



UNIVERSIDADE FEDERAL DE PERNAMBUCO
DEPARTAMENTO DE FÍSICA – CCEN
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

DISSERTAÇÃO DE MESTRADO

ANÁLISE TOPOLÓGICA DE REDES DE LIGAÇÕES DE HIDROGÊNIO EM UM SISTEMA MODELO

por

Juliana Angeiras Batista da Silva

Dissertação apresentada ao Programa de Pós-Graduação em Física do Departamento de Física da Universidade Federal de Pernambuco como parte dos requisitos para obtenção do título de Mestre em Física.

Banca Examinadora:

Prof. Francisco George Brady Moreira (Orientador-UFPE)

Prof. Ricardo Luiz Longo (DQF – UFPE)

Prof. Luiz Carlos Gomide Freitas (CCET - UFSCar)

Prof. Mauro Copelli Lopes da Silva (DF – UFPE)

Recife - PE, Brasil
Agosto - 2006

Silva, Juliana Angeiras Batista da
Análise topológica de redes de ligações de hidrogênio em
um sistema modelo / Juliana Angeiras Batista da Silva. –
Recife : O autor, 2006.
viii 62 folhas : il., fig., tab.

Dissertação (mestrado) – Universidade Federal de
Pernambuco. CCEN. Física, 2006.

Inclui bibliografia.

1. Mecânica estatística. 2. Redes complexas. 3. Transição de
fase. 4. Método de Monte Carlo. I. Título.

530.13

CDD (22.ed.)

FQ2006-009



Universidade Federal de Pernambuco
Departamento de Física – CCEN
Programa de Pós-Graduação em Física
Cidade Universitária - 50670-901 Recife PE Brasil
Fone (++ 55 81) 2126-8449/2126-8450 - Fax (++ 55 81) 3271-0359
<http://www.df.ufpe.br/pg> e-mail: posgrad@df.ufpe.br

Parecer da Banca Examinadora de Defesa de Dissertação de Mestrado

Juliana Angeiras Batista da Silva

ANÁLISE TOPOLOGICA DE REDES DE LIGAÇÕES DE HIDROGÊNIO EM UM SISTEMA MODELO

A Banca Examinadora composta pelos Professores Francisco George Brady Moreira (Presidente e Orientador), Mauro Copelli Lopes da Silva, Ricardo Luiz Longo, todos da Universidade Federal de Pernambuco e Luiz Carlos Gomide Freitas, do Centro de Ciências Exatas e de Tecnologia da Universidade Federal de São Carlos, consideram a candidata:

Aprovada com Distinção () Aprovada () Reprovada

Secretaria do Programa de Pós-Graduação em Física do Departamento de Física do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco em dez de agosto de 2006.

Prof. Francisco George Brady Moreira
Presidente e Orientador

Prof. Ricardo Luiz Longo
Co-orientador

Prof. Mauro Copelli Lopes da Silva

Prof. Luiz Carlos Gomide Freitas

Agradecimentos

A Deus, por tudo.

Aos meus pais, Flávio e Lúcia, pelo incentivo e apoio que sempre me deram.

Ao Prof. Brady, pela ótima orientação e por sua dedicação e paciência durante todo meu curso.

Ao Prof. Ricardo Longo, pela co-orientação, pela orientação não só profissional, mas também pela sua amizade e sempre disponibilidade para me ajudar.

Às minhas grandes amigas, Irlanda e Fernanda, as irmãs que escolhi, por sempre estarem comigo.

Aos amigos Maxwell, David, Róbson, Laura, Germana, Sw, Fabíola, Vanessinha, Kássia, por me ensinarem a ver a vida com outros olhos.

A Ele (ele sabe quem), que tem me ensinado a ser paciente (logo eu! rsrs), que agüentou “alguns ataques” meus na fase final, e que ainda assim, sempre me acalmou e me proporcionou momentos maravilhosos.

A Sidney, que “aperreio” desde antes de entrar no LQC, fazendo meus trabalhos de Análise Multivariada com seus dados emprestados e ainda hoje agüentando meus telefonemas de madrugada perguntando porque meus programas não querem rodar, e com sua imensa paciência sempre me ajudou. Valeu irmãozinho!

A Vivianni, por estar sempre disponível para tirar minhas dúvidas com relação ao programa ANALISYS e pela base para a realização deste trabalho. Também por sua amizade durante todos esses anos.

Aos meus amigos: Cláudia, Flávia (minha irmã química), Fred, Elisa, Carol, Paulinha, Marcus Vinícius, Renaldo, Hélcio, Ênio, Carol, Castro, Vitor, Júlio, Róbson, Wálber, Ivanildo, Rogério, Wagner e Mônica pelos momentos de companheirismo e descontração.

Ao pessoal da Física: Karlla, Priscila, Glendo, Vladimir, Antônio, Miguel, Jairo, Ernesto, Nando, Douglas, Dani, Eroni, Paulo Renato, Eduardo Padrón, Antônio Mário (Cioba), Augusto Marion, Gerson, Tâmara, Patrícia, Erms, Lincoln, Ernesto, Rafael, Fernando (maçã), Hermes, Carlos Eduardo, Marcos, André (zangado), Pedrinho, Ailton, Ferraz, Luis Arturo, Mário Henrique, René, Bernardo, Cleverson, Valderlan, Júnior.

A todos os professores do Departamento de Física que contribuíram para minha formação acadêmica, em especial, aos professores Flávio Aguiar, Sérgio Coutinho, Rios Leite e Sandra Vianna, e à Sarah, secretária da pós-graduação, sempre disposta a resolver os probleminhas burocráticos.

Às Agências Financiadoras CNPq, FACEPE, CAPES-SESu/MEC e RENAMI pela ajuda financeira.

Juliana Angeiras, agosto de 2006.

Resumo

Uma análise das propriedades topológicas locais e globais das redes de ligações de hidrogênio entre as moléculas de água, geradas a partir de simulações em condições próximas da supercrítica com o método de Monte Carlo (MC-NPT), utilizando o modelo de cinco sítios (TIP5P) para descrever as interações intermoleculares, indica o aparecimento de padrões típicos de pequenos-mundos (*small-world*), isto é, alto grau de agrupamento e pequenos tamanhos de ilhas. Em condições normais de temperatura e pressão, estas redes não apresentam tais padrões, pois se observa um alto grau de agregação e a percolação do sistema (Ref. 1). Neste trabalho, realizamos simulações computacionais MC-NPT nas condições normais (298 K e 1 atm) para a água, onde introduzimos no potencial de interação eletrostática um fator de amortecimento que modifica as cargas dos sítios de hidrogênios e de pares isolados, permitindo assim, de forma controlada, reduzir a formação das ligações de hidrogênio. Nossos resultados para o coeficiente de agrupamento (C), a conectividade média ($\langle k \rangle$) e a distância química (L), em função do fator de amortecimento λ ($0,7 \leq \lambda \leq 1$), indicam que o sistema exibe uma transição de fase topológica em $\lambda_c = 0,75$, separando as regiões de pequenos e grandes valores de C e $\langle k \rangle$, para $\lambda < \lambda_c$ e $\lambda > \lambda_c$, respectivamente. Nossa análise de estatística de ilhas mostra que na região $\lambda > \lambda_c$ tem-se a formação de um agregado com cerca de 99% dos componentes, enquanto que para $\lambda < \lambda_c$ ocorre um grande número de monômeros, dímeros e trímeros e não há mais a percolação do sistema. O comportamento do coeficiente de agrupamento e da distância química em função do parâmetro λ , exibe semelhança com o resultado obtido por Watts e Strogatz (Ref. 2), no qual caracterizam uma região de pequenos mundos. A razão C/C_{rand} , onde C_{rand} é o coeficiente de agrupamento de redes aleatórias com o mesmo número de nós (N), é independente do valor de λ , e varia linearmente com N com coeficiente angular dado pela razão $C / \langle k \rangle = 1/6$. O grau de distribuição de ligações para estas redes mostra que, para a região $\lambda > \lambda_c$, a distribuição é muito similar a uma distribuição de Poisson, mostrando que a rede se comporta como uma rede aleatória, mas para valores de $\lambda < \lambda_c$, a distribuição passa a ser assimétrica e bastante diferente da distribuição de Poisson. Cálculos das propriedades termodinâmicas, como densidade, entalpia, e entalpia de vaporização, também são indicativos da ocorrência de uma transição de fase topológica em $\lambda_c = 0,75$.

Palavras-Chaves: mecânica estatística, redes complexas, transição de fase, método de Monte Carlo

[1] V.M.L. Santos, F.G.B. Moreira and R.L. Longo, *Chem. Phys. Lett.* **390**, 157 (2004).

[2] D.J. Watts and S.H. Strogatz, *Nature* (London) **393**, 440 (1998).

Abstract

An analysis of the local and global topological properties of the hydrogen bonds networks between water molecules, generated by Monte Carlo (MC-NPT) simulations near to supercritical conditions and using the TIP5P model potential for water-water interactions, indicates the appearance of small worlds patterns with large clustering coefficient and small path lengths. The H-bonds networks in water at room condition do not present such patterns, and a high degree of clustering is observed with the presence of a giant cluster that percolates through the whole system (Ref. 1). In this work we perform MC-NPT simulations for water in normal conditions (298 K and 1 atm), where we introduce in the electrostatic interaction potential a damping factor that modifies the charges at site of hydrogen and isolated pairs. Therefore we can reduce, in a controlled way, the number of hydrogen bonds. Our results for the clustering coefficient (C), the average degree ($\langle k \rangle$) and chemical distance (L), as a function of the damping factor ($0,7 \leq \lambda \leq 1$), indicates that, at $\lambda_c = 0,75$, the model system shows a topological phase transition separating the regions of small and large values of C and $\langle k \rangle$, for $\lambda < \lambda_c$ and $\lambda > \lambda_c$, respectively. Our statistical analysis of islands in the region $\lambda > \lambda_c$ shows the existence of a giant cluster with about 99% of the components, whereas for $\lambda < \lambda_c$ a large number of monomers, dimers and trimers occurs and we do not have the percolation of the system. The behavior of the clustering coefficient and the chemical distance as a function of the parameter λ shows similarities with the results obtained by Watts and Strogatz (Ref. 2), in which they suggest a region for small worlds patterns. The ratio C/C_{rand} , where C_{rand} is the clustering coefficient for random networks with the same number of nodes (N), is not depend on the parameter λ , and varies linearly with N , with angular coefficient given by ratio $C/\langle k \rangle = 1/6$. The degree distribution for the H-bonds networks in the region $\lambda > \lambda_c$ is very similar to a Poisson distribution, as expected for random networks, but for $\lambda < \lambda_c$ the distribution becomes asymmetrical and different from Poisson distribution. Calculations of the thermodynamics functions, such as density, enthalpy and vaporization enthalpy, are also indicative of the occurrence of a topological phase transition in $\lambda_c = 0,75$.

Keywords: statistical mechanics, complex networks, phase transition, Monte Carlo method.

[1] V.M.L. Santos, F.G.B. Moreira and R.L. Longo, *Chem. Phys. Lett.* **390**, 157 (2004).

[2] D.J. Watts and S.H. Strogatz, *Nature* (London) **393**, 440 (1998).

SUMÁRIO

Resumo.....	iii
Abstract.....	iv
Índice de Figuras.....	vi
Índice de Tabelas.....	viii
1. Introdução.....	1
2. Método de Simulação de Monte Carlo.....	12
2.1. Métodos de Simulação Computacional de Líquidos.....	12
2.2 Método de Monte Carlo.....	13
2.2.1 Método de Monte Carlo com Algoritmo de Metropolis.....	18
2.3 Determinação das Propriedades Termodinâmicas.....	23
2.4 Modelos de Potencial Intermolecular da Água.....	24
2.5 Ligações de Hidrogênio.....	27
2.5.1 Ligações de Hidrogênio entre Moléculas de Água e entre os Sítios com Cargas Modificadas.....	27
2.6 Procedimentos Computacionais.....	28
3. Redes Complexas.....	30
3.1 Grafos Aleatórios.....	30
3.2 Redes de Pequenos Mundos (<i>Small-world Networks</i>)	34
3.3 Redes Livres de Escala (<i>Scale-free Networks</i>)	38
3.4 Procedimentos Computacionais.....	39
4. Resultados e Discussões.....	42
5. Conclusões e Perspectivas	58
Referências Bibliográficas.....	60

Índice de Figuras

Figura 1.1: Rede de ligações de hidrogênio da água líquida em condições ambiente (298 K, 1 bar).....	9
Figura 1.2: Rede de ligações de hidrogênio (esquerda) e o maior agregado da rede (à direita) com os nós em vermelho, para a água em condição supercrítica (673 K e 500 bar).....	9
Figura 1.3: Distribuição de ligações para redes de ligações de hidrogênio. (I) Em condições ambiente (298 K, 1 bar) (II) Em condição supercrítica (673 K, 500 bar)...	10
Figura 2.1: Condições Periódicas de Contorno.....	21
Figura 2.2: Geometria dos monômeros TIP4P (a) e TIP5P (b).....	26
Figura 2.3: Parâmetros geométricos utilizados na determinação de ligações de hidrogênio para o dímero da água.....	28
Figura 3.1: Grau de distribuição que resulta da simulação numérica de um grafo aleatório com $N = 10000$ nós e $p = 0,0015$	33
Figura 3.2: Procedimento de religação aleatório do modelo de Watts e Strogatz.....	35
Figura 3.3: Distância média $L(p)$ e coeficiente de agrupamento para o modelo de Watts e Strogatz.....	36
Figura 3.4: Comparação entre a distribuição de ligações de redes de pequenos mundos, geradas de redes quadradas, e a distribuição de Poisson para $\langle k \rangle = 4$	37
Figura 4.1: Número médio de ligações de hidrogênio para vários valores de N em função do fator de amortecimento λ	43
Figura 4.2: Número médio de ligações de hidrogênio, normalizados pelo tamanho do sistema, em função de λ para as caixas com $N \geq 512$	44
Figura 4.3: Caixas de simulação com 512 moléculas para os valores de $\lambda = 1, 0,76$ e $0,74$	45
Figura 4.4: Coeficiente de agrupamento (média configuracional sob 100 amostras) em função do parâmetro λ , para $N = 512, 640$ e 1000	46
Figura 4.5: Distância química (menor distância entre pares de nós) versus o parâmetro λ , para $N = 512, 640$ e 1000	46
Figura 4.6: Gráfico de $C(\lambda)$ e $L(\lambda)$ normalizados pelos respectivos valores em $\lambda = 1$	47

Figura 4.7: Conectividade média (média configuracional sob 100 amostras) versus o parâmetro λ , para cinco tamanhos de sistemas, $N = 128, 250, 512, 640$ e 1000	48
Figura 4.8: Gráfico de C/C_{rand} versus λ para vários tamanhos de redes ($N = 128, 250, 512, 640$ e 1000).....	51
Figura 4.9: Gráfico de C/C_{rand} versus N para vários valores λ estudados.....	51
Figura 4.10: Gráfico do coeficiente de agrupamento (C) versus N para alguns dos valores de $\lambda > 0,75$	52
Figura 4.11: Gráfico do coeficiente de agrupamento (C) versus N para alguns dos valores de $\lambda < 0,75$	52
Figura 4.12: Grau de distribuição $P(k)$ para os valores de $\lambda = 1$ e $\lambda = 0,74$	54
Figura 4.13: Número médio de ilhas com tamanho no intervalo $[2^{\text{bin}-1}, 2^{\text{bin}}-1]$ para a caixa de simulação com 640 moléculas e valor de $\lambda = 1$ (I) e $\lambda = 0,74$ (II).....	55
Figura 4.14: Densidade em g/cm^3 versus o parâmetro λ para a caixa de simulação com 640 moléculas.....	56
Figura 4.15: Entalpia por molécula, em kcal/mol , versus o parâmetro λ para a caixa de simulação com 640 moléculas.....	56
Figura 4.16: Entalpia de vaporização, em kcal/mol , em função do parâmetro λ para a caixa de simulação com 640 moléculas.....	57

Índice de Tabelas

Tabela 1.1: Características gerais de várias redes reais ($\langle k \rangle$, L e C) e de redes aleatórias (L_{rand} e C_{rand}) com mesmo tamanho e conectividade média.....	5
Tabela 1.2: Comprimento de caminho médio (L e L_{rand}), coeficiente de agrupamento (C e C_{rand}) e conectividade média ($\langle k \rangle$) para redes de ligações de hidrogênio na água.....	8
Tabela 2.1: Geometria e parâmetros para a função potencial TIP5P.....	26
Tabela 4.1: Número médio de ligações de hidrogênio para valores de λ no intervalo $0,7 \leq \lambda \leq 1,0$	42
Tabela 4.2: Comprimento de caminho médio (L e L_{rand}), coeficiente de agrupamento (C e C_{rand}) e conectividade média ($\langle k \rangle$) para as redes de ligações de hidrogênio em alguns valores de λ estudados.....	49

Capítulo 1

Introdução

Redes Complexas descrevem uma grande variedade de sistemas na natureza e sociedade. Por exemplo, as moléculas de água formam uma rede, em que os nós, representando as moléculas, estão ligados por ligações de hidrogênio [1]; uma célula pode ser descrita como uma rede complexa de substâncias químicas conectadas por reações químicas [2, 3]; a Internet é uma rede complexa de roteadores e computadores ligados por várias ligações físicas e/ou sem fio (*wireless*) [4, 5]; a *world wide web* é uma enorme rede virtual de páginas conectadas por hyperlinks [6]; redes de colaborações científicas, em que os nós são os cientistas e dois nós estão conectados se os cientistas escreveram algum artigo juntos [7]; muitas doenças sexualmente transmitidas, incluindo a AIDS, se espalham em uma rede de relações sexuais [8]. Estes sistemas representam apenas uma pequena amostra dos inúmeros exemplos que, nos últimos anos, têm chamado à atenção da comunidade científica na investigação de mecanismos que determinem a topologia e a dinâmica de crescimento de redes complexas.

Tradicionalmente o estudo de redes complexas tem sido assunto da Teoria de Grafos. Inicialmente a Teoria de Grafos se concentrou no estudo de grafos regulares e, desde 1950, redes grandes e aparentemente sem nenhum princípio organizacional têm sido descritas como grafos aleatórios, que são os exemplos mais simples de redes complexas. Paul Erdős e Alfred Rényi [9, 10], matemáticos húngaros, foram os primeiros a estudar grafos aleatórios. O modelo de Erdős-Rényi é construído a partir de N nós (ou vértices) que vão sendo conectados com probabilidade p , criando um grafo com aproximadamente $pN(N - 1)/2$ ligações distribuídas aleatoriamente. Este modelo guiou o pensamento sobre redes complexas por décadas desde sua introdução. Contudo, o crescente interesse por sistemas complexos levou os cientistas a questionar se este modelo era realmente o mais apropriado para descrever redes reais, que caracterizam os sistemas complexos presentes na natureza ou criados artificialmente. Nossa intuição indica que sistemas complexos devem apresentar algum tipo de princípio organizacional, que deve estar codificado, em algum

nível, na sua topologia. Foi a partir da observação da topologia dessas redes que novos modelos, ferramentas e medidas para se entender em termos quantitativos os princípios organizacionais básicos dessas redes foram desenvolvidos.

Desenvolvimentos tecnológicos causaram grande avanço no estudo de redes complexas. A informatização dos processos de aquisição de dados permitiu a construção de grandes bancos de dados com informação sobre a topologia de várias redes reais. O avanço computacional viabilizou o estudo de redes formadas por milhões de sítios, que há alguns anos atrás seriam praticamente intratáveis devido à enorme demanda computacional. Outro aspecto relevante nesse processo foi a crescente interação entre pesquisadores de diferentes áreas, permitindo que os mesmos tivessem acesso a dados sobre redes de diversas naturezas.

