

IFT - UNESP  
INSTITUTO DE FÍSICA TEÓRICA

---

TESE DE DOUTORAMENTO

IFT-T.007/2025

# Métodos de *Machine Learning* Aplicados a Estruturas de Água e Gelo

Lucas Thiago Siqueira de Miranda

Orientador

*Alexandre Reily Rocha*

Novembro de 2025

M672m      Miranda, Lucas Thiago Siqueira de  
             Métodos de machine learning aplicados a estruturas de água e gelo /  
             Lucas Thiago Siqueira de Miranda. – São Paulo, 2025  
             183 f.: il. color.

             Tese (doutorado) – Universidade Estadual Paulista (Unesp), Instituto de  
Física Teórica (IFT), São Paulo  
             Orientador: Alexandre Reily Rocha

             1. Dinâmica molecular. 2. Aprendizado do computador. 3. Gelo. I. Título

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca  
do Instituto de Física Teórica (IFT), São Paulo. Dados fornecidos pelo  
autor(a).

# MÉTODOS DE MACHINE LEARNING APLICADOS EM ESTRUTURAS DE GELO E ÁGUA

Tese de Doutorado apresentada ao Instituto de Física Teórica do Câmpus de São Paulo, da Universidade Estadual Paulista “Júlio de Mesquita Filho”, como parte dos requisitos para obtenção do título Doutor em Física, Especialidade Física Teórica.

## **Comissão Examinadora:**

Prof. Dr. ALEXANRE REILY ROCHA (Orientador)  
Instituto de Física Teórica/UNESP

Prof. Dr. MAURICE DE KONING  
Instituto de Física Gleb Wataghin/UNICAMP

Profa. Dra. ELIZIANE EFIGÊNIA DE MORAES  
Instituto de Física/UFBA

Prof. Dr. DANILO BARBOSA LIARTE  
Instituto de Física Teórica/UNESP

Prof. Dr. GABRIEL RAVANHANI SCHLEDER  
Laboratório Nacional de Nanotecnologia/CNPEM

**Conceito:** Aprovado

São Paulo, 11 de dezembro de 2025.

*Dedico esta tese a todo povo brasileiro.*

# Agradecimentos

Primeiramente à Deus.

Em segundo lugar a todos os brasileiros, que por meio de sua contribuição, tornaram possível que eu obtivesse um novo título acadêmico. Que todo conhecimento produzido seja de impacto a sociedade, uma vez que ela foi determinante para a produção deste trabalho.

A minha família, em especial a minha esposa, Samira, que durante os últimos anos me deu forças para concluir este desafio. Aos meus pais e meu irmão, que sempre se dedicaram para que eu pudesse me tornar o ser humano que sou hoje.

Ao grupo de pesquisa TranSim, liderado pelo professor Alexandre Reily Rocha, a quem devo gratidão por todo suporte teórico e paciência em toda a jornada. A todos os integrantes do nosso grupo, aqueles que estão ou já estiveram, por todas as discussões e aprendizados. Além da rede de colaboradores que nos auxiliou durante estes anos. Em especial aos amigos Graciele e Washington, onde as discussões não se restringiram apenas a ciência. Agradeço também aos professores Mariana Rossi e Marcio Gomes pelo suporte computacional, além do auxílio para o melhor entendimento das propriedades do gelo.

A todos os amigos desde o início em Mogi das Cruzes, passando pelo amadurecimento em Volta Redonda, até chegar a esta etapa novamente em São Paulo.

Um agradecimento a Universidade Estadual Paulista, que entrega a sociedade um doutor em física. Também agradeço as agências de fomento CAPES pela concessão da bolsa de doutorado, a FAPESP e ao INCT que possibilitaram a participação em alguns eventos para aprimoramento profissional e ao CENAPAD-SP, GridUNESP e LNCC pelo suporte computacional para esta pesquisa.

\*\*\*\*\* *finally!*

Kimi Raikkonen, *The Iceman*

# Resumo

Neste trabalho, utilizamos potenciais obtidos por meio de *machine learning* para investigar a influência da inclusão de efeitos quânticos nucleares nas interações interatômicas, bem como na formação e difusão de defeitos na estrutura cristalina do gelo  $I_h$ . Os potenciais foram gerados a partir de energias e forças obtidas em cálculos baseados na teoria do funcional da densidade, empregando diferentes funcionais de correlação e troca: PBE, vdW-cx, SCAN e optB88. Inicialmente, apresentamos a estratégia para a construção de um modelo confiável capaz de reproduzir com precisão as propriedades de sistemas aquosos. Verificamos que um modelo treinado a partir de um sistema contendo 64 moléculas de água é suficiente para reproduzir suas propriedades estruturais na fase líquida. Além disso, para redes neurais de grafos, observamos que a quantidade de configurações utilizadas no treinamento pode ser reduzida para apenas 6% do *dataset* original, sem prejuízo das propriedades estruturais e dinâmicas obtidas com o conjunto completo de dados. Cálculos de dinâmica molecular realizados com potenciais obtidos por *machine learning* (DP-MD) mostram que a densidade do gelo  $I_h$  está em bom acordo com os resultados obtidos por dinâmica molecular *ab initio*. Ademais, a inclusão de efeitos quânticos nucleares leva a um leve aumento no valor médio do comprimento da ligação covalente O–H, promovendo um maior compartilhamento de prótons. Simultaneamente, esse efeito compete com a redução das distâncias intermoleculares decorrente do fortalecimento das ligações de hidrogênio. Como consequência, observa-se um aumento geral da densidade de equilíbrio para todos os funcionais, afastando os resultados dos valores experimentais. No estudo dos defeitos, verificamos que a inclusão de termos não locais reduz a energia de formação dos defeitos intersticiais, alterando o tipo de defeito predominante em todo o intervalo de temperatura analisado. Determinamos também a energia necessária para que um defeito migre de um sítio para outro. Observamos que defeitos intersticiais apresentam barreiras de migração menores que as dos defeitos de vacância. Em contrapartida, a frequência tentativa de migração dos defeitos de vacância é superior à dos intersticiais. Como resultado, os coeficientes de difusão dos diferentes defeitos assumem valores semelhantes. Além disso, constatamos que os defeitos intersticiais apresentam maior tendência a se difundir na direção perpendicular ao eixo  $c$  do que na dire-

ção paralela. Por fim, mostramos que o aumento da temperatura afeta de forma distinta a energia de ligação: intensifica a interação entre a molécula intersticial e o cristal, mas reduz a intensidade das ligações das moléculas no interior da rede cristalina.

**Palavras Chaves:** gelo  $I_h$ ; Efeitos quânticos nucleares; Potenciais de *machine learning*; Dinâmica molecular; Defeitos pontuais em gelo.

**Áreas do conhecimento:** Física; Física da Matéria Condensada;

# Abstract

In this work, we employ machine learning potentials to investigate the influence of nuclear quantum effects on interatomic interactions, as well as on the formation and diffusion of defects in the crystalline structure of ice  $I_h$ . The potentials were generated from energies and forces obtained through density functional theory calculations using different exchange–correlation functionals: PBE, vdW-cx, SCAN, and optB88. Initially, we present the strategy for constructing a reliable model capable of accurately reproducing the properties of aqueous systems. We found that a model trained on a system containing 64 water molecules is sufficient to reproduce the structural properties of liquid water. Furthermore, for graph neural networks, we observed that the number of configurations used for training can be reduced to only 6% of the original dataset without compromising the structural and dynamical properties obtained with the full dataset. Molecular dynamics simulations performed with machine learning potentials (DP-MD) show that the density of ice  $I_h$  is in good agreement with results from *ab initio* molecular dynamics. Moreover, the inclusion of nuclear quantum effects leads to a slight increase in the average O–H covalent bond length, enhancing proton sharing. At the same time, this effect competes with the reduction of intermolecular distances resulting from the strengthening of hydrogen bonds. Consequently, there is an overall increase in the equilibrium density for all functionals, further deviating the results from experimental values. In the study of defects, we found that the inclusion of nonlocal terms lowers the formation energy of interstitial defects, changing the predominant defect type over the entire temperature range analyzed. We also determined the energy required for a defect to migrate from one site to another. Our results show that interstitial defects exhibit lower migration barriers compared to vacancy defects. In contrast, the attempt frequency for vacancy defect migration is higher than that for interstitial defects. As a result, the diffusion coefficients of the different defects display similar values. Additionally, we observed that interstitial defects have a greater tendency to diffuse in the direction perpendicular to the  $c$  axis than parallel to it. Finally, we show that increasing temperature affects the binding energy in different ways: it strengthens the interaction between the interstitial molecule and the crystal, while simultaneously weakening the bonding between molecules in the interior of the crystalline lattice.

**Keywords:** ice  $I_h$ ; Nuclear quantum effects; Machine learning interatomic potential; Molecular Dynamics; Point defects in ice.

**Knowledge Areas:** Physics; Condensed matter physics;

# Índice

<b>Lista de Figuras</b>	<b>xi</b>
<b>Lista de Tabelas</b>	<b>xvi</b>
<b>Introdução</b>	<b>1</b>
Gelo $I_h$ . . . . .	1
Simulação Computacional . . . . .	3
Métodos de <i>Machine Learning</i> . . . . .	4
Efeitos Quânticos Nucleares . . . . .	8
Defeitos em Gelo $I_h$ . . . . .	9
A Tese . . . . .	12
<b>1 Metodologia</b>	<b>15</b>
1.1 Teoria do Funcional da Densidade (DFT) . . . . .	16
1.1.1 Equação de Schrödinger para Muitos Corpos . . . . .	16
1.1.2 Aproximação de Born-Oppenheimer . . . . .	19
1.1.3 Aproximação de elétrons independentes . . . . .	21
1.1.4 Equações de Hartree-Fock . . . . .	24
1.1.5 Equações de Kohn-Sham . . . . .	26
1.1.6 Os Potenciais de Troca e Correlação . . . . .	30
1.1.7 O método de pseudo-potenciais . . . . .	37
1.2 Dinâmica Molecular (MD) . . . . .	39
1.2.1 Campos de força . . . . .	41
1.2.2 Controle de Temperatura . . . . .	44
1.3 Efeitos Quânticos Nucleares (NQE) . . . . .	49
1.3.1 Integrais de Caminho . . . . .	50
1.3.2 Dinâmica Molecular com Integrais de Caminho (PIMD) . . . . .	58
1.3.3 Equações de Langevin . . . . .	62
1.4 Potenciais de <i>Machine Learning</i> . . . . .	64
1.4.1 DeePMD-kit . . . . .	66
1.4.2 MACE . . . . .	70
1.5 Integração Termodinâmica . . . . .	73

1.5.1	Energia de Ligação . . . . .	75
<b>2</b>	<b>Testes de Potenciais</b>	<b>78</b>
2.1	Número de Passos de Treino . . . . .	79
2.2	Raio de Corte . . . . .	82
2.3	Número de Moléculas de Água . . . . .	85
2.4	Inclusão de Configurações de Gelo . . . . .	88
2.5	Número de Configurações . . . . .	89
2.6	Considerações Finais . . . . .	93
<b>3</b>	<b>Efeitos Quânticos Nucleares em Gelo Hexagonal</b>	<b>95</b>
3.1	Métodos Computacionais . . . . .	97
3.1.1	Treinamento das Redes Neurais . . . . .	98
3.1.2	Dinâmica Molecular: Propriedades Estruturais . . . . .	102
3.1.3	Dinâmica Molecular: Propriedades Vibracionais . . . . .	104
3.1.4	Termostato das Simulações DP-PIMD . . . . .	104
3.2	Propriedades Estruturais . . . . .	107
3.3	Propriedades Dinâmicas . . . . .	113
3.4	Considerações Finais . . . . .	114
<b>4</b>	<b>Defeitos em Gelo</b>	<b>116</b>
4.1	Métodos Computacionais . . . . .	118
4.1.1	Treinamento das Redes Neurais . . . . .	119
4.1.2	Análise Energética . . . . .	121
4.1.3	Efeitos de Temperatura na Dinâmica dos Defeitos . . . . .	124
4.2	Análise Energética . . . . .	125
4.2.1	Análise Energética via DFT . . . . .	125
4.2.2	Análise Energética via Redes Neurais . . . . .	128
4.3	Efeitos de Temperatura na Dinâmica dos Defeitos . . . . .	139
4.4	Considerações Finais . . . . .	143
<b>5</b>	<b>Considerações Finais</b>	<b>146</b>
<b>A</b>	<b>Unidades Atômicas</b>	<b>150</b>
<b>B</b>	<b>Dinâmica molecular <i>ab initio</i></b>	<b>152</b>
<b>C</b>	<b>Hipótese ergódica</b>	<b>154</b>

<b>D</b> <i>Script</i> utilizado para treino das redes neurais	<b>157</b>
<b>E</b> <i>Script</i> para cálculos DFT	<b>162</b>
<b>Referências</b>	<b>166</b>

# Lista de Figuras

1	Diagrama de fases da água. . . . .	2
2	Representação das regras de Bernal e Fowler. . . . .	3
3	Representação dos defeitos (a) de vacância e (b) intersticial Tu. . . . .	10
4	Representação do defeito Bc. . . . .	11
5	Visualização do defeito Tu no plano (a) $ab$ e (b) $ac$ . . . . .	12
1.1	Esquema para um átomo hipotético. . . . .	37
1.2	Esquema da substituição do potencial "all-electron" e da função de onda por um pseudo-potencial e uma pseudo-função de onda. . . . .	38
1.3	Fluxograma de uma simulação de dinâmica molecular utilizando o algoritmo de Verlet. . . . .	41
1.4	Esquema da molécula de água representada por diferentes modelos de potencial. . . . .	43
1.5	Representação dos efeitos quânticos nucleares nas ligações de hidrogênio. Imagem retirada da Ref. [1]. . . . .	50
1.6	Representação do isomorfismo entre uma partícula quântica e um polímero cíclico clássico com constante de mola $k = m\omega_p^2$ . Adaptado de Tuckerman [2]. . . . .	59
1.7	Comparação da convergência de quantidades físicas para PIMD e PIGLET em função do número de contas ( $P$ ). Retirado da Referência [3]. . . . .	64
1.8	RSME da energia e treinamento de uma rede neural para água líquida utilizando pesos iniciais para a energia e força sendo, respectivamente, 0,02 e 1000 e os pesos finais como 8 e 1. Retirado da Referência [4]. . . . .	70
1.9	MAE da força em função do número de configurações inclusas no treino para a molécula aspirina. Esquerda: mantém a equivariância como $L = 0$ e varia a correlação. Centro: Mantém a correlação com $\nu = 3$ e varia a equivariância. Direita: Mantém a correlação com $\nu = 1$ e varia a equivariância. Para todos os casos, $s$ representa a inclinação da reta. Retirado da Referência [5]. . . . .	73

1.10	Esquema para o ciclo termodinâmico utilizado no cálculo da energia livre de ligação entre dois sistemas. . . . .	76
1.11	Esquema para o cálculo da energia de ligação em simulações quânticas, utilizado por Rossi <i>et al.</i> [6] O contorno azul representa as simulações clássicas e o contorno vermelho, as simulações quânticas. . . . .	77
2.1	Comparação das energias (painel esquerdo) e forças (painel direito) obtidas com a rede neural treinada com $1 \times 10^6$ passos e com DFT para os funcionais (a) optB88 e (b) vdW-cx. . . . .	81
2.2	Energia total em função do tempo para modelos treinados com (a) $1 \times 10^6$ e (b) $5 \times 10^6$ passos. Dados obtidos para modelos com o funcional optB88. . . . .	82
2.3	RMSE para cada modelo treinado com diferentes raios de corte. . . . .	84
2.4	(a) Densidade de equilíbrio e distribuições (b) do comprimento da ligação covalente e (c) do ângulo H-O-H para a rede treinada com raio de corte de 6.0 Å. . . . .	86
2.5	Evolução temporal da densidade para os sistemas com (a) 64 e (b) 256 moléculas na caixa de simulação. As curvas em verde representam o modelo PBE-64 e, em azul, o modelo PBE-256. . . . .	87
2.6	RMSE para as energias (painel superior) e forças (painel inferior) dos modelos SCAN-W, SCAN-WI e SCAN-WI2 para os dados de treino de . . . . .	89
2.7	Densidade de equilíbrio para os modelos treinados com (a) 540, (b) 1080 e (c) 1620 configurações. . . . .	91
2.8	Evolução temporal do coeficiente de difusão para cada modelo, com diferentes números de configurações no <i>dataset</i> . . . . .	92
2.9	Evolução temporal do coeficiente de difusão e do deslocamento quadrático médio para todas as simulações dos modelos treinados com (a) 540, (b) 1080 e (c) 1620 configurações. As linhas tracejadas coloridas representam as cinco diferentes simulações e a linha sólida em preto representa a média sobre as simulações. . . . .	93
3.1	Vista (a) superior e (b) lateral da estrutura do gelo $I_h$ . As esferas vermelhas representam os átomos de oxigênio e, as cinzas, os átomos de hidrogênio. . . . .	95
3.2	Estrutura do fluxo de trabalho ( <i>workflow</i> ) para a obtenção das propriedades do gelo $I_h$ . . . . .	98

3.3	Comparação das energias e forças entre os cálculos DFT e a rede neural para os funcionais: (a) PBE, (b) vdW-cx, (c) SCAN e (d) optB88. Os <i>insets</i> representam a distribuição das diferenças entre os valores obtidos via DFT e os modelos. . . . .	101
3.4	Comparação das energias e forças entre os cálculos DFT e a rede neural utilizando o método de validação <i>a posteriori</i> para os funcionais: (a) PBE, (b) vdW-cx, (c) SCAN e (d) optB88. Os <i>insets</i> representam a distribuição das diferenças entre os valores obtidos via DFT e os modelos. . . . .	103
3.5	Espectro vibracional para (a) teste de convergência do número de <i>beads</i> e (b) comparação entre uma simulação longa e a média sobre 10 simulações curtas. . . . .	105
3.6	Comparação das propriedades estruturais para cálculos com 48 e 64 <i>beads</i> . Densidade de equilíbrio para (a) 48 e (b) 64 <i>beads</i> e (c-e) função de distribuição radial. . . . .	106
3.7	Comparação das propriedades estruturais para cálculos com os termostatos PILE-L e PIGLET. Densidade de equilíbrio para (a) PIGLET e (b) PILE-L e (c-e) função de distribuição radial. . . . .	106
3.8	Evolução temporal da densidade para os modelos: (a) PBE, (b) vdW-cx, (c) SCAN e (d) optB88. As simulações DP-MD estão à esquerda e as simulações DP-PIMD, à direita. . . . .	108
3.9	Função de distribuição radial para cada funcional. Os funcionais estão descritos no painel esquerdo. O painel esquerdo apresenta a $g_{OO}$ , o central a $g_{OH}$ e o direito a $g_{HH}$ . . . . .	109
3.10	Comparação de cada RDF para todos os funcionais (a) sem e (b) com a inclusão dos NQEs. . . . .	110
3.11	Distribuição da (a) coordenada de transferência de próton $\nu$ e dos comprimentos de ligação (b) intramolecular O-H ( $d_{OH}$ ), (c) intermolecular O-H ( $d_{O...H'}$ ) e (d) intermolecular O-O ( $d_{O...O'}$ ) obtidos via DP-MD e DP-PIMD. A linha tracejada (tracejada e pontilhada) refere-se ao valor mais provável para as simulações DP-PIMD (DP-MD). . . . .	112
3.12	Relação entre a projeção das ligações O-H na direção O-O e a distância O-O. Os valores para <i>clusters</i> de água foram obtidos teoricamente por Li <i>et al.</i> [7] . . . . .	113

3.13	Espectros vibracionais para o gelo $I_h$ a 220 K obtidos com o método TRPMD [8, 9] (linhas contínuas). Para comparação, também mostramos os espectros obtidos via DP-MD no ensemble NVE (linhas tracejadas). . . . .	114
4.1	Geometria dos defeitos: (a) vacância, (b) Tu e (c) Bc. Os círculos pontilhados em azul indicam a posição dos defeitos na estrutura do gelo $I_h$ . As estruturas são apresentadas nos planos (i) $xy$ e (ii) $yz$ . . .	116
4.2	Análise geométrica dos saltos de defeitos. Os círculos representam as posições dos oxigênios da rede cristalina e os triângulos, os oxigênios intersticiais. Vistas sobre os planos (a) $xy$ e (b) $yz$ para o defeito intersticial e (c) $xy$ para o defeito de vacância. . . . .	124
4.3	Energia de formação para cada defeito para os funcionais PW91 e PBE com base <i>double-<math>\zeta</math></i> . . . . .	126
4.4	Energia de formação para cada defeito para o funcional PBE com diferentes abordagens, em comparação com o funcional PW91. . . .	126
4.5	Energia de formação comparando funcionais locais com funcionais que incluem termos não locais. . . . .	127
4.6	Energia de formação dos defeitos para os potenciais treinados com o DeePMD-kit no nível de precisão (a) PBE e (b) SCAN. . . . .	129
4.7	Energia de formação dos defeitos para o potencial treinado com o MACE no nível de precisão SCAN. . . . .	130
4.8	Energia de ligação dos defeitos para o potencial treinado com o MACE no nível de precisão SCAN. . . . .	132
4.9	Energia livre de formação dos defeitos para o potencial treinado com o MACE no nível de precisão SCAN. . . . .	133
4.10	Contribuição vibraional ( $T\Delta S$ ) na energia livre de formação dos defeitos para o potencial treinado com o MACE no nível de precisão SCAN. . . . .	134
4.11	Concentração dos defeitos em função da temperatura, calculada com o potencial treinado com MACE no nível de precisão SCAN. .	135
4.12	(a) Perfil de energia para as imagens ao longo do caminho de migração e (b) caminho de mobilidade para o defeito de vacância. . .	136
4.13	(a) Perfil de energia para as imagens ao longo do caminho de migração e (b) caminho de mobilidade para o defeito intersticial na direção perpendicular ao eixo $c$ . . . . .	137

