Calley's formula^{26,27} tells us that there are n^{n-2} ways to construct an acyclic component with n different labeled nodes, introducing the factor $\prod_{n=1} n^{m_n(n-2)}$. However, since the components are not distinguishable *a priori*, we are still overcounting $\Omega[m_n]$. For instance, if we have m_1 components with just one node, permuting the nodes of these components do not lead to a new partition of the nodes. In the same way, if we take two components of size $n = 2$, and permute the pair of nodes from one component with the two nodes from the other component, this permutation also does not lead to a new partition. The same will happen if we permute any groups of nodes that are in components of the same size. To account for this overcounting we need to divide our enumeration by $\prod_{n=1} m_n!$ obtaining finally

$$\Omega[m_n] = N! \prod_{n=1} \left[\left(\frac{n^{n-2}}{n!} \right)^{m_n} \frac{1}{m_n!} \right], \quad (1)$$

which leads to the entropy in the following form

$$S = \ln \left(\frac{\Omega}{N!} \right) = \sum_{n=1} m_n ((n-2) \ln n - \ln n! - \ln m_n + 1), \quad (2)$$

where we have used the Stirling approximation, $\ln(m_n!) \approx m_n(\ln(m_n) - 1)$.

To obtain the most probable component size distribution m_n , we need to maximize the entropy while taking into account the constraints of the variables m_n . As mentioned, one of these constraints gives the total number of nodes N . However, considering that the average degree of the nodes of the graph is specified, we have another constraint. This happens because in acyclic graphs, each connection between a pair of vertices necessarily reduces the total number of components M by one. As a consequence, we have that the total number of connections is given by $E = N - M$, and the average degree is given by $\mu = 2E/N$, and for acyclic graphs $E/N = (1 - M/N)$. Therefore, in acyclic graphs, fixing the average degree also fixes the total number of components. We have then the following constraints

$$\sum_{n=1} m_n = M \quad \text{and} \quad \sum_{n=1} n m_n = N. \quad (3)$$

Using the method of the Lagrange multipliers,²⁸ we can obtain the maximal entropy by solving

$$\frac{\partial}{\partial m_n} \left[S + \lambda_1 \left(M - \sum m_n \right) + \lambda_2 \left(N - \sum n m_n \right) \right] = 0, \quad (4)$$

which gives us the equilibrium component size distribution

$$m_n^* = \frac{n^{n-2} e^{-\lambda_1} e^{-n\lambda_2}}{n!}. \quad (5)$$

As usual, the multipliers have to be determined from the constraints. Defining $x = e^{-\lambda_2}$ and $y = e^{-\lambda_1}$, we have

$$N = y \sum_n \frac{n^{n-1}}{n!} x^n \quad (6)$$

and

$$M = y \int_0^x \sum_n \frac{n^{n-1}}{n!} x'^{n-1} dx'. \quad (7)$$

The summation in (6) and (7) can be related to the Taylor expansion of a function $g(x) = \sum_{n=1}^{\infty} \frac{n^{n-1}}{n!} x^n$. Using the Lagrange inversion theorem²⁹ it is possible to show that the function implicitly defined by the transcendental equation $g(x)e^{-g(x)} = x$ will have a Taylor series precisely in this form. In fact, our function $g(x)$ can be related to the principal branch of the Lambert- W function³⁰ in the form $g(x) = -W_0(-x)$. From that we have

$$N = yg(x) \quad \text{and} \quad M = y \int_0^x \frac{g(x')}{x'} dx'. \quad (8)$$

Differentiating both sides of the implicit definition $g(x)e^{-g(x)} = x$, one obtains that $\frac{dg}{dx} = \frac{e^g}{1-g}$, and from that one can verify that $M = yg(x)[1 - \frac{g(x)}{2}]$, and that the average degree is given by $\mu = 2(1 - M/N) = g(x)$. This allows us to determine the parameters $x = \mu e^{-\mu}$ and $y = N/\mu$, reaching the final form of the component size distribution³¹

$$m_n^* = N \frac{n^{n-2} (\mu e^{-\mu})^n}{\mu n!}. \quad (9)$$

To assure that the distribution (9) is a stable equilibrium one should observe that $\partial^2 \ln \Omega / \partial m_n^2 = -1/m_n < 0$, indicating that, in fact m_n^* is the most probable value for the component size distribution. One can consider that the probability of observing a number of components deviating from the equilibrium follows $P_n(m_n = m_n^* + \delta_n) \sim \exp(-\delta_n^2/(2\sigma_n^2))$ with $\sigma_n^2 = 1/m_n^*$ leading to $\Omega[m_n] \sim \Omega[m_n^*] \exp(-\delta_n^2/(2\sigma_n^2))$. Considering that, after introducing the Lagrange multipliers, the distributions $P_n(m_n)$ are independent, we have that the total enumeration Ω_T of acyclic graphs, given N and M , follows

$$\Omega_T(N, M) \sim \Omega[m_n^*] \prod_n \int dm_n e^{-\frac{\delta_n^2}{2\sigma_n^2}}. \quad (10)$$

Therefore, $\ln \Omega_T = \ln \Omega[m_n^*] + (1/2) \sum_n \ln m_n^* + \text{const}$. Since $\ln \Omega[m_n^*]$ is extensive with the number of nodes N , while the remaining terms grow at most with $\ln N$, we

can use $S(N, M) = \ln \Omega[m_n^*]$ neglecting the extra terms. Therefore, using the component size distribution, Eq. (9), into the formula for the entropy, Eq. (2), it is possible to determine the entropy under the equilibrium condition

$$S(N, \mu \leq 1) = N \left[\left(1 - \frac{\mu}{2}\right) \left(1 - \ln \frac{N}{\mu}\right) + \mu - \ln \mu \right]. \quad (11)$$

3. Phase Transition in the Ensemble of Random Acyclic Graphs

It is usual to define the probability $P_n = nm_n/N$ of a given node to be found into a component of size n . To obtain the asymptotic behavior of this probability we need to employ the Stirling approximation $n! \sim \sqrt{2\pi n}(\frac{n}{e})^n$, obtaining

$$P_n \approx n^{-3/2} \exp(-(\mu - 1)n + (n - 1) \ln \mu) / \sqrt{2\pi}. \quad (12)$$

The asymptotic behavior described in (12) has a contribution of a power-law decay, $n^{-3/2}$, and an exponential decay, $e^{-(\mu-1-\ln \mu)n}$. For $\mu = 1$ the terms in the exponential decay cancel, and only the power-law $n^{-3/2}$ remains. This point where the exponential decay vanishes defines the critical point $\mu_c = 1$. Note that the critical condition $\mu = 1$ happens when the number of components is half the number of nodes, $M = N/2$. Within percolation theory, the exponent controlling the power-law decay of the component size distribution is called Fisher exponent, and $\tau = 3/2$ is the expected mean-field value,³² and is what is obtained for the random graph model.³¹ For $\mu \lesssim 1$, we have $\mu - 1 - \ln \mu \approx (1 - \mu)^2/2$, and the exponential decay will dominate when $n > 2/(1 - \mu)^2$. A divergence of the size scale with $(\mu - \mu_c)^{-2}$ is also consistent with the expected in the mean-field limit.³²