Motivado por estes fatores, vários conceitos e métodos de medição foram propostos e investigados nos últimos anos. Em particular, três conceitos ocupam um lugar de destaque no pensamento contemporâneo sobre redes complexas. O conceito de *pequenos mundos* (*small-worlds*) que, em termos simples, descreve o fato de que, independente dos seus tamanhos frequentemente grandes, na maioria das redes existe uma distância relativamente curta entre quaisquer dois nós. A distância entre dois nós é definida como o número de ligações (conexões) ao longo do caminho mais curto conectando estes nós. Em sociologia, um trabalho publicado pelo psicólogo social Stanley Milgram [11] popularizou a manifestação do conceito de pequenos mundos, com o conceito de “separação de seis graus”. Neste trabalho, concluiu-se que a distância de conhecimento entre duas pessoas nos Estados Unidos [12] tem o comprimento típico de seis pessoas. A propriedade de pequenos mundos parece caracterizar a maioria das redes complexas. Este conceito não é uma indicação de um princípio organizacional particular. De fato, como Erdős e Rényi demonstraram, a distância típica entre quaisquer dois nós em uma rede aleatória escala com o logaritmo do número de nós.

O segundo dos conceitos citados acima é o de *agrupamento* (*clustering*). O agrupamento é uma propriedade comum em redes sociais, representado por círculos de amigos ou conhecidos em que cada um dos membros de um círculo conhece todos os outros membros. Essa tendência natural de agrupamento pode ser quantificada pelo cálculo do coeficiente de agrupamento (equação 1.1), que tem suas raízes em sociologia. Digamos

que um dado sítio da rede, o sítio i , tenha k_i ligações com outros sítios da rede. Se todos os k_i vizinhos do sítio i , além do próprio sítio, formassem um círculo fechado existiriam $[k_i(k_i - 1)/2]$ ligações entre eles. O coeficiente de agrupamento do sítio i é definido como a razão entre o número de ligações que de fato existem entre seus vizinhos, E_i , e o número de ligações que existiriam se eles formassem um círculo fechado $[k_i(k_i - 1)/2]$. O coeficiente de agrupamento da rede é definido como a média dos coeficientes de todos os sítios da rede, ou seja,

$$C = \frac{1}{N} \sum_{i=1}^N \frac{2E_i}{k_i(k_i - 1)}. \quad 1.1$$

Em grafos aleatórios, uma vez que as ligações estão distribuídas aleatoriamente, o coeficiente de agrupamento é $C = p$ (a probabilidade de ligação entre dois sítios da rede). Contudo, na maioria das redes reais, se não em todas, o coeficiente de agrupamento é tipicamente muito maior que isto (para redes com mesmo número de nós).

O terceiro conceito é o de *distribuição de ligações*. Nem todos os nós em uma rede complexa têm o mesmo número de ligações, isto é, de vizinhos. O número de vizinhos de um dado sítio é definido como sendo a conectividade do sítio k_i . A conectividade dos sítios da rede satisfaz uma distribuição, a distribuição de ligações $P(k)$. Dessa forma, a probabilidade de que um sítio qualquer da rede possua k ligações é dada por $P(k)$. Uma vez que em redes aleatórias as ligações são colocadas aleatoriamente, a maioria dos nós tem aproximadamente o mesmo número de vizinhos, próximo ao grau médio $\langle k \rangle$ da rede. A distribuição de ligações de um grafo aleatório é uma distribuição de Poisson,

$$P(k) = \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle}, \quad 1.2$$

com um máximo em $P(\langle k \rangle)$.

Um dos desenvolvimentos mais interessantes em nosso entendimento de redes complexas foi a descoberta de que para a maioria das grandes redes a distribuição de ligações desvia significativamente de uma distribuição de Poisson. Em particular, para um

grande número de redes, incluindo a *world wide web* [13], a Internet [14], ou redes metabólicas [2], a distribuição de ligações tem um decaimento em lei de potência,

$$P(k) \sim k^{-\gamma} . \quad 1.3$$

Tais redes são chamadas livres de escala (*scale-free*) [15]. Enquanto algumas redes mostram decaimento exponencial, freqüentemente a forma funcional de $P(k)$ ainda desvia significativamente da distribuição de Poisson esperada para um grafo aleatório.

A tabela 1.1 mostra as características gerais para várias redes reais, junto com os respectivos valores, para comparação, das correspondentes redes aleatórias com mesmo tamanho e conectividade média.

Atualmente, três modelos têm sido utilizados no estudo de redes complexas: os grafos aleatórios, derivados do modelo de Erdős e Rényi, que ainda são muito utilizados em diversas áreas; uma classe de modelos, coletivamente chamados de modelos *small-worlds*, que são caracterizadas pelo alto valor do coeficiente de agrupamento e pela pequena distância entre sítios, interpolando entre as redes regulares altamente agrupadas e grafos aleatórios; e, por último, os modelos livres de escala (*scale-free*) devido à descoberta de que muitas redes complexas possuem distribuição de ligações do tipo lei de potência. Estes modelos são particularmente utilizados no estudo da dinâmica de redes, com o objetivo de oferecer uma teoria universal para a evolução de redes. Estes modelos estão melhores detalhados no capítulo 3.

Tabela 1.1: Características gerais de várias redes reais ($\langle k \rangle$, L e C) e de redes aleatórias (L_{rand} e C_{rand}) com mesmo tamanho e conectividade média [16].

Rede	Tamanho	$\langle k \rangle$	L	L_{rand}	C	C_{rand}	Referência
WWW, site	153127	35,21	3,1	3,35	0,1078	0,00023	Adamic, 1999 [6]
Internet (domínio)	3015- 6209	3,52- 4,11	3,7- 3,76	6,36- 6,18	0,18- 0,3	0,001 0,0007	Yook <i>et al.</i> , 2001 [4], Pastor-Satorras <i>et al.</i> , 2001 [5]
Ligações de Hidrogênio (Água Supercrítica)	512	1,3	3,33	23,78	0,2114	0,0025	Santos <i>et al.</i> , 2004 [1]
<i>E. Coli</i> , substratos	282	7,35	2,9	3,04	0,32	0,026	Wagner e Fell, 2000 [3]
<i>E. Coli</i> , reações	315	28,3	2,62	1,98	0,59	0,09	Wagner e Fell, 2000 [3]
Atores de Cinema	225226	61	3,65	2,99	0,79	0,00027	Watts e Strogatz, 1998 [17]
Col. Científica (Matemática)	70975	3,9	9,5	8,2	0,59	$5,5 \times 10^{-5}$	Barabási <i>et al.</i> , 2001 [7]
Col. Científica (Neurociências)	209293	11,5	6	5,01	0,76	$5,5 \times 10^{-5}$	Barabási <i>et al.</i> , 2001 [7]

É bem conhecido que muitas propriedades anômalas da água, tais como contração de volume sob condições de fusão, densidade máxima em condições normais, a ocorrência de pelo menos nove polimorfos cristalinos, alta constante dielétrica, altas temperaturas de transição de fase anômala e temperaturas críticas para um composto não-metálico e não-iônico com massa molecular baixa, são conseqüências das interações específicas de ligações de hidrogênio de suas moléculas e a habilidade de formarem redes tridimensionais. A questão das faixas de temperatura e densidade onde estas interações específicas podem significativamente influenciar as propriedades observáveis da água é muito importante para a construção de modelos estruturais realísticos para este fluido. Evidências experimentais sugerem que algum grau de ligações de hidrogênio ainda persiste na região supercrítica, região na qual a temperatura e a pressão estão acima do ponto crítico. Contudo, investigação experimental direta da estrutura da água supercrítica é um trabalho muito desafiador, e qualquer informação estrutural nova obtida em temperaturas e pressões altas é valiosa.

A vantagem dos “experimentos” computacionais neste contexto está na sua habilidade de gerar e analisar em detalhe o arranjo espacial e energético de cada molécula ou configurações multimoleculares no sistema, fornecendo informações termodinâmicas e estruturais em escala molecular, em geral, inacessíveis experimentalmente. Além disso, simulações computacionais fornecem propriedades que são médias temporais ou médias sobre o ensemble, que podem ser diretamente comparadas com dados experimentais.

Várias fontes de dados experimentais foram analisadas recentemente. Foi demonstrado que existe consistência entre as estimativas do grau de ligações de hidrogênio em água, em diversas condições, inferido de medidas de absorção na região do infravermelho bem como de estudos estruturais por técnicas de difração de raios-X de energia dispersiva.

A análise detalhada das ligações de hidrogênio em água supercrítica, baseada no critério distância-energia, demonstrou boa concordância dos resultados das simulações computacionais com os vários conjuntos de dados experimentais de diferentes fontes [18].

Existem dois obstáculos importantes para a caracterização do comportamento da água em termos de sua estrutura de ligações de hidrogênio. O primeiro é definir o que é ligação de hidrogênio. Tradicionalmente a ligação de hidrogênio tem sido qualitativamente interpretada como interações preferenciais da distribuição de carga positiva no átomo de hidrogênio em uma molécula com a distribuição de carga negativa no átomo de oxigênio em uma segunda molécula. Uma definição quantitativa, não ambígua, se faz necessária para decidir se um par de moléculas de água vizinhas estão ligadas por ligação de hidrogênio, ou não, permitindo relacionar a evolução das propriedades peculiares acima mencionadas da água com a distorção e/ou quebra da rede de ligações de hidrogênio.

O segundo obstáculo está relacionado ao fato de que nenhuma técnica experimental está habilitada para caracterizar completamente a microestrutura (radial e orientacional) da água em detalhes. Não obstante, desenvolvimentos recentes nas técnicas de difração de raios-X e de nêutrons têm proporcionado informações das distribuições de pares oxigênio–oxigênio, hidrogênio–hidrogênio e hidrogênio–oxigênio. Em experimentos de difração, a microestrutura da água tipicamente é obtida em termos das funções de distribuição radial (RDF) de par sítio–sítio, $g_{OO}(r)$, $g_{HH}(r)$, $g_{HO}(r)$. Enquanto as RDF's dão informações microscópicas intuitivas e podem ser utilizadas para obter alguma interpretação das

orientações moleculares (dipolar) relativas, elas não podem sozinhas caracterizar a estrutura do fluido de forma inambígua porque elas são médias orientacionais, isto é, $g_{ij}(r) = \langle g_{ij}(r_{ij}, w_i, w_j) \rangle_{w_i, w_j}$, em que $\langle \dots \rangle$ indica uma média orientacional e w_i representa uma orientação de vetor contendo os ângulos de Euler da molécula i . A combinação de difração de nêutrons com substituição isotópica (NDIS) com um método de reconstrução de imagem, baseado na expansão em harmônicos esféricos de fatores de estrutura parcial sítio-sítio, permite a extração de informações sobre a dependência orientacional sítio-sítio da função de correlação de par.

Esta é uma situação típica na qual a simulação computacional de fluidos pode ajudar na interpretação microscópica de dados experimentais, porque ela trata precisamente sistemas definidos sob controle completo e com total acesso ao ambiente molecular. A determinação das funções de distribuição espaciais por difração de nêutrons envolve três experimentos em diferentes composições isotópicas, cada um dos quais fornece informação estrutural como combinação linear dos fatores de estrutura sítio-sítio no espaço de Fourier. Como os dados no espaço- k devem ser invertidos para o espaço real, a correlação resultante das funções estão sujeitas a incertezas. Em contraste, simulação molecular fornece diretamente as distribuições espaciais e orientacionais.

Além disso, a determinação por métodos de difração de RDF's de líquidos e soluções em condições de alta pressão e temperatura apresenta dificuldades técnicas, quando comparada com as condições ambiente, como, por exemplo, o fraco espalhamento devido ao volume reduzido da amostra. Contudo, recentes avanços permitem realizar experimentos de difração de raios-X sob pressões de até 10,0 GPa e temperaturas de até 1000°C [19]. Entretanto, para a água, a intensidade de espalhamento dos raios-X é tão fraca que não permite uma análise confiável. Os resultados já obtidos com esta técnica são para condições ambiente e para as funções de distribuição radial oxigênio-oxigênio.

Uma análise das propriedades topológicas locais e globais das redes de ligações de hidrogênio entre as moléculas de água geradas a partir de simulações com o método de Monte Carlo (MC-NPT), utilizando o modelo de cinco sítios (TIP5P) para descrever as interações intermoleculares, indica o aparecimento de padrões típicos de pequenos-mundos (*small-world*), isto é, alto grau de agrupamento e pequenos tamanhos de ilhas, para simulações em condições próximas da supercrítica. Em condições normais de temperatura e

pressão, estas redes não apresentam tais padrões, pois se observa um alto grau de agrupamento e a percolação do sistema [1]. A tabela 1.2 apresenta os resultados obtidos para o coeficiente de agrupamento, comprimento de menor caminho médio (distância química) e conectividade média em condições ambiente (298 K, 1 bar) e supercrítica (673 K e 500 bar) obtidos das simulações com 300 e 512 moléculas de água. Os dados correspondem às médias configuracionais em 500 redes.

Tabela 1.2: Comprimento de caminho médio (L e L_{rand}), coeficiente de agrupamento (C e C_{rand}) e conectividade média ($\langle k \rangle$) para redes de ligações de hidrogênio na água [1].

	298 K e 1 bar		673 K e 500 bar	
	$N = 512$	$N = 300$	$N = 512$	$N = 300$
L	20,36	19,65	3,33	3,30
L_{rand}	5,23	4,90	23,78	31,28
L/L_{rand}	3,9	4,0	0,14	0,11
C	0,5456	0,5302	0,2114	0,2057
$C_{rand} = p$	0,0064	0,0106	0,0025	0,0041
C/C_{rand}	85,2	50,0	84,5	50,2
$\langle k \rangle$	3,3	3,2	1,3	1,2

Uma comparação entre os parâmetros calculados para as diferentes condições mostra que, na temperatura ambiente, as redes de ligações de hidrogênio da água líquida têm $L > L_{rand}$ e $C \gg C_{rand}$, que sugere que a distância química da rede aumenta com seu agrupamento, característico de um enorme agregado que percola por todo o sistema. Na condição supercrítica, estas redes têm distâncias químicas muito menores ($L_{rand} > L$), contudo, o agrupamento da rede é ainda grande. ($C \gg C_{rand}$), produzindo, assim, o comportamento de pequenos mundos (*small-world*) [17] das redes de ligações de hidrogênio nesta condição termodinâmica. A conectividade média apresentada na tabela 1.2 concorda com o número de ligações de hidrogênio médio, $\langle n_{HB} \rangle$, das simulações anteriores

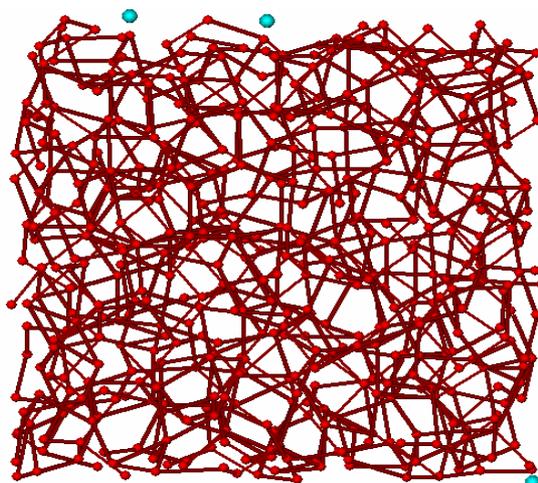


Figura 1.1: Rede de ligações de hidrogênio da água líquida em condições ambiente (298 K, 1 bar). Cada círculo (sítio) representa uma molécula de água e as conexões representam as ligações de hidrogênio. Os nós em vermelho são moléculas que estão ligadas e os nós em azul, moléculas isoladas [1].

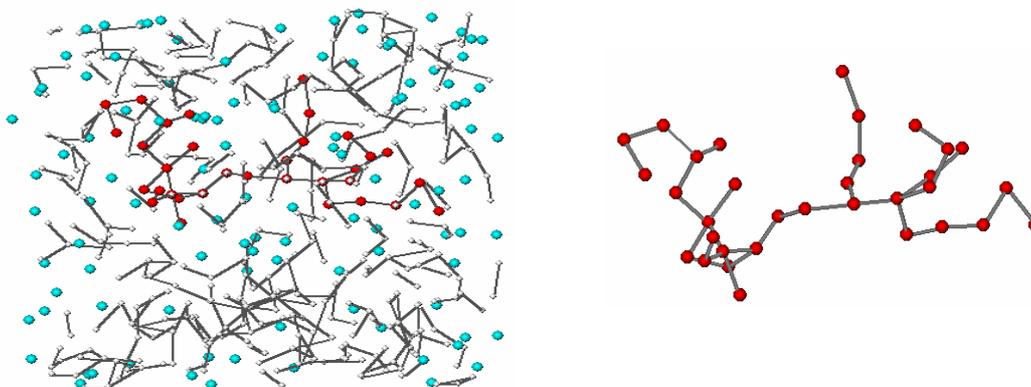


Figura 1.2: Rede de ligações de hidrogênio (esquerda) e o maior agregado da rede (à direita) com os nós em vermelho, para a água em condição supercrítica (673 K e 500 bar). Cada círculo (nó) representa uma molécula de água e as conexões representam as ligações de hidrogênio. Os nós em cinza são moléculas que estão ligadas e os nós em azul, moléculas isoladas [1].

[18]. Em particular, a conectividade média para a condição supercrítica está abaixo do limiar de percolação, $\sim 1,56$ [20], de forma que não existe uma rede contínua de ligações de hidrogênio, enquanto que na condição ambiente a percolação existe, como pode ser constatado pela visualização das caixas de simulação ilustradas nas figuras 1.1 e 1.2.

A distribuição de ligações, ou grau de distribuição, para estas redes está apresentada na figura 1.3. Em temperatura ambiente, esta distribuição é muito similar à distribuição de Poisson, compatível com o comportamento de grafos aleatórios (Figura 1.3 a). Contudo, para a água em condição supercrítica (Figura 1.3 b), a distribuição de ligações desvia bastante do comportamento de grafos aleatórios, se assemelhando às distribuições dadas por uma lei de potência com função de corte [21]. Esta forma da distribuição está relacionada às restrições limitando o número de conexões, sendo as mesmas dificilmente formadas. Isto é consistente com os resultados observados para redes de ligações de hidrogênio da água supercrítica, nas quais a formação de uma interação é muito custosa devido à baixa densidade de moléculas e maiores distâncias intermoleculares.