4.14 (a) Perfil de energia para as imagens ao longo do caminho de migração e (b) caminho de mobilidade para o defeito intersticial na direção paralela ao eixo $c$ . . . . .	138
4.15 Barreira de migração para o defeito Tu (nas direções paralela e perpendicular ao eixo $c$ ) e para o defeito de vacância. As linhas com transparência representam a energia de barreira, enquanto as sem transparência representam a energia de barreira com a adição da contribuição vibracional, enquanto a energia de barreira é . . . . .	139
4.16 Frequência de tentativa de migração dos defeitos intersticiais e de vacância. . . . .	139
4.17 Coeficiente de difusão dos defeitos intersticiais e de vacância em função do parâmetro de rede. . . . .	140
4.18 Energia livre de ligação para o defeito intersticial Tu e de vacância calculados a partir da aproximação harmônica. . . . .	141
4.19 Derivada da energia livre em relação a $\lambda$ para o potencial de confinamento em função do parâmetro $\lambda$ , para a molécula isolada. . . . .	141
4.20 Derivada da energia livre em relação a $\lambda$ para a diferença de energia entre os sistemas em função do parâmetro $\lambda$ , para (a) o defeito intersticial Tu e (b) defeito de vacância. . . . .	142
4.21 Comparação da energia livre dos defeitos intersticiais Tu e de vacância entre os métodos de aproximação harmônica e integração termodinâmica. A linha contínua representa os cálculos obtidos com a aproximação harmônica, enquanto os pontos representam os resultados obtidos por meio da integração termodinâmica. . . . .	144
B.1 Fluxograma de uma simulação de dinâmica molecular <i>ab initio</i> (AIMD). . . . .	153

# Lista de Tabelas

1.1	Parâmetros geométricos, de carga e de Lennard-Jones para diferentes potenciais empíricos para a água. . . . .	43
2.1	Parâmetros da função de perda ( <i>loss function</i> ) empregados no treinamento. . . . .	79
2.2	RMSE para energias e forças dos modelos em comparação com dados DFT para cada número de passos de treino. . . . .	81
2.3	Tabela de sucesso de dinâmicas de 2 ns para cada modelo treinado com diferentes números de passos. O marcador verde (✓) indica que a energia foi conservada durante a simulação, enquanto o marcador vermelho (×) indica o oposto. . . . .	82
2.4	RMSE para energias e forças dos modelos em comparação com dados DFT para cada raio de corte. . . . .	83
2.5	Comparação do comprimento da ligação covalente, ângulo H-O-H, densidade de equilíbrio e precisão da função de distribuição radial para todos os pares para os modelos treinados com diferentes raios de corte em relação aos dados experimentais. O desvio padrão do comprimento de ligação é $\pm 0.03 \text{ \AA}$ , para o ângulo é $\pm 5^\circ$ e para a densidade a incerteza é de $\pm 0.008 \text{ g/cm}^3$ . O valor experimental de $\theta_{HOH}$ refere-se à fase gasosa em condições ambientes. . . . .	85
2.6	Densidade de equilíbrio para os modelos PBE-64 e PBE-256 para diferentes números de moléculas de água na caixa de simulação. . . . .	87
2.7	Comprimento da ligação $d_{OH}$ , ângulo $\theta_{HOH}$ e densidade de equilíbrio $\rho$ para os modelos SCAN-W e SCAN-WI2. O desvio padrão para as simulações de água é $\pm 0,03 \text{ \AA}$ (ligação), $\pm 5^\circ$ (ângulo) e $\pm 0,008 \text{ g/cm}^3$ (densidade). Para o gelo, os desvios são $\pm 0,02 \text{ \AA}$ (ligação), $\pm 4^\circ$ (ângulo) e $\pm 0,008 \text{ g/cm}^3$ (densidade). . . . .	89
2.8	RMSE para energia e força dos modelos treinados com diferentes números de configurações. . . . .	90

2.9	Densidade de equilíbrio e coeficiente de difusão para os modelos MACE treinados com 540, 1080 e 1620 configurações. O desvio padrão é da ordem de $\pm 0.03 \text{ g/cm}^3$ para a densidade e $\pm 0.004 \text{ \AA}^2/\text{ps}$ para a difusão. . . . .	92
3.1	Parâmetros da função de perda ( <i>loss function</i> ) utilizados no processo de treinamento. . . . .	99
3.2	RMSE para cada modelo treinado para os subconjuntos de teste: água, gelo e a combinação de ambos (Água+Gelo). . . . .	100
3.3	RMSE para cada modelo treinado para as configurações de gelo com os dados obtidos para o método de validação <i>a posteriori</i> . . . . .	102
3.4	Densidade de equilíbrio para cada modelo nas simulações DP-MD e DP-PIMD. Os valores entre parênteses denotam os desvios percentuais em relação aos dados experimentais [10]. A incerteza das densidades é da ordem de $\pm 0.009 \text{ g/cm}^3$ . . . . .	107
3.5	Desempenho percentual de cada modelo na reprodução das RDFs experimentais, calculado com a Equação 2.2. . . . .	110
3.6	Distâncias interatômicas ( $d_{OH}$ , $d_{O\dots H}$ e $d_{O\dots O}$ ) para as simulações DP-MD e DP-PIMD. Os valores correspondem ao comprimento de ligação mais provável, apresentado nas Figuras 3.11(b-d). . . . .	111
4.1	RMSE para as energias (em meV/átomo) e forças (em meV/Å) de cada subconjunto para o modelo treinado com dados PBE no DeePMD-kit. . . . .	120
4.2	Número de configurações utilizadas para cada subconjunto para o treinamento da rede utilizando o MACE. . . . .	120
4.3	RMSE para as energias (em meV/átomo) e forças (em meV/Å) de cada subconjunto para o modelo treinado com dados SCAN no MACE. . . . .	120
4.4	Energia de formação para os funcionais PW91, PBE, PBE+D3 e SCAN. . . . .	127
4.5	Energia de ligação para os funcionais PBE e SCAN. . . . .	128
4.6	Diferença entre as energias de formação ( $\Delta E_{form} = E_{form}^{NN} - E_{form}^{DFT}$ ) para todos os defeitos e parâmetros de rede para os modelos treinados com os funcionais PBE e SCAN utilizando o código DeePMD. . . . .	130
4.7	Diferença entre as energias de formação para todos os defeitos e parâmetros de rede para o modelo treinado com o funcional SCAN utilizando o MACE. . . . .	131

4.8	Valores obtidos para a integral do potencial de confinamento ( $I_1$ ) para a molécula isolada. . . . .	142
4.9	Valores obtidos para a integral $I_2$ (interação molécula-sistema) para cada defeito. . . . .	143

# Introdução

A água constitui uma das substâncias mais fundamentais da natureza e, apesar de extensivamente estudada, diversas de suas propriedades permanecem insuficientemente compreendidas. A rede de ligações de hidrogênio na fase líquida, em particular, é determinante para inúmeros fenômenos físicos e aplicações tecnológicas [11]. Sem a água, a vida, na forma que conhecemos, não existiria [12]. Todavia, a formulação de uma descrição microscópica acurada desse sistema revela-se especialmente desafiadora, em virtude da intrincada competição entre interações de curto e de longo alcance. O gelo também desempenha um papel importante em diversas áreas, como química atmosférica e ambiental, criobiologia, ciência dos materiais e engenharia [13]. Quando a água congela, as forças de interação entre as moléculas de  $\text{H}_2\text{O}$  superam o movimento térmico, resultando na formação de um arranjo mais estável com simetria hexagonal, o que explica por que os cristais de neve apresentam sempre essa geometria [14].

## Gelo $I_h$

Ao observar o diagrama de fases da água na Figura 1, notamos a existência de 19 fases de gelo [15]. Apesar dessa miríade de fases, a fase  $I_h$ , também chamada de hexagonal, é a mais encontrada na Terra [16]. O gelo  $I_h$  é formado por moléculas de água fracamente ligadas por meio de ligações de hidrogênio. Para uma molécula livre, a distância de equilíbrio O—H é de  $0,9572 \pm 0,0003 \text{ \AA}$  e o ângulo H—O—H é de  $104,52 \pm 0,05^\circ$ . É de importância crucial para as propriedades da água e do gelo o fato de a molécula ser angular, em vez de ter os três átomos em linha reta. Essa forma angular confere a ela um momento de dipolo e determina como as moléculas se arranjam em um cristal [16].

Suas propriedades únicas devem-se às ligações de hidrogênio, ainda não totalmente compreendidas em escala microscópica [17], que promovem a coesão nas fases sólida e líquida da água. Embora a estrutura seja simples, outro ponto a destacar é o comportamento anômalo da densidade. Em geral, a fase sólida de uma substância é mais densa que sua fase líquida. Contudo, para a água, observamos o oposto, com o gelo sendo menos denso que a água, o que o faz flutuar [12]. Além disso, mesmo na fase líquida, há um intervalo onde é observado

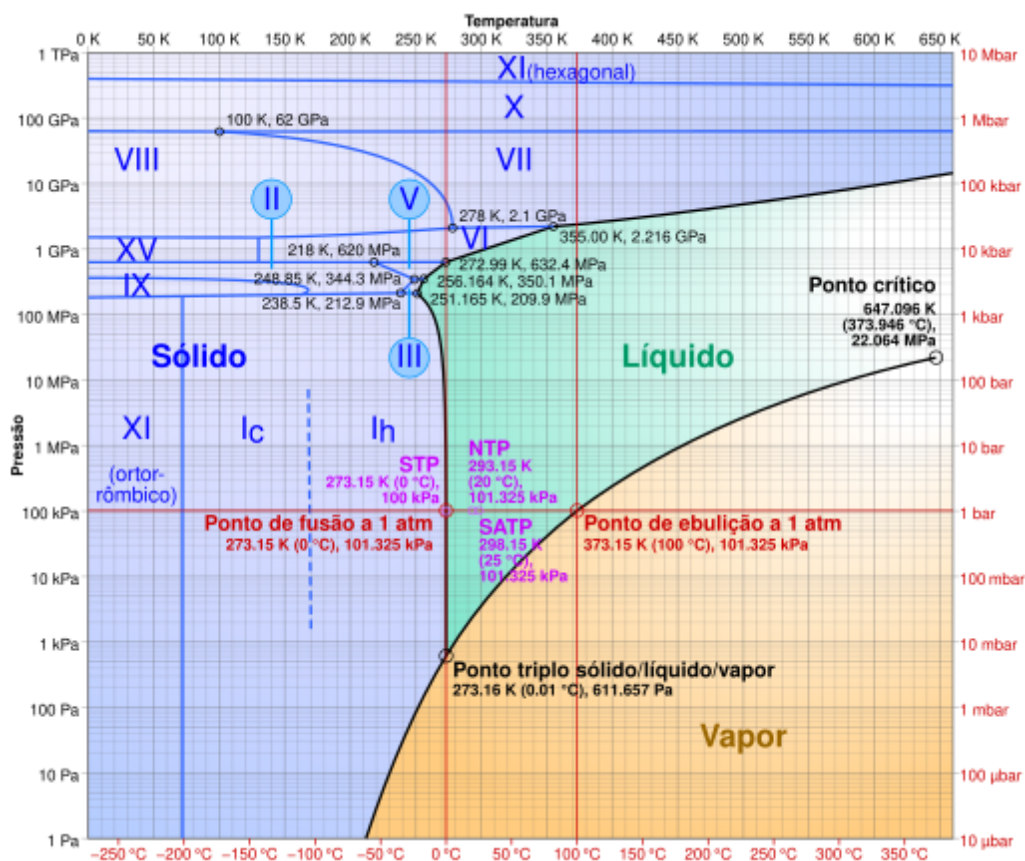


Figura 1: Diagrama de fases da água.

uma anomalia no comportamento da densidade, com a água possuindo máximo valor de densidade em 4 °C [18].

A estrutura cristalográfica do gelo I<sub>h</sub> foi determinada por Barnes [19] utilizando difração de raios X em um monocristal de gelo. A célula unitária hexagonal contém quatro moléculas de água. Na célula unitária, os átomos de oxigênio ocupam posições quádruplas nos sítios de Wyckoff:  $(1/3, 2/3, z_0)$ ;  $(2/3, 1/3, 1/2+z_0)$ ;  $(1/3, 2/3, 1/2-z_0)$ ; e  $(2/3, 1/3, -z_0)$ , onde  $z_0 \approx 1/16$ . De acordo com as regras propostas por Bernal e Fowler [20], na estrutura do gelo, cada átomo de oxigênio está ligado a quatro átomos de hidrogênio: dois por meio de ligações covalentes e outros dois por meio de ligações de hidrogênio. Sendo assim, cada átomo de hidrogênio se liga a dois átomos de oxigênio, um de forma covalente e outro mais fracamente. Essas ligações de hidrogênio são desordenadas, conforme apresentado na Figura 2. A partir dessas regras, Pauling explicou a origem da entropia residual no gelo I<sub>h</sub>, ou seja, mesmo a 0 K, é verificado uma desordem devido à existência de múltiplos arranjos moleculares dos hidrogênios nas ligações de oxigênio [21].

Entender as propriedades microscópica da água é um ativo campo de estudo e

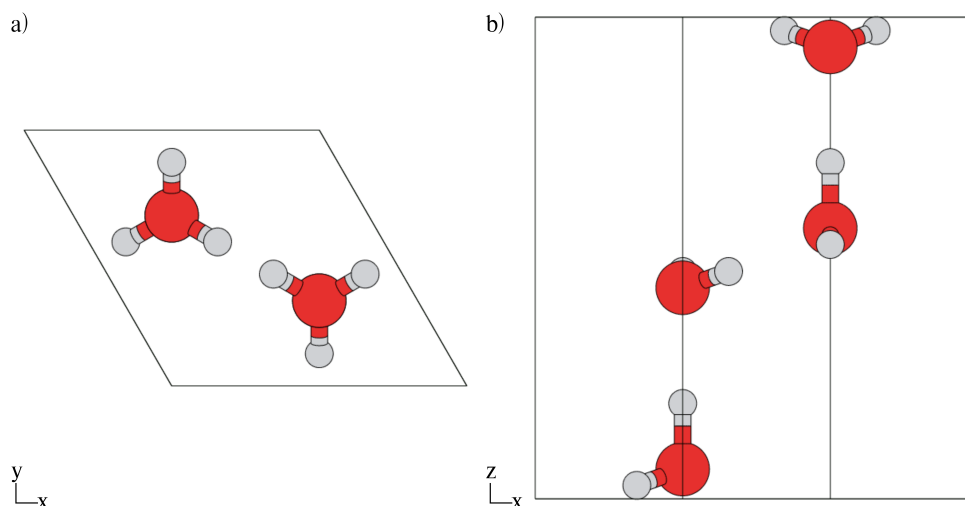


Figura 2: Representação das regras de Bernal e Fowler.

estudos teóricos podem dar uma importante contribuição para melhor entendimento de tais propriedades.

## Simulação Computacional

Simulações atomísticas têm sido fundamentais para o estudo da água, possibilitando a análise microscópica de suas propriedades [17]. A partir da aproximação de Born-Oppenheimer (BO) [22] que nos permite reescrever o Hamiltoniano do sistema em uma parte nuclear e outra eletrônica, definido uma superfície de energia potencial (PES), busca-se o estudo da evolução temporal desses sistemas. A dinâmica molecular clássica de BO, onde, a partir das equações de movimento de Newton, temos a possibilidade de calcular a o movimento dos átomos considerando os núcleos como pontos de massa carregadas, parametrizando a parte eletrônica para a interação entre os átomos é possível obter informações dinâmicas e estruturais de sistemas atomísticos. A parametrização, no entanto, pode ser um problema já que exige o ajuste de um modelo utilizando a introdução de outros cálculos mais precisos ou de resultados experimentais.

Nesse sentido, simulações de dinâmica molecular *ab initio* (AIMD), baseadas na teoria do funcional da densidade (DFT) [23, 24], tornaram-se uma ferramenta fundamental na simulação de sólidos e líquidos, pois oferecem um equilíbrio entre precisão e o tamanho possível dos sistemas estudados [25, 26]. Apesar de a AIMD poder ser altamente precisa, ela é tipicamente limitada a sistemas menores (centenas de átomos) e a tempos de simulação curtos (dezenas de picossegundos),

com resultados que dependem da qualidade do funcional de troca e correlação (XC) utilizado [27]. Isso é particularmente problemático no caso de moléculas de água, onde, devido ao delicado equilíbrio entre ligações de hidrogênio e interações de van der Waals (vdW), a maioria dos funcionais XC não descreve corretamente todas as características da água líquida [27]. Recentemente, Montero *et al.* [28] investigaram uma variedade de funcionais XC para sistemas de água e gelo  $I_h$  e mostraram que não há nenhum funcional que apresente boa concordância com os resultados experimentais para a densidade de equilíbrio de ambos os sistemas.

Foi recentemente proposto que forças de van der Waals (vdW), relativamente fracas e isotrópicas, são essenciais para a compreensão das propriedades da água líquida e do gelo [11]. Essa percepção foi obtida a partir de simulações computacionais *ab initio*, que fornecem uma descrição da água em nível atômico e revelam informações sobre as forças moleculares subjacentes [29]. Para a inclusão das forças de vdW existem diversos funcionais, os quais se destacam optB88-vdW [30] e vdW-cx [31], além as correções feitas por Grimme [32, 33]. Em adição, funcionais meta-GGA, com a inclusão de termos não locais, como o SCAN [34], demonstrou que fornece uma descrição qualitativa correta da água líquida e do gelo, a um custo computacional acessível [35].

Juntamente com a água, o gelo tem sido objeto de extensos estudos teóricos sob uma perspectiva microscópica, utilizando métodos como Monte Carlo e simulações de dinâmica molecular (MD) com modelos empíricos para as interações entre moléculas de água [36, 37, 38, 39]. Embora alguns modelos clássicos possam fornecer bons resultados para fases do gelo [36, 37], como o TIP4P/Ice [39], sua precisão e transferibilidade são frequentemente limitadas, o que representa um desafio para reproduzir com exatidão os dados experimentais em diferentes condições usando um único conjunto de parâmetros. Para os cálculos de dinâmica molecular foram desenvolvidos diversos *softwares*, tais como o LAMMPS [40] e o i-PI [41] que permitem obter propriedades dinâmicas e estruturais de gelo e água, particularmente para sistemas clássicos. Busca-se, no entanto, um novo paradigma: realizar cálculos precisos – com precisão *ab initio* – e que permitam simular sistemas grandes e com tempo de simulação semelhantes a cálculos clássicos.

## Métodos de *Machine Learning*

Redes neurais artificiais constituem uma subárea da inteligência artificial e são inspiradas na forma como o cérebro realiza cálculos: por meio de unidades

simples, conectadas em redes complexas. No caso de inteligência artificial, estas redes são treinadas para alcançar um determinado objetivo [42]. Do ponto de vista puramente matemático, as redes neurais são formas funcionais cujos parâmetros são otimizados pela minimização de uma *loss function*, podendo aproximar qualquer função suave e apresentando grande capacidade de interpolação [43]. A *loss function* é calculada a partir dos erros das quantidades físicas utilizadas para treinar o modelo. A otimização da *loss function* faz com que o modelo seja preciso, diminuindo a diferença entre os valores gerados pela rede e pelo predito.

A partir da modelagem da superfície de energia potencial utilizando dados de simulações DFT [44, 45], Behler e Parrinello [44] propuseram uma representação geral de superfícies de energia potencial utilizando redes neurais *feed-forward* introduzindo elementos fundamentais que viabilizaram o uso amplo de redes neurais para descrever a PES. Eles decompueram a energia total do sistema em somas de energias atômicas, atribuindo uma rede neural distinta para cada espécie atômica, e, de forma crucial, introduziram funções de simetria centradas no átomo (ACSF) [46] para descrever a geometria local da configuração atômica dentro de um raio de corte, capturando de forma eficiente o ambiente químico ao redor de cada átomo.

Na última década, foram publicados trabalhos utilizando redes neurais para determinar as propriedades de sistemas complexos de mecânica quântica [47, 48, 49, 50, 51, 52, 53, 42]. Esses avanços abrem possibilidades não apenas para a construção de novos campos de força a partir da PES, mas também para o aumento do tamanho do sistema e da escala temporal das simulações. No entanto, por se tratar de uma área relativamente nova, ainda é necessário sistematizar o processo de treinamento de uma boa rede neural [49, 50], o que envolve a determinação do melhor conjunto de parâmetros que produza um potencial preciso.

No início de 2018, um novo conceito para o cálculo de dinâmica molecular foi introduzido, de modo a combinar a precisão dos cálculos *ab initio* com a eficiência computacional comparável à da dinâmica molecular clássica. Este conceito foi denominado DeePMD (*Deep Potential Molecular Dynamics*) [4]. Ele se baseia no surgimento de interfaces que possibilitam modelar a superfície de energia potencial de sistemas por meio de informações coletadas de cálculos obtidos via Teoria do Funcional da Densidade (DFT) [45, 54]. Nesse conceito, redes neurais profundas (DNN) são treinadas utilizando apenas informações sobre as espécies, posições, forças e energias do sistema por meio de descritores.

Mais recentemente, outra forma de treinar modelos também foi empregada

para obter potenciais aplicáveis em simulações de dinâmica molecular: as redes neurais de grafos (GNN). Dentre os possíveis tipos de GNN, destaca-se o uso de redes neurais de passagem de mensagens (MPNN), que parametrizam um mapeamento de um grafo rotulado para um espaço-alvo, seja ele um grafo ou um espaço vetorial [5]. *Softwares* como o MACE [5, 55], o NequIP [56] e o Allegro [57] foram desenvolvidos utilizando GNN equivariantes, ou seja, as características internas se transformam de uma maneira específica sob a ação de um determinado grupo. Desse modo, é possível reduzir o número de configurações utilizadas no processo de treinamento dos modelos.

Outro ponto destacado dos modelos GNN é a diminuição nos erros obtidos para forças e energias quando comparados aos modelos treinados por meio de DNN [5]. Modelos obtidos através de treinamento utilizando o MACE forneceram erros em torno de 0,1 meV/átomo e 6,6 meV/Å para a molécula de aspirina. O modelo treinado conta com 950 configurações de dez diferentes moléculas orgânicas e apresenta baixos erros para todas as moléculas contidas em seu conjunto de treino. Com estes dados iniciais sobre os métodos, tanto o DeePMD-kit quanto o MACE surgem como candidatos a serem utilizados para o treinamento de potenciais para cálculos de dinâmica molecular, uma vez que apresentam baixos erros na descrição de sistemas em comparação ao encontrado por simulações AIMD.