To illustrate our results we proceed to perform Monte-Carlo simulations of our ensemble of acyclic graphs. To do this we create an acyclic graph of the prescribed average degree μ and, at each Monte-Carlo step, we remove a random connection of the graph and include a new random connection that does not create a cycle. It is possible to see that the probability to transit from a graph **A** to another graph **B** is same as to transit back from **B** to **A**. Detailed balance³³ ensures that in this condition every possible graph in the ensemble has the same chance to appear in the Monte-Carlo process. Our Monte-Carlo results are presented in Fig. 1. In the main panel of this figure we show the component size distribution of systems with average degree $\mu \leq 1$. One may see that our numerical results follow closely the predictions of Eq. (9). In the inset of Fig. 1 we present results for $\mu > 1$. The form of the distribution in this regime will be discussed in the following.

As mentioned before, using Stirling formula for the factorial, one can note that the component distribution m_n defined in (9) has exponential decay with the leading term in the form $\exp(-n(\mu - 1)^2/2)$, and at $\mu = 1$ only the power-law dependence remains. As we will see, above the critical condition, $\mu > 1$ one observes the emergence of a giant component that takes a nonvanishing portion P_∞ of the graph.

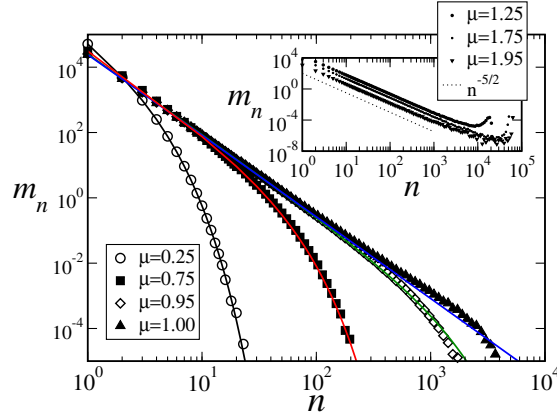


Fig. 1. Component size distribution m_n for different values of average degree μ . In the main panel we show results for $\mu \leq 1$. The symbols are results from Monte-Carlo simulations for a system with $N = 2^{16}$ nodes, while the continuous lines are the predictions of Eq. (9). There is good agreement between theory and simulation, with the exception of the case $\mu = 1$, where one may observe a small deviation in the tail of the distribution. The reason of this deviation is the finite size of the system. As $\mu \rightarrow 1$ the scale of the exponential crossover diverges, however, the largest components are limited by the finite size of the system, thus the deviation. In the inset, we present results for $\mu > 1$. In this regime the component size distribution follows the form described by Eq. (18) that decays as a power-law with exponent $-5/2$. The spikes observed in the tail of the distribution when $\mu > 1$ indicate the presence of a giant component in this regime.

To obtain the fraction P_∞ when $\mu > 1$ we have to investigate the entropy of a composite system, comprised of a sub-system being the giant component, and another subsystem that includes the remaining components. The enumeration Ω of the number of possible acyclic graphs for this composite system is given by

$$\Omega(N_\infty, N) = \Omega_R(N_R)\Omega_\infty(N_\infty) \frac{N!}{N_\infty!(N_R)!}, \quad (13)$$

where Ω_R is the enumeration of the possible configurations for the N_R nodes outside the giant component. Similarly Ω_∞ is the enumeration of the possible configurations for the giant component, and $N_\infty = N - N_R$ is the number of nodes inside the giant component. The binomial factor needs to be introduced to account for the possible ways the N distinguishable nodes can be split among the giant component and the remaining components.²⁴ We need to ensure that the entropy is additive, that is, the entropy of the composite system should be the sum of the entropies of the giant component and the entropy of the rest of the graph, $S = S_\infty + S_R$. This leads to $S = \ln[\Omega/N!]$, $S_\infty = \ln[\Omega_\infty/N_\infty!]$ and $S_R = \ln[\Omega_R/N_R!]$. Next we proceed to obtain the equilibrium condition between these two subsystems.

The enumeration of the possible configurations for the giant component can be readily obtained from Calley's formula²⁷ leading to

$$S_\infty = \ln \left(\frac{N_\infty^{N_\infty-2}}{N_\infty!} \right) \approx N_\infty - \frac{5}{2} \ln(N_\infty), \quad (14)$$

where the approximation can be deduced from the Stirling formula for the factorial. From this we have $dS_\infty/dN_\infty \approx 1 - 5/(2N_\infty) \approx 1$.

We have also to obtain the variation of the entropy of the remaining components. Note that nodes can move inside and outside the giant component, but the number of components should remain constant as long as μ remains constant. To obtain the entropy S_R of the subsystem of nodes outside the giant component we take advantage of formula (11) defining

$$S_R(N_R, M-1) = S\left(N_R, \mu_R = 2\left[1 - \frac{M-1}{N_R}\right]\right), \quad (15)$$

where N_R and μ_R are the number and average degree of the nodes outside the giant component, respectively. The number of remaining components is $M-1$ as one of the components is the giant component. Differentiating with respect to N_R while the number of components is kept constant one can show that

$$\left(\frac{\partial S_R}{\partial N_R}\right)_M = \mu_R - \ln \mu_R, \quad (16)$$

where one should remember that $\mu_R = 2(1 - \frac{N_R}{M-1})$. We have finally

$$\left(\frac{\partial S}{\partial N_\infty}\right)_\mu = \frac{dS_\infty}{dN_\infty} - \left(\frac{\partial S_R}{\partial N_R}\right)_M \approx 1 - \mu_R + \ln \mu_R \leq 0. \quad (17)$$

Note that the entropy variation vanishes when $\mu_R = 1$, that is, $[(\partial S / \partial N_\infty)]_{\mu_R=1} = 0$. However, the second derivative also vanishes at this point, $[(\partial^2 S / \partial N_\infty^2)]_{\mu_R=1} = 0$. This seems to suggest that $\mu_R = 1$ is not a point of maximum entropy, but rather an inflection point. However, one should remember that $\mu_R \leq 1$ since the remaining components do not have a giant component and have to be below the critical condition. Since the entropy S will always increase as N_∞ decreases, the maximum entropy will happen when N_∞ is minimum, or when N_R is maximum. As we mentioned, the number of nodes outside the giant component is limited to be less than twice the number of components, $N_R \leq 2(M-1)$. To maximize the entropy S , therefore, N_R should reach its maximum possible value that leads to $\mu_R = 1$. We conclude that, similar to what happens in the Bose–Einstein condensation,^{34,35} where the particles outside the condensate present a critical distribution with a chemical potential going to zero, the equilibrium condition of the ensemble of acyclic graphs will happen when the component size distribution outside the giant component is critical, $\mu_R = 1$. Therefore, above the critical point $\mu > 1$ we have

$$m_n^* = (1 - P_\infty)N \frac{n^{n-2}e^{-n}}{n!}, \quad (18)$$

with the remaining $P_\infty N$ nodes placed into the giant component. This prediction is consistent with the Monte-Carlo results shown in the inset of Fig. 1.