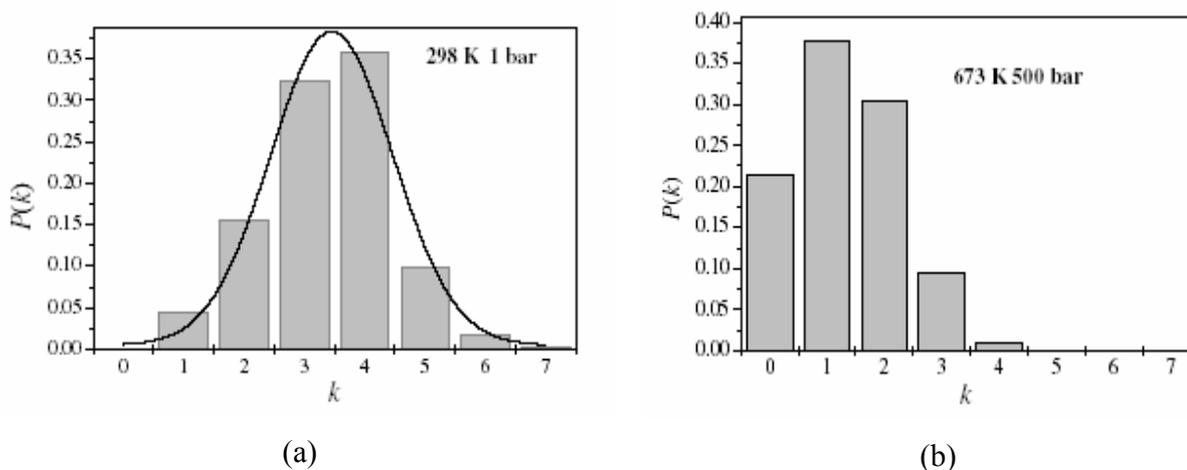


Figura 1.3: Distribuição de ligações para redes de ligações de hidrogênio. (a) Em condições ambiente (298 K, 1 bar) (b) Em condição supercrítica (673 K, 500 bar) [1].

Devido à sua aparente universalidade, padrões de pequenos mundos podem produzir informações sobre as propriedades de redes complexas, tais como desenvolvibilidade, adaptabilidade e robustez com relação à flutuações externas, podendo produzir novas

informações e perspectivas sobre a relação entre as propriedades da água e as redes de ligação de hidrogênio.

Motivados pelos resultados obtidos por Santos *et al.* [1], decidimos introduzir na parte eletrostática do potencial de interação,

$$E_{ab} = \sum_{i \in a} \sum_{j \in b} \frac{q_i q_j e^2}{4\pi\epsilon_0 r_{ij}} + 4\epsilon_0 \left[\left(\frac{\sigma_o}{r_{o-o}} \right)^{12} - \left(\frac{\sigma_o}{r_{o-o}} \right)^6 \right], \quad 1.4$$

um fator de amortecimento, λ , que modifica as cargas dos sítios (hidrogênio e pares isolados), permitindo assim, de forma controlada, reduzir a formação das ligações de hidrogênio. Consideramos neste trabalho, o modelo de potencial dado por

$$E_{ab} = \sum_{i \in a} \sum_{j \in b} \frac{\lambda q_i \lambda q_j e^2}{4\pi\epsilon_0 r_{ij}} + 4\epsilon_0 \left[\left(\frac{\sigma_o}{r_{o-o}} \right)^{12} - \left(\frac{\sigma_o}{r_{o-o}} \right)^6 \right]. \quad 1.5$$

Variações no parâmetro λ , no intervalo [0,70, 1], fazem com que o sistema deixe de ser água líquida nas condições ambiente ($\lambda = 1$), e passe a ser um sistema modelo, no qual a interação entre as moléculas passa a ser mais fraca, reduzindo, assim, a formação de ligações de hidrogênio.

Esperamos que a análise das propriedades topológicas das redes de ligação de hidrogênio em função deste fator de amortecimento forneça informações adicionais sobre os regimes de percolação, de pequenos-mundos e aleatório, bem como a transição entre estes regimes.

Capítulo 2

Método de Simulação de Monte Carlo

Neste capítulo serão abordados o método de simulação computacional Monte Carlo, os modelos de potenciais utilizados para a água, a definição de ligação de hidrogênio utilizada para a descrição das redes e os procedimentos computacionais adotados para gerar as redes de ligação de hidrogênio.

2.1 Métodos de Simulação Computacional de Líquidos

Em décadas recentes, o desenvolvimento de novas técnicas computacionais e experimentais permitiu uma nova abordagem qualitativa e quantitativa para o estudo de líquidos. Historicamente, podemos relacionar o início dessa nova etapa com a publicação dos trabalhos pioneiros que introduziram os métodos de Monte Carlo [22] e de Dinâmica Molecular [23, 24]. Essas técnicas computacionais têm sido amplamente utilizadas para a investigação de líquidos moleculares complexos, em particular, no caso da água líquida os métodos mais utilizados para a simulação computacional são Dinâmica Molecular (MD – Molecular Dynamics) e Monte Carlo (MC). Estes métodos fornecem informações termodinâmicas, estruturais e, no caso de MD, dinâmicas de líquidos, a partir de modelos moleculares de suas interações.

Em ambos os casos, cria-se uma caixa de simulação, geralmente cúbica, com condições periódicas de contorno, contendo entre 10^2 e 10^4 moléculas. O tamanho do sistema é limitado principalmente pelos recursos computacionais disponíveis. A energia do sistema (intramolecular + intermolecular) é, em geral, calculada a partir de funções de potencial ou campos de força. A escolha criteriosa de um modelo do potencial de interação pode ser descrita somente a partir dos potenciais intermoleculares, visto que as propriedades de interesse sofrem poucos efeitos dos graus de liberdade intramoleculares. O sucesso desses métodos reside, principalmente, na qualidade dos modelos de potencial empregados, sem comprometer a viabilidade computacional dos cálculos. A aplicação

desses modelos a sistemas de tamanho reduzido, comparados aos sistemas macroscópicos, tem sido suficiente para permitir medidas de propriedades observáveis em muito bom acordo com dados experimentais.

Os métodos de Dinâmica Molecular (MD) e Monte Carlo (MC) diferenciam-se pela forma de exploração do espaço de fase, contudo, devem tender aos mesmos resultados quando o tempo de simulação MD for muito longo e o número de configurações MC muito grande. A escolha de uma ou outra metodologia depende das propriedades de interesse e do sistema a ser simulado. Por exemplo, quando se deseja calcular propriedades dependentes do tempo, como coeficientes de transporte, opta-se pelo método MD. Para investigar sistemas em equilíbrio, com pressão e temperatura bem definidas, opta-se pelo método MC.

No ensemble NpT , adotado nas simulações deste trabalho, o número de moléculas (N), a temperatura (T) e a pressão (p) são mantidas fixas, enquanto se permite variar o volume da caixa de simulação. Este ensemble é particularmente apropriado para simulação de misturas, visto que os dados experimentais são frequentemente obtidos a pressão constante e as teorias de mistura formuladas com esta suposição. O método de Monte Carlo foi adotado neste trabalho e, portanto, será detalhado a seguir [25, 26, 27].

2.2 Método de Monte Carlo

Através da função de partição podem-se obter relações entre as propriedades termodinâmicas e os níveis de energias de uma molécula. Métodos de estrutura eletrônica permitem a determinação da função de partição para gases ideais, através da determinação de seus níveis de energia utilizando, em geral, aproximações de oscilador harmônico e rotor rígido. Para sistemas compostos por moléculas interagentes, no entanto, um tratamento clássico é adequado. Nessa aproximação, a energia do sistema é representada pela Hamiltoniana, \mathcal{H} , que depende somente das posições e dos momentos das partículas constituintes. A função de partição é, então, transformada de uma soma sobre os níveis de energia para uma integral multidimensional sobre todas as posições e momentos de todas as partículas, ou seja,

$$Q_{NVT} = \frac{1}{N!} \frac{1}{h^{3N}} \int \int d\mathbf{p}^N d\mathbf{r}^N \exp\left[-\frac{\mathcal{H}(\mathbf{p}^N, \mathbf{r}^N)}{k_B T}\right] \quad 2.2.1$$

em que o fator $N!$ representa a indistinguibilidade das partículas, $\mathcal{H}(\mathbf{p}^N, \mathbf{r}^N)$ é a Hamiltoniana que corresponde à energia total do sistema, $\mathbf{p}^N \equiv \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N\}$ e $\mathbf{r}^N \equiv \{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N\}$, N é o número total de partículas, k_B a constante de Boltzmann e T a temperatura. A Hamiltoniana pode ser escrita como a soma das energias cinética e potencial do sistema.

$$\mathcal{H}(\mathbf{p}^N, \mathbf{r}^N) = \sum_{i=1}^N \frac{|\mathbf{p}_i|^2}{2m} + \mathcal{V}(\mathbf{r}^N) \quad 2.2.2$$

As integrais na função de partição podem ser separadas em integrais sobre as posições e sobre os momentos:

$$Q_{NVT} = \frac{1}{N!} \frac{1}{h^{3N}} \int d\mathbf{p}^N \exp\left[-\frac{|\mathbf{p}|^2}{2mk_B T}\right] \int d\mathbf{r}^N \exp\left[-\frac{\mathcal{V}(\mathbf{r}^N)}{k_B T}\right] \quad 2.2.3$$

Esta separação é possível apenas se a função energia potencial, $\mathcal{V}(\mathbf{r}^N)$, não for dependente das velocidades, condição satisfeita por quase todas as funções potencial normalmente utilizadas. A integral sobre os momentos pode agora ser calculada analiticamente, resultando em:

$$\int d\mathbf{p}^N \exp\left[-\frac{|\mathbf{p}|^2}{2mk_B T}\right] = (2\pi mk_B T)^{\frac{3N}{2}} \quad 2.2.4$$

A função de partição pode, assim, ser escrita como

$$Q_{NVT} = \frac{1}{N!} \left(\frac{2\pi mk_B T}{h^2}\right)^{\frac{3N}{2}} Z_{NVT} \quad 2.2.5$$

em que a integral sobre as posições,

$$Z_{NVT} = \int d\mathbf{r}^N \exp\left[-\frac{\mathcal{V}(\mathbf{r}^N)}{k_B T}\right] \quad 2.2.6$$

é freqüentemente chamada de integral configuracional.

Em um gás ideal não existem interações entre as partículas e, então, a função energia potencial, $\mathcal{V}(\mathbf{r}^N)$, é igual a zero. Logo, a integral configuracional é igual a V^N , em que V é o volume. Isto leva ao seguinte resultado, para a função de partição canônica de um gás ideal:

$$Q_{NVT}^{ideal} = \frac{V^N}{N!} \left(\frac{2\pi m k_B T}{h^2}\right)^{\frac{3N}{2}} \quad 2.2.7$$

Da discussão acima, pode-se concluir que a função de partição para um sistema real tem uma contribuição devido ao comportamento de gás ideal (os momentos) e uma contribuição devido às interações entre as partículas. Quaisquer desvios do comportamento do gás ideal são devidos às interações dentro do sistema. Isto nos habilita a escrever a função de partição como:

$$Q_{NVT} = Q_{NVT}^{ideal} Q_{NVT}^{ex} \quad 2.2.8$$

A parte em excesso (ex) da função de partição é dada por:

$$Q_{NVT}^{ex} = \frac{1}{V^N} \int d\mathbf{r}^N \exp\left[-\frac{\mathcal{V}(\mathbf{r}^N)}{k_B T}\right] \quad 2.2.9$$

Uma consequência de escrever a função de partição como o produto das funções de partição de um excesso (interações) e de um gás ideal, é a correspondente separabilidade das propriedades termodinâmicas. Por exemplo, a energia de Helmholtz está relacionada à função de partição canônica por:

$$F = -k_B T \ln Q_{NVT} \quad 2.2.10$$

Escrevendo a função de partição como um produto, equação 2.2.8 , implica em

$$F = F^{ideal} + F^{ex} , \quad 2.2.11$$

em que a contribuição de gás ideal pode ser determinada analiticamente por integração sobre os momentos.

A conclusão importante é que todos os desvios do comportamento de gás ideal são devidos à presença de interações entre as partículas no sistema. Se a função potencial utilizada for independente dos momentos, uma simulação Monte Carlo pode ser feita para calcular as contribuições do excesso que dão origem aos desvios do comportamento ideal.

Tendo estabelecido que se pode explorar o espaço de fase configuracional e obter as propriedades termodinâmicas, vamos considerar como se pode alcançar isto na prática. A média sobre o ensemble (NVT) de uma dada propriedade A é dada por,

$$\langle A \rangle = \frac{\int d\mathbf{r}^N A(\mathbf{r}^N) \exp[-\beta \mathcal{H}(\mathbf{r}^N)]}{\int d\mathbf{r}^N \exp[-\beta \mathcal{H}(\mathbf{r}^N)]} \quad 2.2.12$$

em que $\beta = 1/k_B T$, ou seja, uma média no espaço de fase configuracional de $3N$ dimensões.

Para o ensemble NpT esta média torna-se

$$\langle A \rangle = \frac{\int dV d\mathbf{r}^N A(\mathbf{r}^N, V) \exp[-\mathcal{H}(\mathbf{r}^N) + pV]/k_B T}{\int dV d\mathbf{r}^N \exp[-\mathcal{H}(\mathbf{r}^N) + pV]/k_B T} \quad 2.2.13$$

em que a integral adicional é sobre o volume do sistema. Esta equação contém o peso de Boltzmann modificado, $\exp[-\beta(\mathcal{H}(\mathbf{p}^N, \mathbf{r}^N) + pV)]$, e um espaço de integração de $(3N + 1)$ dimensões, incluindo o volume como variável adicional. O denominador de 2.2.13 é a

função de partição configuracional Z_{NpT} no ensemble NpT . As propriedades obtidas desta forma, portanto, aparecem como média sobre algum espaço de fase.

Por exemplo, a energia potencial média pode, pelo menos em princípio, ser determinada pelo cálculo da integral:

$$\langle \mathcal{V}(\mathbf{r}^N) \rangle = \int d\mathbf{r}^N \mathcal{V}(\mathbf{r}^N) \rho(\mathbf{r}^N) \quad 2.2.14$$

em que $\rho(\mathbf{r}^N)$ é a probabilidade de obter a configuração \mathbf{r}^N , isto é,

$$\rho(\mathbf{r}^N) = \frac{\exp[-\mathcal{V}(\mathbf{r}^N)/k_B T]}{Z} \quad 2.2.15$$

As configurações espaciais que farão parte da amostragem são obtidas através de um gerador de números aleatórios para modificar as coordenadas espaciais das partículas que compõem o sistema. Estas configurações são escolhidas de modo probabilístico, de acordo com um peso estatístico, proporcional à distribuição de equilíbrio termodinâmico do sistema, a distribuição de Boltzmann, portanto, ao processo de gerar configurações não está associada uma lei dinâmica de evolução temporal. A cinética de que trata o método de MC é uma cinética estocástica, em que configurações sucessivas, ou passos MC, não são, necessariamente, temporalmente consecutivas.

Como existem muitas configurações que não têm energia muito próxima à energia média de equilíbrio termodinâmico, estas configurações têm uma contribuição desprezível para o cálculo dos valores médios. Portanto, é necessário o uso de um algoritmo que amostra configurações com relativa contribuição para o cálculo dos valores médios, acelerando a convergência. O algoritmo de amostragem por importância (*importance sampling*) foi introduzido em 1953 por Metropolis *et al.* [22], juntamente com a introdução de condições periódicas de contorno no método. Com este algoritmo, as médias são calculadas sobre o espaço configuracional amostrado, constituído de um número suficientemente grande de configurações espaciais, e escolhidas de acordo com os pesos estatísticos de Boltzmann. Na próxima seção, o método para o cálculo de médias de propriedades utilizando este algoritmo será detalhado.

2.2.1 Método de Monte Carlo com Algoritmo de Metropolis

Com as condições periódicas de contorno, a amostra macroscópica, nesse caso o líquido a ser modelado, é considerada como infinitas réplicas de uma célula básica de referência. Essas condições de contorno são válidas para sistemas homogêneos. Nesses sistemas, pode-se considerar que os movimentos feitos pelas moléculas nas células réplicas são os mesmos que ocorrem na célula de referência. A célula de referência contém um número N de moléculas, usualmente entre $10^2 - 10^4$, cuja densidade média é finita e a mesma da amostra macroscópica. A vantagem desse procedimento é reduzir um problema de alta dimensionalidade a um problema tratável computacionalmente, ao mesmo tempo em que permite alcançar o limite termodinâmico.

A idéia básica do método de MC convencional é calcular numericamente as integrais envolvidas na equação 2.2.12, ou 2.2.13, aproximando-as por somas sobre um número n de pontos escolhidos aleatoriamente no espaço de fase τ ,

$$\langle A \rangle = \frac{\sum_{v=1}^n A(\mathbf{r}_v^N) \exp[-\mathcal{V}(\mathbf{r}_v^N)/k_B T]}{\sum_{v=1}^n \exp[-\mathcal{V}(\mathbf{r}_v^N)/k_B T]}, \quad 2.2.16$$

em que a energia potencial, $\mathcal{V}(\mathbf{r}_v^N)$ corresponde à energia potencial total das N moléculas na v -ésima configuração.

Entretanto, como já foi discutido na seção anterior, muitas configurações têm contribuições desprezíveis para o cálculo dessas médias e, portanto, um procedimento mais eficiente para a seleção das configurações deve estar restrito às regiões mais densas do espaço de fase e termodinamicamente mais prováveis. O algoritmo de amostragem por importância proposto por Metropolis e colaboradores (*importance sampling*) consiste em evitar a amostragem excessiva de configurações que tem pequena contribuição (peso de Boltzmann) para o cálculo das médias. Este algoritmo calcula a probabilidade de que um

movimento pode ser feito, caracterizando a transição de uma configuração s para uma configuração s' , através de um processo probabilístico, denominado processo de Markov.

Em um processo de Markov, uma mudança aleatória do estado do sistema $\{s\}$ de acordo com uma regra de transição fixa, $P(s \rightarrow s')$, gera uma trajetória aleatória (a cadeia de Markov) através do espaço de fase, $\{s_0, s_1, s_2, \dots\}$. Os elementos de matriz $\pi(s \rightarrow s')$ devem satisfazer uma condição óbvia: eles não devem destruir a distribuição de equilíbrio uma vez que esta já tenha sido alcançada. Isto significa que, no equilíbrio, o número médio de movimentos tentativos aceitos saindo da configuração s para s' deve ser exatamente igual ao número de movimentos tentativos aceitos saindo da configuração s' para s . É conveniente impor condição muito mais forte, que no equilíbrio o número médio de movimentos aceitos de s para qualquer outra configuração s' seja exatamente cancelado pelo número de movimentos reversos. Este é o balanço detalhado, que significa que a taxa de transição de s para s' é igual à taxa de transição de s' para s ,

$$\pi(s)P(s \rightarrow s') = \pi(s')P(s' \rightarrow s) \quad 2.2.17$$

em que $P(s \rightarrow s')$ é a regra de transição e $\pi(s)$ é proporcional à distribuição de Boltzmann, $\pi(s) \propto e^{-\beta E(s)}$, sendo $E(s)$ a energia do estado s . Se as funções $\pi(s)$ e $P(s \rightarrow s')$ satisfazem a condição do balanço detalhado e se $P(s \rightarrow s')$ é ergódica (isto é, todo ponto acessível no espaço de configurações pode ser alcançado em um número finito de passos Monte Carlo de qualquer outro ponto), então a trajetória aleatória irá convergir para π . No método introduzido por Metropolis, a regra de transição é separada em duas contribuições tal que o balanço detalhado seja preservado,

$$P(s \rightarrow s') = D(s \rightarrow s')a(s \rightarrow s'), \quad 2.2.18$$

sendo $D(s \rightarrow s')$ uma distribuição da amostragem *a priori* e $a(s \rightarrow s')$ a probabilidade de aceitação, com $0 \leq a \leq 1$. De acordo com Metropolis, esta probabilidade de aceitação é dada por

$$a(s \rightarrow s') = \min[1, q(s' \rightarrow s)], \quad 2.2.19$$

em que

$$q(s \rightarrow s') = \frac{\pi(s')D(s' \rightarrow s)}{\pi(s)D(s \rightarrow s')} = e^{-\beta[E(s')-E(s)]}, \quad 2.2.20$$

assumindo que a distribuição de amostragem *a priori* seja uniforme e simétrica, $D(s' \rightarrow s) = D(s \rightarrow s')$.