Os potenciais treinados através de *machine learning* foram utilizados para diversas aplicações, como baterias recarregáveis [58], adsorção de gás nitrogênio na superfície de ferro [59] e diversas interfaces [60, 61]. Além das diversas aplicações em diferentes áreas, existe uma extensa literatura de aplicações em sistemas com água [62, 35, 63, 64]. Torres *et al.* [42] mostraram a precisão de potenciais treinados com 18.000 configurações de água para quatro funcionais de correlação e troca diferentes (PBE [65], vdW-cx [31], optB88 [30] e SCAN [34]). Apesar disto, um estudo complementar mostrou que com apenas 30% desses dados é possível obter os mesmos resultados estruturais e dinâmicos para a água líquida [66]. A construção dos dados é determinante para a qualidade do modelo. Quanto mais diversos os dados, menor a limitação da rede neural, fazendo com que ela possa prever mais pontos no espaço de fase.

Como citado anteriormente, o diagrama de fases da água é extremamente rico, uma vez que existem 19 fases de gelo, além da fase líquida e gasosa. Desse modo, Zhang *et al.* [62] construíram um modelo de forma a abranger todo o espaço de fases do diagrama. No domínio para temperaturas abaixo de 400 K e pressão inferior a 50 GPa, existem nove fases estáveis de gelo (Ih, II, III, V, VI,

VII, VIII, XI e XV), cinco metaestáveis (IV, IX, XII, XIII e XIV) e a fase líquida. Para a construção da superfície de energia potencial, utilizaram configurações de água e de todas as fases estáveis e metaestáveis experimentalmente conhecidas a pressões abaixo de 50 GPa (I<sub>h</sub>, I<sub>c</sub>, II, III, IV, V, VI, VII, VIII, IX, XI, XII, XIII, XIV, XV). As forças e energias utilizadas no processo de treinamento foram obtidas com o funcional SCAN [34]. De modo geral, os erros obtidos pelo modelo, comparados aos resultados DFT para todas as fases incluídas no treinamento, estão em torno de 1 meV por molécula de água e 50 meV/Å.

De modo geral, o diagrama de fases obtido com este modelo está de acordo com o diagrama experimental, com exceção das fases III e XV, que se tornam metaestáveis com esta abordagem. A linha de coexistência entre o gelo I<sub>h</sub> e a fase líquida está deslocada em 40 K acima do resultado experimental. A estabilidade do gelo I<sub>h</sub> é superestimada devido ao fato de o funcional SCAN superestimar as ligações de hidrogênio [62]. Com esta aplicação, é possível verificar que os potenciais obtidos por meio de *machine learning* conseguem obter bons resultados para diferentes fases da água em comparação com os resultados experimentais e, principalmente, com resultados obtidos via AIMD, porém com um custo computacional muito menor.

Recentemente, potenciais obtidos através de *machine learning* foram utilizados para investigar algumas propriedades do gelo I<sub>h</sub> [35, 67] e do gelo superiônico [15]. Um estudo sobre o equilíbrio entre o gelo I<sub>h</sub> e a água líquida, utilizando um potencial treinado com configurações de água líquida e 15 fases de gelo, aponta a transição de fase em  $310 \text{ K} < T_m < 312,5 \text{ K}$  [35]. Neste trabalho, o potencial foi treinado no nível SCAN [34]. Para obter esse resultado, os autores iniciam a simulação com a interface de água e gelo, realizam 4 simulações de pelo menos 3 ns e medem o número de moléculas do tipo gelo. Para temperaturas abaixo de 310 K, todas as moléculas permanecem na estrutura de gelo, enquanto para temperaturas acima de 312,5 K, não há ocorrência de estruturas de gelo, com o sistema tornando-se água líquida. Apesar do desvio em torno de 40 K entre o valor teórico e experimental para a temperatura de fusão, o modelo acerta a tendência da densidade em função da temperatura para ambas as fases, além de prever o fato de o gelo flutuar na água.

## Efeitos Quânticos Nucleares

Finalmente, quando falamos de uma descrição atomística precisa, para água e gelo, falta um ingrediente: a natureza quântica dos núcleos pode ser relevante em diversos sistemas, especialmente no caso de átomos leves. Átomos de hidrogênio, por exemplo, exibem comportamento quântico mesmo à temperatura ambiente [68, 69, 70]. Assim, para sistemas aquosos, já foi demonstrado que os efeitos quânticos nucleares (NQE) alteram quantitativamente sua estrutura e dinâmica, afetando a intensidade das ligações de hidrogênio [71, 72, 73, 74, 1]. Também foi mostrado que a inclusão dos NQE na descrição melhoram as propriedades estruturais de diferentes fases do gelo [73, 75, 76]. Esses efeitos são, na maioria das vezes, incluídos em simulações por meio da dinâmica molecular de integrais de caminho (PIMD) [77]. Contudo, ainda considera-se que os núcleos pontuais, de acordo com a aproximação de Born-Oppenheimer [22].

Cheng *et al.* investigaram as características termodinâmicas da água líquida empregando DFT no nível de funcionais híbridos (revPBE0-D3), levando em conta os NQE. Seus resultados indicaram que esses efeitos resultam em um aumento de aproximadamente 1% na densidade da água líquida, do gelo hexagonal e do gelo cúbico, em comparação com simulações que tratam os núcleos de forma clássica [78]. Resultados semelhantes também foram observados para a água líquida calculada utilizando os funcionais meta-GGA XC SCAN [53] e SCAN0 [79]. No entanto, a combinação de AIMD e PIMD é computacionalmente custosa e tipicamente limitada a curtas escalas de tempo e sistemas pequenos. Importante salientar que para potenciais clássicos, a inclusão de NQE apresenta um comportamento oposto [80]. Isto significa que há uma controvérsia nesse sentido e, mais importante, o mecanismo não está bem definido.

De forma análoga ao seu efeito sobre a densidade de equilíbrio, os NQE também influenciam a rede de ligações de hidrogênio. Como exemplo, cálculos de PIMD baseados em DFT [81] revelam uma densidade diferente de zero para valores positivos da coordenada de transferência de prótons, indicando compartilhamento de prótons e maior aproximação entre as moléculas de água, o que leva a um alargamento da distribuição do momento de dipolo.

Torres *et al.* [42] mostraram que a inclusão dos NQE melhora as propriedades estruturais e dinâmicas independentemente da escolha do funcional utilizado para treinar um modelo de *machine learning*. Neste trabalho, os autores obtiveram as energias e forças de 18.000 configurações de água em diferentes condições

de temperatura e densidade para quatro diferentes funcionais de correlação e troca e treinaram redes neurais para cada conjunto de dados [42]. Verificou-se que, sem a inclusão dos NQE, as funções de distribuição radial para os pares que contêm átomos de hidrogênio (O-H e H-H) apresentam um grande desvio em comparação com o resultado experimental. Contudo, quando se incluem tais efeitos, a descrição estrutural dos pares O-H e H-H se aproxima da curva obtida experimentalmente. Outro ponto a destacar é o aumento significativo do coeficiente de difusão da água para os potenciais treinados a partir de dados obtidos para os funcionais que incluem termos não locais. O uso de potenciais obtidos por meio de *machine learning* é vantajoso, especialmente ao se considerar simulações que incluem NQE em sistemas aquosos [42, 82, 79, 83, 53], permitindo cálculos mais longos e em maior escala — tanto em número de átomos quanto em tempo total de simulação — em comparação com a AIMD tradicional combinada com PIMD.

## Defeitos em Gelo $I_h$

Tal como qualquer cristal, o gelo  $I_h$  pode apresentar defeitos em sua estrutura. Esses defeitos podem ser observados em diversas formas, como defeitos pontuais, 1D e 2D [84]. Em relação aos defeitos pontuais, temos quatro possibilidades: os defeitos de vacância, os defeitos intersticiais, os defeitos de Bjerrum e os defeitos iônicos. O defeito mais simples em cristais é a ausência de um átomo em um sítio da rede. No caso do gelo, o defeito de vacância decorre da ausência de uma molécula de água na estrutura perfeita. Resultados teóricos indicam que a energia de formação dos defeitos de vacância obtida a partir de cálculos DFT é de aproximadamente 0,7 eV [85, 86]. Contudo, a energia de formação obtida por meio de relatos experimentais, utilizando técnicas de aniquilação de pósitrons [87, 88], indica uma energia de formação na faixa de 0,2 eV a 0,35 eV. Há fortes indícios de que os valores inferidos indiretamente a partir das medições com pósitrons sejam subestimadas significativamente [84]. Na Figura 3a apresentamos a representação do defeito de vacância.

Em contraste com o defeito de vacância, outro defeito observado no gelo  $I_h$  é a adição de uma molécula de água no interstício da estrutura cristalina. De modo geral, há duas posições possíveis que a molécula intersticial pode ocupar, referidas, respectivamente, como sítios *trigonal uncapped* (Tu) e *trigonal capped* (Tc). O defeito intersticial Tu (Figura 3b) está localizado no centro dos canais

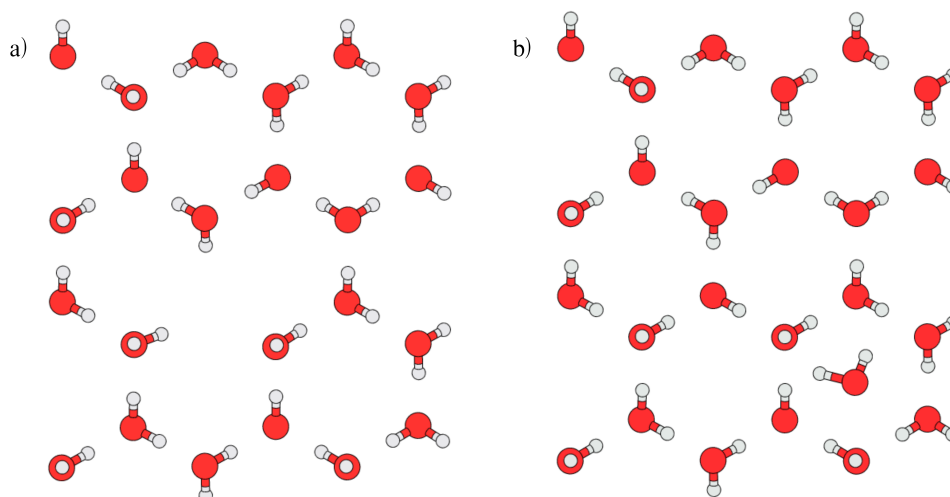


Figura 3: Representação dos defeitos (a) de vacância e (b) intersticial Tu.

hexagonais que se estendem ao longo da direção do eixo  $c$ , distorcendo pouco a estrutura e não afetando a rede de ligações de hidrogênio [84]. O sítio Tc, por outro lado, provoca perturbações substanciais. Sua posição está localizada dentro das fileiras moleculares que se estendem ao longo do eixo  $c$ . Assim, as duas moléculas vizinhas ao intersticial nessa fileira são significativamente deslocadas para os lados, distorcendo as ligações de hidrogênio das quais participam. Devido a essas distorções, a energia de formação do defeito Tc é substancialmente maior do que a do defeito Tu, com cálculos de DFT [85] fornecendo estimativas de aproximadamente 1,0 eV e 0,8 eV, respectivamente. Resultados experimentais apontam que a energia de formação para os defeitos intersticiais é de 0,4 eV [89, 90]. Esses resultados indicam ainda que, para temperaturas próximas ao ponto de fusão, a concentração de defeitos intersticiais domina sobre os de vacância, apesar de os defeitos de vacância apresentarem menor valor para a energia de formação.

Uma vez que a molécula intersticial difunde na estrutura cristalina do gelo [91], podemos estudar uma configuração observada durante o processo de difusão da molécula intersticial. A esta configuração damos o nome de *bond center* (Bc), que apresentamos na Figura 4. Neste caso, a molécula extra está posicionada entre um par de moléculas de água originalmente unidas por uma ligação de hidrogênio. Para essa configuração, devido à ligação parcial com as moléculas do cristal, a energia do defeito é reduzida, como indicado por cálculos DFT [85] com energia abaixo dos demais defeitos intersticiais e comparável à do defeito de vacância, porém ainda distante dos valores experimentais.

É amplamente aceito que a molécula de água se difunde como uma unidade

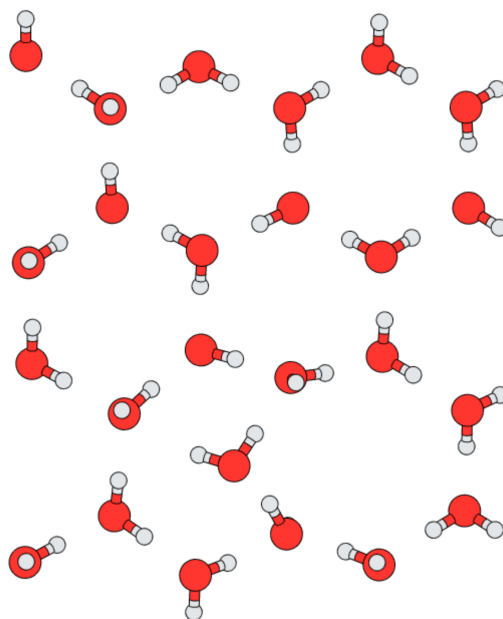


Figura 4: Representação do defeito Bc.

através do gelo [92]. A barreira de migração de uma molécula intersticial  $T_u$  de um sítio para outro é de 0,16 eV, e o coeficiente de difusão sendo de  $1,8 \times 10^{-6} \text{ m}^2\text{s}^{-1}$  [92]. A difusão de moléculas intersticiais ocorre nas direções paralela e perpendicular ao eixo  $c$ , cuja representação dos eixos está presente na Figura 5. A difusão perpendicular ao eixo  $c$  ocorre entre dois hexágonos diferentes, enquanto a difusão paralela faz com que a molécula mantenha-se no mesmo hexágono, porém movendo-se livremente nesta direção. Enquanto a difusão paralela ao eixo  $c$  não altera a natureza das ligações de hidrogênio, a difusão perpendicular não segue este comportamento. A molécula que ocupa o interstício expulsa uma molécula da rede cristalina, ocupa seu espaço e faz com que a molécula expulsa ocupe outro interstício. De acordo com Goto *et al.* [92], a frequência de salto na direção paralela deve ser maior devido ao caminho mais livre que a molécula percorre, em comparação com a direção normal ao eixo  $c$ , porém não é possível afirmar qual a magnitude da diferença entre as frequências. Contudo, cálculos de dinâmica molecular mostram que o coeficiente de difusão perpendicular ao eixo  $c$  é superior ao da difusão paralela, mesmo possuindo uma maior barreira de migração.

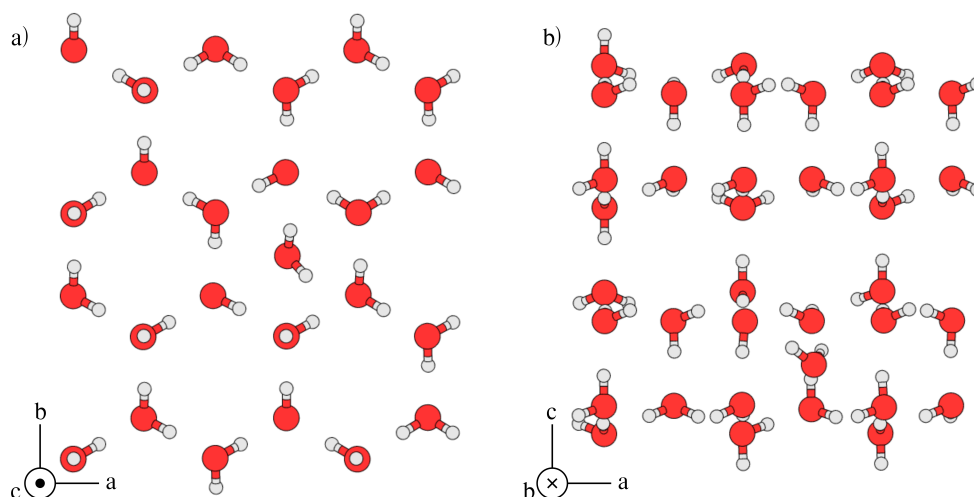


Figura 5: Visualização do defeito Tu no plano (a)  $ab$  e (b)  $ac$ .

## A Tese

A partir de todas as informações apresentadas, há a necessidade de responder algumas questões em aberto: quais os parâmetros e dados necessários para a construção de uma rede neural capaz de obter bons resultados para sistemas aquosos? Apesar de saber da melhoriana descrição teórica por meio da inclusão dos efeitos quânticos nucleares nas propriedades estruturais e dinâmicas de água e gelo, qual a influência destes efeitos nas distâncias interatômicas e intermoleculares e como isto altera de tais propriedades? Por fim, qual defeitos apresentam maior concentração na estrutura cristalina e qual o mecanismo de difusão dos defeitos na rede do gelo?

Para responder estas questões, esta tese é dividida em cinco partes: uma revisão sobre as teorias aplicadas, um manual para o treinamento de potenciais acurados, a melhoria na descrição das propriedades estruturais e vibracionais do gelo  $I_h$  ao incluir os NQE, uma análise sobre a formação e a mobilidade de defeitos pontuais na estrutura cristalina do gelo  $I_h$  e, finalizando, as conclusões da presente tese. No Capítulo 1, apresentamos a teoria utilizada nesta tese. Iniciamos com a Teoria do Funcional da Densidade (DFT), apresentando a teoria proposta por Kohn e Sham [24], que transforma a equação de Schrödinger de muitos corpos em um sistema de equações para elétrons independentes, ou seja, transformando um problema de  $3N$  variáveis em  $N$  problemas de três variáveis. A aproximação de Born-Oppenheimer [22] separa o hamiltoniano total na soma da parte nuclear e da parte eletrônica. Enquanto a DFT resolve o hamiltoniano da parte eletrônica, a Dinâmica

Molecular (MD) considera os átomos como partículas pontuais que se movem de acordo com as leis de Newton, sob um potencial (parametrizado) devido aos elétrons. Para átomos leves, como o hidrogênio, a descrição de partículas pontuais não é uma boa aproximação, e a natureza quântica do núcleo deve ser considerada. Desse modo, apresentamos os efeitos quânticos nucleares e as integrais de caminho, uma forma de solucionar um problema quântico através de uma abordagem clássica. Em seguida, abordaremos os potenciais obtidos por meio de *machine learning*, que aliam a precisão de cálculos *ab initio* com o baixo custo computacional e a escalabilidade das simulações de dinâmica molecular. Existem diversos tipos de redes neurais utilizadas para o treinamento dos potenciais. Dois tipos serão apresentados: as redes neurais profundas e as redes neurais de grafos. Discutiremos a teoria por trás dessas redes. Por fim, introduziremos o conceito da integração termodinâmica, principalmente seu uso para o cálculo da energia livre de formação de um sistema considerando efeitos de temperatura.

No Capítulo 2, apresentaremos um manual para a construção de um bom potencial obtido através de *machine learning*. Para a construção de um bom modelo para sistemas aquosos a partir do DeePMD-kit [93], observaremos os seguintes parâmetros: número de passos de treino, raio de corte, e tamanho do sistema no *dataset*. Além disto, observaremos os efeitos da inclusão de configurações de gelo na qualidade do modelo, identificando se há necessidade de aumentar o *dataset*. De forma complementar, observaremos a quantidade de dados necessária para se obter um bom potencial através das redes GNN, utilizando o *software* MACE [55, 5], que diminui consideravelmente o número de configurações utilizadas no processo de treinamento. Por fim, apresentaremos uma análise sobre o número de canais e camadas. Esses parâmetros influenciam a precisão do modelo, contudo, podem implicar um maior custo computacional.

Em continuidade, no Capítulo 3, elucidaremos o papel da inclusão dos NQE nas propriedades estruturais do gelo  $I_h$ . Realizaremos uma análise detalhada da densidade de equilíbrio, das distâncias inter e intramoleculares e das frequências vibracionais das moléculas de água. Para isso, foram conduzidas simulações utilizando Dinâmica Molecular com os potenciais obtidos através de *machine learning* (DP-MD) e Dinâmica Molecular de Integrais de Caminho (DP-PIMD), em diferentes modelos DP treinados com distintos funcionais de correlação e troca. Os resultados de DP-MD mostram boa concordância com os cálculos de AIMD. Com a inclusão dos NQE, observamos um ligeiro aumento no valor médio do comprimento da ligação covalente O-H, o que leva a um desvio para o vermelho

(*redshift*) nas frequências de estiramento. Consequentemente, ocorre o compartilhamento de prótons, em consonância com resultados anteriores para a água líquida [94, 1]. Ao mesmo tempo, esse efeito compete diretamente com a redução das distâncias intermoleculares devido ao fortalecimento das ligações de hidrogênio. Isso resulta em um aumento geral da densidade de equilíbrio para todos os funcionais, afastando ainda mais os resultados dos valores experimentais. Isso indica que todos os funcionais, sejam do tipo GGA ou SCAN, tendem a superligar as moléculas de água no gelo, e que os NQE aumentam ainda mais a força das ligações de hidrogênio.

No Capítulo 4, investigaremos a formação e a mobilidade de defeitos pontuais na estrutura do gelo  $I_h$ . Apresentamos uma análise da energia de formação de cada defeito através de cálculos de minimização DFT e, também, com novos potenciais treinados com a inclusão de configurações de defeitos. Para esta análise, treinamos potenciais tanto DNN quanto GNN para encontrar o mais preciso em comparação aos resultados *ab initio*. Os resultados indicam que a inclusão de termos não locais diminui a energia de formação dos defeitos intersticiais e, consequentemente, altera o defeito que aparece em maior concentração em todo o intervalo de temperatura analisado. Mostramos também a energia necessária para o defeito migrar de um sítio para outro. Encontramos que os defeitos intersticiais possuem menor barreira de migração em comparação aos defeitos de vacância. Em contrapartida, a frequência de tentativa de migração do defeito de vacância é superior à dos defeitos intersticiais. Desse modo, os coeficientes de difusão dos defeitos apresentam valores próximos entre si. Além disto, verificamos que o defeito intersticial apresenta uma tendência maior a se difundir na direção perpendicular ao eixo  $c$  do que na direção paralela a este eixo. Por fim, verificamos que o aumento da temperatura influencia de forma distinta a energia de ligação, aumentando a interação entre a molécula intersticial e o cristal, porém diminuindo a intensidade da ligação das moléculas no interior do cristal.

Por fim, no Capítulo 5, apresentaremos um resumo do que foi abordado nesta tese, elucidando ponto a ponto, desde a construção dos potenciais até a mobilidade dos defeitos na estrutura do gelo  $I_h$ , passando pelas análises em relação à importância da inclusão dos NQE nas propriedades estruturais e vibracionais do gelo.