One can now obtain in a simple way the dependence of P_∞ with the average degree μ . This is done by noting that the average degree is the weighted average of the expected degree of nodes inside and outside the giant component

$$\mu = P_\infty \mu_\infty + (1 - P_\infty) \mu_R, \quad (19)$$

where μ_∞ is the average degree of the nodes inside the giant component, and we have mentioned that μ_R is the average degree of the remaining nodes. For acyclic components, the average degree of a component of size n is given by $2(1 - 1/n)$, and for a giant component, with the size going to infinity, we have $\mu_\infty = 2$. Since the remaining nodes will display a critical distribution, we have $\mu_R = 1$. From that we gather

$$P_\infty = \mu - 1. \quad (20)$$

The order parameter above the critical point follows $P_\infty = (\mu - \mu_c)^\beta$, and we have the critical exponent $\beta = 1$. Once more, this value of the critical exponent is what is expected in the mean-field limit,³² and agrees with what is observed in the random graph model.³¹ Equation (20) should be valid for any average degree in the range $1 \leq \mu \leq 2$. At $\mu = 2$ we have $P_\infty = 1$, this limit indicates a single acyclic component that comprises all the nodes in the graph. These predictions are illustrated by the Monte-Carlo results presented in Fig. 2, where we show that as the system size grows the fraction occupied by the largest component P_∞ approaches the behavior predicted by Eq. (20).

Finally, we can derive the entropy of the ensemble of acyclic graphs above the critical condition, $\mu > 1$. As mentioned, this can be done by summing the entropy of the giant component and the entropy of the remaining components. The entropy of the giant component should be given by Eq. (14), with $N_\infty = (\mu - 1)N$, while the entropy of the remaining can be obtained from Eq. (11) with a number of nodes $N_R = (2 - \mu)N$ and a critical average degree for the remaining nodes $\mu_R = 1$. With that one can derive

$$S(N, \mu \geq 1) = N \left\{ 2 - \frac{\mu}{2} - \left(1 - \frac{\mu}{2} \right) \ln[N(2 - \mu)] \right\}, \quad (21)$$

where we have neglected the terms that are not linear in N . Note that Eqs. (11) and (21) are identical in the limit $\mu = 1$, where $S(N, \mu = 1) = N[3/2 - \ln(N)]$.

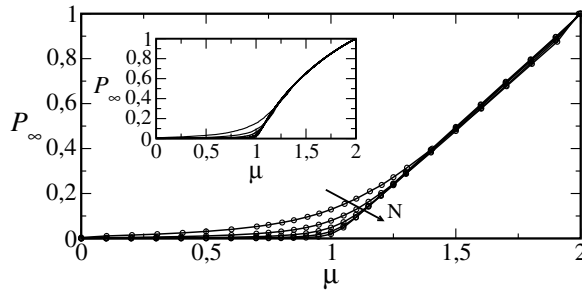


Fig. 2. The fraction occupied by the largest component P_∞ as a function of the average degree μ . This figure presents Monte-Carlo results for acyclic graphs of sizes $N = 2^8, 2^{10}, 2^{12}, 2^{14}$ and 2^{16} . As the system sizes increase the behavior approaches the expected by Eq. (20), with $P_\infty \rightarrow 0$ for $\mu < 1$ and $P_\infty = \mu - 1$, for $\mu > 1$. In the inset, we show results for a model where the acyclic graphs are grown by the successive addition of connections which does not create cycles. This model of acyclic graphs differs from the unbiased ensemble of acyclic graphs.²²

4. Comparison to Other Models of Acyclic Graphs

It should be noted that, above the critical condition, random graphs should present cycles,³¹ and our approach, that is exact for the ensemble of random acyclic graphs, should not be suitable for the general random graphs model. In fact, the random graph model has been shown to be similar to a different model of acyclic graphs (see Ref. 22 and references therein), where an acyclic graph with N nodes and E connections is obtained by adding a connection that does not create a cycle to an acyclic graph with N nodes and $E - 1$ connections. We will refer to this modified model as the growth model. The growth model of acyclic graphs does not sample the possible graphs with the same probability, and therefore produces averages different from those of the unbiased ensemble of acyclic graphs. This can be seen in the inset of Fig. 2, where we present numerical results for the growth model. Curiously, the growth model has a closer relation with the model of random graphs.²² Let us define $\xi = 2(1 - M/N)$, remembering that N and M are the number of nodes and components, respectively. In acyclic graphs, ξ can be related to the average degree, $\xi = \mu$. For random graphs, however, due to the possible presence of cycles the same relation does not hold, and $\xi < \mu$. This distinction is particularly relevant above the critical condition, since below criticality random graphs should not present cycles. The relation between the growth model and the random graph model is that the behavior observed in the inset of Fig. 2 is exactly what would be observed for random graphs in a plot of P_∞ vs ξ .²²

5. Conclusion

We have shown how to obtain an exact combinatorial enumeration for the ensemble of random acyclic graphs, given a particular component size distribution. This enumeration leads to a definition of an entropy for this ensemble, that when maximized gives us the component size distribution in the equilibrium condition, and from that we obtain the equilibrium entropy as function of the number of nodes and average degree of the graph. Our results indicate the onset of a phase transition, where a giant component emerges. Treating the graph after the critical condition as a composite system, we were able to determine the size of the giant component by searching for the equilibrium condition of the composite system. Our approach allows us to determine the onset of the critical transition, as well as the critical exponents associated. Although the results we obtain are known for quite some time,^{8–10,31} the methods and language we use should be more familiar to researchers in the fields of statistical physics. We hope that this solution to the problem of random acyclic graphs may be expanded in novel solutions to other general problems in the field of network science.

Acknowledgments

We acknowledge the financial support from the Brazilian agencies CNPq, CAPES and FUNCAP, and the National Institute of Science and Technology for Complex Systems (INCT-SC) in Brazil for financial support.