Assim, os pontos na trajetória evoluirão de acordo com um processo Markoviano segundo a probabilidade de transição 2.2.20. A ordem sequencial de estados gerados nesse processo não está, necessariamente, relacionada ao tempo físico e só dependem unicamente do estado atual do sistema. Dito de outra forma, as configurações geradas serão selecionadas de acordo com o peso de Boltzmann 2.2.20, as quais podem ou não ser temporalmente consecutivas.

Para gerar as configurações tentativas que passarão pelo critério dos pesos de Boltzmann, inicialmente colocam-se as N moléculas em uma configuração arbitrária em uma célula de referência (caixa de simulação). Uma configuração tentativa pode ser gerada fazendo-se um movimento de translação ou rotação, para uma molécula escolhida aleatoriamente. Se os sistemas estão sob condições do ensemble NpT , variações do volume da célula também levam a novas configurações.

Movimentos de translação da n -ésima molécula ao longo de cada um dos três eixos cartesianos fixos no espaço pode ser feito de acordo com as equações abaixo

$$x_n^v \rightarrow x_n^{v'} = x_n^v + (2\xi_1 - 1)\Delta r_{\max}, \quad 2.2.21a$$

$$y_n^v \rightarrow y_n^{v'} = y_n^v + (2\xi_2 - 1)\Delta r_{\max}, \quad 2.2.21b$$

$$z_n^v \rightarrow z_n^{v'} = z_n^v + (2\xi_3 - 1)\Delta r_{\max}, \quad 2.2.21c$$

em que ξ_k são números aleatórios sorteados uniformemente no intervalo $[0, 1]$ e Δr_{\max} é o máximo deslocamento permitido, ou seja, a molécula poderá deslocar-se de $-\Delta r_{\max}$ a $+\Delta r_{\max}$ na caixa de referência. De acordo com as condições periódicas de contorno, se a molécula, como resultado dos movimentos acima, sair da caixa de referência, uma outra molécula na

caixa réplica entrará pelo lado oposto, de modo a manter constante o número N de moléculas.

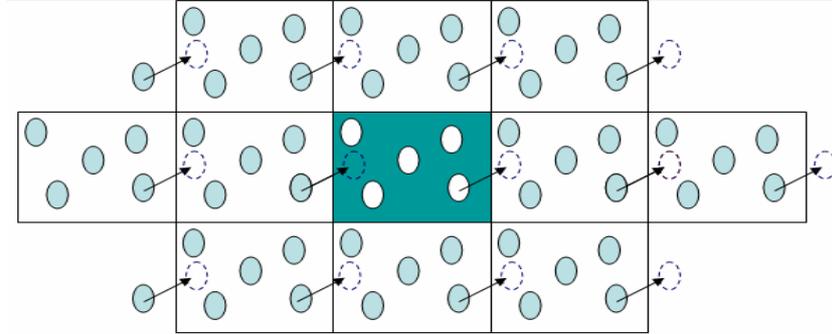


Figura 2.1: Condições Periódicas de Contorno

A rotação ao redor do centro de massa (ou centro de simetria da molécula) da n -ésima molécula é feita em um dos três eixos cartesianos fixos no espaço e escolhido aleatoriamente. Uma vez escolhido um eixo para a rotação, por exemplo x , os deslocamentos angulares para os outros dois eixos cartesianos, y e z , podem variar entre $-\Delta\Theta_{max}$ e $+\Delta\Theta_{max}$ de acordo com:

$$\theta_n^v \rightarrow \theta_n^{v'} = \theta_n^v + (2\xi_5 - 1)\Delta\Theta_{max} \quad 2.2.22a$$

$$\psi_n^v \rightarrow \psi_n^{v'} = \psi_n^v + (2\xi_6 - 1)\Delta\Theta_{max} \quad 2.2.22b$$

Para o ensemble NpT , o sistema também pode expandir ou contrair seu volume. Para uma caixa de referência cúbica as três dimensões são iguais, de modo que $V = L^3$. Os movimentos de contração ou expansão são tentados de $-\Delta V_{max}$ a $+\Delta V_{max}$ de acordo com,

$$V^v \rightarrow V^{v'} = V^v + (2\xi_7 - 1)\Delta V_{max}. \quad 2.2.23$$

Após o movimento de volume, todas as coordenadas das moléculas na caixa de referência são multiplicadas pelo fator $\alpha = L'/L$. Em geral, esses movimentos são tentados com uma frequência controlada, ao longo da simulação. Isso implica em uma redução de tempo computacional, uma vez que em movimento de volume, todos os $\frac{1}{2}[N(N-1)]$

elementos da matriz de interação entre as N moléculas devem ser recalculados, ao passo que quando só uma molécula é movimentada por translação ou rotação somente uma linha desta matriz deve ser calculada.

Os deslocamentos dos movimentos em cada iteração são governados pelos deslocamentos máximos, que são parâmetros ajustáveis cujos valores são normalmente escolhidos de modo que razão de aceitação fique em torno de 40-50%. Se os deslocamentos máximos forem muito pequenos, muitos movimentos serão aceitos e os estados serão muito similares, de modo que o espaço de fase será explorado muito lentamente. E se os deslocamentos máximos forem muito grandes, muitos movimentos tentativos serão rejeitados, porque pode ocorrer mais facilmente a sobreposição de moléculas, aumentando muito a energia do par. Valores típicos para esses parâmetros para líquidos nas condições normais são $\Delta r_{max} = 0,15 \text{ \AA}$, $\Delta \Theta_{max} = 15^\circ$ e $\Delta V_{max} = 200-600 \text{ \AA}^3$.

Assim, pode-se resumir o algoritmo de Metropolis como segue. Inicialmente especifica-se uma configuração inicial para o sistema. Gera-se, a partir desta configuração, uma nova configuração de acordo com as equações 2.2.21 ou 2.2.22, dependendo se o movimento é de translação ou de rotação, respectivamente, ou ainda, se o ensemble é o NpT , se vai ser feita uma contração ou expansão do volume. Calcula-se, então, a diferença de energia entre a configuração nova e a antiga, $\Delta \mathcal{H} = \mathcal{H}(\mathbf{r}^N_{v'}) - \mathcal{H}(\mathbf{r}^N_v)$. Se a energia da nova configuração é menor do que a energia da configuração anterior ($\Delta \mathcal{H} < 0$) o movimento é aceito, caso contrário, é sorteado um número aleatório no intervalo $[0, 1]$ e comparado ao peso de Boltzmann (equação 2.2.20). Para o ensemble NpT , este número é comparado ao peso de Boltzmann modificado, ou seja, $\exp[\beta(-\mathcal{H}(\mathbf{r}^N) + pV)]$. Se este número for menor que o peso de Boltzmann, o movimento é aceito pelo algoritmo de Metropolis, caso contrário, esse movimento é rejeitado e retornamos a configuração anterior. A configuração anterior é, então, contada novamente para o cálculo das médias. A seguir, um novo movimento é tentado.

As configurações aceitas por este procedimento serão, então, utilizadas para o cálculo das médias com a equação 2.2.16 ou, no caso do ensemble NpT , a equação 2.2.16 com o peso de Boltzmann modificado, $\exp[\beta(-\mathcal{H}(\mathbf{r}^N) + pV)]$.

2.3 Determinação das Propriedades Termodinâmicas

Propriedades termodinâmicas, como densidade, entalpia de vaporização e energia interna são obtidas a partir de médias sobre as configurações.

A densidade, ρ , é expressa em termos do volume médio (V) pela equação

$$\rho = \frac{M}{0,6022V / N}, \quad 2.3.1$$

onde M é a massa molecular expressa em g/mol, N é o número de moléculas na caixa de simulação, V é o volume médio calculado em \AA^3 e o fator 0,6022 de conversão de unidades é introduzido de modo que a densidade seja dada em g/cm^3 [28].

A entalpia de vaporização para modelos rígidos de moléculas é aproximada pela energia de coesão do líquido e do termo $\Delta(pV)$ considerando o vapor como gás ideal [28],

$$\begin{aligned} \Delta H_{\text{vap}} &= \Delta E + \Delta(pV) \\ \Delta H_{\text{vap}} &\approx -\frac{E_{\text{liq}}}{N} + RT, \end{aligned} \quad 2.3.2$$

em que R é a constante dos gases ideais e T é a temperatura absoluta. A capacidade calorífica, o coeficiente de expansão térmica e a compressibilidade isotérmica foram calculados a partir de médias sobre flutuações, respectivamente [29],

$$C_p = (\delta H / \delta T)_p \quad 2.3.3$$

$$\alpha_p = V^{-1} (\delta V / \delta T)_p \quad 2.3.4$$

$$\beta_T = -V^{-1} (\delta V / \delta P)_T. \quad 2.3.5$$

A determinação destas propriedades através das flutuações implica num grande número de configurações e logo, alta demanda computacional. Para contornar este problema, pode-se optar pelo calculado destas propriedades através de diferenças finitas. Por exemplo, para a capacidade calorífica à pressão constante, temos

$$C_p = \left(\frac{\partial H}{\partial T} \right)_p = \left(\frac{\partial (E + pV)}{\partial T} \right)_p = \left(\frac{\partial E}{\partial T} \right)_p + p \left(\frac{\partial V}{\partial T} \right)_p \approx \frac{\Delta E}{\Delta T} + p \frac{\Delta V}{\Delta T} \quad 2.3.6$$

Dessa forma, basta executar as simulações em temperaturas acima e abaixo da condição desejada.

2.4 Modelos de Potencial Intermolecular da Água

Existem inúmeros modelos para a descrição das interações entre as moléculas de água, que são, em geral, agrupados em três categorias. No modelo de sítios de interação simples cada molécula de água é mantida com uma geometria fixa e a interação entre as moléculas é descrita usando as expressões de Lennard-Jones e coulombianas, modelos flexíveis, que permitem mudanças na estrutura da molécula, e modelos que incluem os efeitos de polarização e de muitos corpos explicitamente.

Antes de 1983, em cálculos Monte Carlo (MC) e de dinâmica molecular (MD) para a água líquida geralmente eram usadas as funções potenciais de quatro sítios de Bernal e Fowler [30], agora muito pouco utilizado, mas importante por razões históricas (data de 1933), BNS [31], MCY [31] e ST2 [32]. A partir de 1983, os modelos para água mais comumente utilizados têm sido os modelos TIP3P (*Transferable Intermolecular Potential 3-polarization*) [33], TIP4P [34] (figura 2.2 (a)), SPC (*Simple Point Charge*) [35] e SPC/E (*Extended Simple Point Charge*) [36]. Estes modelos foram desenvolvidos em conjunto com cálculos de estado líquido, geralmente a 25°C, e têm sido utilizados com sucesso para o estudo de inúmeras propriedades da água líquida frequentemente em condições longe daquelas utilizadas na parametrização original.

Não obstante, nenhum dos modelos acima conseguia reproduzir o comportamento da densidade da água em função da temperatura e pressão, $\rho(T, P)$. Água líquida sob pressão atmosférica exibe uma temperatura de densidade máxima (TMD), a 4,0°C, e sua densidade é aproximadamente constante entre -15°C e +25°C. Nenhum dos modelos comumente utilizados para a água conseguia reproduzir $\rho(T)$ bem na região de temperatura

de interesse. Vários modelos produzem uma densidade máxima mas a uma TMD diferente do valor correto. O modelo TIP4P para a água fornece uma temperatura de densidade máxima próxima de -15°C além de uma região de densidade aproximadamente constante entre -35°C e 0°C . Assim, embora a característica qualitativa do máximo da densidade existir para vários modelos simples da água, ainda faltava concordância quantitativa, bem como a forma de $\rho(T)$ que ainda não tinha sido bem reproduzida.

Em 2000, foi proposto o modelo de cinco sítios, TIP5P [28], com a mesma forma funcional (Coulomb + Lennard-Jonnes) que resultou em significativo progresso em relação aos outros modelos rígidos com cargas fixas para a água líquida. Por exemplo, a densidade e a energia em função da temperatura são reproduzidas com um erro médio inferior a 1%, no intervalo de $-37,5^{\circ}\text{C}$ a $62,5^{\circ}\text{C}$, com o valor correto da TMD; a dependência correta da constante dielétrica com a temperatura é obtida e, em particular, seu valor a 25°C reproduz o resultado correto (80); o erro médio na densidade em função da pressão até 9000 atm, a 25°C , é menor que 2% e o TMD desloca-se corretamente para uma temperatura mais baixa com o aumento da pressão; as funções de distribuição radiais e suas temperaturas diferenciais concordam bem com o experimento. O problema principal está na capacidade calorífica ser muito alta e a densidade aumentar muito rapidamente com o aumento da pressão ou diminuição da temperatura acima da TMD.

A geometria do modelo TIP5P está ilustrada na figura 2.2 (b), na qual o comprimento da ligação O–H, r_{OH} , e o ângulo de ligação H–O–H, θ_{HOH} , são os valores experimentais para a fase gasosa, isto é, $0,9572 \text{ \AA}$ e $104,52^{\circ}$, os sítios de interação carregados negativamente (denotados por L) estão localizados simetricamente ao longo da direção dos pares isolados do átomo de oxigênio na molécula de água formando um ângulo $\theta_{\text{LOL}} = 109,47^{\circ}$. Uma carga de $+0,241e$ é colocada em cada sítio de hidrogênio (denotados por H) e cargas negativas, de mesma magnitude, são colocadas nos sítios de interação dos pares isolados (L). Não existe carga no oxigênio (sítio O), portanto, apenas o potencial Lennard-Jones opera entre oxigênios. Os parâmetros para este modelo estão na tabela 2.1. Os valores dos parâmetros são denominados de OPLS (*Optimized Parameters for Liquid Simulations*), pois são obtidos através da reprodução de propriedades experimentais do líquido (densidade, entalpia de vaporização, etc.) por simulações MC.

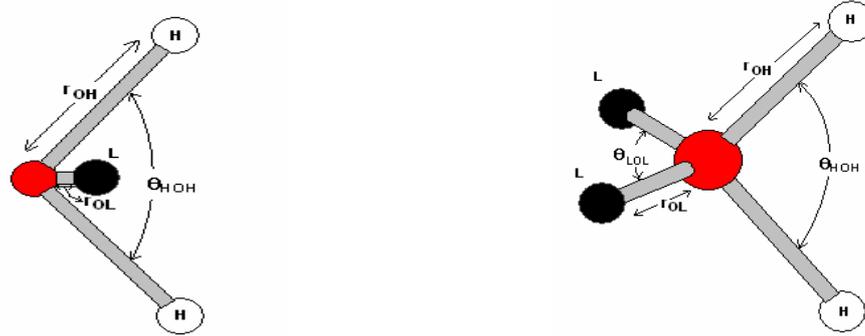


Figura 2.2 (a): Geometria do monômero TIP4P Figura 2.2 (b): Geometria do monômero TIP5P

Tabela 2.1: Geometria e parâmetros para as funções potencial TIP4P [34] e TIP5P [28]

Parâmetros	TIP4P (Figura 2.2 a)	TIP5P (Figura 2.2 b)
q_H (e)	0,52	0,2410
σ_O (Å)	3,15365	3,12
ε_O (kcal/mol)	0,1550	0,16
r_{O-L} (Å)	0,15	0,7
r_{O-H} (Å)	0,9572	0,9572
q_L (e)	-1,04	-0,2410
θ_{HOH} (°)	104,52	104,52
θ_{LOL} (°)	-	109,47

A energia potencial entre duas moléculas de água, a e b , é dada por

$$E_{ab} = \sum_{i \in a} \sum_{j \in b} \frac{q_i q_j e^2}{4\pi \varepsilon_0 r_{ij}} + 4\varepsilon_O \left[\left(\frac{\sigma_O}{r_{O-O}} \right)^{12} - \left(\frac{\sigma_O}{r_{O-O}} \right)^6 \right], \quad 2.4.1$$

onde i e j denota os sítios carregados (H e L), nas moléculas a e b respectivamente, r_{O-O} é a distância oxigênio–oxigênio, $\sigma_O = 3,12$ Å e $\varepsilon_O = 0,16$ kcal/mol.

2.5 Ligação de Hidrogênio

Um dos obstáculos importantes para a caracterização do comportamento de moléculas em termos de sua estrutura de ligações de hidrogênio é a definição do que é ligação de hidrogênio. Tradicionalmente a ligação de hidrogênio tem sido qualitativamente interpretada como interações preferenciais da distribuição de carga positiva no átomo de hidrogênio em uma molécula com a distribuição de carga negativa no átomo de oxigênio em uma segunda molécula, mas esta interação tem carência de uma definição quantitativa não ambígua. Esta definição é essencial para decidir se um par de moléculas vizinhas estão ligadas por ligação de hidrogênio ou não e para fenomenologicamente conectar a evolução das propriedades peculiares de líquidos e misturas com as características da rede de ligações de hidrogênio.

Duas abordagens mais frequentemente utilizadas são aquelas que descrevem quantitativamente ligações de hidrogênio baseadas em critérios geométricos e energéticos. Estes critérios devem, em princípio, ser consistentes, isto é, a energia configuracional associada com a ligação de hidrogênio deve ser uma função explícita e única dos parâmetros que definem o critério geométrico. Na prática, contudo, esta conexão pode ser não ambígua e faz surgir situações na qual a definição energética prevê a ligação de hidrogênio quando a correspondente geometria está em desacordo com a estrutura qualitativa aceita da ligação de hidrogênio.

2.5.1 Ligações de Hidrogênio entre Moléculas de Água e entre Sítios com Cargas Modificadas

Para o estudo do comportamento das ligações de hidrogênio adotamos um critério com as restrições geométrica e energética. As restrições geométricas adotadas foram aquelas utilizada por Chialvo e Cummings [37], ilustrada na figura 2.3. Estas são: a distância entre dois sítios de oxigênios, R_{O-O} , tem que ser $R_{O-O} \leq 3,3 \text{ \AA}$, o ângulo entre um sítio de hidrogênio (H) em uma molécula e o sítio de oxigênio (O) da outra molécula, θ_{H-O} , deve ser $\theta_{H-O} \leq 70,53^\circ$, o ângulo entre um sítio de par isolado de elétrons (L) em uma

molécula e o sítio de oxigênio (O) da outra molécula, θ_{L-O} , deve ser $\theta_{L-O} \leq 70,53^\circ$ e o ângulo diédrico, δ_D , deve ser $\delta_D \leq 180,0^\circ$. Neste modelo os sítios de pares isolados de elétrons (L) estão localizados perpendicularmente ao triângulo H–O–H e tem as mesmas dimensões dele em uma disposição tetraédrica.

As restrições nos ângulos foram transformadas para a maior distância possível entre os sítios constituintes da ligação da molécula doadora de hidrogênios e a receptora, sendo a distância entre dois sítios de oxigênios (O), $R_{O-O} \leq 3,3 \text{ \AA}$ e a distância entre o sítio de hidrogênio (H) em uma molécula e o sítio de par isolado de elétrons (L), $R_{H-L} \leq 2,74 \text{ \AA}$.