# Capítulo 1

## Metodologia

Neste trabalho, serão utilizados diversos métodos e teorias para a solução dos problemas relacionados às propriedades eletrônicas, estruturais e energéticas das estruturas de água e gelo. Deste modo, antes da discussão de qualquer resultado, é necessário apresentar a metodologia a ser empregada.

Neste capítulo, serão apresentadas duas formas de se solucionar o problema de muitos corpos, sob duas óticas: uma abordagem quântica [23, 24], que resolve a parte eletrônica a partir da equação de Schrödinger; e uma abordagem clássica [95], que considera cada átomo como uma massa com carga pontual. Apesar de observarem o mesmo problema sob diferentes pontos de vista, cada método é aplicado a fim de obtermos diferentes propriedades do sistema.

De início, será apresentada a Teoria do Funcional da Densidade (DFT) [23, 24], que transforma o problema de muitos corpos em um problema de elétrons não interagentes. Serão discutidos, além da matemática para se obter a equação de Kohn-Sham, as diferentes abordagens para a aproximação do funcional de energia de troca e correlação — elemento importante para a aplicação da DFT — e também o método dos pseudo-potenciais, um aspecto relevante para o estudo, uma vez que o potencial Coulombiano é substituído por um potencial suave e finito na origem.

Uma vez que a DFT soluciona a equação de Schrödinger para um sistema de muitos corpos com núcleos estáticos, a Dinâmica Molecular (MD) é a alternativa para se medirem as propriedades dinâmicas do sistema, incluindo os efeitos de temperatura. Nesta abordagem, na qual o problema é tratado de forma clássica, com os átomos representados por massas pontuais carregadas, existem diversas formas de se descreverem as interações interatômicas: desde o uso de potenciais empíricos [39] até interações mais sofisticadas, como as descritas por métodos *ab initio* (calculadas a cada passo) [96] ou por modelos gerados por aprendizado de máquina a partir dos cálculos *ab initio* [97, 5].

Apesar de a dinâmica molecular ser uma boa representação do problema, alguns átomos leves, como o de hidrogênio, não podem ser descritos de forma

puramente clássica, precisando ter sua natureza quântica considerada. Deste modo, é necessário incluir os chamados efeitos quânticos nucleares nos cálculos. Uma possível forma de resolver este problema é utilizar-se da formulação de integrais de caminho para se resolver, por meio de um isomorfismo clássico, um problema de natureza quântica.

Como citado anteriormente, a interação entre os átomos pode ser descrita de diversas formas. Contudo, dada a precisão das interações obtidas por métodos *ab initio*, busca-se uma forma de utilizar esta descrição acurada para aprimorar a resolução das equações de movimento. Neste contexto, os potenciais obtidos por meio de redes neurais se tornam uma alternativa para gerar um potencial com precisão em nível de DFT, porém com menor custo computacional. Serão apresentadas duas abordagens para a obtenção destes potenciais: as redes neurais profundas [97, 93, 98] e as redes neurais de passagem de mensagens [5, 55].

Por fim, será apresentada uma breve discussão sobre o método de integração termodinâmica, uma abordagem para se calcular a diferença de energia livre entre um estado  $\mathcal{A}$  e um estado  $\mathcal{B}$ .

## 1.1 Teoria do Funcional da Densidade (DFT)

No contexto de sistemas atômicos costumeiramente deparamos com problemas de muitos corpos. Independentemente de se tratar o problema de modo quântico ou clássico, a resolução de um problema de  $N$  corpos não é trivial e, em geral, é impossível de ser obtida de forma analítica. Sendo assim, para o caso quântico, a Teoria do Funcional da Densidade (DFT) — baseada nos teoremas de Hohenberg-Kohn [23] — colocadas em prática pelas equações de Kohn-Sham [24], emerge como uma alternativa para solucionar a equação de Schrödinger de muitos corpos [99, 100, 101] para sistemas atômicos, a partir de uma abordagem de elétrons não interagentes em um potencial efetivo.

### 1.1.1 Equação de Schrödinger para Muitos Corpos

A equação de Schrödinger é a equação que rege a Mecânica Quântica em baixas energias. Ao estudar a matéria em nível atômico, podemos considerá-la como composta apenas por núcleos e elétrons. Visto que estes corpos formam os materiais, observa-se uma interação Coulombiana entre eles, seja ela repulsiva ou atrativa, dada a natureza de suas cargas. As equações 1.1 a 1.3 apresentam a

energia potencial associada a cada interação Coulombiana

$$E_{ee} = \frac{e^2}{4\pi\epsilon_0 d_{ee}} \quad (1.1)$$

$$E_{nn} = \frac{Z^2 e^2}{4\pi\epsilon_0 d_{nn}} \quad (1.2)$$

$$E_{en} = \frac{-Ze^2}{4\pi\epsilon_0 d_{en}} \quad (1.3)$$

em que  $Z$  é o número atômico,  $d_{ee}$ ,  $d_{nn}$  e  $d_{en}$  são as distâncias elétron-elétron, núcleo-núcleo e elétron-núcleo, respectivamente, e  $\epsilon_0$  é a permissividade elétrica no vácuo.

De modo geral, equação de Schrödinger independente do tempo é dada por

$$\hat{H}\Psi(\mathbf{r}) = E\Psi(\mathbf{r}), \quad (1.4)$$

na qual o operador Hamiltoniano ( $\hat{H}$ ) é composto pelas contribuições de energia cinética e potencial. Sua solução é mais simples ao se tratar de apenas uma única partícula. Porém, para um sistema mais geral, a função de onda  $\Psi(\mathbf{r})$  deve ser reescrita para  $N$  elétrons e  $K$  núcleos

$$\Psi = \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, \mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_K), \quad (1.5)$$

onde  $\mathbf{r}$  representa as posições dos elétrons e  $\mathbf{R}$  as posições dos núcleos. Neste ponto, é preciso introduzir a ideia de que o sistema se encontrará em equilíbrio, em seu estado de menor energia. Este estado é representado por  $\Psi_0$ . Sabe-se que a probabilidade de se encontrar uma partícula em um dado ponto  $\mathbf{r}$  é dada por  $|\Psi|^2$ ; com isto, o estado de menor energia fornece informação sobre a distribuição de carga eletrônica, ou seja

$$n(\mathbf{r}) = N \int |\Psi_0(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N, \mathbf{R}_1, \dots, \mathbf{R}_K)|^2 d\mathbf{r}_2 \dots d\mathbf{r}_N d\mathbf{R}_1 \dots d\mathbf{R}_K. \quad (1.6)$$

A equação 1.6 nos mostra que a densidade eletrônica é dada pela probabilidade de se encontrar um elétron em uma posição  $\mathbf{r}$ . Desta maneira, com  $\Psi$  normalizada, a partir da equação 1.6 temos que

$$\int n(\mathbf{r}) d\mathbf{r} = N. \quad (1.7)$$

Retornando à equação de Schrödinger, precisamos definir os termos de energia cinética e potencial. Sabe-se que a energia cinética de uma partícula é dada por

$$\hat{T} = \frac{p^2}{2m'}, \quad (1.8)$$

sendo o operador momento  $\hat{\mathbf{p}} = -i\hbar\nabla$ , em que  $\hbar$  é a constante de Planck reduzida, o operador de energia cinética total do sistema pode ser escrito na representação das posições em função do operador  $\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$ . Podemos, assim, definir a primeira parte do Hamiltoniano do sistema

$$\hat{T} = -\sum_{i=1}^N \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_{I=1}^K \frac{\hbar^2}{2M_I} \nabla_I^2, \quad (1.9)$$

sendo  $m_e$  a massa do elétron e  $M_I$  a massa de cada núcleo. Vale ressaltar que os índices em minúsculo fazem referência aos elétrons e os maiúsculos, aos núcleos, assim como na representação das posições.

Uma vez conhecida a parcela cinética do Hamiltoniano, precisamos definir a parte potencial para que se possa ter a equação de Schrödinger total do sistema. Assim, retornemos ao início desta seção, em particular às equações 1.1 a 1.3. Desta forma, as energias potenciais devido aos pares de cargas são

$$V_{ee} = \frac{1}{2} \sum_{i \neq j} \frac{e^2}{4\pi\epsilon_0} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad (1.10)$$

$$V_{nn} = \frac{1}{2} \sum_{I \neq J} \frac{e^2}{4\pi\epsilon_0} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}, \quad (1.11)$$

$$V_{en} = -\sum_{i,I} \frac{e^2}{4\pi\epsilon_0} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|}, \quad (1.12)$$

o fator 1/2 presente nos potenciais  $V_{ee}$  e  $V_{nn}$  tem a finalidade de corrigir a dupla contagem das interações, sendo que as somas excluem os termos  $i = j$  e  $I = J$ , uma vez que uma partícula não interage consigo mesma. Os índices  $i$  e  $j$  variam de 1 a  $N$ , enquanto os índices  $I$  e  $J$  variam de 1 a  $K$ .

A partir das equações 1.9 a 1.12, apresentamos a Equação de Schrödinger para muitos corpos, o ponto de partida para a construção da teoria do funcional da

densidade

$$\left[ -\sum_{i=1}^N \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_{I=1}^K \frac{\hbar^2}{2M_I} \nabla_I^2 + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{4\pi\epsilon_0} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2} \sum_{I \neq J} \frac{e^2}{4\pi\epsilon_0} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} - \sum_{i,I} \frac{e^2}{4\pi\epsilon_0} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} \right] \Psi = E_{tot} \Psi. \quad (1.13)$$

Esta é a formulação mais simples para a equação de Schrödinger estacionária para muitos corpos. Outras interações — como as devido a um campo magnético externo — também podem gerar termos de potencial a serem adicionados ao Hamiltoniano total, porém os termos da equação 1.13 são suficientes para descrever os materiais que serão estudados.

Ao se resolver a equação para muitos corpos, encontram-se todos os estados possíveis, inclusive o estado fundamental, que é o estado de menor energia no qual o sistema se encontra em equilíbrio. A partir deste estado, é possível calcular as principais propriedades do material. Contudo, com exceção do átomo de hidrogênio, não há solução analítica para o problema. A DFT emerge como uma alternativa para a solução de um sistema de muitos corpos (apesar de ainda exigir aproximações).

### 1.1.2 Aproximação de Born-Oppenheimer

A partir deste momento, precisamos começar a buscar soluções para a equação 1.13. Como apresentado no Apêndice A, a equação de muitos corpos, pode ser reescrita em unidades atômicas como

$$\left[ -\sum_{i=1}^N \frac{\nabla_i^2}{2} - \sum_{I=1}^K \frac{\nabla_I^2}{2M_I} + \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} - \sum_{i,I} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} \right] \Psi = E_{tot} \Psi. \quad (1.14)$$

Os materiais estudados, em geral, são sólidos, gases e líquidos, e a equação de muitos corpos consegue descrever basicamente tais sistemas. Neste caso, para que a solução não seja tão complexa, há a necessidade de se buscar aproximações a fim de simplificar a solução. Quando trazemos uma visão da estrutura molecular dos líquidos e gases, podemos observar que os núcleos possuem um maior espaço para percorrer. Já nos sólidos, os núcleos permanecem na vizinhança de uma posição de equilíbrio. Portanto, iremos considerar, então, que os núcleos estão

fixos e em posições já conhecidas.

Uma outra abordagem para esta solução é a aproximação de Born-Oppenheimer [22]. Imagine o sistema Sol-Terra. Para este sistema, temos que o Sol possui uma massa muito maior que a massa da Terra, e o planeta tem uma órbita ao redor da estrela. Ao fazer todos os cálculos, verificamos que o centro de massa do sistema está muito próximo ao Sol, fazendo com que a Terra tenha uma velocidade bem superior à do Sol. Por esta ótica, consideramos o Sol como parado e apenas o planeta em movimento. Para o caso da interação entre núcleo e elétrons, também se considera esta relação, uma vez que  $M_I \gg m_e$ . Há de se ressaltar também que, por este motivo, consideramos que os núcleos e os elétrons possuem escalas temporais distintas, ou seja, os elétrons se movem mais rapidamente que os núcleos. Isto tem algumas implicações na análise do problema de muitos corpos.

A primeira implicação é que o termo de energia cinética nuclear pode ser desprezado, o que é justificado tanto pelo fato de os núcleos serem considerados fixos quanto por sua massa ser muito maior que a do elétron ( $M_I \gg m_e$ , que no limite  $M_I \rightarrow \infty$  anula o termo). A segunda é que a interação entre os núcleos acaba se tornando uma função escalar parametrizada pela posição dos núcleos. A terceira é que a interação elétron-núcleo entra como um potencial externo na parte nuclear. Desta forma, é possível separar o Hamiltoniano total do sistema entre a parte eletrônica e a parte nuclear

$$\hat{H} = \hat{H}_{el}(\hat{R}) + \hat{H}_{nucl}, \quad (1.15)$$

onde o termo de interação elétron-núcleo é introduzido como um potencial externo em função das posições nucleares.

Com esta aproximação, podemos definir uma nova função de onda para o sistema, presente na equação 1.16

$$\Psi(\mathbf{r}, \mathbf{R}) = \psi(\mathbf{r}; \mathbf{R})\Phi(\mathbf{R}), \quad (1.16)$$

sendo  $\psi(\mathbf{r}; \mathbf{R})$  a parte eletrônica da função de onda e  $\Phi(\mathbf{R})$  a contribuição nuclear. Devido a separação do hamiltoniano, a DFT descreve apenas a superfície de energia correspondente ao estado eletrônico fundamental, não considerando transições entre diferentes superfícies eletrônicas. Logo, toda a dinâmica ou termodinâmica obtida a partir da DFT está restrita a uma única superfície de energia potencial (PES). Notamos a aparição do ponto e vírgula na função de onda para os elétrons devido ao fato de as posições dos núcleos se tornarem parâmetros para a parte

eletrônica. Então, por esta definição apresentada pela equação 1.16, começamos a escrever a equação 1.14 em função apenas da contribuição eletrônica do sistema. Para isto, definimos as seguintes quantidades

$$E = E_{tot} - \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}, \quad (1.17)$$

$$V_{ext}(\mathbf{r}) = - \sum_I \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|}. \quad (1.18)$$

Desta maneira, a equação 1.17 leva a constante para o lado direito da equação 1.14. Conseguimos, ainda, remover qualquer dependência explícita dos graus de liberdade nucleares, sendo a equação 1.18 o potencial externo dos núcleos experimentado pelos elétrons. Substituindo, então, as equações 1.17 e 1.18 em 1.14 e trazendo as considerações da aproximação de Born-Oppenheimer, teremos

$$\left[ - \sum_{i=1}^N \frac{\nabla_i^2}{2} + \sum_i V_{ext}(\mathbf{r}_i) + \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right] \psi = E\psi. \quad (1.19)$$

Já que a equação 1.15 separa o Hamiltoniano em termos eletrônico e nuclear e a equação 1.19 retira a dependência nuclear do Hamiltoniano, podemos buscar uma expressão para um caso específico, uma partícula simples, cujo Hamiltoniano de uma partícula é dado por

$$\hat{h}_0(\mathbf{r}) = -\frac{1}{2}\nabla^2 + V_{ext}(\mathbf{r}), \quad (1.20)$$

sendo assim, como  $\hat{h}_0$  descreve um elétron que não interage com os demais, podemos reescrever o Hamiltoniano para muitos elétrons como:

$$\hat{H}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_i \hat{h}_0(\mathbf{r}_i) + \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (1.21)$$

### 1.1.3 Aproximação de elétrons independentes

Uma tentativa de simplificar o problema eletrônico é desprezar a interação Coulombiana, como se um elétron não interagisse com os outros. Assim, temos a aproximação de elétrons independentes, na qual a expressão simplificada para a

equação de Schrödinger é dada por

$$\left( \sum_i \hat{h}_0(\mathbf{r}_i) \right) \psi = E\psi. \quad (1.22)$$

Esta aproximação faz com que possamos reescrever a função de onda total como o produto das soluções para cada elétron, uma vez que o Hamiltoniano é separável. Conhecida como aproximação de Hartree, podemos escrever a solução como

$$\psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \phi_1(\mathbf{r}_1) \dots \phi_N(\mathbf{r}_N), \quad (1.23)$$

onde temos que  $\hat{h}_0$  é igual para todos os elétrons, então

$$\hat{h}_0\phi_i = \epsilon_i\phi_i. \quad (1.24)$$

No caso de apenas um elétron, podemos encontrar autovalores  $\epsilon_i$  e, ao substituir a equação 1.23 em 1.22, notamos que  $\hat{h}_0(\mathbf{r}_1)$  atuará apenas em  $\phi_1(\mathbf{r}_1)$ , e assim sucessivamente. Assim, temos ainda que a energia da configuração de mais baixa energia é dada pela soma dos auto valores de cada elétron (respeitando o princípio da exclusão de Pauli).

Uma consequência da aproximação de elétrons independentes é que a equação 1.23 não é antissimétrica, ou seja, ao se permutar as coordenadas de dois elétrons, o sinal da função de onda deveria se alterar, uma vez que os elétrons são férmions.

Por exemplo, para um par de elétrons com soluções distintas no caso não-interagente, a solução mínima que satisfaz esta condição é

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} [\phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2) - \phi_1(\mathbf{r}_2)\phi_2(\mathbf{r}_1)]. \quad (1.25)$$

Uma segunda maneira de reescrever a mesma função de onda é o conhecido determinante de Slater, utilizando-se o determinante de uma matriz, com a forma mais compacta para apresentar a função de onda antissimétrica<sup>1</sup>

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_1(\mathbf{r}_1) & \phi_1(\mathbf{r}_2) \\ \phi_2(\mathbf{r}_1) & \phi_2(\mathbf{r}_2) \end{vmatrix}. \quad (1.26)$$

Para sistemas com  $N$  elétrons, o fator  $1/\sqrt{2}$  é substituído por  $1/\sqrt{N!}$  para

<sup>1</sup>Para o caso onde consideramos *spin*, temos que  $\Psi(r_1, r_2) = \phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)\chi^s$ , onde o produto das funções de onda é simétrico e o termo de singlete  $\chi^s$  é responsável pela antissimetria.

normalizar corretamente a função de onda. Ao utilizar a equação 1.26, chegamos a um resultado importante para o andamento do problema. A equação 1.6 diz que a densidade eletrônica é dada pelo quadrado do módulo da função de onda. Porém, com os elétrons independentes, temos que

$$|\psi|^2 = \sum_{i,j} \phi_i^* \phi_j. \quad (1.27)$$

Ao mesmo tempo sabemos também que

$$\int \phi_i^*(\mathbf{r}) \phi_j(\mathbf{r}) d\mathbf{r} = \delta_{ij}. \quad (1.28)$$

Portanto, utilizando as equações 1.6, 1.27 e 1.28, chegamos à seguinte expressão para a densidade eletrônica

$$n(\mathbf{r}) = \sum_i^{occup} |\phi_i(\mathbf{r})|^2, \quad (1.29)$$

válido tanto para a aproximação de Hartree quanto para um determinante de Slater.

Este resultado nos mostra que a densidade eletrônica é obtida como a soma da probabilidade de se encontrar um elétron em uma dada posição  $\mathbf{r}$  para cada estado ocupado  $i$ , o que é intuitivo, uma vez que a probabilidade de encontrar o  $i$ -ésimo elétron no ponto  $\mathbf{r}$  é exatamente  $|\phi_i(\mathbf{r})|^2$ .

Ignorar a interação Coulombiana é uma ação drástica, uma vez que ela é da mesma ordem de grandeza dos outros potenciais. Assim, apesar da simplificação do problema, é necessário adicionar uma contribuição desta natureza a fim de obter resultados mais precisos. Uma solução é relembrar a eletrostática, uma vez que a densidade eletrônica (equação 1.29) gera um potencial  $\varphi(\mathbf{r})$  colocado por meio da equação de Poisson:  $\nabla^2 \varphi(\mathbf{r}) = 4\pi n(\mathbf{r})$ .

No sistema de unidades de Hartree, o potencial eletrostático é dado por  $V_H(\mathbf{r}) = -\varphi(\mathbf{r})$ , chamado de potencial de Hartree, e que satisfaz a equação de Poisson

$$\nabla^2 V_H(\mathbf{r}) = -4\pi n(\mathbf{r}). \quad (1.30)$$

Este problema tem como solução geral

$$V_H(\mathbf{r}) = \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \quad (1.31)$$

Cada elétron, agora, "sente" o potencial de Hartree, fazendo com que a equação para os orbitais  $\phi_i$  ganhe um novo termo

$$\left[ -\frac{\nabla^2}{2} + V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}). \quad (1.32)$$

Assume-se que o potencial  $V_H(\mathbf{r})$  substitui a interação coulombiana por uma média do potencial devido a tal interação; por isto, dá-se a esta abordagem o nome de aproximação de campo médio. Pela primeira vez, mudamos a maneira de resolver o sistema, onde, no início, tínhamos uma única equação de  $3N$  variáveis e, agora, temos  $N$  equações tridimensionais, algo que facilita a resolução do problema.

Vamos nos atentar agora para outras duas equações descritas nesta seção

$$n(\mathbf{r}) = \sum_i^{occup} |\phi_i(\mathbf{r})|^2. \quad (1.33)$$

$$\nabla^2 V_H(\mathbf{r}) = -4\pi n(\mathbf{r}), \quad (1.34)$$

Qual a relação entre as equações 1.32 a 1.34? Observa-se uma dependência entre elas, ou seja, ao se obter a função de onda  $\phi_i(\mathbf{r})$  na equação 1.32, podemos encontrar a densidade eletrônica por meio da equação 1.33 e, com seu resultado, obter  $V_H(\mathbf{r})$  com a equação 1.34. Porém, podemos também, a partir de  $V_H(\mathbf{r})$ , obter a função de onda. Nota-se que as equações são interdependentes, formando um ciclo que deve ser resolvido iterativamente, no que é conhecido como método auto-consistente.

Apesar de estarmos em um ótimo caminho para a equação final do nosso problema, ainda falta a adição de mais dois potenciais: o potencial de troca e o potencial de correlação. Estas correções serão discutidas em breve, nas seções seguintes.