References

1. J. Cserti, G. Dávid and A. Piróth, *Am. J. Phys.* **70**, 153 (2002).
2. B. Berkowitz and R. P. Ewing, *Surv. Geophys.* **19**, 23 (1998).
3. R. Zwanzig, *J. Chem. Phys.* **87**, 4870 (1987).
4. S. Lawrence and C. L. Giles, *Nature* **400**, 107 (1999).
5. H. Jeong *et al.*, *Nature* **407**, 651 (2000).
6. M. Faloutsos, P. Faloutsos and C. Faloutsos, *Comput. Commun. Rev.* **29**, 251 (1999).
7. M. Serafino *et al.*, *PLoS Comput. Biol.* **18**, e1009865 (2022).
8. P. Erdos and A. Rényi, *Publ. Math. Debrecen* **6**, 290 (1959).
9. P. Erdos and A. Rényi, *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17 (1960).
10. P. Erdos and A. Rényi, *Acta Math. Hung.* **12**, 261 (1961).
11. D. S. Callaway *et al.*, *Phys. Rev. Lett.* **85**, 5468 (2000).
12. M. Li *et al.*, *Phys. Rep.* **907**, 1 (2021).
13. A. Bunde and S. Havlin (eds.), *Fractals and Disordered Systems* (Springer Science & Business Media, 2012).
14. M. E. J. Newman, *Phys. Rev. E* **76**, 045101 (2007).
15. M. Molloy and B. Reed, *Random Struct. Algor.* **6**, 161 (1995).
16. R. Cohen, D. Ben-Avraham and S. Havlin, *Phys. Rev. E* **66**, 036113 (2002).
17. M. E. J. Newman, S. H. Strogatz and D. J. Watts, *Phys. Rev. E* **64**, 026118 (2001).
18. M. E. J. Newman, *Phys. Rev. Lett.* **103**, 058701 (2009).
19. P. Mann, V. A. Smith, J. B. O. Mitchell and S. Dobson, *Phys. Rev. E* **103**, 012313 (2021).
20. J. Park and M. E. J. Newman, *Phys. Rev. E* **70**, 066117 (2004).
21. G. Bianconi, *Europhys. Lett.* **81**, 28005 (2008).
22. T. Luczak and B. Pittel, *Comb. Probab. Comput.* **1**, 35 (1992).
23. E. T. Jaynes, *The Gibbs Paradox*, eds. C. R. Smith, G. J. Erickson and P. O. Neudorfer (Kluwer Academic Publishers, Dordrecht, 1992).
24. R. H. Swendsen, *Entropy* **10**, 15 (2008).
25. N. G. van Kampen, *The Gibbs Paradox*, ed. W. E. Parry (Pergamon, New York, 1984).
26. J. Pitman, *J. Comb. Theory A* **85**, 165 (1999).
27. M. Aigner and G. M. Ziegler, *Proofs from The Book* (Springer, Berlin, 1999).
28. J. Stewart, *Calculus*, 7nd edn. (Cengage Learning, 2011).
29. M. Abramowitz and A. I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Dover, New York, 1972).
30. R. M. Corless, G. H. Gonnet, D. E. G. Hare *et al.*, *Adv. Comput. Math.* **5**, 329 (1996).
<https://doi.org/10.1007/BF02124750>.
31. B. Bollobás, *Random Graphs*, 2nd edn. (Academic Press, New York, 2001).
32. D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (CRC Press, 2018).
33. W. Krauth, *Statistical Mechanics: Algorithms and Computations*, Vol. **13** (OUP Oxford, 2006).
34. F. Reif, *Fundamentals of Statistical and Thermal Physics* (Waveland Press, 2009).
35. K. Huang, *Statistical Mechanics*, 2nd edn. (John Wiley & Sons, New York, 1987).

ANEXO B – ARTIGO: *SCALING LAWS OF CHEMICAL AND EUCLIDEAN DISTANCES IN CRITICAL PERCOLATION TREES*

Scaling Laws of Chemical and Euclidean Distances in Critical Percolation Trees

E. A. Soares* and A. A. Moreira

*Departamento de Física, Universidade Federal do Ceará,
Caixa Postal 6030, Campus do Pici, 60455-760 Fortaleza, Ceará, Brazil*

We study the scaling behavior of chemical distances in spanning trees of critical percolation clusters. Precisely, we analyze spanning trees constructed using four different algorithms: Minimal Path Trees (MPT), Depth-First Search (DFS), Right-First Search (RFS), and Minimum Spanning Trees (MST). In each case, the chemical distance ℓ is not defined in the original percolation graph, but rather on the specific spanning tree generated by the chosen traversal rule. We analyze the conditional probability distribution $P(r|\ell)$, where r is the Euclidean distance given a chemical distance ℓ , and show that it follows a universal scaling form. Using scaling arguments and Bayes' theorem, we derive and numerically verify general relations between the characteristic exponents that govern these scaling laws. Each spanning-tree construction reveals the geometry of a different self-similar structure within the percolation cluster: MPTs yield the fractal dimension of the minimal path, MSTs recover the optimal path dimension, and DFS and RFS give the fractal dimension of the external hull in two dimensions. In three dimensions, the DFS tree reveals a previously unreported exponent, $\theta_1 = 1.575$, suggesting a new universality class for this type of self-similar structure. These results establish a general framework for extracting fractal dimensions of paths embedded in disordered media via spanning-tree induced chemical distances.

Keywords: percolation theory, critical phenomena, scaling laws, fractal dimensions, chemical distance

I. INTRODUCTION

Percolation theory is a fundamental model in statistical physics [1], widely used to describe disordered systems such as porous materials, resistor networks, and epidemic spreading [1–6]. In regular lattices, the model is defined by randomly activating sites or bonds with probability p [7, 8]. At a critical threshold, $p = p_c$, a continuous phase transition occurs, marked by the emergence of a spanning cluster and governed by a set of scaling laws and critical exponents that characterize its geometry [7, 9–11].

Among the many observables in percolation, a key geometric quantity is the chemical distance, denoted by ℓ , which measures the shortest path between two sites. This metric plays a crucial role in characterizing transport phenomena, such as diffusion, fluid flow, and signal propagation in complex media. The average chemical distance between two points separated by Euclidean distance r typically scales as [12–22]

$$\bar{\ell} \sim r^{d_{\min}}, \quad (1)$$

where $d_{\min}$ is the fractal dimension of the minimal path.