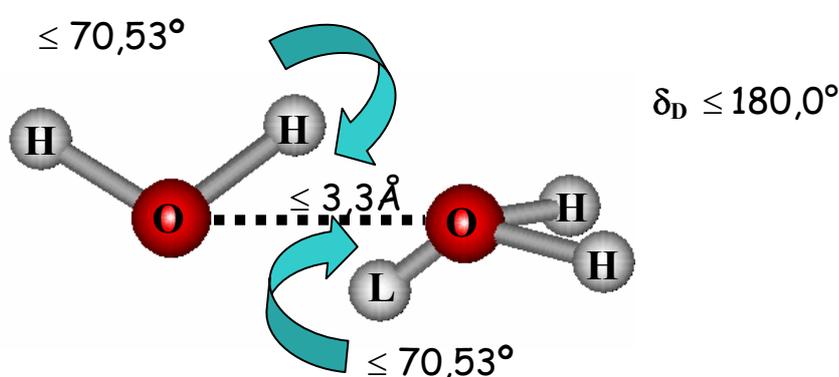


Figura 2.3: Parâmetros geométricos utilizados na determinação de ligações de hidrogênio para o dímero da água.

A existência de ligação de hidrogênio é aceita quando as restrições geométricas descritas acima são satisfeitas e ainda a energia calculada do par é negativa ($E_{ab} < 0$).

2.6 Procedimentos Computacionais

Simulações computacionais de Monte Carlo foram executadas para a água em condições normais de pressão e temperatura (1 atm e 298 K) com o modelo de potencial TIP5P [28], pois é um dos mais adequados para o estudo de propriedades estruturais da água, tendo em vista os resultados obtidos das RDF's por Santos *et al.* [44]. Foi introduzido no potencial de interação eletrostático (equação 2.4.1) um fator de amortecimento, λ , que modifica as cargas dos sítios dos hidrogênios e dos pares isolados de

elétrons, permitindo assim, de forma controlada, reduzir a formação das ligações de hidrogênio (equação 2.7.1). Variamos o parâmetro λ no intervalo $[0,7, 1]$.

$$E_{ab} = \sum_{i \in a} \sum_{j \in b} \frac{(\lambda q_i)(\lambda q_j)e^2}{4\pi\epsilon_0 r_{ij}} + 4\epsilon_0 \left[\left(\frac{\sigma_o}{r_{o-o}} \right)^{12} - \left(\frac{\sigma_o}{r_{o-o}} \right)^6 \right] \quad 2.7.1$$

As simulações computacionais foram realizadas com o programa DIADORIM [38], mantendo-se o número de moléculas (N), a temperatura (T) e a pressão (p) constantes e controladas.

Para cada simulação se utilizaram caixas cúbicas com raio de corte esférico igual a 9 Å, tentativa de variação de volume a cada 2000 configurações e razão de aceitação próxima de 40%, de acordo com trabalhos prévios para este tipo de potencial e fluidos [28].

Foram feitas simulações com caixas cúbicas com 128, 250, 512, 640 e 1000 moléculas. Para a equilibração, geraram-se mais de 10^7 configurações.

Capítulo 3

Redes Complexas

Redes são representadas matematicamente por grafos. Grafos são pares de conjuntos $G = \{P, E\}$, em que P é um conjunto de N nós (ou vértices) e E é um conjunto de ligações (ou arestas) que conectam dois elementos de P . A representação destes grafos é feita através de pontos (representando os nós) e linhas (representando as conexões) conectando os pontos.

Nas próximas seções introduzimos e discutimos algumas propriedades dos grafos aleatórios, redes de pequenos mundos e redes livres de escala, que são os modelos mais utilizados no estudo de redes complexas.

3.1 Grafos Aleatórios

Grafos aleatórios são grafos nos quais as ligações estão distribuídas aleatoriamente. Redes com uma topologia complexa e com princípios organizacionais desconhecidos frequentemente parecem ser aleatórias, assim, a teoria de grafos aleatórios [9, 10] é normalmente usada no estudo de redes complexas.

Paul Erdős e Alfred Rényi, em seu artigo original [9], propuseram um modelo no qual um grafo aleatório com N vértices é obtido ligando-se cada par de vértices com probabilidade p . Desta forma o número total de ligações é uma variável com valor esperado igual a $p[N(N-1)/2]$. Em um grafo aleatório, cada nó i possui um determinado número de ligações k_i . O número médio de ligações por sítio, ou a conectividade média $\langle k \rangle$, do grafo é dada por

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i = \frac{1}{N} 2 \frac{p[N(N-1)]}{2} = p(N-1) \approx pN. \quad 3.1.1$$

onde a aproximação final é válida para N muito grande, ou seja, $N \rightarrow \infty$.

A teoria de grafos aleatórios estuda propriedades do espaço de probabilidades associado com grafos com N nós, no limite $N \rightarrow \infty$. Muitas propriedades de tais grafos aleatórios podem ser determinadas usando argumentos probabilísticos. Normalmente o que se deseja determinar são questões relacionadas com a conectividade do grafo, agrupamento e dependência do diâmetro com o tamanho do grafo.

O principal objetivo da teoria de grafos é determinar a probabilidade de conexão p , na qual uma propriedade particular do grafo irá surgir. Erdős e Rényi mostraram que muitas propriedades importantes de grafos aleatórios aparecem subitamente, ou seja, em uma dada probabilidade ou quase todo o grafo tem alguma propriedade (por exemplo, todo par de nós está conectado por um caminho de ligações consecutivas) ou quase nenhum não a tem, sendo a transição muito rápida. Qual dos dois regimes prevalece, depende apenas de se ter p menor ou maior que p_c , a probabilidade crítica na qual a transição ocorre. O valor de p_c é aquele do limite $p_c = p_c(N \rightarrow \infty)$, que é obtido por escalonamento de tamanho finito. Este procedimento é aceito baseado na existência deste limite, refletindo o fato de que o limiar de percolação é independente do tamanho do sistema. Este é normalmente o caso de sistemas de dimensão finita, que incluem a maioria dos sistemas físicos de interesse para teoria de percolação e fenômenos críticos.

Uma das quantidades de interesse é o diâmetro do grafo, definido como a distância máxima entre quaisquer dois de seus nós. No caso particular de um grafo desconectado, que é constituído de vários agregados isolados, o diâmetro do grafo é definido como o diâmetro máximo de seus agregados. Grafos aleatórios tendem a ter diâmetros pequenos, desde que p não seja muito pequeno. Pode-se mostrar que o diâmetro apresenta apenas uma dependência logarítmica com o número de nós.

O diâmetro de um grafo aleatório tem sido estudado por muitos autores [39]. Uma conclusão geral é que, para a maioria dos valores de p , quase todos os grafos com mesmos N e p têm precisamente o mesmo diâmetro, dado por

$$d_{rand} = \frac{\ln(N)}{\ln(\langle k \rangle)}. \quad 3.1.2$$

Uma outra forma de caracterizar um grafo aleatório é através do cálculo da distância média entre qualquer par de nós, ou o comprimento de caminho médio. A distância média entre vértices varia com N e $\langle k \rangle$ da mesma forma que o diâmetro

$$l_{rand} \sim \frac{\ln(N)}{\ln(\langle k \rangle)}. \quad 3.1.3$$

Verifica-se que o comprimento de caminho médio de redes reais está próximo do valor para grafos aleatórios com o mesmo número de nós.

A partir da definição de coeficiente de agrupamento, dada na equação 1.1, e de que o número de ligações entre os k_i vizinhos de um nó i é dado pelo produto $p[k_i(k_i - 1)/2]$, obtém-se o coeficiente de agrupamento de um grafo aleatório

$$C = \frac{1}{N} \sum_{i=1}^N \frac{p k_i (k_i - 1)/2}{k_i (k_i - 1)/2} \quad 3.1.4$$

$$C_{rand} = p = \frac{\langle k \rangle}{N}. \quad 3.1.5$$

De fato, a probabilidade de que dois vizinhos de um nó estejam ligados entre si é igual à probabilidade de que dois nós quaisquer do grafo estejam ligados entre si.

De acordo com a relação acima, redes aleatórias apresentam uma dependência da razão $C_{rand}/\langle k \rangle$ com N^{-1} . Por outro lado, redes reais não apresentam esta dependência, e alguns trabalhos sugerem que esta razão é independente do valor de N . Esta propriedade é característica de grandes redes ordenadas, cujos coeficientes de agrupamento dependem apenas do número de coordenação da rede e não do seu tamanho [17].

O estudo da conectividade máxima e mínima de um grafo aleatório também foi feito por Erdős e Rényi, enquanto que distribuição de probabilidade de ligações foi obtida por completo alguns anos mais tarde por Bollobás [40]. A probabilidade de que um nó qualquer de um grafo aleatório tenha k ligações (isto é, k vizinhos) é dada pela distribuição binomial

$$P(k) = C_{N-1}^k p^k (1-p)^{N-1-k}. \quad 3.1.6$$

Esta expressão representa o número de arranjos em que k ligações podem ser feitas a partir de um dado nó, p^k é a probabilidade que um dado nó tenha k ligações, $(1 - p)^{N-1-k}$ é a probabilidade de que não existam outras ligações e C_{N-1}^k é a quantidade de formas de se fazer as k ligações.

Para pequenos valores de p ($p \ll 1$) e grandes valores de N ($N \gg k$), a distribuição binomial pode ser aproximada por uma distribuição de Poisson. Utilizando-se a fórmula de Stirling ($\ln N! \approx N \ln N - N$), pode-se mostrar que $C_{N-1}^k \approx N^k/k!$, e utilizando-se a aproximação $\ln(1 - p) \approx -p$, obtemos $(1 - p)^{N-1-k} \approx e^{-pN}$. Substituindo estas aproximações na equação 3.1.6, tem-se

$$P(k) \approx \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle}, \quad 3.1.7$$

o que mostra que $P(k)$ é aproximadamente uma distribuição de Poisson com média $\lambda = pN = \langle k \rangle$, ilustrada na figura 3.1

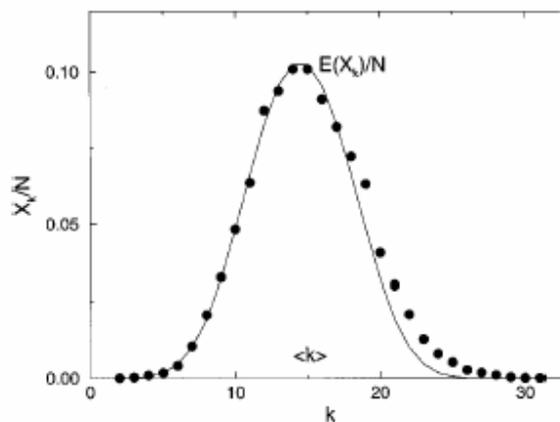


Figura 3.1: Grau de distribuição que resulta da simulação numérica de um grafo aleatório com $N = 10000$ nós e $p = 0,0015$ [16]. A linha cheia corresponde à distribuição de Poisson.

3.2 Redes de Pequenos Mundos (*Small-world Networks*)

O segundo modelo mais utilizado para descrever redes complexas são as redes de pequenos mundos. Muitas redes do mundo real têm características de pequenos mundos, apresentando topologias altamente agrupadas (característica de redes regulares) e pequenas distâncias médias (característica de redes aleatórias). Além disso, o coeficiente de agrupamento parece ser independente do tamanho da rede. Esta última propriedade é característica de redes ordenadas, cujo coeficiente de agrupamento é independente do tamanho e depende apenas do número de coordenação (ou ligações ou conexões).

D. J. Watts e S. H. Strogatz em 1998 [17], propuseram um modelo com as características acima. Este modelo tem suas raízes em sistemas sociais, no qual a maioria das pessoas é amiga dos seus vizinhos imediatos (vizinhos da mesma rua, colégios, etc.). Contudo, todo mundo tem um ou mais amigos que estão longe (pessoas em outros países, conhecimentos antigos), que são representados por ligações de longo alcance obtidas por religação no modelo Watts-Strogatz (WS).

O algoritmo para construção do modelo é o seguinte:

1. Inicia-se com uma rede regular que consiste de N nós e K ligações por nó, com $N \gg K \gg \ln(N) \gg 1$. A condição $K \gg \ln(N)$ garante que o grafo aleatório esteja conectado [41].
2. Escolhe-se aleatoriamente um nó e uma de suas ligações e , com probabilidade p (probabilidade de religação), a ligação sorteada é transferida para um outro nó também escolhido aleatoriamente. As autoligações e ligações duplicatas são excluídas.

No artigo original [17], Watts e Strogatz ilustram o processo para uma cadeia com $N = 20$ nós, onde cada nó possui $K = 4$ ligações. Se o valor da probabilidade de religação for zero, $p = 0$, a rede não recebe nenhuma ligação de longo alcance e permanece como uma rede regular. Se $p = 1$, todas as ligações de curto alcance são substituídas por ligações de longo alcance e a rede aproxima-se ao máximo de um grafo aleatório. Quando $0 < p < 1$, tem-se uma fração p de ligações de longo alcance e a topologia da rede é um meio termo entre uma rede regular e um grafo aleatório. A figura 3.2 representa o processo de religação e a transição de uma rede regular para uma rede aleatória, passando pelo regime *small-world*. Observe que esse processo de criação de uma rede *small-world* a partir de uma rede regular, proposto por Watts e Strogatz, não altera o número de ligações nem a

conectividade média da rede, uma vez que o processo de religação altera apenas o sítio final da ligação.

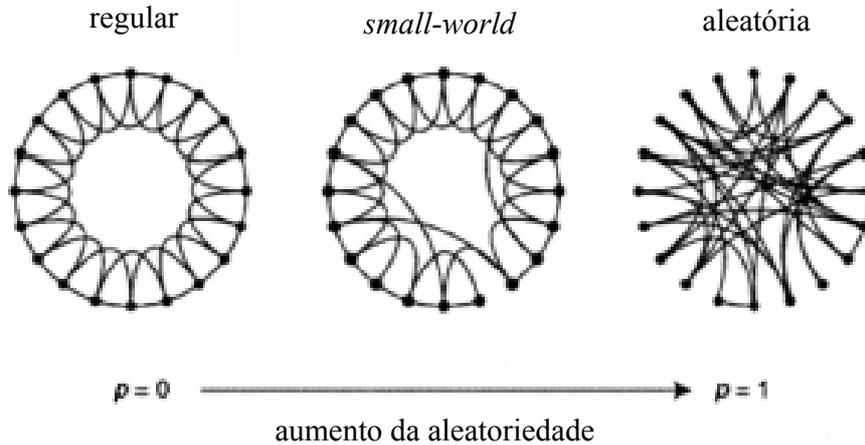


Figura 3.2: Procedimento de religação aleatório do modelo de Watts e Strogatz. Este modelo interpola entre uma rede regular e uma rede aleatória sem alterar o número de nós ou de ligações [17].

Para este modelo foi estudado o comportamento do coeficiente de agrupamento $C(p)$ e da distância média $L(p)$ como função da probabilidade de religação, p . Watts e Strogatz observaram que o comportamento dessas grandezas é bastante robusto, apresentando variações semelhantes para diversos tipos de redes regulares iniciais.

No limite $p \rightarrow 0$ tem-se

$$l(0) \sim \frac{N}{2K} \gg 1 \text{ e } C(0) \approx \frac{3}{4} \quad 3.2.1$$

assim, l depende linearmente do tamanho do sistema, e o coeficiente de agrupamento é grande. Por outro lado, quando $p \rightarrow 1$, o modelo converge para um grafo aleatório, onde se tem

$$l(1) \sim \frac{\ln(N)}{\ln(K)} \text{ e } C(1) \sim \frac{K}{N} \quad 3.2.2$$

assim, l depende logaritmicamente de N e o coeficiente de agrupamento diminui com N .

Estes casos limites podem sugerir que grandes valores de C estejam sempre associados com grandes valores de l e pequenos valores de C com pequenos valores de l , mas ao contrário, como se pode observar no gráfico da figura 3.3, existe uma grande faixa de intervalo de p no qual $l(p)$ é quase tão pequeno quanto $l(1)$, que é o próprio l_{rand} , enquanto que $C(p) \gg C(1)$. Este regime origina-se de uma rápida queda do valor de $l(p)$ para pequenos valores de p , por causa da introdução de algumas poucas ligações de longo alcance. Tais atalhos conectam nós que antes eram distantes. Para pequenos p , cada atalho tem um efeito altamente não linear em l , diminuindo a distância não somente entre pares de nós que eles conectam, como também entre seus vizinhos imediatos, vizinhos dos vizinhos, etc. Entretanto, uma ligação removida de uma vizinhança agrupada faz um atalho ter um efeito linear em C ; conseqüentemente $C(p)$ permanece praticamente inalterado para pequenos valores de p , mesmo embora $l(p)$ diminua rapidamente. Portanto, localmente, considerando-se apenas $C(p)$, a transição para *small-world* quase não é detectada.

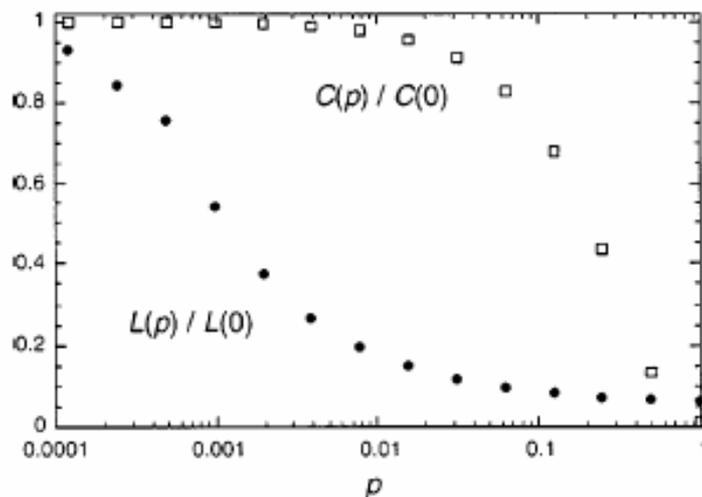


Figura 3.3: Distância média $L(p)$ e coeficiente de agrupamento para o modelo de Watts e Strogatz. Os dados estão normalizados pelos valores de $L(0)$ e $C(0)$, para uma rede regular [17].

O modelo Watts e Strogatz com $p = 0$, corresponde a uma rede regular onde cada nó da rede tem $k = \langle k \rangle$ ligações. Logo, a distribuição de ligações é uma função delta centrada em $\langle k \rangle$, $P(k) = \delta(k - \langle k \rangle)$. Para valores não nulos de p o processo de religação retira ligações de alguns nós adicionando-se a outros, provocando um alargamento da distribuição $P(k)$ em torno de $\langle k \rangle$. A forma da distribuição de ligações de uma rede *small-world* é semelhante à de um grafo aleatório, possuindo um máximo em $k = \langle k \rangle$ e decaindo exponencialmente para valores de k maiores que $\langle k \rangle$, como mostrado na figura 3.4. A linha contínua é a distribuição de Poisson, com média $\lambda = \langle k \rangle = 4$ (equação 3.1.8).