### 1.1.4 Equações de Hartree-Fock

Até agora, em nosso estudo, foi feita uma série de aproximações que resultou em uma Equação de Schrödinger para muitos corpos com dependência apenas eletrônica. No entanto, a interação elétron-elétron foi tratada de maneira puramente clássica. Vamos, então, considerar o valor esperado da energia  $E_0$  para o

estado fundamental. Sabe-se, então, que

$$E_0 = \langle \Psi | \hat{H} | \Psi \rangle. \quad (1.35)$$

Adota-se, então, uma função de onda tentativa  $|\Psi\rangle$  na forma de um determinante de Slater. Ao aplicar o princípio variacional para minimizar a energia fundamental  $E_0$ , e impondo a ortonormalidade das funções de onda, obtêm-se os orbitais  $\phi_i$  que compõem o determinante. Dessa forma, resultam as chamadas equações de Hartree-Fock [102].

$$\left[ -\frac{\nabla^2}{2} + V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) \right] \phi_i(\mathbf{r}) + \hat{V}_X \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}), \quad (1.36)$$

$$n(\mathbf{r}) = \sum_i |\phi_i(\mathbf{r})|^2, \quad (1.37)$$

$$\nabla^2 V_H(\mathbf{r}) = -4\pi n(\mathbf{r}). \quad (1.38)$$

A diferença entre as equações de campo médio (equações 1.32 a 1.34) e as equações de Hartree-Fock (1.36 a 1.38) é o potencial de troca,  $\hat{V}_X$ , que surge da natureza antissimétrica da função de onda e ao considerarmos a interação Coulombiana como um operador em  $\mathbf{r}_i$  e  $\mathbf{r}_j$ . Sua ação sobre um orbital  $\phi_i(\mathbf{r})$  é não local e dada por

$$\hat{V}_X \phi_i(\mathbf{r}) = - \sum_j^{\text{ocup}} \int d\mathbf{r}' \frac{\phi_j^*(\mathbf{r}') \phi_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \phi_j(\mathbf{r}). \quad (1.39)$$

Dessa forma, o modelo avança para além da aproximação de campo médio puramente clássica, reintroduzindo o efeito quântico da troca. Porém, ao se utilizar um potencial não local, os cálculos ficam novamente mais complexos, dada a natureza da integral na equação 1.39.

A principal dificuldade do método de Hartree-Fock está relacionada ao custo computacional envolvido na sua formulação. Ao expandir os orbitais moleculares em uma base finita, a resolução do problema requer a diagonalização de matrizes de dimensão  $N \times N$ , cujo custo típico é da ordem de  $\mathcal{O}(N^3)$ . Ainda mais onerosa, porém, é a construção da matriz, que depende do cálculo dos integrais de dois elétrons, cujo número cresce como  $\mathcal{O}(N^4)$ . Dessa forma, tanto o tamanho das matrizes quanto a ordem das operações necessárias impõem um limite prático à aplicação do método em sistemas de grande porte.

Após todas as correções e refinamentos da equação de Schrödinger para muitos corpos, falta apenas um único potencial a ser inserido: o potencial de correlação.

Mas qual seria a natureza deste potencial? Apesar da dificuldade em compreendê-lo conceitualmente, sua natureza pode ser explicada de forma simples. Pense em um sistema com dois elétrons em uma caixa. Devido à interação coulombiana, é menos provável encontrar o segundo elétron na vizinhança do primeiro. Ou seja, este termo descreve a forma como a posição de um elétron afeta a probabilidade de se encontrar outro elétron nas proximidades, para além do Princípio de Exclusão de Pauli. A adição deste potencial torna-se, então, intuitiva. Na abordagem de Kohn-Sham, o potencial de troca não local de Hartree-Fock é substituído, juntamente com o de correlação, por um único potencial local de troca-correlação.

Com esta inserção, estamos cada vez mais próximos da equação final estudada pela chamada DFT. Começamos com a aproximação de Born-Oppenheimer, passando, depois, a considerar os elétrons independentes, transformando uma equação de  $3N$  variáveis em  $N$  equações tridimensionais independentes, culminando na adição de potenciais para corrigir a interação coulombiana que havia sido desprezada. Daqui em diante, iremos refinar nosso estudo e facilitar a resolução do nosso problema. A forma exata dos potenciais  $V_x$  e  $V_c$  não é explicitada, pois, em geral, ela é desconhecida e precisa ser aproximada.

### 1.1.5 Equações de Kohn-Sham

#### Energia total do estado fundamental

Nas seções anteriores, apresentamos alguns resultados que serão de suma importância neste momento. Notamos na equação 1.35 que temos:

$$E = \langle \Psi | \hat{H} | \Psi \rangle. \quad (1.40)$$

Temos também que, para um dado sistema (ou seja, um dado Hamiltoniano), a energia total se altera conforme a função de onda  $\Psi$  muda. Usando uma nomenclatura matemática,  $E$  é um funcional de  $\Psi$ , sendo indicado por:

$$E = \mathcal{F}[\Psi]. \quad (1.41)$$

O cerne da DFT é que a energia,  $E$ , é um funcional único da densidade eletrônica do estado fundamental e que a energia do estado fundamental corresponde ao valor do funcional para a densidade do estado fundamental,  $n_0(\mathbf{r})$

$$E_0 = \mathcal{F}[n_0]. \quad (1.42)$$

Como consequência, o problema de se determinar a energia do estado fundamental é reformulado: em vez de se buscar uma função de onda de  $3N$  variáveis, busca-se uma densidade de apenas 3 variáveis, desde que o funcional seja conhecido.

### Teorema de Hohenberg-Kohn

A afirmação de que a energia é um funcional da densidade eletrônica é conhecida como Teorema de Hohenberg-Kohn [23]. O teorema está baseado em três premissas, nas quais nos basearemos e que justificaremos a seguir:

1. A densidade eletrônica  $n(\mathbf{r})$  determina unicamente o potencial externo  $V_{ext}(\mathbf{r})$  (a menos de uma constante).
2. Em qualquer estado quântico, o potencial  $V_{ext}(\mathbf{r})$  gera uma única função de onda  $\Psi$ .
3. Em qualquer estado quântico, a energia total é um funcional da função de onda  $\Psi$ .

Combinando as afirmações, obtém-se a seguinte relação:  $n \rightarrow V_{ext} \rightarrow \Psi \rightarrow E$ , da qual se conclui que a densidade determina unicamente a energia, ou seja,  $E = \mathcal{F}[n]$ .

A segunda e a terceira premissas são claras e intuitivas de se provar. Ao mudar as posições dos núcleos (e, portanto,  $V_{ext}$ ), fica evidente que se altera a função de onda. Ao mesmo tempo, a equação 1.40 justifica a terceira afirmação.

A primeira afirmação do teorema também é simples de se mostrar. Vamos nos atentar à equação 1.19 e definir os seguintes operadores

$$\hat{T} = -\sum_i \frac{\nabla_i^2}{2}, \quad \hat{V}_{ee} = \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (1.43)$$

Com a equação 1.40, temos que

$$E = \langle \Psi | \sum_i V_{ext}(\mathbf{r}_i) | \Psi \rangle + \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle. \quad (1.44)$$

Usando a relação entre a função de onda e a densidade eletrônica, chegamos a uma equação final para a energia do estado fundamental

$$E = \int d\mathbf{r} n(\mathbf{r}) V_{ext}(\mathbf{r}) + \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle. \quad (1.45)$$

Vamos supor agora que haja um novo potencial  $V'_{ext} \neq V_{ext}$  que gere a mesma densidade  $n$ . Teremos então uma nova função de onda  $\Psi'$  com  $\hat{H}'$  e  $E'$  sendo o novo Hamiltoniano e energia do estado fundamental, respectivamente. Partindo da premissa de que  $\Psi$  não é o estado fundamental de  $\hat{H}'$ , o princípio variacional nos diz que

$$\langle \Psi | \hat{H}' | \Psi \rangle > E'. \quad (1.46)$$

Expandindo  $\langle \Psi | \hat{H}' | \Psi \rangle$  e usando a equação 1.45, chegamos a

$$E - E' > \int d\mathbf{r} n(\mathbf{r}) [V_{ext}(\mathbf{r}) - V'_{ext}(\mathbf{r})]. \quad (1.47)$$

Se o mesmo raciocínio fosse aplicado partindo de  $E'$ , teríamos um resultado semelhante

$$E' - E > \int d\mathbf{r} n(\mathbf{r}) [V'_{ext}(\mathbf{r}) - V_{ext}(\mathbf{r})]. \quad (1.48)$$

Ao somar as inequações 1.47 e 1.48, chegamos a um absurdo. Logo, podemos observar que a densidade eletrônica gera um único potencial externo  $V_{ext}$ .

Uma vez que podemos escrever a energia como

$$E[n] = \langle \Psi[n] | \hat{T} + \hat{V}_{ee} + \hat{V}_{ext} | \Psi[n] \rangle. \quad (1.49)$$

Se  $n$  é uma densidade diferente da densidade do estado fundamental  $n_0$  no potencial  $v(r)$ , então as funções de onda  $\Psi$  que produzem esse  $n$  são diferentes da função de onda do estado fundamental  $\Psi_0$ . De acordo com o princípio variacional, o mínimo obtido de  $E[n]$  é maior (ou igual) à energia do estado fundamental  $E_0 = E[n_0]$ . Assim, o funcional  $E[n]$  é minimizado pela densidade do estado fundamental  $n_0$ , e seu valor no mínimo é  $E_0$ . Logo

$$E[n] = \int d\mathbf{r} n(\mathbf{r}) V_{ext} + \langle \Psi[n] | \hat{T} + \hat{V}_{ee} | \Psi[n] \rangle, \quad (1.50)$$

onde a energia  $F[n] = \langle \Psi[n] | \hat{T} + \hat{V}_{ee} | \Psi[n] \rangle$  independe do potencial externo, o que nos permite definir a função de onda do estado fundamental  $\Psi_0$ , que fornece o mínimo de  $F[n]$  e reproduz  $n_0$ . Se o estado fundamental for não degenerado, esse duplo requisito determina unicamente  $\Psi_0$  em termos de  $n_0(r)$ , sem necessidade de especificar  $V_{ext}$  explicitamente [103].

Observando a equação 1.50, apenas o primeiro termo da soma apresenta uma dependência explícita da densidade  $n$ , enquanto os termos de energia cinética e potencial têm uma dependência implícita. A ideia de Kohn e Sham, então, foi

mapear um problema interagente em um problema não interagente com a mesma densidade, mantendo a "ignorância" em um único termo, então:

$$E[n] = \overbrace{T_s[n] + \int d\mathbf{r} n(\mathbf{r}) V_{ext}(\mathbf{r}) + E_H[n]}^{\text{Termos conhecidos ou de fácil aproximação}} + \underbrace{E_{xc}[n]}_{\text{Energia de Troca-Correlação}}. \quad (1.51)$$

O termo  $E_{xc}$  representa o que chamamos de energia de troca e correlação. Desta forma, a equação 1.51 contempla tudo o que desconhecíamos sobre o funcional da densidade, a partir da soma de quantidades conhecidas, desde a aproximação de elétrons independentes até a contribuição da energia de troca e correlação. Se soubéssemos a forma exata de  $E_{xc}[n]$ , poderíamos calcular a energia total do estado fundamental. Aqui, falta apenas determinar o valor de  $n$ .

No estado fundamental,  $n_0$  é o valor da densidade que minimiza  $E$ . Desta maneira, utilizando-se o princípio variacional

$$\left. \frac{\delta E[n]}{\delta n} \right|_{n_0} = 0. \quad (1.52)$$

Aplicando o princípio variacional ao funcional de energia, sob a restrição de que os orbitais sejam ortonormais e que a integral da densidade total seja igual ao número de elétrons, chega-se às equações que regem o método de Kohn-Sham e nas quais nos apoiaremos neste trabalho

$$\left[ -\frac{\nabla^2}{2} + V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}), \quad (1.53)$$

onde  $V_{xc}$  é dado pela derivada funcional de  $E_{xc}[n]$

$$V_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})}, \quad (1.54)$$

e é chamado de potencial de troca e correlação. As equações do tipo da equação 1.53 são conhecidas como equações de Kohn-Sham [24]. Estas equações constituem uma ferramenta de grande poder para calcular várias propriedades dos materiais a partir de primeiros princípios da mecânica quântica. Com o potencial de troca e correlação, podemos definir um único potencial efetivo, denominado potencial de

Kohn-Sham

$$V_{KS}(\mathbf{r}) = V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}), \quad (1.55)$$

de modo que

$$\left[ -\frac{\nabla^2}{2} + V_{KS}(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}). \quad (1.56)$$

A DFT é, em princípio, uma teoria exata. Toda a complexidade do problema de muitos corpos é encapsulada no termo de troca-correlação,  $E_{xc}[n]$ . Conhecendo-se a forma exata deste funcional, seria possível descrever o sistema de forma completa. Porém, não temos um modo de precisar tal potencial. Segundo o próprio Walter Kohn, a praticidade da DFT depende de aproximações para  $E_{xc}[n]$ , pois, uma vez que uma boa aproximação é encontrada, o problema tem uma solução simples e acurada [100].

### 1.1.6 Os Potenciais de Troca e Correlação

Como já dito, é preciso encontrar a forma para  $V_{xc}$ . Para tal, nesta seção, apresentamos quatro aproximações para o potencial: LDA (Aproximação de Densidade Local) [104, 105], GGA (Aproximação de Gradientes Generalizados) [65], meta-GGA [34] e funcionais com correção de van der Waals [31, 30].

#### LDA - Aproximação de Densidade Local

A Aproximação de Densidade Local (LDA) postula que a energia de troca-correlação em um ponto  $\mathbf{r}$  pode ser aproximada pela de um gás de elétrons homogêneo com a mesma densidade  $n(\mathbf{r})$ . A equação 1.57 nos apresenta a energia de troca e correlação de acordo com a LDA

$$E_{xc}^{LDA}[n] = \int d^3r n(\mathbf{r}) \epsilon_{xc}^{\text{hom}}(n(\mathbf{r})), \quad (1.57)$$

onde  $\epsilon_{xc}^{\text{hom}} = \epsilon_x^{\text{hom}} + \epsilon_c^{\text{hom}}$  é a energia de troca e correlação por partícula do gás homogêneo.

O termo de correlação é mais complexo. Porém, Ceperley e Alder [104], por meio de simulações de Monte Carlo Quântico, obtiveram este termo com alta precisão para vários valores de densidade. Assim,  $\epsilon_c$  é parametrizado em função do raio de Wigner-Seitz,  $r_s$ , obtendo-se, desta forma,  $\epsilon_c$  para qualquer valor atribuído

à densidade. A relação entre  $n$  e  $r_s$  é dada por

$$n(\mathbf{r}) = \frac{3}{4\pi r_s^3}. \quad (1.58)$$

Chegamos então às equações que definem os termos  $\epsilon_x$  (equação 1.59) e  $\epsilon_c$  (equação 1.60), de acordo com a parametrização de Perdew e Zunger [105]

$$\epsilon_x = -\frac{0,4582}{r_s/a_0}, \quad (1.59)$$

$$\epsilon_c = \begin{cases} -\frac{0,1423}{(1+1,9529\sqrt{\frac{r_s}{a_0}}+0,3334\frac{r_s}{a_0})'}, & r_s \geq 1 \\ -0,0480 + 0,0311 \ln \frac{r_s}{a_0} - 0,0116\frac{r_s}{a_0} + 0,0020r_s \ln \frac{r_s}{a_0}, & r_s < 1 \end{cases}, \quad (1.60)$$

como  $\epsilon_{xc} = \epsilon_x + \epsilon_c$ , temos que a equação 1.61 nos apresenta a forma do potencial de troca e correlação de acordo com a LDA

$$V_{xc}[n] \equiv \frac{d}{dn(\mathbf{r})} \left( n(\mathbf{r})\epsilon_{xc}[n(\mathbf{r})] \right). \quad (1.61)$$

Caso o sistema não possua uma densidade uniforme, como ocorre na maioria dos casos, a aproximação LDA não é capaz de descrever de forma correta a energia de troca e correlação, apesar de ser capaz de descrever propriedades do estado fundamental de sólidos, em especial sistemas metálicos e covalentes. Por isso, partiremos para uma aproximação que descreve melhor sistemas em que a densidade não se comporta uniformemente, como a maioria dos sistemas reais.

### GGA - Aproximação de Gradientes Generalizados

De modo a refinar o potencial de troca e correlação, apresentamos a GGA (Aproximação de Gradientes Generalizados), em que, além da densidade eletrônica local do sistema, consideramos também o gradiente da densidade. Assim, a energia em torno de um ponto  $\mathbf{r}$  é dada não somente pela densidade neste ponto, mas também pelo seu gradiente. Utiliza-se a seguinte forma funcional para  $E_{xc}$  [65]

$$E_{xc}^{GGA}[n] = \int f(n(\mathbf{r}), |\nabla n(\mathbf{r})|) d^3r. \quad (1.62)$$

Os funcionais GGA possuem dois tipos: os semi-empíricos, nos quais os parâ-

metros advêm de experimentos a partir de materiais ou moléculas; e os não empíricos, obtidos por meio de vínculos teóricos. Dentro deste trabalho, utilizaremos o funcional do segundo tipo proposto por Perdew, Burke e Ernzerhof (PBE) [65]. Para este funcional, a energia de troca é dada por  $E_x^{\text{PBE}}[n] = \int n(\mathbf{r})\epsilon_x^{\text{unif}}(n)F_x(s)d^3r$ , onde  $F_x(s)$  é o fator de aprimoramento

$$F_x(s) = 1 + \kappa - \frac{\kappa}{1 + \frac{\mu}{\kappa}s^2}, \quad (1.63)$$

onde  $\kappa = 0,804$ ,  $\mu = 0,21951$  e  $s = |\nabla n|/(2k_F n)$  é o gradiente reduzido da densidade, com  $k_F$  sendo o vetor de onda de Fermi.

### Funcionais meta-GGA

Apesar de melhorarem a descrição eletrônica, é possível aprimorar os funcionais GGA [106], levando-se em conta também a densidade de energia cinética local. Nestes casos, os funcionais recebem o nome de funcionais meta-GGA. A energia de troca-correlação para um funcional meta-GGA pode ser escrita como [107]

$$E_{xc}[n, \tau] = \int f(n(\mathbf{r}), \nabla n(\mathbf{r}), \tau(\mathbf{r}))d^3r, \quad (1.64)$$

onde a densidade de energia cinética  $\tau(\mathbf{r})$  é dada pelos orbitais de Kohn-Sham ocupados

$$\tau(\mathbf{r}) = \frac{1}{2} \sum_i^{\text{ocup}} |\nabla \phi_i(\mathbf{r})|^2. \quad (1.65)$$

Neste trabalho, será dado enfoque ao funcional SCAN [34], que recentemente foi utilizado para o estudo de sistemas aquosos [12, 79]. O funcional possui dois pontos-chave para sua maior acurácia: o primeiro diz respeito à melhor descrição das ligações atômicas, enquanto o segundo trata da não uniformidade da densidade. Deste modo, definimos o parâmetro  $\alpha$  que mede o quão localizado é o elétron

$$\alpha = \frac{\tau(\mathbf{r}) - \tau^W(\mathbf{r})}{\tau^{\text{unif}}(\mathbf{r})}, \quad (1.66)$$

onde  $\tau^W(\mathbf{r}) = \frac{|\nabla n(\mathbf{r})|^2}{8n(\mathbf{r})}$  e  $\tau^{\text{unif}}(\mathbf{r}) = \frac{3}{10}(3\pi^2)^{2/3}n^{5/3}(\mathbf{r})$ . No caso  $\alpha \approx 0$ , os elétrons estão fortemente localizados (ligação covalente); para  $\alpha \approx 1$ , tem-se o gás de elétrons interagentes; e para  $\alpha \gg 1$ , está-se em uma região com baixa densidade eletrônica.

Analogamente, para a não uniformidade da densidade, apresentamos o parâ-

metro de gradiente reduzido  $s(\mathbf{r})$

$$s(\mathbf{r}) = \frac{|\nabla n(\mathbf{r})|}{2(3\pi^2)^{1/3}n^{4/3}(\mathbf{r})}. \quad (1.67)$$

Com isso, a energia de troca do funcional pode ser escrita como

$$E_x[n] = \int n(\mathbf{r})\varepsilon_x^{\text{unif}}(n(\mathbf{r}))F_x(s, \alpha)d^3r, \quad (1.68)$$

onde  $\varepsilon_x^{\text{unif}} = -(3/4\pi)(3\pi^2n)^{1/3}$  é a energia de troca por partícula e  $F_x(s, \alpha)$  é o fator de aprimoramento de troca.

O SCAN usa funções de interpolação para satisfazer diferentes restrições:

- Para  $s \rightarrow 0$  e  $\alpha \rightarrow 1$ , o funcional recupera o limite do gás de elétrons uniforme, base para funcionais como o PBE;
- Para  $\alpha \rightarrow 0$ , o funcional se aproxima do comportamento correto para elétrons fortemente localizados;
- Melhora a descrição de energias de ligação e *gaps* de energia.

Este funcional satisfaz um total de 17 restrições exatas conhecidas [34].

### Funcionais com correção de van der Waals

Para finalizar as discussões em torno dos funcionais de troca e correlação, será apresentada de forma breve a correção de van der Waals. A primeira publicação de um funcional de densidade que inclui nativamente as interações de van der Waals foi feita por Dion e colaboradores, em 2004 [108].

Inicialmente, podemos pensar em um funcional onde o termo de co correlação e troca é escrito da seguinte forma

$$E_{xc}[n] = E_x^{\text{GGA}} + E_c^{\text{LDA}}. \quad (1.69)$$

Deste modo, podemos adicionar um novo termo não local,  $E_c^{\text{nl}}$ , relacionado à correlação, logo

$$E_{xc}[n] = E_x^{\text{GGA}} + E_c^{\text{LDA}} + E_c^{\text{nl}}. \quad (1.70)$$

De forma geral, podemos assumir que este termo não local pode ser escrito como

$$E_c^{\text{nl}}[n] = \frac{1}{2} \iint n(\mathbf{r}_1)\phi(\mathbf{r}_1, \mathbf{r}_2)n(\mathbf{r}_2)d\mathbf{r}_1d\mathbf{r}_2, \quad (1.71)$$

onde o núcleo  $\phi(\mathbf{r}_1, \mathbf{r}_2)$  é um funcional da densidade que garante o comportamento assintótico correto da interação de dispersão ao decair como  $1/|\mathbf{r}_1 - \mathbf{r}_2|^6$  para longas distâncias  $|\mathbf{r}_1 - \mathbf{r}_2|$ .