In this work, we take a different approach by defining the chemical distance ℓ not on the original percolation cluster, but in spanning-trees built over the cluster using specific traversal rules. Each spanning tree provides a unique network of paths, and the chemical distance is defined as the number of steps along this tree. Since different algorithms result in different spanning trees, each leads to a distinct definition of ℓ , highlighting different

substructures within the same percolation cluster. We focus on four types of spanning trees: Minimal Path Trees (MPT), Depth-First Search (DFS), Right-First Search (RFS), and Minimum Spanning Trees (MST). Our main object of study is the conditional probability distribution $P(r|\ell)$, which gives the probability that a site at chemical distance ℓ is located at Euclidean distance r . We show that this distribution follows a universal scaling form, governed by a characteristic exponent whose values depend on the type of spanning tree used. The number of nodes $M(\ell)$ accessed in ℓ steps or less through the spanning tree follows the form $M(\ell) \sim \ell^{d_\ell}$. Using scaling arguments and Bayes' theorem, we prove that $d_\ell = d_f/d_{\min}$. We show that this identity follows directly from the scaling hypothesis and Bayes' theorem. As such, it is a necessary consequence of the theoretical framework. Our numerical results are fully consistent with this relation and provide further support. This clarifies a question raised in [7, 23], where it was suggested that the identity might fail in two and three dimensions. Our numerical simulations in two and three dimensions show that each tree construction captures a known or new fractal structure. For MPTs, we recover the minimal path exponent $d_{\min}$. For MSTs, we obtain the fractal dimension of the optimal path. For DFS and RFS in two dimensions, we retrieve the external hull dimension. In three dimensions, the DFS tree reveals a new exponent, $\theta_1 = 1.575$, indicating a distinct universality class for this type of path. This framework provides a general method for extracting the fractal dimensions of self-similar paths embedded in disordered systems. By designing suitable spanning tree constructions, new critical exponents may be discovered and systematically analyzed.

* E-mail: edson.soares@fisica.ufc.br

II. MODELS, SAMPLED QUANTITIES AND NUMERICAL RESULTS

We use the Leath-Alexandrowicz algorithm [13, 24] to obtain samples of critical clusters in two and three dimensions. This algorithm grows a percolation cluster starting from an occupied seed site at the origin of a lattice. With a probability P , each neighbor of the seed is activated and added to the growing cluster, while with a probability $1 - P$, the neighbor is disabled and excluded from further growth. Once all neighbors of the current active sites are processed, the growth continues from the newly activated sites until no more sites can be added, i.e., the cluster growth process ends when the cluster of active sites is confined by an external hull of inactive sites.

The Leath algorithm is an implementation of a Breadth-First Search. We can view this algorithm as building a spanning tree considering that during the search process we include a bond between the current active site and the neighbor that it is being activated. We will call this a Minimal Path Tree (MPT). In this tree, each site is connected to the lattice center by the minimal path in the original graph. A key quantity of interest is the conditional probability $P(r|\ell)$, which describes the probability of reaching a Euclidean separation r given a chemical distance ℓ [25]. Since the definition of the MPT does not introduce any characteristic length, the scale invariance of the critical cluster should be preserved in the MPT. Therefore, we expect that $P(r|\ell)$ should obey a scaling law where the chemical distance ℓ is the only relevant scale [26],

$$P(r|\ell) = \ell^{-\theta} g_x\left(\frac{r}{\ell^\theta}\right), \quad (2)$$

where $g(x)$ is a universal scaling function and the exponent θ should be identified with $1/d_{min}$ [26]. This scaling relation is supported by the numerical results shown in Fig. (1). We have also measured the number of activated sites $N(\ell)$ as a function of ℓ , and in the critical regime we expect that the behavior of the statistical average $\langle N(\ell) \rangle$ follow the scaling form

$$N(\ell) = \ell^{d_\ell - 1}, \quad (3)$$

where d_ℓ is the spreading dimension [27], and it should be related to the fractal dimension d_f of the cluster and the fractal dimension of the minimal path by $d_\ell = d_f/d_{min}$ [12]. This power-law dependence is shown in the inset of the Fig.(1). The results of Fig. (1) are consistent with these relations. In ([7, 23]) it was suggested that the relation $d_\ell = d_f/d_{min}$ should not hold. In fact, as we show in the following, this relation may be deduced from the hypothesis of a universal scaling relation, in the form of Eq. (2).

Besides the breadth-first strategy, other strategies may be used to build spanning trees of the critical percolation cluster. Evidently, employing different strategies should lead to different sets of chemical distances ℓ in the spanning trees. For instance, we could build a spanning tree

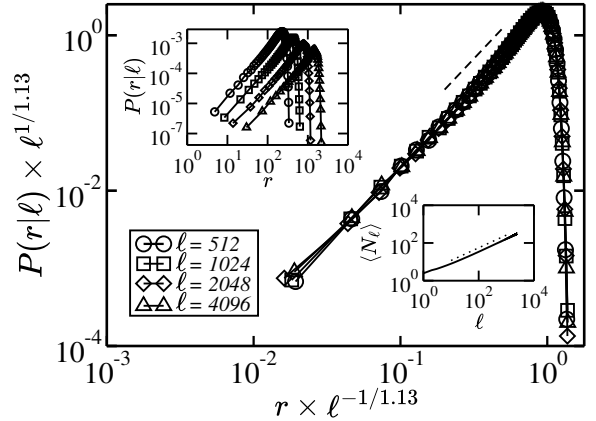


Figure 1. For $d = 2$, the main panel shows the rescaled probability $P(r|\ell)\ell^{1/\theta}$ plotted against the rescaled distance $r \times \ell^{-1/\theta}$ in the percolating cluster at criticality ($p_c = 0.592$) [28]. The cluster was grown using the Leath-Alexandrowicz method, starting from an active seed site at the lattice center, and the MPT was extracted. The fractal dimension is $d_f = 1.89$ and the scaling exponent obtained is $\theta = 1.130$. Symbols correspond to different chemical distances in the spanning trees: 512 (circle), 1024 (square), 2048 (diamond), and 4096 (triangle). The dashed line has a slope of 2, matching the increasing phase of the rescaled probability. The left inset shows the original unscaled probability $P(r|\ell)$. The right inset displays the mean number of sites at a chemical distance ℓ , with a slope $\alpha = 0.672 \pm 0.005$ (slope of the dotted line). This value was used to determine the scaling exponent θ by using $\theta = (d_f/\alpha + 1) = 1.130 \pm 0.003$. These results are based on 50000 cluster configurations on a square lattice with side $L = 10^4$.

through the Depth-First Search (DFS), which visits all sites in the cluster from the origin by exploring each branch as far as possible before backtracking to the last fork. We perform the depth-first search in two different ways. When our search reaches a fork, in our DFS we either choose a random direction, or we choose always the unexplored path that is most to the right. We call this latter strategy Right-First Search (RFS). Each of these strategies leads to the construction of a tree with typical topological characteristics. An instance of the MPT, DFS and RFS trees in a small percolation cluster are shown in the Fig. (2).