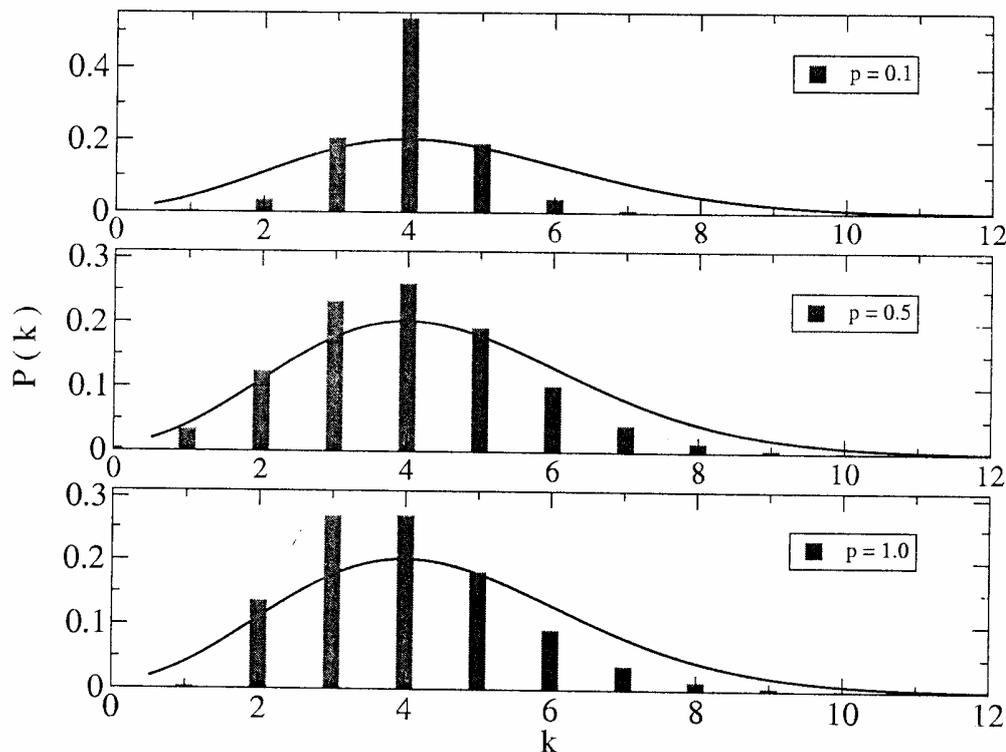


Figura 3.4: Comparação entre a distribuição de ligações de redes de pequenos mundos, geradas de redes quadradas, e a distribuição de Poisson para $\langle k \rangle = 4$ [42].

Em geral, também para redes de pequenos mundos, estamos interessados nas propriedades que caracterizam estas redes, tais como distância média, coeficiente de agrupamento e distribuição de ligações, ou grau de distribuição.

Como discutido anteriormente, há uma mudança na dependência da distância média quando a probabilidade de religação p é aumentada. Os atalhos criados aleatoriamente, conectando nós que antes estavam distantes, causa um significativo impacto no valor da distância mínima média entre os sítios. Mesmo uma fração pequena de atalhos é suficiente para drasticamente diminuir esse valor, ainda que, localmente, a rede permanece altamente ordenada. Além das pequenas distâncias médias, redes de pequenos mundos possuem um coeficiente de agrupamento relativamente alto.

3.3 Redes Livres de Escala (*Scale-free Networks*)

Por fim, têm-se as redes livres de escala (*scale-free networks*) que possuem uma distribuição de ligações do tipo lei de potência para grandes valores de k . A análise de redes reais mostrou que várias delas são do tipo livre de escala. A teoria de grafos aleatórios e o modelo Watts e Strogatz não conseguem reproduzir esta característica. Barabási e Albert [15] foram os primeiros a estudar as origens deste tipo de distribuição de ligações para redes complexas reais e propuseram possíveis explicações para este comportamento.

A primeira causa provável seria que os modelos de redes discutidos até agora assumem que a rede inicia e permanece com o número fixo de nós N , que são, então conectados ou religados aleatoriamente, sem modificar N . Na realidade, a maioria das redes reais descreve sistemas abertos que crescem pela adição contínua de novos nós. Por exemplo, a *world wide web* cresce constantemente pela adição de novos servidores e páginas na *web*, a bibliografia científica também cresce pela publicação de novos artigos, etc.

A segunda causa, seria que, nos modelos anteriores a probabilidade que dois nós estejam conectados (ou suas ligações sejam religadas) é independente do número de ligações que cada nó possui, isto é, as novas ligações são colocadas aleatoriamente. Contudo, a maioria das redes reais possui a propriedade de ligação preferencial. Isto é, a

probabilidade de um novo nó se conectar a outro nó já existente depende do número de ligações que este nó possui. Um exemplo bastante claro dessa ligação preferencial em uma rede real é o caso da rede de citações de artigos científicos. A probabilidade de que um novo artigo cite um artigo já muito conhecido, e conseqüentemente muito citado, é muito maior do que a probabilidade de que ele cite um artigo menos conhecido (menos citado).

Com base nessas duas características principais, foi proposto o modelo Barabási-Albert para redes com distribuição de ligações do tipo lei de potência. Resultados de simulações numéricas mostram que a distância média l entre nós, em uma rede Barabási-Albert, é menor que a de um grafo aleatório com os mesmo valores de N e $\langle K \rangle$ [16]. Resultados analíticos recentes [43] mostram que l depende de N da seguinte forma:

$$l \sim \frac{\ln(N)}{\ln[\ln(N)]}. \quad 3.3.1$$

Resultados numéricos [16], mostram que uma rede de Barabási-Albert possui um coeficiente de agrupamento aproximadamente cinco vezes maior que o de um grafo aleatório com os mesmos N e $\langle K \rangle$, e que $C \sim N^{-0,75}$ na rede Barabási-Albert, enquanto em um grafo aleatório $C \sim N^{-1}$ e em uma rede *small-world* C é independente de N .

Por fim, resultados analíticos [15, 16] mostram que a distribuição de ligações de uma rede Barabási-Albert tem a forma

$$P(k) \sim k^{-\gamma}. \quad 3.3.2$$

com o valor do parâmetro $\gamma = 3$.

3.4 Procedimentos Computacionais

Para a análise topológica das redes de ligações de hidrogênio, foi utilizada uma versão modificada do programa ANALYSIS desenvolvido por Santos *et al.* [44].

Este programa inicialmente lê informações como número de moléculas, número de sítios das moléculas, coordenadas da caixa de simulação e as restrições para a determinação

das ligações de hidrogênio. Depois lê as coordenadas dos átomos da caixa de simulação, equilibrada pelo programa DIADORIM [38], que estão em relação ao centro de massa das moléculas, e as transforma em coordenadas em relação à origem. Feito isto, os átomos recebem as informações necessárias para o cálculo do potencial, como carga, σ e ϵ .

A partir daí, o programa calcula as distâncias entre os átomos de oxigênio e entre os átomos de oxigênio e hidrogênio de todos os possíveis pares de moléculas e compara-as com as restrições destas distâncias necessárias para a determinação da ligação de hidrogênio. Os pares de moléculas que obedecem a estas restrições são ditos conectados e uma matriz de conexão ($N \times N$) é formada, em que se a molécula i e a molécula j estão conectadas, os correspondentes elementos nesta matriz são 1, caso contrário, são zero. Então, até aqui, tem-se uma matriz que obedece apenas ao critério geométrico para a formação da ligação de hidrogênio.

Agora, a partir dos pares conectados na rede, pelo critério geométrico, calcula-se a energia de cada par a partir das coordenadas do par e das informações de carga, σ e ϵ de cada átomo (ou sítio). Se esta energia calculada for menor que zero, o par está conectado, caso contrário, este par não satisfaz os dois critérios necessários para a formação de ligações de hidrogênio na rede e, portanto, não está conectado. Conta-se, então, o número de pares que satisfazem estes critérios e o número de ligações de hidrogênio que cada molécula faz. Agora, constrói-se uma nova matriz de conexão ($N \times N$), que obedece aos dois critérios.

Como agora a rede de ligações de hidrogênio já está formada, podem-se calcular as propriedades de redes, como coeficiente de agrupamento, conectividade média e distância química (ou comprimento de menor caminho médio).

Para calcular o coeficiente de agrupamento, conta-se com quantas moléculas uma dada molécula está conectada, k_i , a partir da matriz de conexão, verificando-se cada elemento de matriz, e calcula-se o coeficiente de agrupamento para esta molécula por

$$C_i = \frac{2E_i}{k_i(k_i - 1)}, \quad 3.4.1$$

onde E_i é o número de ligações que existem entre os k_i vizinhos, e $[k_i(k_i - 1)/2]$ é o número de ligações que existiriam se eles formassem um círculo fechado.

Faz-se isto para todas as moléculas e calcula-se, então, a média para uma dada configuração. A média final é realizada sobre 100 configurações descorrelacionadas.

A conectividade média é calculada a partir da média do número de conexões que cada molécula tem, também através da matriz de conexão.

A distância química é calculada verificando-se dentro do maior agregado a menor distância entre dois sítios, os sítios i e j , e tomando-se a média sobre todos os sítios do agregado.

$$l(i) = \frac{1}{N} \sum_{j=1}^N l_{\min}(i, j) \quad 3.4.2$$

A distribuição de ligações é calculada através de outro programa que lê um dos arquivos de saída gerado pelo programa ANALYSIS que contém informação do número de conexões que cada molécula faz. Este número é, então, armazenado em um vetor que conta as moléculas que fazem o mesmo número de conexões, o gráfico para $P(k)$.

O histograma com a distribuição dos tamanhos das ilhas nas redes foi feito a partir de outro programa que lê um dos arquivos gerado pelo programa ANALYSIS, que contém informações sobre o número de componentes em cada ilha.

Os valores destas propriedades apresentados no capítulo referente aos resultados obtidos neste trabalho correspondem às médias configuracionais em 100 amostras. Essas 100 configurações foram obtidas a partir da caixa equilibrada pelo programa DIADORIM de acordo com a seção 2.6. Entre cada uma dessas amostras há uma diferença de 10^4 configurações MC, garantindo que estas sejam configurações descorrelacionadas.

Capítulo 4

Resultados e Discussões

Neste capítulo serão apresentados e discutidos os resultados para as simulações computacionais, usando o método de Monte Carlo, das redes de ligações de hidrogênio geradas a partir da água em condições normais de temperatura e pressão. Para a descrição dos nós das redes de ligações de hidrogênio, foi utilizado o modelo de cinco sítios TIP5P com as restrições de Chialvo e Cummings [37] para formação de ligações de hidrogênio. As simulações computacionais foram realizadas com o programa DIADORIM [38] mantendo-se o número de moléculas (os nós) (N), a temperatura (T) e a pressão (p) constantes. Para a análise topológica das redes de ligação de hidrogênio foi utilizado o programa ANALYSIS, inicialmente desenvolvido por Santos [44] e modificado neste trabalho.

A tabela 4.1 apresenta o número médio de ligações de hidrogênio para alguns valores do parâmetro λ , obtido das simulações em caixas com 128, 250, 512, 640 e 1000 moléculas. Estes valores correspondem às médias configuracionais em 100 amostras.

Tabela 4.1: Número médio de ligações de hidrogênio para valores de λ no intervalo $0,7 \leq \lambda \leq 1,0$.

λ	$N = 128$	$N = 250$	$N = 512$	$N = 640$	$N = 1000$
1	192	390	834	1064	1695
0,94	182	368	791	1007	1598
0,80	127	254	551	693	1103
0,76	79	193	429	526	856
0,74	18	96	241	291	560
0,70	6	35	117	89	185

Analisando esta tabela, verifica-se que o maior número de ligações de hidrogênio está presente para o sistema com $\lambda = 1$, ou seja, quando a interação entre os sítios é mais

forte, como era de se esperar. O gráfico do número de ligações de hidrogênio, N_{HB} versus λ (figura 4.1) mostra que N_{HB} é uma função monotonicamente crescente do parâmetro λ e do número de nós (N). Tem-se ainda que para $\lambda \leq 0,75$ o número médio de ligações de hidrogênio é bem menos sensível ao tamanho do sistema (N). Na figura 4.2 mostramos a razão N_{HB}/N versus λ para redes com $N \geq 512$. Vemos que a razão N_{HB}/N é muito pouco sensível em relação ao valor de N e só depende do parâmetro λ , ou seja,

$$N_{HB} = Nf(\lambda). \quad 4.1$$

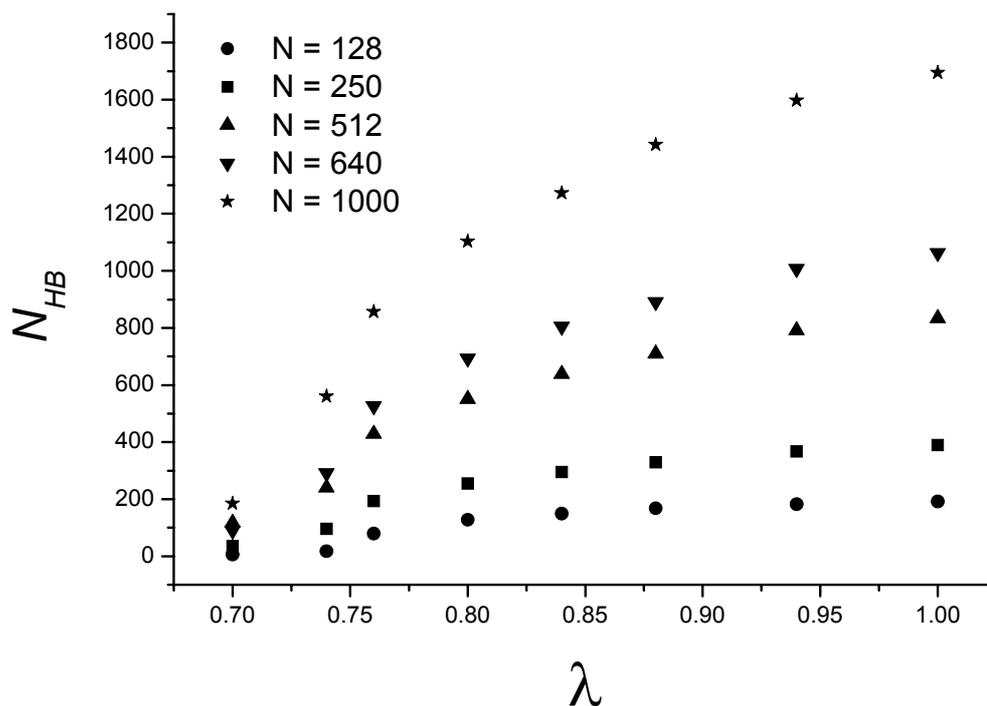


Figura 4.1: Número médio de ligações de hidrogênio para vários valores de N em função do fator de amortecimento λ .

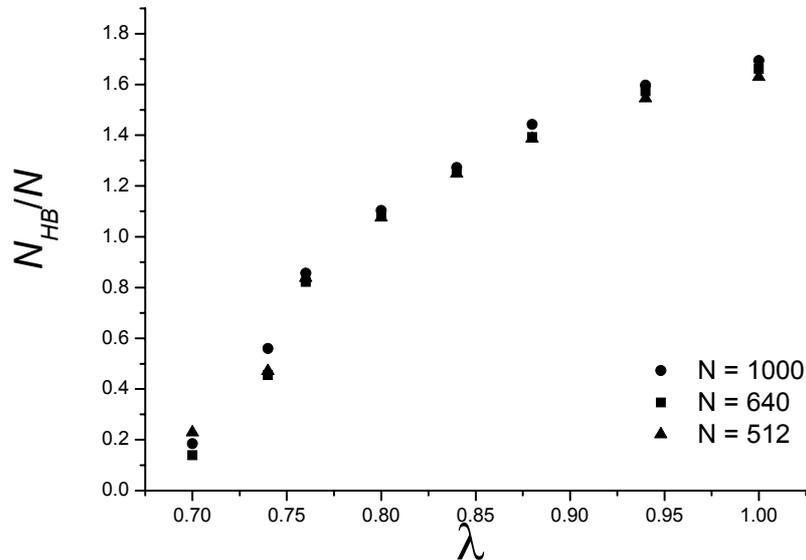


Figura 4.2: Número médio de ligações de hidrogênio, normalizados pelo tamanho do sistema, em função de λ para as caixas com $N \geq 512$.

A topologia das redes de ligações de hidrogênio mostra que o parâmetro λ no potencial de interação é determinante na formação das ligações de hidrogênio. Existe uma clara transição entre os regimes de percolação (aproximadamente $1,00 \geq \lambda \geq 0,75$) e o de agregados ou ilhas ($\lambda < 0,75$) quando as interações eletrostáticas diminuem.

No estudo da topologia das redes de ligações de hidrogênio com valores de λ menores que um (1) e, portanto, em valores de força de interação menores, é razoável esperar que ocorra a formação de agregados isolados e espaços vazios. Isso foi verificado de forma qualitativa pela visualização das caixas de simulação, algumas ilustradas na figura 4.3, e quantitativamente pelos cálculos de propriedades topológicas dos agregados, verificando-se a presença ou não de padrões *small-world*.

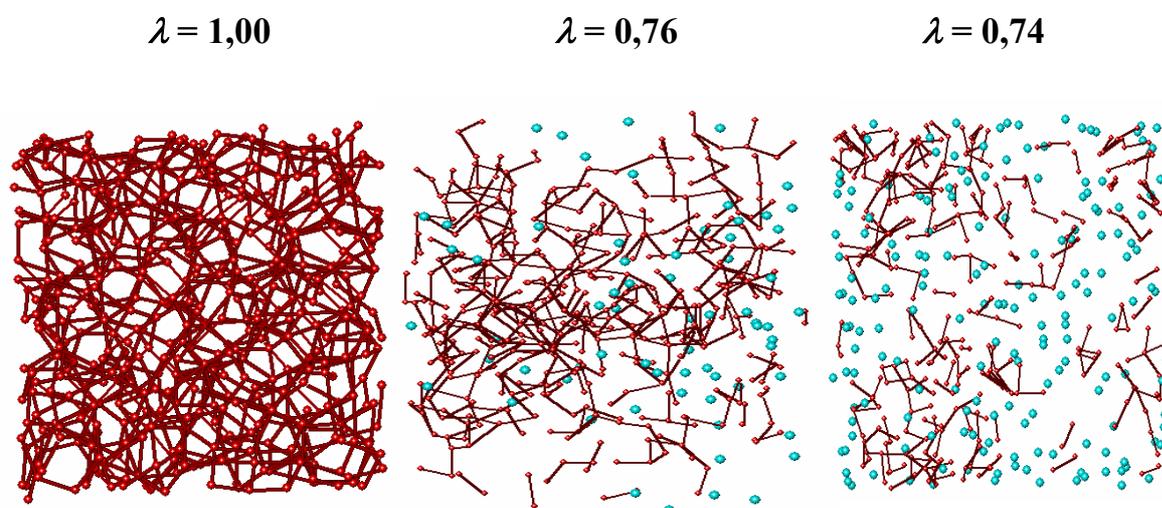


Figura 4.3: Caixas de simulação com 512 moléculas para os valores de $\lambda = 1, 0,76$ e $0,74$. Cada círculo (nó) representa uma molécula e as linhas representam ligações de hidrogênio. Os nós conectados e os nós sem conexão alguma com outros nós da rede são mostrados em vermelho e azul, respectivamente.

Foram feitos cálculos de propriedades topológicas locais, como o coeficiente de agrupamento (C), a distribuição de ligações ($P(k)$), além da distância química (comprimento de menor caminho médio) (L) que é uma propriedade global da rede. As definições destas propriedades foram apresentadas no capítulo 3. As figuras 4.4 e 4.5 mostram a dependência de C e L com o parâmetro λ para os sistemas de tamanhos $N = 512$, 640 e 1000 . Os pontos com os respectivos erros foram calculados com 100 amostras.