Os termos de longo alcance (como interações de van der Waals) são adicionados como um termo de dispersão na energia total. Uma abordagem mais simples para o tratamento destas interações foi publicada por Grimme e atualmente é conhecida como DFT-D2 [32]. Posteriormente, Grimme [33] refinou sua abordagem para um modelo de três corpos (DFT-D3), onde a energia de dispersão é dada por um termo de dois corpos somado a um termo de três corpos. O termo de dois corpos é dado por

$$E^{(2)} = - \sum_{A < B} \sum_{n=6,8,10,\dots} s_n \frac{C_n^{AB}}{r_{AB}^n} f_{d,n}(r_{AB}), \quad (1.72)$$

onde o primeiro somatório é sobre todos os pares atômicos  $AB$  no sistema,  $C_n^{AB}$  denota o coeficiente de dispersão isotrópico médio de ordem  $n$  para o par, e  $r_{AB}$  é sua distância internuclear. O termo  $f_{d,n}(r_{AB})$  é uma função de amortecimento dada por

$$f_{d,n}(r_{AB}) = \frac{1}{1 + 6 \left( \frac{r_{AB}}{s_{r,n} R_0^{AB}} \right)^{-\alpha_n}}, \quad (1.73)$$

onde  $s_{r,n}$  é um fator de escala dependente da ordem  $n$  para os raios de corte  $R_0^{AB}$ .

Enquanto isso, o termo de três corpos é dado por

$$E^{(3)} \approx \sum_{A,B,C} E_{ABC}^{(3)}, \quad (1.74)$$

onde

$$E_{ABC}^{(3)} = \frac{C_9^{ABC} (1 + 3 \cos \theta_A \cos \theta_B \cos \theta_C)}{(r_{AB} r_{BC} r_{CA})^3}, \quad (1.75)$$

$$C_9^{ABC} \approx -\sqrt{C_6^{AB} C_6^{AC} C_6^{BC}}, \quad (1.76)$$

$$C_6^{KL} = \frac{3}{\pi} \int_0^\infty d\omega \alpha_K(i\omega) \alpha_L(i\omega) \quad (1.77)$$

em que  $\theta_i$  representa cada ângulo interno do triângulo formado pelos átomos  $A$ ,  $B$  e  $C$ ;  $r_{ij}$  é a distância entre dois átomos; e  $\alpha(i\omega)$  é a polarizabilidade de dipolo média com frequência imaginária  $\omega$ .

Outra abordagem para a inclusão dos termos não locais é a feita por Berland e Hyldgaard [31]. O funcional *vdW-cx* (*consistent exchange*) é uma variante da

família vdW-DF que busca consistência entre a descrição plasmônica usada no termo de correlação não local e o parceiro de troca GGA. A ideia central é alinhar o *inner functional*, que parametriza a resposta plasmônica usada em  $E_c^{\text{nl}}$ , e o *outer functional*, que corresponde à troca GGA efetivamente somada ao funcional total, ao menos no regime de gradientes reduzidos pequenos a moderados. Quando esse alinhamento ocorre, o termo de correção  $\delta E_{xc}$  do acoplamento adiabático tende a zero e a conservação do buraco de troca e correlação é respeitada [31].

A energia de troca e correlação pode ser escrita a partir da fórmula de conexão adiabática e de uma expansão de segunda ordem em  $S = 1 - \epsilon^{-1}$

$$E_{xc}^{\text{vdW-DF}} = \int_0^\infty \frac{du}{2\pi} \text{Tr}[\ln(\nabla\epsilon \cdot \nabla G)] - E_{\text{self}} - \delta E_{xc}, \quad (1.78)$$

$$\delta E_{xc} = \int_0^\infty \frac{du}{2\pi} \text{Tr} \epsilon - E_{\text{self}} - (E_x^{\text{GGA}} + E_c^{\text{LDA}}). \quad (1.79)$$

A correlação não local assume a forma de integral dupla densidade-densidade

$$E_c^{\text{nl}}[n] = \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' n(\mathbf{r}) \Phi(q_0(\mathbf{r}), q_0(\mathbf{r}'), R) n(\mathbf{r}'), \quad R = |\mathbf{r} - \mathbf{r}'|, \quad (1.80)$$

onde o comprimento local  $q_0(\mathbf{r})$  vem do mesmo modelo plasmônico que define o *inner functional*. Para funcionais desta categoria, tem-se

$$d(\mathbf{r}) = q_0(\mathbf{r}) R, \quad q_0(\mathbf{r}) = -4\pi(\epsilon_x^{\text{LV}} + \epsilon_c^{\text{LDA}}). \quad (1.81)$$

Assim, regiões de grande gradiente reduzido  $s$  fazem  $q_0$  crescer como  $\sim s^2$ , desligando suavemente o termo  $E_c^{\text{nl}}$ .

Define-se o gradiente reduzido

$$s = \frac{|\nabla n|}{2k_F n}, \quad k_F = (3\pi^2 n)^{1/3}. \quad (1.82)$$

O vdW-cx utiliza um parceiro de troca que segue a expansão de Langreth-Vosko (LV) [108] no regime  $s \lesssim 2.5$ , coincidindo com o *inner functional*, e faz uma transição suave para a PW86r [109, 110] quando  $s$  cresce. Isso é implementado via

o fator de realce  $F_x(s)$

$$F_x^{\text{LV}}(s) = 1 + \mu_{\text{LV}} s^2, \quad \mu_{\text{LV}} = -\frac{Z_{ab}}{9}, \quad (1.83)$$

$$F_x^{\text{PW86r}}(s) = (1 + as^2 + bs^4 + cs^6)^{1/15}, \quad (1.84)$$

$$F_x^{\text{cx}}(s) = \frac{1}{1 + \alpha s^6} [1 + \mu_{\text{LV}} s^2] + \frac{\alpha s^6}{\beta + \alpha s^6} F_x^{\text{PW86r}}(s), \quad (1.85)$$

com os parâmetros ajustados por mínimos quadrados

$$\alpha = 0.02178, \quad \beta = 1.15. \quad (1.86)$$

Assim, a troca  $cx$  herda a física plasmônica no domínio que mais contribui a  $E_c^{\text{nl}}$ , garantindo consistência interna do esquema vdW-DF.

Logo, o funcional total é dado por

$$E_{xc}^{\text{cx}}[n] = E_x^{\text{GGA}[F_x^{\text{cx}}]}[n] + E_c^{\text{LDA}}[n] + E_c^{\text{nl}}[n]. \quad (1.87)$$

O ganho conceitual é que, como  $F_x^{\text{cx}}$  reproduz o inner-LV para  $s \lesssim 2.5$ , tem-se  $\delta E_{xc} \approx 0$  nesse regime, garantindo maior robustez física. Em testes, o  $vdW\text{-DF}\text{-}cx$  mostrou melhor desempenho para constantes de rede de sólidos e separações de camadas, sem sacrificar a descrição de interações de van der Waals em moléculas e superfícies [108, 31].

Enquanto o  $vdW\text{-}cx$  foca na consistência física entre troca e correlação plasmônica, o  $optB88\text{-}vdW$  adota a mesma correlação não local  $E_c^{\text{nl}}$  da  $vdW\text{-}cx$ , mas combina com uma versão otimizada da troca B88 [111]. No  $optB88\text{-}vdW$ , os parâmetros da troca são ajustados empiricamente para reproduzir com precisão energias de ligação de dímeros moleculares [30].

A forma geral do funcional  $optB88\text{-}vdW$  é

$$E_{xc}^{\text{optB88-vdW}}[n] = E_x^{\text{optB88}}[n] + E_c^{\text{LDA}}[n] + E_c^{\text{nl}}[n], \quad (1.88)$$

onde  $E_x^{\text{optB88}}$  possui o fator de realce  $F_x^{\text{optB88}}(s)$  reparametrizado para melhorar a descrição de sistemas dispersivos, onde  $F_x^{\text{optB88}}(s)$  segue a mesma ideia do funcional B88 [111]

$$F_x^{\text{B88}}(s) = 1 + \mu s^2 / (1 + \beta s \arcsin(cs)), \quad (1.89)$$

onde  $c = 2^{4/3}(3\pi^2)^{1/3}$ ,  $\mu \approx 0,2743$  e  $\beta = 9\mu(6/\pi)^{1/3}/(2c)$ . Para o funcional

optB88-vdW, os parâmetros  $\mu$  e  $\beta$  são otimizados para fortalecer a ligação. Deste modo, temos as seguintes relações:  $\mu = 0,22$  e  $\mu/\beta = 1,2$ .

### 1.1.7 O método de pseudo-potenciais

A justificativa para o método dos pseudo-potenciais parte da configuração eletrônica de um átomo: ao centro temos o núcleo, formado por prótons e nêutrons, circundado por uma nuvem de elétrons de caroço (ou *core*), mais fortemente ligados ao núcleo. Por fim, consideramos uma nuvem de elétrons de valência, com menor ligação ao núcleo, sendo esta a camada na qual os elétrons contribuem de fato para as ligações químicas. Um esquema do átomo descrito é apresentado na figura 1.1.

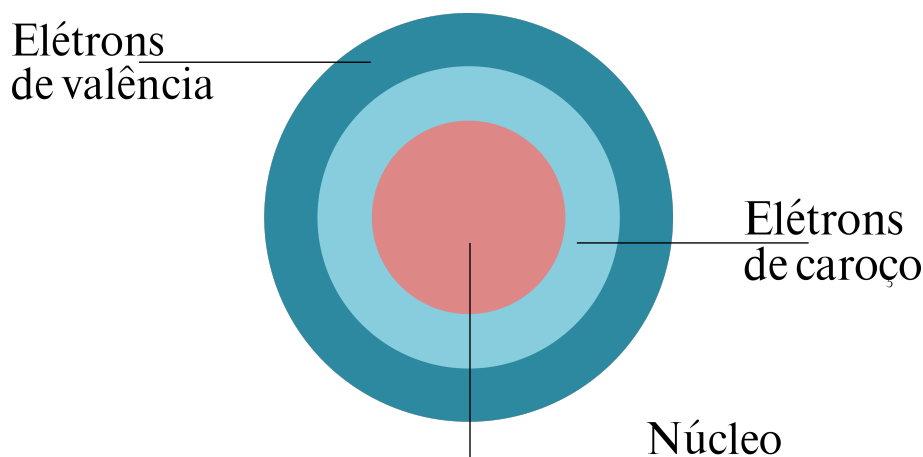


Figura 1.1: Esquema para um átomo hipotético.

Sabendo disso, o método emprega pseudo-funções de onda apenas para os elétrons de valência. Esta abordagem foi originalmente concebida para simplificar os cálculos de estrutura eletrônica, substituindo-se o forte potencial iônico (núcleo e elétrons de caroço) e os orbitais de valência por um pseudo-potencial mais suave e por pseudo-orbitais correspondentes. Neste ponto, os orbitais de Kohn-Sham necessitam ser expandidos em uma base; para tal, uma escolha comum é a expansão em uma base de ondas planas, que são soluções naturais para sistemas periódicos devido ao teorema de Bloch.

Os orbitais de valência verdadeiros devem oscilar rapidamente na região do caroço para manter a ortogonalidade com os orbitais de caroço. Isto torna o uso de uma base de ondas planas computacionalmente inviável, já que a representação dessas oscilações exigiria um número muito grande de componentes. Contudo,

ao se utilizarem pseudo-potenciais, os pseudo-orbitais de valência são suaves e podem ser representados eficientemente por uma base de ondas planas.

Com isso, justifica-se a retirada dos elétrons de caroço dos cálculos explícitos, substituindo-se o potencial Coulombiano por um potencial suave e finito na origem. Além disso, temos a substituição da função de onda oscilante por uma função de onda sem nós na região do caroço. A figura 1.2 ilustra esquematicamente estas substituições.

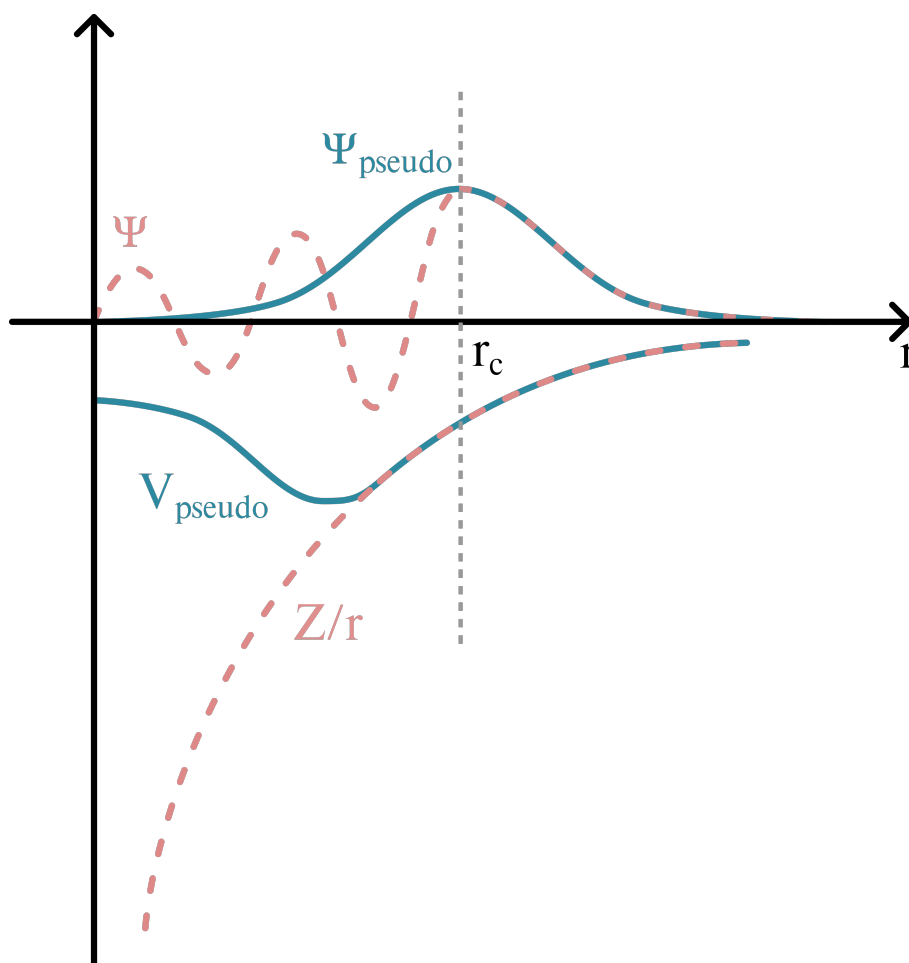


Figura 1.2: Esquema da substituição do potencial "all-electron" e da função de onda por um pseudo-potencial e uma pseudo-função de onda.

Além desta abordagem, podemos utilizar orbitais localizados centrados nos átomos. Os orbitais de Kohn-Sham são expandidos em funções de base numérica  $\varphi$ , chamados de orbitais pseudo-atômicos,

$$\varphi = \sum_i c_i \zeta_i, \quad (1.90)$$

que correspondem a diferentes soluções para o átomo isolado.

A quantidade de funções  $\zeta$  dá o nome da base atômica utilizada. No caso de apenas uma função  $\zeta$  teremos a base *single- $\zeta$* ; com duas funções temos a base *double- $\zeta$*  e assim consecutivamente. O número de funções  $\zeta$  pode trazer melhorias significantes no resultado, mas é importante ressaltar que, neste caso, a solução não é variacional na base. Por um outro lado, quanto maior o número de funções  $\zeta$ , maior o custo computacional, porém ainda menor que os orbitais de Kohn-Sham expandidos em funções ondas. Para todas as bases podemos acrescentar funções de polarização, representando melhor distorções da densidade de carga devido a campos elétricos internos do sólido.

Existem diversos tipos de pseudo-potenciais; dentre eles, podemos destacar os pseudo-potenciais de Phillips e Kleinman [112], de Kerker [113], além dos pseudo-potenciais de Troullier e Martins [114] e o método PAW (Projector-Augmented Wave) [115].

## 1.2 Dinâmica Molecular (MD)

Até o presente momento, discutimos apenas sobre o problema eletrônico não dependente do tempo. Deste modo, a dinâmica molecular — baseada na separação de Born-Oppenheimer, proposta inicialmente por Alder e Wainwright [95] e que se apropria das equações de Newton para analisar o movimento das moléculas — surge como um candidato para estudar a evolução temporal de um sistema atômico, no qual os átomos podem ser tratados como massas pontuais. Nesta abordagem, trataremos a interação entre os núcleos, fitando a parte eletrônica como o potencial de interação entre os núcleos.

Começamos com a análise de um sistema no qual a energia total  $E$ , o volume  $V$  e o número de partículas  $N$  são constantes (ensemble NVE). Para este caso, as equações de movimento para as partículas seguem a segunda lei de Newton

$$M_I \ddot{\mathbf{R}}_I = \mathbf{F}_I, \quad (1.91)$$

sendo a força dada por

$$\mathbf{F}_I = -\nabla_I U(\mathbf{R}), \quad (1.92)$$

onde  $M_I$  é a massa da  $I$ -ésima partícula,  $\mathbf{R}_I$  é sua coordenada,  $\mathbf{F}_I$  é a força que atua sobre esta partícula e  $U(\mathbf{R})$  é o potencial de interação do sistema.

É possível parametrizar o potencial de interação por meio de uma equação de

muitos corpos com diversos termos, tais como potenciais externos, interações de pares, tripletos, etc. [116]

$$U = \sum_I u_1(\mathbf{R}_I) + \sum_I \sum_{J>I} u_2(\mathbf{R}_I, \mathbf{R}_J) + \sum_I \sum_{J>I} \sum_{K>J>I} u_3(\mathbf{R}_I, \mathbf{R}_J, \mathbf{R}_K) + \dots \quad (1.93)$$

Por simplicidade, pode-se remover os termos de interação de quatro ou mais corpos. Além disto, poucos trabalhos utilizam interações de tripletos, uma vez que estas operações são altamente custosas e as interações de pares já descrevem o problema com boa precisão [116]. Sendo assim, um potencial com boa aproximação pode ser escrito apenas em função de um potencial externo e das interações de pares

$$U = \sum_I u_1(\mathbf{R}_I) + \sum_I \sum_{J>I} u_2(\mathbf{R}_I, \mathbf{R}_J). \quad (1.94)$$

Uma vez conhecidas as equações de movimento, a solução do problema pode ser encontrada computacionalmente. Um dos algoritmos para obter a evolução temporal do problema mais utilizados é o de Verlet [117]. Este método numérico permite que, conhecendo-se as variáveis dinâmicas das espécies em um instante de tempo  $t$ , seja possível determinar com precisão suficiente as mesmas variáveis em um instante posterior  $t + \Delta t$ . Dada a natureza iterativa da solução das equações diferenciais, é necessário que o passo de tempo  $\Delta t$  seja muito menor que o período da oscilação mais rápida das espécies em estudo [99].

Inicia-se a construção do algoritmo expandindo-se em série de Taylor a posição de um átomo para os instantes  $t - \Delta t$  e  $t + \Delta t$

$$\mathbf{R}_I(t - \Delta t) = \mathbf{R}_I(t) - \Delta t \mathbf{v}_I(t) + \frac{(\Delta t)^2}{2} \mathbf{a}_I(t) - \dots \quad (1.95)$$

$$\mathbf{R}_I(t + \Delta t) = \mathbf{R}_I(t) + \Delta t \mathbf{v}_I(t) + \frac{(\Delta t)^2}{2} \mathbf{a}_I(t) + \dots \quad (1.96)$$

Somando-se as duas equações e utilizando a relação  $\mathbf{a}_I(t) = \mathbf{F}_I/m_I$ , obtém-se

$$\mathbf{R}_I(t + \Delta t) = 2\mathbf{R}_I(t) - \mathbf{R}_I(t - \Delta t) + (\Delta t)^2 \frac{\mathbf{F}_I}{M_I}. \quad (1.97)$$

Do mesmo modo, subtraindo-se a primeira equação da segunda, obtém-se a velocidade

$$\mathbf{v}_I(t) = \frac{\mathbf{R}_I(t + \Delta t) - \mathbf{R}_I(t - \Delta t)}{2\Delta t}. \quad (1.98)$$

O algoritmo de Verlet possui um erro de truncamento de ordem  $(\Delta t)^4$  nas

posições e de ordem  $(\Delta t)^2$  nas velocidades. A figura 1.3 mostra o fluxograma de uma simulação de dinâmica molecular com o algoritmo de Verlet.

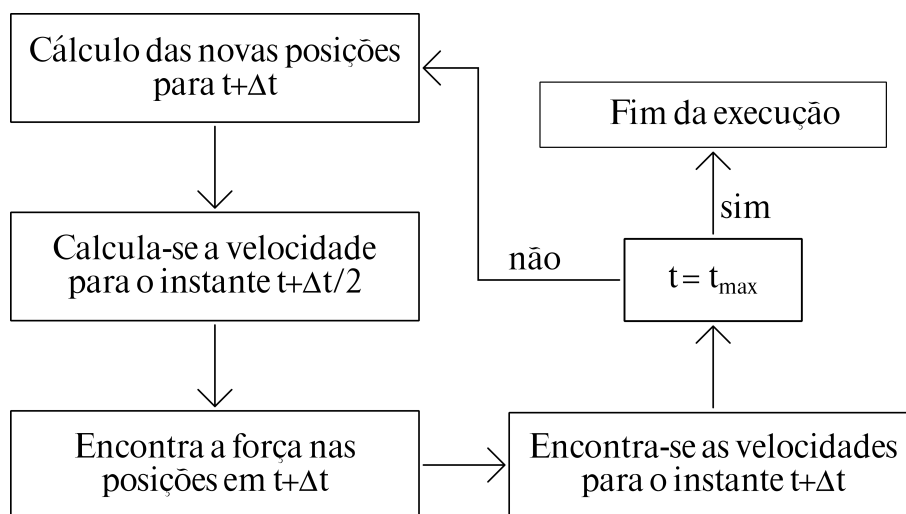


Figura 1.3: Fluxograma de uma simulação de dinâmica molecular utilizando o algoritmo de Verlet.

### 1.2.1 Campos de força

Apesar das equações de movimento clássicas serem conhecidas, ainda se faz necessário definir a forma do potencial  $U(\mathbf{R})$  para que as variáveis dinâmicas possam ser calculadas. Comumente chamados de campos de força, estes potenciais podem ser descritos tanto por métodos empíricos quanto por métodos *ab initio* e, mais recentemente, podem ser obtidos por meio de técnicas de aprendizado de máquina (*machine learning*) [118, 119, 120].