Another possible way to build a spanning tree for the critical percolation cluster is identifying the so-called Minimum Spanning Tree (MST)[29, 30]. In this procedure, random weights are assigned to each bond of the percolating cluster, and the spanning tree is grown starting from the seed site located at the center of the lattice. The growth proceeds iteratively incorporating neighboring sites, by selecting the bond with the smallest weight at each step, ensuring that the tree expands along the path of least total weight. This process continues until

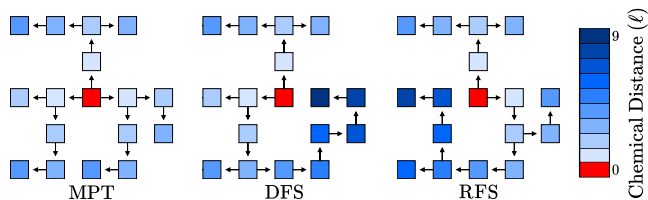


Figure 2. Illustration of a pictorial percolating cluster in a square lattice with side $L = 5$. For each case, the same cluster was traversed from the seed site (red site) using one of three strategies, represented by the directions of the arrows. Each strategy leads to a different tree and a different set of chemical distances ℓ , according to the color bar. The first case is a Breadth-First Search (BFS), a strategy that visits all sites at the current depth before moving on to sites at the next depth level. Using this strategy, we obtain the Minimal Path Tree (MPT). In the second case, the cluster was traversed using a Depth-First Search (DFS), which visits all sites by exploring each branch as far as possible before backtracking to the last fork. In the DFS, when faced with more than one path option, a random decision is made to choose which branch to follow first. Finally, the third case is a Right First Search (RFS), a biased DFS where, unlike the normal DFS, when faced with more than one path option, the branch chosen is always the rightmost unexplored path.

all the sites in the percolating cluster form a spanning tree.

None of the different strategies mentioned above introduce characteristic scales. Thus, for all cases, the distribution of Euclidean distances r should present a scaling relation similar to Eq. (2). In fact, the same should hold for the distribution $P(\ell|r)$, of chemical distances through the spanning tree given an Euclidean distance. Therefore, we should have the following scaling relations

$$P(\ell|r) = r^{-\theta_1} g_\ell \left(\frac{\ell}{r^{\theta_1}} \right) \quad \text{and} \quad P(r|\ell) = \ell^{-\theta_2} g_r \left(\frac{r}{\ell^{\theta_2}} \right). \quad (4)$$

From these relations we have that the average chemical distance follows $\bar{\ell}(r) = c_\ell r^{\theta_1}$, with $c_\ell = \int x g_\ell(x) dx$. Therefore, for the case of the MPT, defined before, the exponent θ_1 will be the fractal dimension of the minimal path d_{min} . In the same way $\bar{r}(\ell) = c_r \ell^{\theta_2}$, and $c_r = \int x g_r(x) dx$. An informed reader may be surprised that at this point we assume that the exponents θ_1 and θ_2 are independent. As we will demonstrate in the following, these two exponents in fact need to be related to each other.

We define the number $\mathcal{N}(\ell, r)$ of sites that are simultaneously at a given chemical and Euclidean distance of the origin. We have that

$$\mathcal{N}(\ell, r) = P_\ell(\ell|r) N_r(r) = P_r(r|\ell) N_\ell(\ell), \quad (5)$$

were $N_r(r)$ and $N_\ell(\ell)$ are the number of sites at a given Euclidean and chemical distances in the spanning trees, respectively. This equation is an expression of the Bayes theorem, and, as we will show, this identity together with

the scaling forms (4) can be used to relate the exponents d_f , d_ℓ , θ_1 , and θ_2 . We consider that $N_r(r) = r^{d_f-1}$ and $N_\ell(\ell) = \ell^{d_\ell-1}$. The exponent d_ℓ here has the same role as in Eq. (3), however, we should note that the value of this exponent depends on the particular spanning tree that is being considered. From Eqs. (4) and (5), we have

$$r^{d_f-\theta_1-1} g_\ell \left(\frac{\ell}{r^{\theta_1}} \right) = \ell^{d_\ell-\theta_2-1} g_r \left(\frac{r}{\ell^{\theta_2}} \right). \quad (6)$$

We now define $f_r(x) = g_r(x^{-\theta_2})$, which leads to $g_r \left(\frac{r}{\ell^{\theta_2}} \right) = f_r \left(\frac{\ell}{r^{1/\theta_2}} \right)$. Note that the function $g_r(x)$ shapes the distribution $P_r(r|\ell)$ that has a well defined mode and is non-monotonic. Therefore, we can be confident that the function $f_r(x)$ is also non-monotonic. We can also define $\zeta = \frac{\ell}{r^{\theta_1}}$, and rewrite Eq. (6) as

$$\frac{g_\ell(\zeta)}{f_r \left(\zeta r^{\theta_1 - \frac{1}{\theta_2}} \right)} = \zeta^{d_\ell - \theta_2 - 1} r^\gamma, \quad (7)$$

where $\gamma = \theta_1(d_\ell - \theta_2 - 1) - (d_f - \theta_1 - 1)$. Given their definitions, the variables ζ and r are independent, and one can consider a case where r varies as ζ remains constant. The left hand suggests a non-monotonic dependence on r through $f_r(\zeta r^{\theta_1 - 1/\theta_2})$, while the right hand side presents a monotonic power law dependence in r . Consequently, for Eq. (7) to be consistent we need to set the exponents of r to be both zero. One may verify that this leads to

$$\theta_1 = \frac{1}{\theta_2}, \quad d_\ell = \frac{d_f}{\theta_1}. \quad (8)$$

For the case of MPT, we have $\theta_1 = d_{min}$. It has been claimed [7, 23] that for the MPT in dimension $d = 2$, Eq. (8) is not valid, with $d_\ell < d_f/d_{min}$. Moreover, [7, 23] propose a range for the exponent d_ℓ and indicate that numerical results for $d = 4$ and $d = 5$ place d_{min} outside this range. As we have shown, in any dimension, a breakdown of the relations (8) would be incompatible with the scaling hypothesis for the conditional probabilities (4).

The Eq. (8) show us that the exponents controlling the scaling of the conditional distributions in Eq. (4) are in fact related. Also, these exponents may be obtained from the function $N_\ell(\ell) \sim \ell^{d_\ell-1}$, with $\theta_1 = d_f/d_\ell$. This gives us an indirect way to obtain numerically the fractal dimension of the average chemical distance in the spanning tree, $\bar{\ell}(r) \sim r^{\theta_1}$, by examining the form of $N_\ell(\ell)$. In the following, we do this for each of the possible strategies to build a spanning tree.

We start by generating a sample of critical clusters using the Leath-Alexandrowicz algorithm [13, 24]. Each sample grows up to the minimal distance $\ell_{min} = 512$ and samples that do not reach this size are discarded. The lattice size used exceeds 512, so, in practice, this maximum chemical distance is the only scale constraint of the growing cluster. For each cluster we obtain the multiple spanning trees following each of the previously described strategies. Averages were obtained from 10^5 samples.