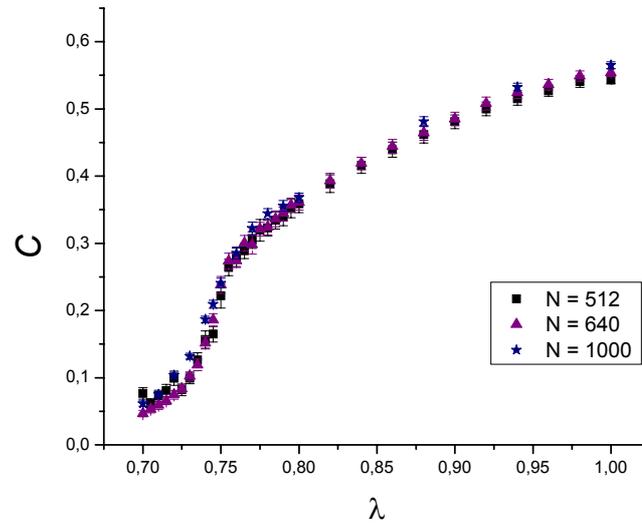


Figura 4.4: Coeficiente de agrupamento (média configuracional sob 100 amostras) em função do parâmetro λ , para $N = 512$, 640 e 1000.

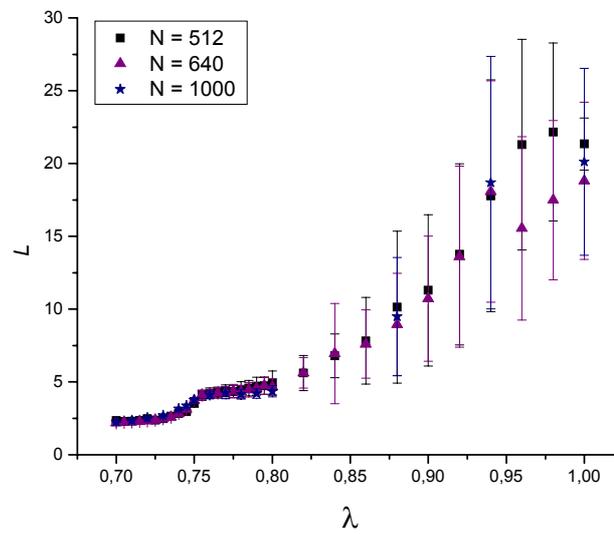


Figura 4.5: Distância química (menor distância entre pares de nós) versus o parâmetro λ , para $N = 512$, 640 e 1000.

Na figura 4.6 mostramos os resultados para o coeficiente de agrupamento e a distância química são normalizados pelos respectivos valores em $\lambda = 1$, onde a região de pequenos mundos está circundada pela linha vermelha para uma caixa de simulação com 640 nós (moléculas). Observamos que nesta região as redes possuem alto coeficiente de agrupamento, característico de redes regulares, e pequena distância química, característico de redes aleatórias. Observe a semelhança dessa figura com a figura 3.3 para a modelo de Watts e Strogatz [17]. Em ambos os gráficos há uma região de pequenos comprimentos de caminho médio e altos coeficientes de agrupamento, caracterizando a formação dos padrões de pequenos mundos nestas regiões.

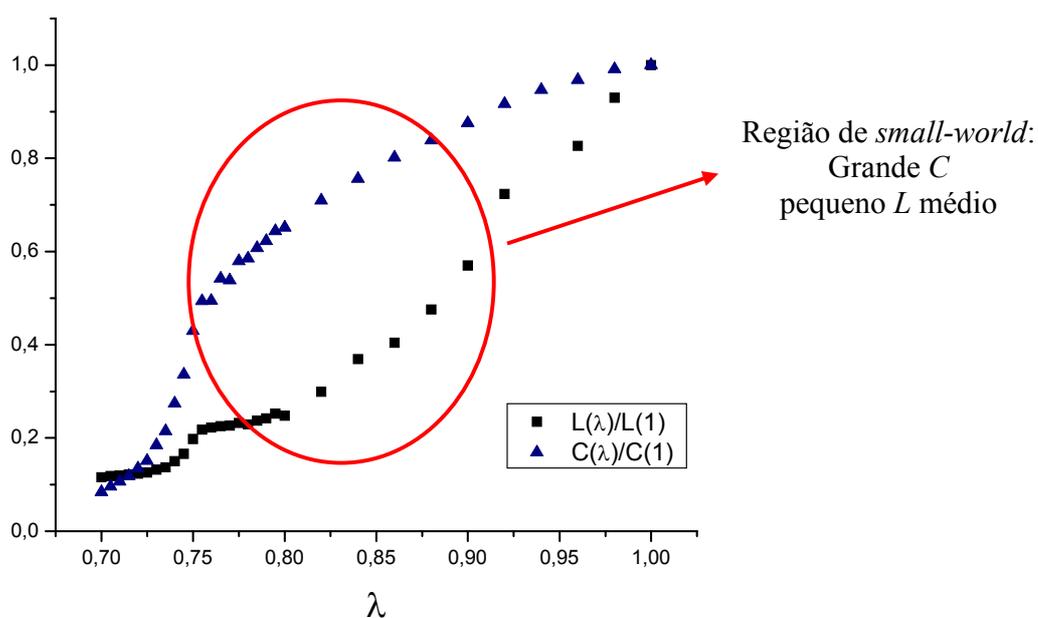


Figura 4.6: Gráfico de $C(\lambda)$ e $L(\lambda)$ normalizados pelos respectivos valores em $\lambda = 1$. A região circundada representa a região onde há a formação de padrões de pequenos mundos (*small-world*). Os dados são para $N = 640$ moléculas. Curvas semelhantes são obtidas para outros valores de N .

Na figura 4.7 mostramos os correspondentes resultados para a conectividade média, $\langle k \rangle$, igual ao número médio de ligações de hidrogênio de uma dada molécula, em função do parâmetro λ .

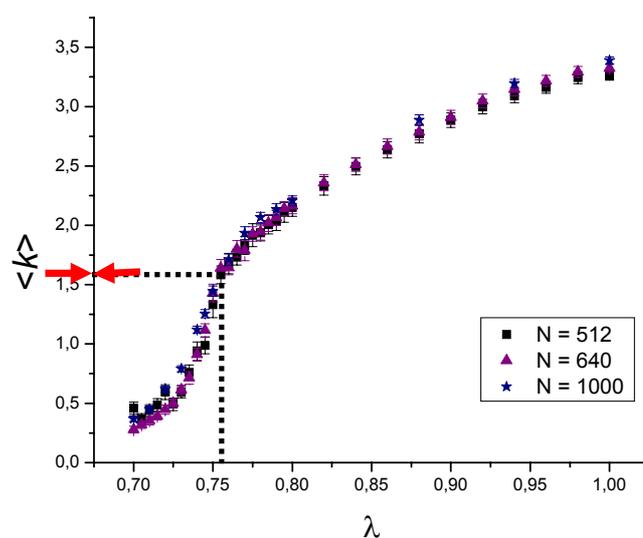


Figura 4.7: Conectividade média (média configuracional sob 100 amostras) versus o parâmetro λ , para cinco tamanhos de sistemas, $N = 512$, 640 e 1000. O valor $\langle k \rangle = 1,6$ [20] para o limiar de percolação está indicado por uma seta.

Observamos que para valores de λ abaixo de 0,76 a conectividade média é menor do que o limiar de percolação, $\sim 1,6$ [20], de modo que não se observa mais a formação de redes contínuas de ligações de hidrogênio. A ausência de percolação pode ser observada visualmente na figura 4.3, para $\lambda = 0,74$.

A tabela 4.2 apresenta alguns resultados para o comprimento de caminho médio, coeficiente de agrupamento e conectividade média, obtidos como uma média para 100 caixas de simulação com 128, 250, 512, 640 e 1000 moléculas (nós) em diferentes valores de λ . Observa-se da tabela 4.2 que a análise de apenas uma propriedade local, como o coeficiente de agrupamento, não é suficiente para a decisão de padrões de pequenos mundos, que neste caso seria um indicativo para todos os valores de λ estudados. Para efeitos de comparação, incluímos os valores de L_{rand} e C_{rand} das redes aleatórias com o mesmo número de nós.

Tabela 4.2: Comprimento de caminho médio (L e L_{rand}), coeficiente de agrupamento (C e C_{rand}) e conectividade média ($\langle k \rangle$) para as redes de ligações de hidrogênio em alguns valores de λ estudados.

$\lambda = 1$					
N	128	250	512	640	1000
L	13,73	17,46	21,34	18,81	20,12
L_{rand}	4,42	4,86	5,28	5,38	5,66
L/L_{rand}	3,11	3,65	4,04	3,50	3,55
C	0,50	0,52	0,54	0,55	0,56
$C_{rand} = p$	0,023	0,012	0,006	0,005	0,003
C/C_{rand}	21,33	41,67	85,32	106,66	166,69
$\langle k \rangle$	2,99	3,12	3,26	3,32	3,39

$\lambda = 0,94$					
N	128	250	512	640	1000
L	11,99	13,47	17,79	18,08	18,69
L_{rand}	4,65	5,12	5,53	5,64	5,95
L/L_{rand}	2,58	2,63	3,22	3,20	3,14
C	0,47	0,49	0,51	0,52	0,53
$C_{rand} = p$	0,022	0,012	0,006	0,005	0,003
C/C_{rand}	21,33	41,67	85,32	106,67	166,69
$\langle k \rangle$	2,84	2,94	3,09	3,15	3,20

$\lambda = 0,76$					
N	128	250	512	640	1000
L	3,48	3,96	4,22	4,18	4,11
L_{rand}	23,44	12,68	12,09	13,01	12,83
L/L_{rand}	0,15	0,31	0,35	0,32	0,32
C	0,21	0,26	0,28	0,27	0,29
$C_{rand} = p$	0,010	0,006	0,003	0,003	0,002
C/C_{rand}	21,34	41,66	85,32	106,67	166,63
$\langle k \rangle$	1,23	1,55	1,68	1,64	1,71

$\lambda = 0,74$					
N	128	250	512	640	1000
L	2,2	2,62	2,84	2,83	3,15
L_{rand}	-3,75	-20,92	-100,82	-68,51	60,95
L/L_{rand}	0,59	-0,13	-0,03	-0,04	0,05
C	0,05	0,13	0,16	0,15	0,19
$C_{rand} = p$	0,002	0,003	0,002	0,001	0,001
C/C_{rand}	21,30	41,68	85,34	106,71	166,69
$\langle k \rangle$	0,27	0,77	0,94	0,91	1,12

Observamos que a razão C/C_{rand} é independente do valor de λ , como mostra a figura 4.8. De fato, mostramos na figura 4.9 que esta razão depende linearmente do tamanho do sistema N com coeficiente angular da reta igual a $1/6$, ou seja,

$$\frac{C}{C_{rand}} = \frac{1}{6}N. \quad 4.1$$

Note também que a razão entre o coeficiente de agrupamento e a conectividade média das redes de ligações de hidrogênio se mantém constante ($= 1/6$), independentemente do valor do parâmetro λ . Temos que $C_{rand} = \langle k \rangle / N = p$ e, portanto,

$$C = \frac{1}{6}\langle k \rangle. \quad 4.2$$

Qualitativamente, das figuras 4.4 e 4.7, vemos que C e $\langle k \rangle$ têm uma relação de proporcionalidade entre eles. Esta relação está de acordo com o que foi discutido no capítulo 3, onde vimos que redes aleatórias apresentam uma dependência de $C_{rand} / \langle k \rangle$ com N^{-1} , enquanto que para redes reais esta razão é independente de N .

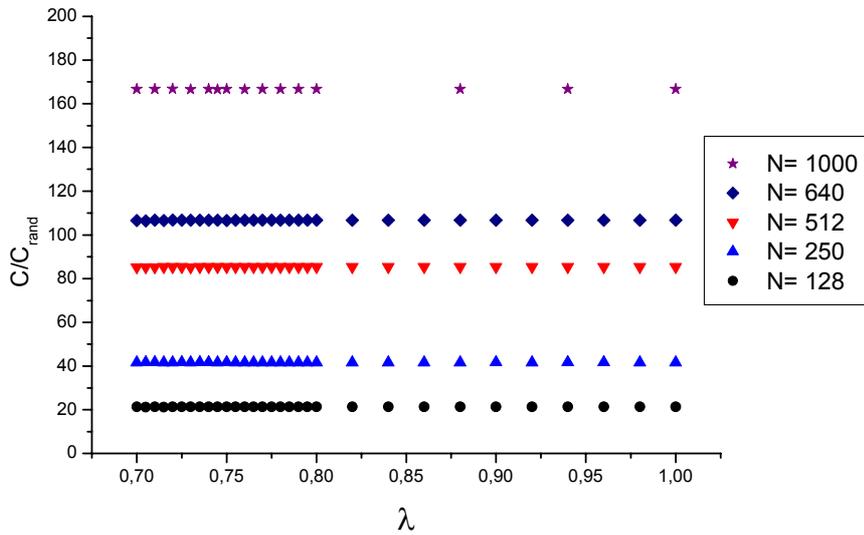


Figura 4.8: Gráfico de C/C_{rand} versus λ para vários tamanhos de redes ($N = 128, 250, 512, 640$ e 1000).

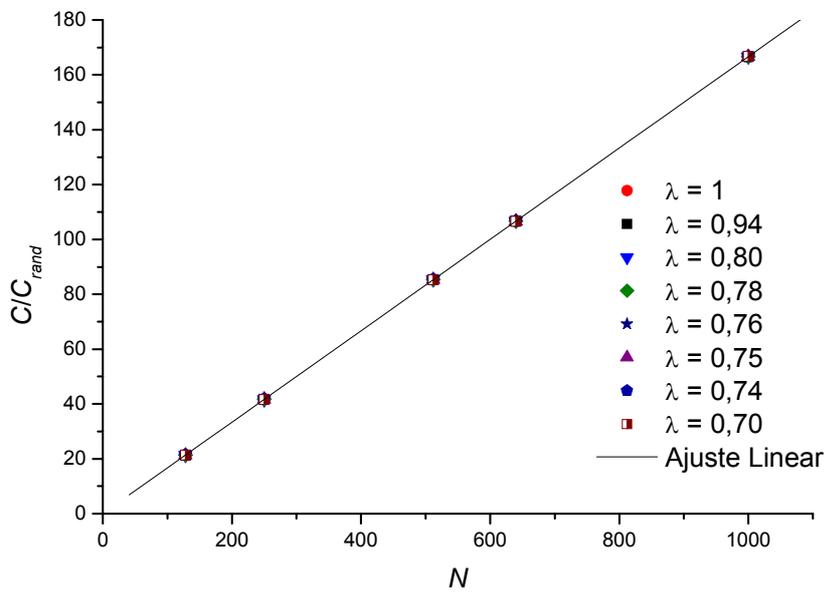


Figura 4.9: Gráfico de C/C_{rand} versus N para vários valores λ estudados. A reta representa o melhor ajuste linear dado pela equação 4.1.

Ainda, vemos nas figuras 4.10 e 4.11 dependências distintas entre o coeficiente de agrupamento (C) e o número de moléculas (N) para $\lambda > 0,75$ (linear, quase constante) e para $\lambda < 0,75$ (oscilatório).

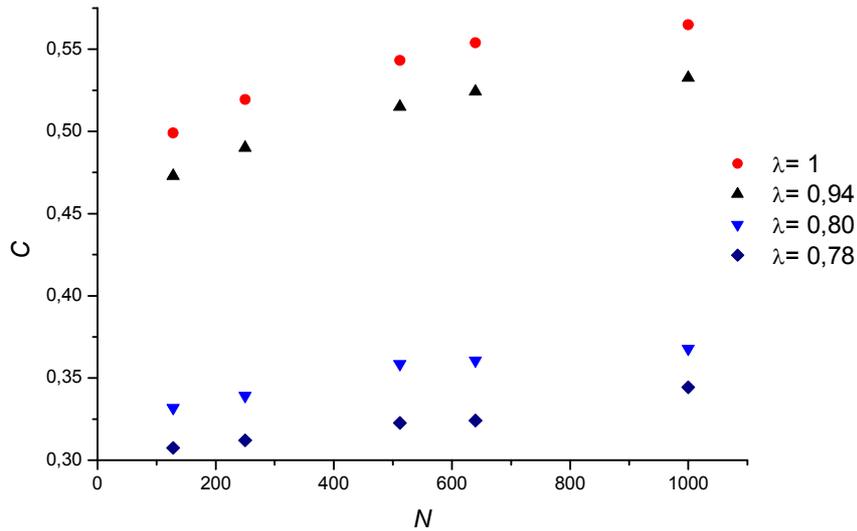


Figura 4.10: Gráfico do coeficiente de agrupamento (C) versus N para alguns dos valores de $\lambda > 0,75$.

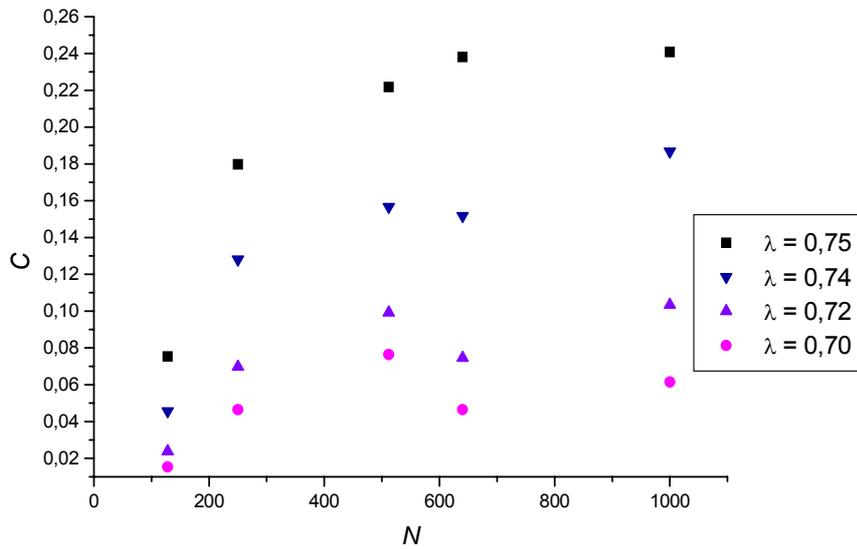


Figura 4.11: Gráfico do coeficiente de agrupamento (C) versus N para alguns dos valores de $\lambda \leq 0,75$.

Estes regimes e comportamentos são similares aos de redes com probabilidades de conexão dependentes da distância, estudadas por Sen *et. al.* [45]. Neste trabalho, uma rede é gerada onde ligações de longo alcance estão presentes com uma probabilidade dependente da distância euclidiana separando os nós,

$$P(l) = l^{-\delta}$$

onde δ é um parâmetro ajustável, mostrando transições de fases contínua ocorrendo em dois valores finitos do parâmetro δ , onde o coeficiente de agrupamento se comporta como um parâmetro de ordem. Os diferentes comportamentos de rede aleatória, pequenos mundos e regular podem ser obtidos nesta rede ajustando-se o valor do parâmetro δ . Além disso, ambas as transições de fase podem ser detectadas do comportamento dos comprimentos de ligação médios, que mostram diferentes dependências com o tamanho do sistema (N) nas três regiões. No nosso caso, a rede formada também é dependente da distância euclidiana, que depende da interação entre os sítios, que é controlada pelo valor do parâmetro λ .

O grau de distribuição $P(k)$ para as redes de ligações de hidrogênio em dois valores de λ estão apresentados na figura 4.12. Concluímos que para valores de λ no intervalo $0,75 < \lambda \leq 1$ a distribuição é muito similar a uma distribuição de Poisson, mostrando que a rede se comporta como uma rede aleatória. Por outro lado, para valores de $\lambda < 0,75$, a distribuição passa a ser assimétrica e bastante diferente da distribuição de Poisson. Neste caso, o comportamento é similar ao de distribuições dadas por uma lei de potência com função de corte [21]. Esta forma da distribuição é esperada quando ocorrem restrições limitando o número de conexões, sendo as mesmas dificilmente formadas. Isto é consistente com os resultados para estas redes de ligações de hidrogênio, tendo em vista que a formação da interação para valores de λ menores que um determinado valor é muito custosa devido à fraca interação entre os nós da rede, refletindo na forma do grau de distribuição. Vale ressaltar que o grau de distribuição é independente do número de nós da rede.