A energia potencial de um sistema inclui as interações covalentes (ligadas, ou *bonded*) e as interações intermoleculares (não ligadas, ou *nonbonded*). A decomposição de cada termo depende do campo de força escolhido, porém, de uma forma geral

$$U_{tot} = U_{bonded} + U_{nonbonded}, \quad (1.99)$$

onde tipicamente

$$U_{bonded} = U_{ligacao} + U_{angulo} + U_{diedro} \quad (1.100)$$

e

$$U_{nonbonded} = U_{eletrostatico} + U_{van\ der\ Waals}. \quad (1.101)$$

Geralmente, os modelos mais simples para estes campos de força são os pri-

meiros a serem testados [121]. Uma primeira abordagem para se tratar a energia de ligação é utilizar um modelo similar à interação de um sistema massa-mola, ou seja, um potencial harmônico

$$U(R_{IJ}) = \frac{k}{2}(R_{IJ} - R_0)^2, \quad (1.102)$$

onde  $R_{IJ}$  corresponde à distância entre duas partículas e  $R_0$  é a distância de equilíbrio.

Em sequência, pode-se utilizar a lei de Coulomb para iniciar o tratamento das interações não ligadas. Segundo esta lógica, a interação entre as partículas se dá pela energia potencial eletrostática em função de suas respectivas cargas parciais  $q_i$  e  $q_j$

$$U(R_{IJ}) = \frac{1}{4\pi\epsilon_0} \frac{q_I q_J}{R_{IJ}}. \quad (1.103)$$

Existem outros exemplos de modelos simples que podem ser utilizados para compor a interação entre os átomos e moléculas [122], tais como o potencial gravitacional

$$U(R_{IJ}) = -G_{grav} \frac{m_I m_J}{R_{IJ}}, \quad (1.104)$$

o potencial de van der Waals (termo atrativo)

$$U(R_{IJ}) = -a \left( \frac{1}{R_{IJ}} \right)^6, \quad (1.105)$$

o potencial de Morse

$$U(R_{IJ}) = D(1 - \exp(-a(R_{IJ} - R_0)))^2 \quad (1.106)$$

e o potencial de Lennard-Jones

$$U(R_{IJ}) = 4\epsilon \left[ \left( \frac{\sigma}{R_{IJ}} \right)^n - \left( \frac{\sigma}{R_{IJ}} \right)^m \right], \quad m < n. \quad (1.107)$$

Em particular, o potencial de Lennard-Jones, que foi empregado com sucesso para cálculos com argônio líquido [123] é um modelo parametrizado por  $\epsilon$ , que define a profundidade do poço de potencial, e por  $\sigma$ , que corresponde à distância em que o potencial é nulo.

Para sistemas aquosos, existe uma grande variedade de potenciais empíricos, destacando-se entre eles os potenciais TIP3P, TIP4P [36] e suas variações, TIP5P

[124] e SPC [125]. Estes potenciais combinam as interações eletrostáticas e de Lennard-Jones para os sistemas. Nos modelos TIP3P e SPC, cada átomo corresponde a um sítio de carga. Para o potencial TIP4P e suas variações, é adicionado um sítio de carga fictício M próximo à bissetriz do ângulo H-O-H, tornando-se, assim, um modelo com quatro sítios, o que aprimora a descrição da distribuição eletrostática da molécula de água. Por fim, nos modelos de cinco sítios, como o TIP5P, duas cargas negativas fictícias L são adicionadas, representando os pares de elétrons não ligantes do átomo de oxigênio. A Figura 1.4 mostra um esquema da molécula de água para cada potencial.

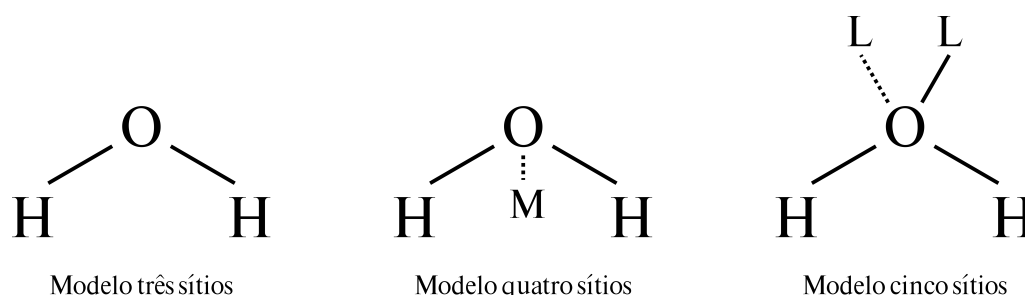


Figura 1.4: Esquema da molécula de água representada por diferentes modelos de potencial.

Além da geometria, os potenciais diferem em outros aspectos. Os modelos também possuem diferentes valores para os parâmetros de Lennard-Jones, distâncias interatômicas e cargas. Apresentamos alguns destes parâmetros na tabela 1.1.

Parâmetro	TIP3P [36]	SPC [125]	TIP4P [36]	TIP4P/Ice [39]	TIP5P [124]
Carga O ( $e$ )	-0.834	-0.820	-1.040	-1.1794	0.00
Carga H ( $e$ )	0.417	0.410	0.520	0.5897	0.241
Carga L ( $e$ )	-	-	-	-	-0.241
$\epsilon_{OO}$ (kcal/mol)	0.1521	0.1553	0.1550	0.21084	0.1600
$\sigma_{OO}$ (Å)	3.1507	3.166	3.1536	3.1668	3.1200
$r(OH)$ (Å)	0.9572	1.0	0.9572	0.9572	0.9572
$\angle(HOH)$ (°)	104.52	109.47	104.52	104.52	104.52
$r(OM)$ (Å)	-	-	0.15	0.1577	-
$r(OL)$ (Å)	-	-	-	-	0.70
$\angle(LOL)$ (°)	-	-	-	-	109.47

Tabela 1.1: Parâmetros geométricos, de carga e de Lennard-Jones para diferentes potenciais empíricos para a água.

Outro ponto relevante a se destacar é que, para os modelos supracitados, é

necessário que, durante os cálculos, as distâncias interatômicas e os ângulos sejam mantidos fixos, uma vez que a parametrização é feita para geometrias rígidas.

Verifica-se que tais modelos são dependentes de parâmetros que ainda não foram explicados. Estes dados são, em geral, provenientes de experimentos, tornando os modelos apropriados apenas para determinadas condições, diferentemente dos cálculos *ab initio*. Estes potenciais são chamados de potenciais empíricos. Apesar de a precisão destes potenciais não corresponder à dos cálculos DFT, seu custo computacional é menor, possibilitando o estudo de sistemas com mais átomos e por uma escala de tempo mais longa.

### 1.2.2 Controle de Temperatura

Em sistemas termodinamicamente isolados, a energia total é constante. Porém, em experimentos, também há a possibilidade de se verificar o efeito da temperatura sobre determinada propriedade, tal como uma transição de fase. Nestes casos, pode-se considerar que o sistema está imerso em um banho térmico, cuja temperatura é mantida constante, uma vez que este reservatório é suficientemente grande para não sofrer influência do sistema analisado.

Devido ao movimento dos átomos do sistema, a interação destes com o reservatório afeta sua energia cinética, podendo tanto aumentá-la quanto diminuí-la. Como a energia cinética está diretamente relacionada com a temperatura, verifica-se uma oscilação na temperatura do sistema analisado. Para simularmos este acoplamento, é possível adicionar ou remover energia do sistema de forma controlada até que a temperatura desejada seja atingida.

Este processo de atingir a temperatura desejada é chamado de fase de equilíbrio ou termalização. Inicialmente, atribui-se uma velocidade inicial aos átomos, e a temperatura é ajustada por um termostato. Ao se atingir a temperatura desejada, passa-se para a próxima etapa, na qual se coletam as informações para a produção dos resultados. Esta etapa é denominada de fase de produção [122].

Classicamente, o teorema da equipartição de energia relaciona a energia cinética média,  $E_{cin}$ , com a temperatura,  $T$ , do seguinte modo

$$\langle E_{cin} \rangle = \frac{3N}{2} k_B T, \quad (1.108)$$

onde  $N$  corresponde ao número de átomos do sistema e  $k_B$  é a constante de

Boltzmann. Então, a temperatura instantânea pode ser dada pela seguinte relação

$$T(t) = \frac{2}{3Nk_B} E_{cin}(t) = \frac{2}{3Nk_B} \sum_I \frac{M_I}{2} \mathbf{v}_I(t)^2, \quad (1.109)$$

onde  $M_I$  e  $\mathbf{v}_I$  são a massa e a velocidade de cada partícula, respectivamente. Para sistemas em equilíbrio térmico, as velocidades das partículas satisfazem a distribuição de Maxwell-Boltzmann

$$f(v) = 4\pi \left( \frac{M_I}{2\pi k_B T} \right)^{3/2} v^2 \exp \left( -\frac{m_I v^2}{2k_B T} \right). \quad (1.110)$$

Durante experimentos, a temperatura pode ser controlada por um termostato, um dispositivo que mantém o sistema em uma certa temperatura. Nos cálculos de MD, pode-se adaptar este conceito utilizando-se, por exemplo, o reescalonamento das velocidades das partículas [126]. Este reescalonamento das velocidades altera a energia cinética e, conseqüentemente, afeta a temperatura do sistema. Para se reescalonar a temperatura, um fator  $\lambda$  é introduzido, usando-se o fato de que a energia cinética é uma função homogênea de segunda ordem em relação às velocidades. Deste modo, escreve-se a expressão para a nova temperatura

$$T_{nova} = \frac{2}{3Nk_B} \sum_i \frac{m_i}{2} (\lambda \mathbf{v}_i)^2, \quad (1.111)$$

$$T_{nova} = \lambda^2 T_{antiga}. \quad (1.112)$$

O fator  $\lambda$  está diretamente relacionado com o fluxo de calor entre o reservatório e o sistema analisado.

### O Método de Nosé-Hoover

Existem outros métodos que podem ser utilizados como termostato [127, 128]. Dentre eles, será apresentado, de forma breve, o método de Nosé-Hoover [129, 130].

A ideia deste método é adicionar uma variável dinâmica fictícia ao sistema, cujo significado físico corresponde a um atrito,  $\zeta$ . Este termo, apesar do nome, pode tanto acelerar quanto frear o movimento das partículas até que a temperatura desejada seja atingida. Com isso, reescreve-se a equação de movimento do seguinte modo

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i - \zeta m_i \mathbf{v}_i, \quad (1.113)$$

onde

$$\frac{d\zeta(t)}{dt} = \frac{1}{Q} \left[ \sum_i \frac{m_i \mathbf{v}_i^2}{2} - \frac{3N}{2} k_B T_{ref} \right]. \quad (1.114)$$

O termo  $Q$  é um parâmetro com dimensão de massa que representa a "inércia" ou o tempo de relaxação da dinâmica do atrito  $\zeta$ . Caso  $\zeta > 0$ , a energia é retirada do sistema, enquanto para  $\zeta < 0$ , energia é injetada no sistema.

Com isso, as equações de movimento podem ser modificadas para introduzir o termostato de Nosé-Hoover. As etapas de integração, em uma formulação do tipo Verlet, são dadas por

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \left( \frac{\mathbf{f}(t)}{m} - \zeta(t)\mathbf{v}(t) \right) \frac{(\Delta t)^2}{2} \quad (1.115)$$

$$\mathbf{v}(t + \Delta t/2) = \mathbf{v}(t) + \frac{\Delta t}{2} \left( \frac{\mathbf{f}(t)}{m} - \zeta(t)\mathbf{v}(t) \right) \quad (1.116)$$

$$\mathbf{f}(t + \Delta t) = \mathbf{f}(\mathbf{r}(t + \Delta t)) \quad (1.117)$$

Para manter a estrutura do algoritmo, o cálculo do atrito também é feito em duas etapas, de forma similar à velocidade.

$$\zeta(t + \Delta t/2) = \zeta(t) + \frac{\Delta t}{2Q} \left[ \sum_i \frac{m_i \mathbf{v}(t)^2}{2} - \frac{3N}{2} k_B T_{ref} \right] \quad (1.118)$$

$$\zeta(t + \Delta t) = \zeta(t + \Delta t/2) + \frac{\Delta t}{2Q} \left[ \sum_i \frac{m_i \mathbf{v}(t + \Delta t/2)^2}{2} - \frac{3N}{2} k_B T_{ref} \right] \quad (1.119)$$

$$\mathbf{v}(t + \Delta t) = \frac{\mathbf{v}(t + \Delta t/2) + \frac{\Delta t}{2} \frac{\mathbf{f}(t + \Delta t)}{m_i}}{1 + \frac{\Delta t}{2} \zeta(t + \Delta t)}. \quad (1.120)$$

Deste modo, em todos os passos de integração calcula-se as velocidades com os atritos, afetando todas as quantidades derivadas de momento e energia cinética. Isto garante que o sistema oscile em torno da temperatura desejada de forma suave e contínua.

### Equações Generalizadas de Langevin

Outra alternativa de termostato é utilizar as equações generalizadas de Langevin. Em resumo, realizamos simulações de dinâmica molecular de estruturas de gelo utilizando um potencial profundo treinado com diferentes funcionais de

troca-correlação. A inclusão dos efeitos quânticos nucleares leva a uma melhora geral na descrição das propriedades estruturais do gelo em comparação com experimentos. De fato, as funções de distribuição radial obtidas com a inclusão de NQE foram, em geral, mais precisas do que as obtidas sem a inclusão, devido a uma melhor descrição da distribuição O–H. A distribuição O–O não apresentou mudanças significativas com a inclusão dos efeitos quânticos nucleares para grandes distâncias; entretanto, para distâncias abaixo de 4 Å, a posição do pico desloca-se ligeiramente para distâncias menores.

Observamos que a densidade, para todos os funcionais de troca-correlação, aumentou quando os efeitos quânticos nucleares foram considerados. Ao inspecionar as diferentes ligações no sistema, identificamos que a ligação intramolecular O–H enfraquece ligeiramente, enquanto as ligações de hidrogênio tornam-se mais fortes, fazendo com que as distâncias O–O diminuam, levando a densidades maiores, mais distantes dos valores experimentais. Isso nos leva a concluir que os funcionais do tipo GGA e o SCAN tendem a superligar as moléculas de água no gelo, resultando em uma forte rede de H-bonds, que é ainda mais reforçada pela inclusão de interações de van der Waals e efeitos quânticos nucleares [131]. O termostato baseado na Equação de Langevin Generalizada (GLE) é uma extensão do formalismo clássico de Langevin para o controle da temperatura em simulações de dinâmica molecular [132, 133]. Na forma usual, a equação de Langevin para a coordenada  $q$  e o momento  $p$  de uma partícula de massa  $m$  pode ser escrita como

$$\dot{q} = \frac{p}{m}, \quad \dot{p} = -\nabla V(q) - \gamma p + R(t), \quad (1.121)$$

onde  $\gamma$  é o coeficiente de atrito e  $R(t)$  é uma força aleatória de ruído branco com média nula e correlação  $\langle R(t)R(t') \rangle = 2m\gamma k_B T \delta(t - t')$ .

No caso do GLE, esses termos são generalizados para incluir memória temporal e ruído colorido. Assim, a equação de movimento assume a forma

$$\dot{p}(t) = -\nabla V(q(t)) - \int_0^t K(t-s) p(s) ds + R(t), \quad (1.122)$$

em que  $K(t)$  é o *kernel* de memória e  $R(t)$  é um processo estocástico com correlação dada pela relação de flutuação-dissipação:

$$\langle R(t)R(t') \rangle = k_B T K(|t - t'|). \quad (1.123)$$

Para tornar a implementação viável em simulações, o GLE é usualmente reescrito em um espaço aumentado, no qual variáveis auxiliares lineares  $s_i$  são introduzidas, de modo que o ruído colorido e a memória sejam representados por um sistema de equações diferenciais ordinárias [132]

$$\dot{\mathbf{p}} = -\nabla V(q) - \mathbf{A}_{pp} \mathbf{p} - \mathbf{A}_{ps} \mathbf{s} + \mathbf{B}_p \boldsymbol{\zeta}(t), \quad (1.124)$$

$$\dot{\mathbf{s}} = -\mathbf{A}_{sp} \mathbf{p} - \mathbf{A}_{ss} \mathbf{s} + \mathbf{B}_s \boldsymbol{\zeta}(t), \quad (1.125)$$

onde  $\boldsymbol{\zeta}(t)$  é um vetor de ruídos gaussianos não correlacionados. As matrizes  $\mathbf{A}$  e  $\mathbf{B}$  determinam o espectro efetivo de dissipação e são ajustadas para reproduzir o acoplamento desejado com o banho térmico.

Esse formalismo permite controlar seletivamente o acoplamento em diferentes faixas de frequência, o que é útil tanto para a melhora da eficiência de amostragem quanto para a introdução de efeitos quânticos aproximados, como no esquema *Quantum Thermal Bath* (QTB) [133]. Dessa forma, o GLE se apresenta como um dos termostatos mais versáteis e poderosos para simulações modernas.

A implementação de um termostato GLE é relativamente direta. Consideramos o caso do algoritmo de Verlet, que atualiza posições e momentos por um passo de tempo  $\Delta t$ , de acordo com

$$p \leftarrow p - V'(q)\Delta t/2 \quad (1.126)$$

$$q \leftarrow q + p\Delta t \quad (1.127)$$

$$p \leftarrow p - V'(q)\Delta t/2 \quad (1.128)$$

É possível introduzir nosso termostato GLE realizando dois passos de partícula livre de  $\Delta t/2$  sobre as variáveis  $(p, \mathbf{s})$

$$(p, s) \leftarrow \mathcal{P}[(p, s), \Delta t/2](p, s) \quad (1.129)$$

$$p \leftarrow p - V'(q) \frac{\Delta t}{2} \quad (1.130)$$

$$q \leftarrow q + p \Delta t \quad (1.131)$$

$$p \leftarrow p - V'(q) \frac{\Delta t}{2} \quad (1.132)$$

$$(p, s) \leftarrow \mathcal{P}[(p, s), \Delta t/2](p, s) \quad (1.133)$$

Em contraste com os termostatos baseados em equações de movimento de segunda ordem, como o Nosé-Hoover, nos quais é necessário um método de múltiplos passos de tempo para obter trajetórias precisas [134], este passo de partícula livre pode ser realizado sem introduzir erros adicionais de amostragem. O propagador exato de tempo finito para  $(p, \mathbf{s})$  é dado por

$$\mathcal{P}[(p, \mathbf{s}), \Delta t]^T = \mathbf{T}(\Delta t)(p, \mathbf{s})^T + \mathbf{S}(\Delta t)\xi^T, \quad (1.134)$$

onde  $\xi$  um vetor de  $n + 1$  números gaussianos não correlacionados, e as matrizes  $\mathbf{T}$  e  $\mathbf{S}$  podem ser calculadas uma única vez no início da simulação e para todos os graus de liberdade. As relações entre  $\mathbf{T}$ ,  $\mathbf{S}$ ,  $\mathbf{A}_p$ ,  $\mathbf{C}_p$  e  $\Delta t$  são dadas por

$$\mathbf{T} = e^{-\Delta t \mathbf{A}_p}, \quad \mathbf{S}\mathbf{S}^T = \mathbf{C}_p - e^{-\Delta t \mathbf{A}_p} \mathbf{C}_p e^{-\Delta t \mathbf{A}_p^T} \quad (1.135)$$

As matrizes  $\mathbf{A}_p$  e  $\mathbf{C}_p$  são parametrizáveis e podem ser obtidas para diversas condições iniciais no *site* GLE4MD [135].

### 1.3 Efeitos Quânticos Nucleares (NQE)

A aproximação de Born-Oppenheimer [22], já discutida anteriormente, retira da parte nuclear toda a informação quântica do problema, tratando os átomos apenas como "massas pontuais carregadas". Apesar da diminuição do custo computacional, negligenciar tais efeitos quânticos impossibilita a resolução de alguns problemas [68]. A aproximação clássica impossibilita a descrição da energia de ponto zero (ZPE), da quantização dos níveis vibracionais e também do tunelamento.

Embora os fenômenos não descritos pela aproximação clássica sejam, em sua maioria, relevantes apenas para baixas temperaturas, para casos específicos como ligações de alta frequência, a energia de ponto zero (ZPE), na qual  $\hbar\omega/2$  excede a energia térmica, pode ser significativa. A inclusão dos chamados efeitos quânticos nucleares (NQEs) pode afetar de forma significativa as propriedades dinâmicas e estruturais. Além disso, efeitos provenientes de isótopos também são capturados com a inclusão dos NQEs.

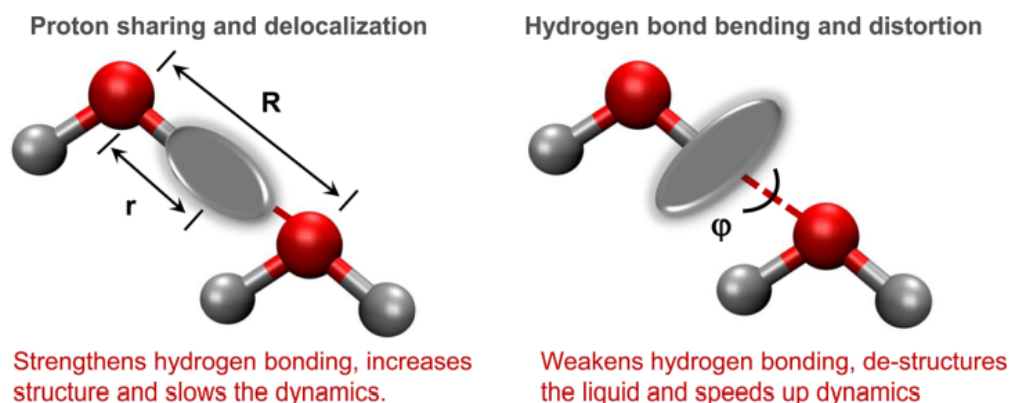


Figura 1.5: Representação dos efeitos quânticos nucleares nas ligações de hidrogênio. Imagem retirada da Ref. [1].

Embora sistemas aquosos pareçam simples, a água líquida apresenta diversas anomalias em suas propriedades, tal como a diminuição de sua densidade quando congelada. Todos esses efeitos são de suma importância para a regulação da vida e do clima na Terra [16]. Como os NQEs afetam diretamente as ligações de hidrogênio [71, 72, 1, 73, 74], é necessário apresentar ferramentas para melhorar a descrição dos fenômenos de interesse, uma vez que esses efeitos são fundamentais na descrição destes fenômenos. A Figura 1.5 apresenta os efeitos quânticos nucleares nas ligações de hidrogênio, elucidando sua importância, principalmente na deslocalização do átomo de hidrogênio, o que afeta a intensidade da ligação. Estes efeitos geram uma competição entre as ligações de hidrogênio e as ligações covalentes e o balanço entre essas forças é sutil [1].