Table I. Numerical estimative of the fractal dimension of the strands of the spanning tree (θ_1) in 2 and 3 dimensions.

	2 d	3 d
MPT	1.130 ± 0.003	1.375 ± 0.003
MST	1.227 ± 0.004	1.482 ± 0.004
DFS	1.331 ± 0.004	1.575 ± 0.008
RFS	1.331 ± 0.004	

Fig. (3) presents our results for $N_\ell(\ell)$ for each strategy, in 2 and 3 dimensions. The slopes of $N_\ell(\ell)$ in the log-log plot gives us the exponents d_ℓ , and using the values of $d_f = 1.89$, for $d = 2$, and $d_f = 2.52$ for $d = 3$, we get the values of $\theta_1 = 1/\theta_2 = d_f/d_\ell$ presented in Table I. In every case the exponent θ_1 gives the fractal dimension of the strands of the spanning tree. For the MPT, both in $d = 2$ and $d = 3$, the values obtained for the scaling exponent θ_1 agree with the expected fractal dimension of the minimum path $d_{min} = 1.130$ [12] and 1.375 [12], respectively. In the same way, for the MST we obtain values consistent with the dimension of the optimum path, $d_{opt} = 1.22$ [31] and 1.482 ± 0.004 . In the case of the DFS and RFS, in $d = 2$ we obtain $\theta_1 = 1.331$, a value consistent with the dimension as the external hull $d_{hull} = 4/3$ [32]. In three dimensions the external hull has the same fractal dimension of the critical cluster. The Right-First Search is not defined in $d = 3$, but for the DFS we obtain an exponent $\theta_1 = 1.575 \pm 0.008$. This seems to be a new exponent, that coincides with d_{hull} only in two dimensions.

To illustrate the usefulness of our approach, In Fig (4), we show the result for the scaling plot of the probability $P(r|\ell) \times \ell^{\theta_2}$ vs $r \times \ell^{\theta_2}$, where we used the values of Table I and the fact that $\theta_2 = 1/\theta_1$. These scaling give further support to the scaling exponents obtained.

III. CONCLUSIONS

In this work, we investigated the relationship between chemical and Euclidean distances in critical percolation clusters by analyzing spanning trees constructed through four traversal strategies: Minimal Path Trees (MPT), Depth-First Search (DFS), Right-First Search (RFS), and Minimum Spanning Trees (MST).

Our central observable was the conditional probability distribution $P(r|\ell)$, which describes the Euclidean distance r given a chemical distance ℓ defined along a spanning tree. This formulation allowed us to probe the geometric properties of different self-similar structures embedded in the percolation cluster. Each tree construction produced a distinct notion of chemical distance and, correspondingly, a distinct scaling behavior.

The scaling hypothesis means that conditional distributions should follow universal scaling forms $P(\ell|r) = r^{-\theta_1} g_\ell(\frac{\ell}{r^{\theta_1}})$ and $P(r|\ell) = \ell^{-\theta_2} g_r(\frac{r}{\ell^{\theta_2}})$. We derived and verified the relation $\theta_1 = 1/\theta_2$, which follows from scaling

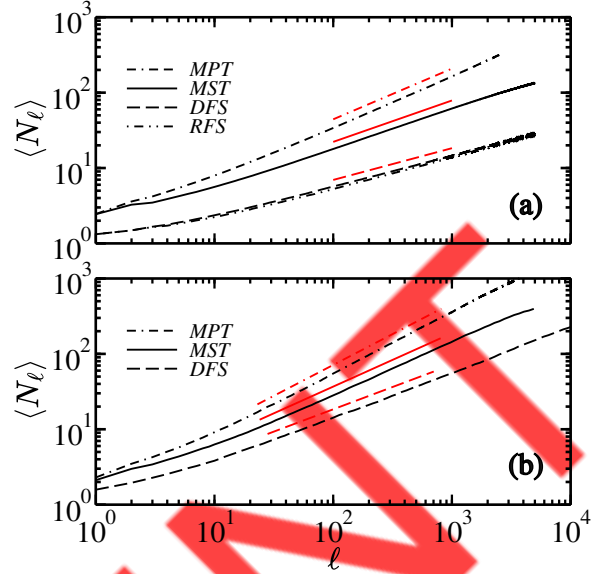


Figure 3. For $d = 2$ (a) and $d = 3$ (b), the average number of sites in the ℓ -th shell $\langle N_\ell \rangle$ versus chemical distance ℓ . (a) The slope α obtained from the numerical data in $d = 2$ is 0.672 ± 0.005 for MPT, 0.540 ± 0.005 for MST, and 0.419 ± 0.005 for DFS and for the RFS. Thus, the values obtained for $\theta_1 = d_f/\alpha + 1$ were 1.130 ± 0.003 for MPT, 1.227 ± 0.004 for MST, and 1.331 ± 0.004 for the DFS and the RFS. The exponents for RF and DFS are consistent with the exponent of the external hull in $d = 2$. (b) For $d = 3$, the slope obtained from the numerical data is 0.832 ± 0.005 for MPT, 0.700 ± 0.005 for MST and 0.600 ± 0.005 for DFS. Thus, the values obtained for the scaling exponent θ_1 were 1.375 ± 0.003 for MPT, 1.482 ± 0.004 for MST, and 1.575 ± 0.008 for the DFS. These means are obtained over 10^5 cluster configurations for MST, and 50000 cluster configurations for DFS.

hypotheses and Bayes' theorem. The value of the exponent θ_1 depends on the type of spanning tree structure: MPTs yield the minimal path exponent d_{min} , MSTs the optimal path dimension, and DFS and RFS the external hull dimension (in two dimensions). In three dimensions, the DFS tree yields a new exponent, $\theta_1 = 1.575$, suggesting the existence of a previously unrecognized universal class for such tree-induced paths.