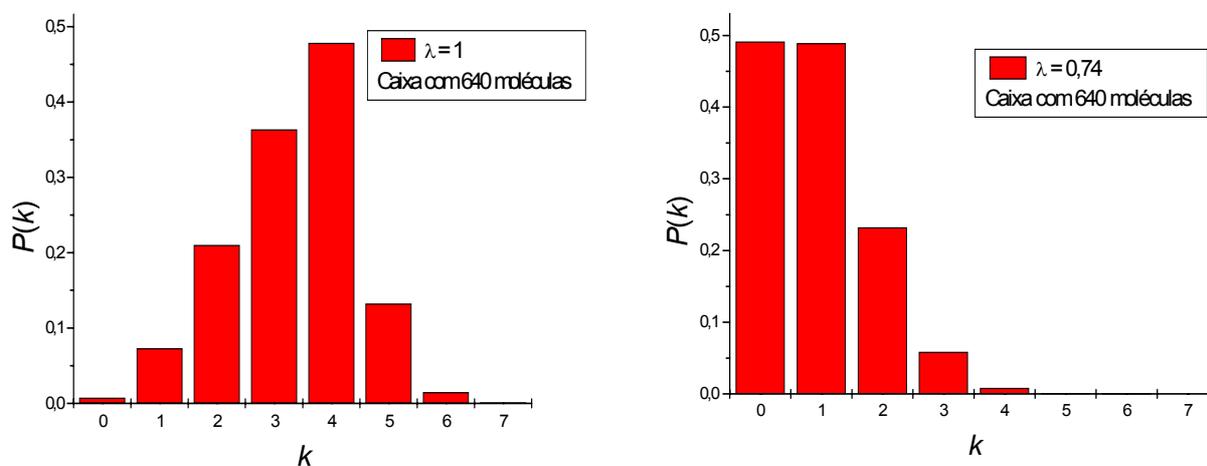


Figura 4.12: Grau de distribuição $P(k)$ para os valores de $\lambda = 1$ e $\lambda = 0,74$.

Uma análise de agregados utilizando estatística de ilhas mostra que o cálculo do número de ilhas, n_s , com s componentes, resulta em um comportamento claramente distinto para as regiões analisadas. Para uma melhor visualização gráfica, considera-se que para cada “bin” têm-se ilhas com número de componentes no intervalo $[2^{\text{bin}-1}, 2^{\text{bin}}-1]$. Para valores de λ acima de 0,75 tem-se a formação de um grande agregado (contendo em média, 99% dos componentes) com a existência de uma pequena quantidade de monômeros ou dímeros e o gráfico tem a forma de blocos opostos com pequenos n_s (figura 4.13 (I)). Com a diminuição de λ observa-se que há um maior número de monômeros, dímeros e trímeros (Figura 4.13 (II)).

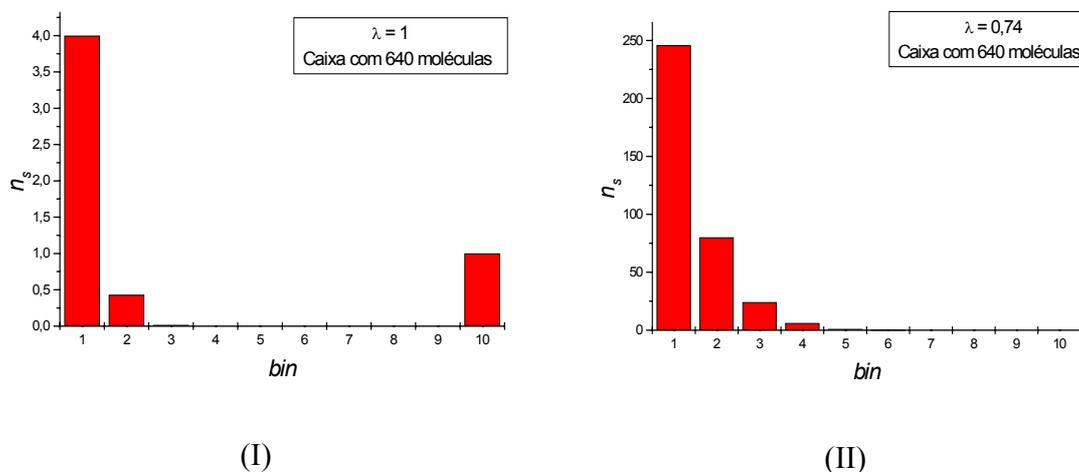


Figura 4.13: Número médio de ilhas com tamanho no intervalo $[2^{\text{bin}-1}, 2^{\text{bin}}-1]$ para a caixa de simulação com 640 moléculas e valor de $\lambda = 1$ (I) e $\lambda = 0,74$ (II).

Os resultados para as propriedades termodinâmicas também são indicativos da transição de fase em $\lambda = \lambda_c = 0,75$, como podemos observar nas figuras 4.14, 4.15 e 4.16, respectivamente, para a densidade (ρ), entalpia por molécula (H) e entalpia de vaporização ($\Delta_{\text{vap}}H$). Observe que o comportamento nas propriedades termodinâmicas em $\lambda = \lambda_c$ decorre da transição de fase topológica que resulta da análise das redes de ligação de hidrogênio.

Cálculos da capacidade calorífica, coeficiente de expansão térmica e compressibilidade isotérmica foram realizados a partir de médias sobre flutuações, considerando 10^7 configurações e não mostraram resultados satisfatórios. Cálculos deste tipo requerem cerca de 10^8 – 10^9 configurações e, portanto, uma demanda computacional muito grande ou então cálculos em diferentes temperaturas, como foi discutido na seção 2.3.

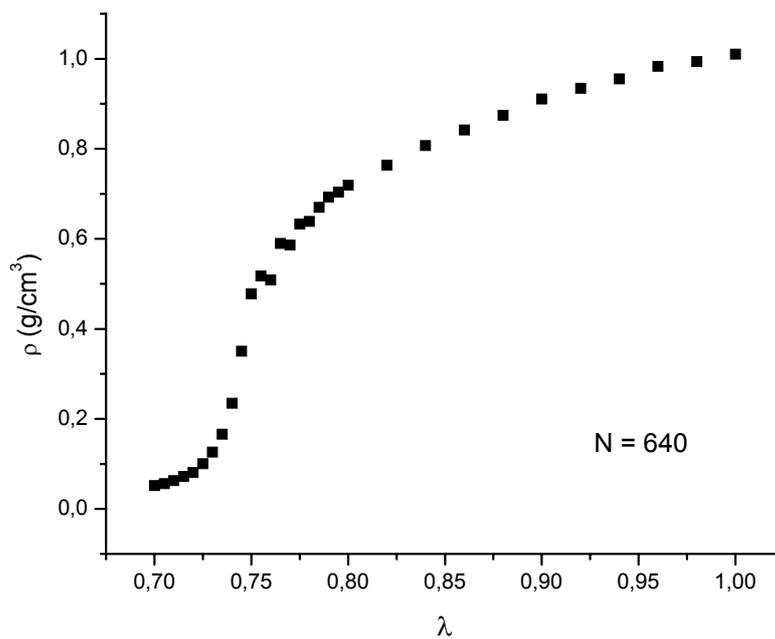


Figura 4.14: Densidade em g/cm^3 versus o parâmetro λ para a caixa de simulação com 640 moléculas.

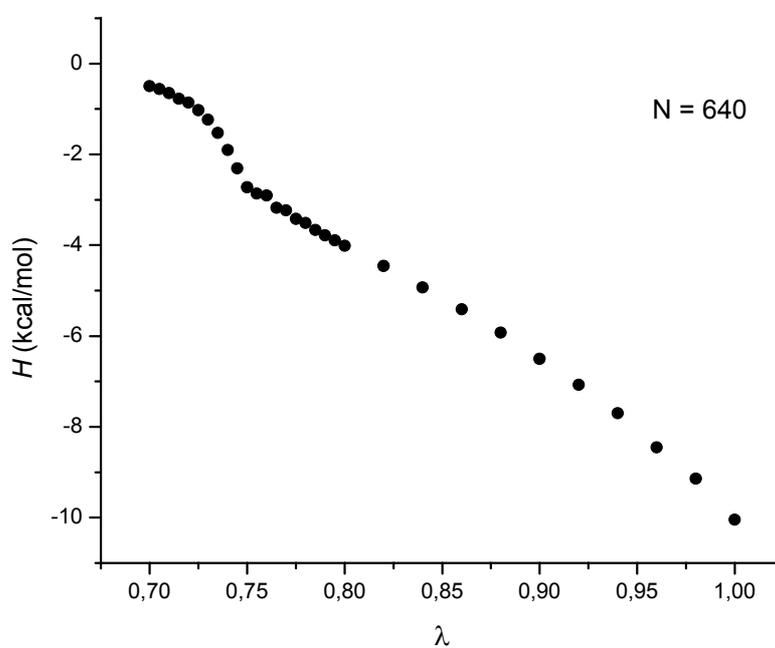


Figura 4.15: Entalpia por molécula, em kcal/mol , versus o parâmetro λ para a caixa de simulação com 640 moléculas.

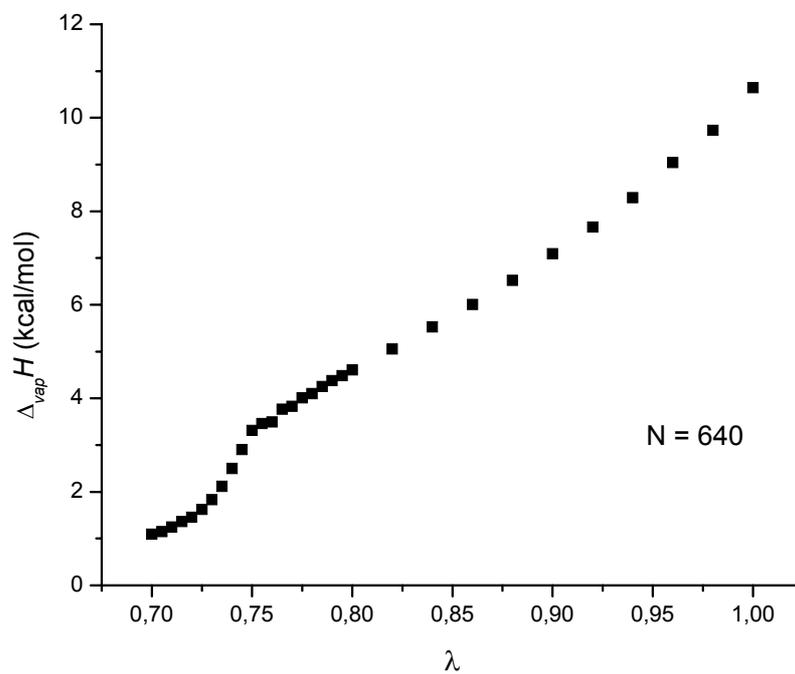


Figura 4.16: Entalpia de vaporização, em kcal/mol, em função do parâmetro λ para a caixa de simulação com 640 moléculas.

Capítulo 5

Conclusões e Perspectivas

Nossos resultados para o coeficiente de agrupamento (C), a conectividade média $\langle k \rangle$ e a distância química (L), em função do fator de amortecimento λ ($0,7 \leq \lambda \leq 1$), indicam que o sistema exibe uma transição de fase topológica em $\lambda_c = 0,75$, separando as regiões de pequenos e grandes valores de C e $\langle k \rangle$, para $\lambda < \lambda_c$ e $\lambda > \lambda_c$, respectivamente. Nossa análise de estatística de ilhas mostra que na região $\lambda > \lambda_c$ tem-se a formação de um agregado com cerca de 99% dos componentes, enquanto que para $\lambda < \lambda_c$ ocorre um grande número de monômeros, dímeros e trímeros e não há mais a percolação do sistema. O comportamento do coeficiente de agrupamento e da distância química em função do parâmetro λ , exibe semelhança com o resultado obtido por Watts e Strogatz [17], no qual caracterizam uma região de pequenos mundos. A razão C/C_{rand} , onde C_{rand} é o coeficiente de agrupamento de redes aleatórias com o mesmo número de nós (N), é independente do valor de λ , e varia linearmente com N com coeficiente angular dado pela razão $C / \langle k \rangle = 1/6$. O grau de distribuição de ligações para estas redes mostra que, para a região $\lambda > \lambda_c$, a distribuição é muito similar a uma distribuição de Poisson, mostrando que a rede se comporta como uma rede aleatória, mas para valores de $\lambda < \lambda_c$, a distribuição passa a ser assimétrica e bastante diferente da distribuição de Poisson. Cálculos das propriedades termodinâmicas, como densidade, entalpia, e entalpia de vaporização, também são indicativos da ocorrência de uma transição de fase topológica em $\lambda_c = 0,75$.

Como primeira perspectiva deste trabalho, pretende-se estendê-lo para o sistema formado pela mistura água-metanol, pois, como observado no trabalho de Santos *et al.* [44], o metanol puro em condições ambiente apresenta a formação de padrões de pequenos mundos, diferentemente da água pura nestas mesmas condições. Assim, uma investigação da formação destes padrões pela mudança da concentração do metanol na água líquida, trará novas informações com relação ao aparecimento de padrões de pequenos mundos em redes de ligações de hidrogênio.

Também como perspectiva deste trabalho, pretende-se fazer um estudo detalhado de percolação considerando a estatística de ilhas para as caixas geradas pelas simulações Monte Carlo.

Também, pretende-se modificar o programa ANALYSIS para que fique mais rápido, por exemplo, transformando as variáveis simples, que ficam armazenadas em arquivos, em vetores ou arrays.

Referências Bibliográficas

- [1] V.M.L. Santos, F.G.B. Moreira and R.L. Longo, *Chem. Phys. Lett.* **390**, 157 (2004).
- [2] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai and A.-L. Barabási, *Nature* (London), **407**, 651 (2000).
- [3] A. Wagner and D. Fell, *Technical Report* N° 00-07-041, Santa Fe Institute (2000).
- [4] S. Yook, H. Jeong and A.-L. Barabási, *Proceedings Of The National Academy Of Sciences Of The United States Of America* **99** (21): 13382-13386 OCT 15 (2002).
- [5] R. Pastor-Satorras, A. Vazquez and A. Vespignani, *Physical Review Letters* **87** (25), 258701 (2001).
- [6] L. A. Adamic, Proceedings of the Third European Conference, ECDL'99 (Springer-Verlag, Berlin), p. 443 (1999).
- [7] A.-L. Barabási, H. Jeong, E. Ravasz, Z. Néda, A. Schubert and T. Vicsek, *Physica A* **311** 590 (2002).
- [8] F. Liljeros, C.R. Edling, L.A.N. Amaral, H.E. Stanley and Y. Aberg, 2001, *Nature* (London) **411**, 907.
- [9] P. Erdős and A. Rényi, *Publ. Math. (Debrecen)*, **6**, 290 (1959).
- [10] P. Erdős and A. Rényi, *Publ. Math. Inst. Hung. Acad. Sci.*, **5**, 17 (1960).
- [11] S. Milgram, The Small-World Problem, *Psychology Today*, **2** (1967).
- [12] M. Kochen, *The Small World* (Ablex, Norwood, NJ) (1989).
- [13] R. Albert, H. Jeong and A.-L. Barabási, *Nature* (London) **401**, 130 (1999).
- [14] M. Faloutsos, P. Faloutsos and C. Faloutsos, *Comput. Commun. Rev.* **29**, 251 (1999).
- [15] A.-L. Barabási and R. Albert, *Science*, **286**, 509 (1999).
- [16] R. Albert and A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).
- [17] D.J. Watts and S.H. Strogatz, *Nature* (London) **393**, 440 (1998).
- [18] A.G. Kalinichev and J.D. Bass, *J. Phys. Chem. A* **101**, 9720 (1997).
- [19] H. Ohtaki, T. Radnai, T. Yamaguchi, *Chem. Soc. Rev.* 41-51 (1997).
- [20] R.L. Blumberg, H.E. Stanley, A. Geiger and P. Mausbach, *J. Chem. Phys.* **80**, 5230 (1984).

- [21] L.A.N. Amaral, A. Scala, M. Barthélémy and H.E. Stanley, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11149 (2000).
- [22] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [23] B.J. Alder and T.E. Wainwright, *J. Chem. Phys.* **27**, 1208 (1957).
- [24] B.J. Alder and T.E. Wainwright, *J. Chem. Phys.* **31**, 459 (1959).
- [25] V.E. Barlette and L.C.G. Freitas, *Química Nova*, **22(2)**, 254 (1999).
- [26] D. Frenkel and B. Smit, *Understanding Molecular Simulation*, Academic Press (1996).
- [27] A.R. Leach, *Molecular Modelling. Principles and Applications*, Longman. England. (1996).
- [28] M.W. Mahoney and W.L. Jorgensen, *J. Chem. Phys.* **112**, 8910 (2000).
- [29] W.L. Jorgensen, Encyclopedia of Computational Chemistry, edited by P.v.R. Schleyer (Wiley, New York, 1998).
- [30] J.D. Bernal and R.H. Fowler, *The Journal of Chemical Physics* **1**, 515 (1933).
- [31] D.L. Beveridge, M. Mezei, P.K. Mehrotra, F.T. Marchese, G. Ravi-Shankar, T. Vasu and S. Swaminathan, in *Molecular-Based Study of Fluids*, ACS Advances in Chemistry Series 204, edited by J.M. Haile and G.A. Mansoori (American Chemical Society, Washington, DC, 1983) p. 297.
- [32] F.H. Stillinger and A. Rahman, *J. Chem. Phys.* **60**, 1545 (1974).
- [33] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey and M.L. Klein, *J. Chem. Phys.* **79**, 926 (1983).
- [34] W.L. Jorgensen and J.D. Madura, *Mol. Phys.* **56**, 1381 (1985).
- [35] H.J.C. Berendsen, J.P.M. Postman, W.F. van Gunsteren and J. Hermans, *Intermolecular Forces* (1981).
- [36] H.J.C. Berendsen, J.R. Grigera and T.P. Straatsma, *J. Phys. Chem.* **91**, 6269 (1987).
- [37] A.A. Chialvo and P.T. Cummings, *J. Phys. Chem.* **100**, 1309 (1996).
- [38] L.C.G. Freitas, DIADORIM Program, UFScar São Carlos, SP, Brazil (1999).
- [39] F. Chung and L. Lu, *Adv. Appl. Math.* **26**, 257 (2001).
- [40] B. Bollobás, Degree Sequences of Random Graphs. *Discrete Mathematics*, **33**, 1 (1981).
- [41] B. Bollobás, *Random Graphs*, Academic Press, London (1985).

- [42] L.F.C. Pereira, *Diagrama de Fases e Expoentes Críticos do Modelo do Voto da Maioria em Grafos Aleatórios*. Dissertação de Mestrado, Departamento de Física – UFPE (2005).
- [43] B. Bollobás and O. Riordan, The diameter of a scale-free random graph. *Combinatorica*, **24** (2004).
- [44] V. M.L. Santos, *Estruturas e Propriedades de Fluidos Supercríticos e Líquidos Iônicos*. Tese de Doutorado, Departamento de Química Fundamental – UFPE (2004).
- [45] P. Sen, K. Banerjee and T. Biswas, *Phys. Rev. E* **66**, 037102 (2002).