We have also settled a question concerning the relation between the spreading dimension d_ℓ , the fractal dimension d_f of the cluster, and the minimal path exponent d_{min} . We emphasized that the identity $d_\ell = d_f/d_{min}$ is not merely an empirical observation but follows directly from the scaling forms and Bayes' theorem. It is a strict consequence of the scaling hypothesis itself. Therefore, any violation of this identity would imply an inconsistency in the theoretical framework. Taken together, these results establish a flexible and general framework for extracting fractal dimensions associated with self-similar paths in disordered media. By varying

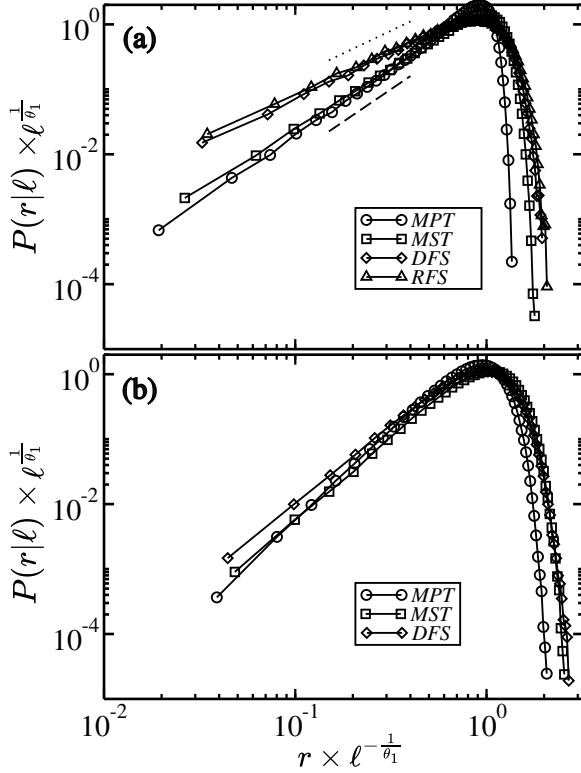


Figure 4. Rescaled plot of the conditional probability $P(r|\ell)$ for spanning trees of the critical percolating cluster. These results concern the Minimal Path Tree (MPT) (circle), Minimal Spanning Tree (MST) (square), Breadth First Search (DFS) (diamond), and right-biased search Right First Search (RFS) (triangle). Results for $d = 2$ are presented in (a) and for $d = 3$ are presented in (b). The Right First search cannot be defined in $d = 3$. The sampled critical clusters are obtained using Leath algorithm. In both $d = 2$ and $d = 3$, only clusters that grow to minimal distance 512 are sampled, and no constrained lattice size is imposed. Our averages are from 10^5 clusters. The values used for scaling exponent θ_1 for $d = 2$ were 1.130 for MPT, 1.227 for MST, and 1.331 for the DFS and the RFS. The increasing phase of the rescaled probability for the MPT and MST has a slope equal to approximately 2, represented by the dotted line, and for the DFS and RFS the slope is 1.4, as the dashed line. For $d = 3$ the values used for the scaling exponent θ_1 were 1.375 for MPT, 1.482 for MST, and 1.575 for the DFS. For $d = 3$, the increasing phase of the rescaled probability for the MPT has a slope equal to approximately 2.7, 2.53 for MST, and 2.31 for the DFS.

the spanning tree construction, one can selectively isolate different substructures and access new critical exponents. This approach can be extended to other types of clusters or growth models and may offer a path toward discovering new universality classes in geometrically constrained transport.

ACKNOWLEDGMENTS

We acknowledge the financial support from the Brazilian agencies CNPq, CAPES and FUNCAP, and the National Institute of Science and Technology for Complex Systems (INCT-SC) in Brazil.

- [1] D. Stauffer, A. Aharony, Introduction to Percolation Theory, Taylor & Francis, Philadelphia (1994).
- [2] M. Cieplak, A. Maritan, J. R. Banavar, Phys. Review L 72.15 (1994).
- [3] J. S. Andrade Jr., S. V. Buldyrev, N. V. Dokholyan, S. Havlin, P. R. King, Y. Lee, G. Paul, H. E. Stanley, Phys.

Review E 62.6 (2000).

- [4] A. A. Moreira, C. L. N. Oliveira, A. Hansen, N. A. M. Araújo, H. J. Herrmann, J. S. Andrade Jr., Phys. Review L 109.255701 (2012).
- [5] A. D. Araújo, A. A. Moreira, H. A. Makse, H. E. Stanley, J. S. Andrade Jr., Phys. Review E 66.046304 (2002).

- [6] E.López, R. Parshani, R. Cohen, S. Carmi, S. Havlin. Phys. Review letters, 99(18), 188701 (2007).
- [7] A. S. Balankin, Chaos, Solitons & Fractals 184.115044 (2024).
- [8] G. Paul, R. M. Ziff, H. E. Stanley, Phys. Review E 64.026115 (2001).
- [9] H. E. Stanley, J. Phys. A: Math. Gen. 10.L211 (1977).
- [10] D. Stauffer, Phys. Rep. 54.1, 1–74 (1979).
- [11] M. B. Isichenko, Rev. Mod. Phys. 64, 961 (1992).
- [12] Z. Zhou, J. Yang, R. M. Ziff, Y. Deng, Phys. Review E 86.021102 (2012).
- [13] Z. Alexandrowicz, Phys. Lett. A 80, 284 (1980).
- [14] R. Pike, H. E. Stanley, J. Phys. A 14, L169 (1981).
- [15] S. Havlin, R. Nossal, J. Phys. A 17, L427 (1984).
- [16] R. Rammal, J. C. Angles d'Auriac, A. Benoit, J. Phys. A 17, L491 (1984).
- [17] P. Grassberger, J. Phys. A 18, L215 (1985).
- [18] H. J. Herrmann, H. E. Stanley, J. Phys. A 21, L829 (1988).
- [19] P. Grassberger, J. Phys. A: Math. Gen. 32, 6233–6238 (1999).
- [20] Y. Deng, W. Zhang, T. M. Garoni, A. D. Sokal, A. Sportiello, Phys. Rev. E 81, 020102(R) (2010).
- [21] P. Grassberger, J. Phys. A: Math. Gen. 25, 5475 (1992).
- [22] H. E. Roman, J. Drager, A. Bunde, S. Havlin, D. Stauffer, Phys. Rev. E 52, 6 (1995).
- [23] A. S. Balankin, B. Mena, M. A. Martínez Cruz, Phys. Lett. A 381.33, 2665–2672 (2017).
- [24] P. L. Leath, Phys. Rev. B 14, 5046 (1976).
- [25] M. Porto, S. Havlin, H. E. Roman, A. Bunde, Phys. Rev. E 58, R5205(R) (1998).
- [26] M. Porto, A. Bunde, S. Havlin, H. E. Roman, Phys. Review E 56.2 (1997).
- [27] J. Vannimenus, J. P. Nadal, H. Martin, J. Phys. A: Math. Gen. 17.L351 (1984).
- [28] M. E. J. Newman, R. B. Ziff, Phys. Review L 85.19 (2000).
- [29] R. Dobrin, P. M. Duxbury, Phys. Review L 86.22 (2001).
- [30] D. Wilkinson, J. F. Willemsen, J. Phys. A: Math. Gen. 16, 3365–3376 (1983).
- [31] M. Porto, N. Schwartz, S. Havlin and A. Bunde, Phys. Rev. E 60.3 (1999).
- [32] H. Saleur, B. Duplantier, Phys. Rev. L 58.22 (1987).