### 1.3.1 Integrais de Caminho

Uma importante ferramenta a ser apresentada para a inclusão dos NQEs é a dinâmica molecular por integrais de caminho (PIMD). As integrais de caminho foram propostas inicialmente por Richard Feynman [136, 137], e podemos utilizar

esta abordagem para apresentar, de forma elegante, os efeitos quânticos nucleares. No formalismo Lagrangiano, uma partícula se move de um ponto  $\mathcal{A}$  para um ponto  $\mathcal{B}$  pelo caminho que minimiza a ação. Enquanto isso, o formalismo proposto por Feynman investiga as implicações de a partícula percorrer caminhos diferentes daquele considerado na abordagem clássica. Ao se somar sobre todos os caminhos possíveis, Feynman mostrou que a mecânica quântica emerge de maneira intuitiva [68].

A seguir, será apresentada uma abordagem matemática das integrais de caminho e de como elas são aplicadas no contexto da dinâmica molecular. Esta abordagem é feita com maiores detalhes por Tuckerman [2].

Considere, então, uma partícula que parte de um ponto  $\mathcal{A}$  e chega até um ponto  $\mathcal{B}$ . Como não podemos inferir o caminho que esta partícula percorreu entre os pontos, temos infinitas possibilidades, podendo ela percorrer todos os caminhos simultaneamente. De modo simples, a amplitude deste movimento está associada à soma das amplitudes de todos os caminhos possíveis entre os pontos.

Agora, expandindo esse conceito para o experimento da fenda dupla, podemos associar cada caminho percorrido a uma amplitude  $A_i(y)$ , onde  $i$  representa cada caminho possível pelo qual a partícula pode chegar a um ponto  $y$  do detector. De maneira explícita, temos que  $A(y) = A_1(y) + A_2(y) + \dots$  e, de forma direta,  $P(y) = |A(y)|^2$ , onde  $P(y)$  é a probabilidade de a partícula ser observada em um determinado ponto  $y$ . Para uma partícula que percorre apenas dois caminhos, a probabilidade é facilmente obtida. Porém, para uma partícula que percorre infinitos caminhos, temos a seguinte expressão

$$P(y) = \left| \sum_{\text{caminhos}} A_{\text{caminho}}(y) \right|^2. \quad (1.136)$$

Com o conceito teórico da soma sobre os caminhos já elucidado, partiremos para a aplicação matemática.

Inicialmente, focaremos em uma partícula simples se movendo em uma dimensão. Logo, sabe-se que o Hamiltoniano é dado por

$$\hat{H} = \frac{\hat{p}^2}{2m} + \hat{U}(\hat{x}). \quad (1.137)$$

Apresentamos anteriormente que as integrais de caminho descrevem o processo em que uma partícula se move entre os pontos  $\mathcal{A}$  e  $\mathcal{B}$ . Desta forma, a

amplitude de transição para essa partícula evoluir no tempo pode ser dada por

$$K(\mathcal{A}, \mathcal{B}; t) = \langle \mathcal{B} | e^{-i\hat{H}t/\hbar} | \mathcal{A} \rangle. \quad (1.138)$$

Projetando-se o propagador na base de coordenadas, chega-se a

$$\begin{aligned} \Psi(\mathcal{B}, t) &= \langle \mathcal{B} | \Psi(t) \rangle = \langle \mathcal{B} | e^{-i\hat{H}t/\hbar} | \Psi(0) \rangle \\ &= \int d\mathcal{A} \langle \mathcal{B} | e^{-i\hat{H}t/\hbar} | \mathcal{A} \rangle \Psi(\mathcal{A}, 0). \end{aligned} \quad (1.139)$$

Novamente, a abordagem de Feynman fornece uma técnica na qual os elementos de matriz do propagador podem ser obtidos por meio de uma soma sobre todos os caminhos no tempo  $t$ . Antes da exposição da derivação formal da integral de caminho de Feynman, é fundamental estabelecer a relação existente entre o propagador temporal e o operador densidade do ensemble canônico. Definindo-se o operador densidade canônico como  $\hat{\rho}(\beta) = e^{-\beta\hat{H}}$ , temos que

$$\hat{\rho}(\beta) = \hat{K}(-i\beta\hbar), \quad \hat{K}(t) = \hat{\rho}(it/\hbar). \quad (1.140)$$

Com esta equação, percebe-se que a matriz densidade pode ser obtida a partir da evolução do propagador no tempo imaginário  $t = -i\beta\hbar$ . Por apresentar maior tratabilidade matemática, será desenvolvido inicialmente o formalismo de integrais de caminho para a matriz densidade canônica, obtendo-se, posteriormente, a expressão para o propagador quântico. Vamos definir, então, os elementos da matriz na base de coordenadas para  $\hat{\rho}(\beta)$

$$\rho(\mathcal{A}, \mathcal{B}; \beta) = \langle \mathcal{B} | e^{-\beta\hat{H}} | \mathcal{A} \rangle. \quad (1.141)$$

Note que o Hamiltoniano é a soma dos operadores de energia cinética e potencial, para os quais  $[\hat{K}, \hat{U}] \neq 0$ . Com isso, o operador  $\hat{\rho}(\beta)$  não pode ser avaliado diretamente. Contudo, aplicando-se o teorema de Trotter, expressamos o operador da seguinte maneira

$$e^{-\beta\hat{H}} = \lim_{P \rightarrow \infty} \left[ e^{-\beta\hat{U}/2P} e^{-\beta\hat{K}/P} e^{-\beta\hat{U}/2P} \right]^P, \quad (1.142)$$

onde  $P$  é definido, por ora, como um inteiro.

Substituindo a equação 1.142 na equação 1.141, temos

$$\rho(\mathcal{A}, \mathcal{B}; \beta) = \lim_{P \rightarrow \infty} \langle \mathcal{B} | \left[ e^{-\beta \hat{U}/2P} e^{-\beta \hat{K}/P} e^{-\beta \hat{U}/2P} \right]^P | \mathcal{A} \rangle. \quad (1.143)$$

Para simplificar a notação, definimos o operador  $\hat{\Omega}$  como

$$\hat{\Omega} = e^{-\beta \hat{U}/2P} e^{-\beta \hat{K}/P} e^{-\beta \hat{U}/2P}. \quad (1.144)$$

Para tornar os cálculos tratáveis, introduz-se o operador identidade entre cada operador  $\hat{\Omega}$ . Como temos  $P$  fatores do operador  $\hat{\Omega}$ , são introduzidos  $P - 1$  operadores identidade. A matriz densidade é dada por

$$\begin{aligned} \rho(\mathcal{A}, \mathcal{B}; \beta) &= \lim_{P \rightarrow \infty} \int d\mathcal{A}_2 \cdots d\mathcal{A}_P \\ &\times \langle \mathcal{B} | \hat{\Omega} | \mathcal{A}_P \rangle \langle \mathcal{A}_P | \hat{\Omega} | \mathcal{A}_{P-1} \rangle \langle \mathcal{A}_{P-1} | \cdots | \mathcal{A}_2 \rangle \langle \mathcal{A}_2 | \hat{\Omega} | \mathcal{A} \rangle. \end{aligned} \quad (1.145)$$

A aplicação do teorema de Trotter possibilita que os elementos de matriz sejam calculados em forma fechada. Considere o elemento de matriz

$$\langle \mathcal{A}_{k+1} | \hat{\Omega} | \mathcal{A}_k \rangle = \langle \mathcal{A}_{k+1} | \left[ e^{-\beta \hat{U}/2P} e^{-\beta \hat{K}/P} e^{-\beta \hat{U}/2P} \right] | \mathcal{A}_k \rangle. \quad (1.146)$$

De forma direta, como  $\hat{U}$  depende das coordenadas, então  $\langle \mathcal{A}_{k+1} |$  e  $| \mathcal{A}_k \rangle$  são autoestados com autovalores  $e^{-\beta U(\mathcal{A}_{k+1})/2P}$  e  $e^{-\beta U(\mathcal{A}_k)/2P}$ , respectivamente. Com esta abordagem, simplificamos a expressão anterior para

$$\langle \mathcal{A}_{k+1} | \hat{\Omega} | \mathcal{A}_k \rangle = e^{-\beta U(\mathcal{A}_{k+1})/2P} \langle \mathcal{A}_{k+1} | e^{-\beta \hat{K}/P} | \mathcal{A}_k \rangle e^{-\beta U(\mathcal{A}_k)/2P}. \quad (1.147)$$

Uma vez que  $\hat{K}$  é uma função do operador momento, seu elemento de matriz na base de coordenadas é menos trivial de se calcular. Porém, inserindo-se o operador identidade na base de momento, temos

$$\langle \mathcal{A}_{k+1} | e^{-\beta \hat{K}/P} | \mathcal{A}_k \rangle = \int dp \langle \mathcal{A}_{k+1} | e^{-\beta \hat{K}/P} | p \rangle \langle p | \mathcal{A}_k \rangle. \quad (1.148)$$

Agora, o operador relacionado ao momento atua no autoestado  $| p \rangle$ , logo

$$\langle \mathcal{A}_{k+1} | e^{-\beta \hat{K}/P} | \mathcal{A}_k \rangle = \int dp \langle \mathcal{A}_{k+1} | p \rangle \langle p | \mathcal{A}_k \rangle e^{-\beta p^2/2mP}. \quad (1.149)$$

Uma vez que:

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar}, \quad (1.150)$$

podemos reescrever a equação 1.149 do seguinte modo

$$\langle \mathcal{A}_{k+1}|e^{-\beta\hat{K}/P}|\mathcal{A}_k\rangle = \frac{1}{2\pi\hbar} \int dp e^{-\beta p^2/2mP} e^{ip(\mathcal{A}_{k+1}-\mathcal{A}_k)/\hbar}. \quad (1.151)$$

Como o intervalo de integração do momento é de  $(-\infty, \infty)$ , ao se completar o quadrado no expoente, a integral se torna uma integral Gaussiana. Para tal, temos

$$\begin{aligned} & \frac{\beta p^2}{2mP} - \frac{ip(\mathcal{A}_{k+1} - \mathcal{A}_k)}{\hbar} \\ &= \frac{\beta}{2mP} \left[ p^2 - \frac{2imPp(\mathcal{A}_{k+1} - \mathcal{A}_k)}{\beta\hbar} \right] \\ &= \frac{\beta}{2mP} \left\{ \left[ p - \frac{imP(\mathcal{A}_{k+1} - \mathcal{A}_k)}{\beta\hbar} \right]^2 + \frac{m^2P^2(\mathcal{A}_{k+1} - \mathcal{A}_k)^2}{\beta^2\hbar^2} \right\} \\ &= \frac{\beta}{2mP} \left[ p - \frac{imP(\mathcal{A}_{k+1} - \mathcal{A}_k)}{\beta\hbar} \right]^2 + \frac{mP}{2\beta\hbar^2} (\mathcal{A}_{k+1} - \mathcal{A}_k)^2. \end{aligned} \quad (1.152)$$

Utilizando as últimas duas linhas da equação 1.152 na equação 1.151, chegamos ao seguinte resultado

$$\langle \mathcal{A}_{k+1}|e^{-\beta\hat{K}/P}|\mathcal{A}_k\rangle = \left( \frac{mP}{2\pi\beta\hbar^2} \right)^{1/2} e^{-\frac{mP}{2\beta\hbar^2}(\mathcal{A}_{k+1}-\mathcal{A}_k)^2}, \quad (1.153)$$

onde, combinando com a equação 1.147, temos

$$\langle \mathcal{A}_{k+1}|\hat{\Omega}|\mathcal{A}_k\rangle = \left( \frac{mP}{2\pi\beta\hbar^2} \right)^{1/2} e^{-\frac{\beta}{2P}(U(\mathcal{A}_{k+1})+U(\mathcal{A}_k))} e^{-\frac{mP}{2\beta\hbar^2}(\mathcal{A}_{k+1}-\mathcal{A}_k)^2}. \quad (1.154)$$

Finalmente, multiplicando para os  $P$  termos necessários, chega-se à expressão final para a matriz densidade

$$\begin{aligned} \rho(\mathcal{A}, \mathcal{B}; \beta) &= \lim_{P \rightarrow \infty} \left( \frac{mP}{2\pi\beta\hbar^2} \right)^{P/2} \int d\mathcal{A}_2 \cdots d\mathcal{A}_P \\ &\times \exp \left\{ - \sum_{k=1}^P \left[ \frac{mP}{2\beta\hbar^2} (\mathcal{A}_{k+1} - \mathcal{A}_k)^2 + \frac{\beta}{2P} (U(\mathcal{A}_{k+1}) + U(\mathcal{A}_k)) \right] \right\} \Bigg|_{\mathcal{A}_1=\mathcal{A}}^{\mathcal{A}_{P+1}=\mathcal{B}}. \end{aligned} \quad (1.155)$$

Observe que o termo relacionado à energia cinética pode ser comparado à forma de um acoplamento harmônico entre vizinhos mais próximos ao longo do caminho. Este acoplamento é dado por uma mola cuja constante é  $mP/(\beta\hbar)^2$ .

Importante ressaltar que as integrais sobre  $\mathcal{A}_k$  representam a soma sobre todos os caminhos, e  $P$  representa o número de "fatias de tempo" ou contas (*beads*) entre os pontos  $\mathcal{A}$  e  $\mathcal{B}$ .

Fazendo-se a substituição  $\beta \rightarrow it/\hbar$ , recupera-se a expressão para o propagador  $K(\mathcal{A}, \mathcal{B}; t)$

$$K(\mathcal{A}, \mathcal{B}; t) = \lim_{P \rightarrow \infty} \left( \frac{mP}{2\pi i t \hbar} \right)^{P/2} \int d\mathcal{A}_2 \cdots d\mathcal{A}_P \times \exp \left\{ \frac{i}{\hbar} \sum_{k=1}^P \left[ \frac{m}{2} \left( \frac{\mathcal{A}_{k+1} - \mathcal{A}_k}{t/P} \right)^2 - \frac{t}{2P} (U(\mathcal{A}_{k+1}) + U(\mathcal{A}_k)) \right] \frac{t}{P} \right\} \Bigg|_{\mathcal{A}_1 = \mathcal{A}}^{\mathcal{A}_{P+1} = \mathcal{B}}. \quad (1.156)$$

Partindo da equação 1.155, podemos encontrar a função de partição  $\mathcal{Z}(L, T)$  para um sistema confinado em  $x \in [0, L]$ . Uma vez que  $\mathcal{Z}(L, T) = \text{Tr} [e^{-\beta\hat{H}}]$ , temos que

$$\mathcal{Z}(L, T) = \int_0^L dx \langle x | e^{-\beta\hat{H}} | x \rangle = \int_0^L dx \rho(x, x; \beta). \quad (1.157)$$

Para se computar a função de partição, requer-se o conhecimento dos elementos diagonais da matriz densidade na representação de coordenadas, os quais são determinados mediante a imposição das condições  $\mathcal{A}_1 = \mathcal{A}_{P+1} = x$  na equação 1.155. Conclui-se o procedimento com a execução de uma integração sobre tais elementos diagonais. Como  $\mathcal{A}_1 = \mathcal{A}_{P+1} = x$ , a integração sobre  $x$  transforma o problema em uma integral sobre  $P$  variáveis,  $\mathcal{A}_1, \dots, \mathcal{A}_P$ . Logo, a expressão para a função de partição é dada por

$$\mathcal{Z}(L, T) = \lim_{P \rightarrow \infty} \left( \frac{mP}{2\pi\beta\hbar^2} \right)^{P/2} \int_{D(L)} d\mathcal{A}_1 \cdots d\mathcal{A}_P \times \exp \left\{ - \sum_{k=1}^P \left[ \frac{mP}{2\beta\hbar^2} (\mathcal{A}_{k+1} - \mathcal{A}_k)^2 + \frac{\beta}{P} U(\mathcal{A}_k) \right] \right\} \Bigg|_{\mathcal{A}_{P+1} = \mathcal{A}_1}. \quad (1.158)$$

onde foi usada a propriedade de que, para um caminho cíclico,  $\sum (1/2)[U(\mathcal{A}_{k+1}) + U(\mathcal{A}_k)] = \sum U(\mathcal{A}_k)$ .

Importante notar que, para  $T \rightarrow \infty$  (ou seja,  $\beta \rightarrow 0$ ), as constantes das molas

tendem também a infinito, o que provoca o colapso dos caminhos cíclicos presentes na função de partição para um único ponto, correspondente a uma partícula pontual clássica. Ou seja, o formalismo de integrais de caminho mostra que, para altas temperaturas, recupera-se o limite clássico. Apresenta-se a função de partição de modo simplificado

$$\mathcal{Z}(L, T) \propto \lim_{P \rightarrow \infty} \int_{D(L)} d\mathcal{A}_1 \cdots d\mathcal{A}_P e^{-\beta\phi(\mathcal{A}_1, \dots, \mathcal{A}_P)}, \quad (1.159)$$

onde

$$\phi(\mathcal{A}_1, \dots, \mathcal{A}_P) = \sum_{k=1}^P \left[ \frac{1}{2} m \omega_P^2 (\mathcal{A}_{k+1} - \mathcal{A}_k)^2 + \frac{1}{P} U(\mathcal{A}_k) \right], \quad (1.160)$$

com  $\omega_P = \sqrt{P}/(\beta\hbar)$ .

Até o momento, verificamos a função de partição de forma analítica, porém sua aplicação computacional necessita, apesar da evolução das máquinas, que  $P$  assuma um valor finito. Contudo, é necessário que  $P$  seja suficientemente grande para que o erro de Trotter seja negligenciável.

Com a expressão da função de partição do problema, devemos buscar um modo de calcular os valores dos observáveis através das integrais de caminho. Para tal, partimos da seguinte relação

$$\langle \hat{A} \rangle = \frac{1}{\mathcal{Z}(L, T)} \text{Tr} \left[ \hat{A} e^{-\beta\hat{H}} \right]. \quad (1.161)$$

Se o operador  $\hat{A}$  é uma função apenas das coordenadas, então  $|x\rangle$  é um autoestado do operador, ou seja

$$\hat{A}|x\rangle = a(x)|x\rangle, \quad (1.162)$$

onde  $a(x)$  é o autovalor associado ao operador.

Portanto, o valor esperado do observável pode ser escrito da seguinte forma

$$\begin{aligned} \langle \hat{A} \rangle &= \frac{1}{\mathcal{Z}(L, T)} \lim_{P \rightarrow \infty} \left( \frac{mP}{2\pi\beta\hbar^2} \right)^{P/2} \int_{D(L)} d\mathcal{A}_1 \cdots d\mathcal{A}_P a(\mathcal{A}_1) \\ &\times \exp \left\{ - \sum_{k=1}^P \left[ \frac{mP}{2\beta\hbar^2} (\mathcal{A}_{k+1} - \mathcal{A}_k)^2 + \frac{\beta}{P} U(\mathcal{A}_k) \right] \right\} \Bigg|_{\mathcal{A}_{P+1}=\mathcal{A}_1}. \end{aligned} \quad (1.163)$$

A equação acima aparenta privilegiar apenas uma posição, em detrimento das demais, uma vez que  $a(x)$  é avaliada apenas em  $\mathcal{A}_1$ . Porém, como os caminhos

são cíclicos, todos os pontos são equivalentes. Uma alternativa é fazer uma média sobre todas as posições. Deste modo, a equação pode ser reescrita como

$$\begin{aligned} \langle \hat{A} \rangle &= \frac{1}{\mathcal{Z}(L, T)} \lim_{P \rightarrow \infty} \left( \frac{mP}{2\pi\beta\hbar^2} \right)^{P/2} \int_{D(L)} d\mathcal{A}_1 \cdots d\mathcal{A}_P \left[ \frac{1}{P} \sum_{k=1}^P a(\mathcal{A}_k) \right] \\ &\times \exp \left\{ - \sum_{k=1}^P \left[ \frac{mP}{2\beta\hbar^2} (\mathcal{A}_{k+1} - \mathcal{A}_k)^2 + \frac{\beta}{P} U(\mathcal{A}_k) \right] \right\} \Big|_{\mathcal{A}_{P+1}=\mathcal{A}_1}. \end{aligned} \quad (1.164)$$

Podemos compactar essa equação. O primeiro passo é definir a função de densidade de probabilidade  $f(\mathcal{A}_1, \dots, \mathcal{A}_P)$

$$\begin{aligned} f(\mathcal{A}_1, \dots, \mathcal{A}_P) &= \frac{1}{\mathcal{Z}_P(L, T)} \left( \frac{mP}{2\pi\beta\hbar^2} \right)^{P/2} \\ &\times \exp \left\{ -\beta \sum_{k=1}^P \left[ \frac{1}{2} m\omega_P^2 (\mathcal{A}_{k+1} - \mathcal{A}_k)^2 + \frac{1}{P} U(\mathcal{A}_k) \right] \right\} \Big|_{\mathcal{A}_{P+1}=\mathcal{A}_1}, \end{aligned} \quad (1.165)$$

onde  $\mathcal{Z}_P(L, T)$  é a função de partição para um valor finito de  $P$

$$\begin{aligned} \mathcal{Z}_P(L, T) &= \left( \frac{mP}{2\pi\beta\hbar^2} \right)^{P/2} \int_{D(L)} d\mathcal{A}_1 \cdots d\mathcal{A}_P \\ &\times \exp \left\{ -\frac{1}{\hbar} \sum_{k=1}^P \left[ \frac{mP}{2\beta\hbar} (\mathcal{A}_{k+1} - \mathcal{A}_k)^2 + \frac{\beta\hbar}{P} U(\mathcal{A}_k) \right] \right\} \Big|_{\mathcal{A}_{P+1}=\mathcal{A}_1}. \end{aligned} \quad (1.166)$$

Além disso, a função de probabilidade deve satisfazer a relação de normalização

$$\int d\mathcal{A}_1 \cdots d\mathcal{A}_P f(\mathcal{A}_1, \dots, \mathcal{A}_P) = 1. \quad (1.167)$$

Em cálculos de integral de caminho, os valores esperados no equilíbrio podem ser aproximados utilizando um estimador que depende das coordenadas  $\mathcal{A}_1, \dots, \mathcal{A}_P$ . O estimador para o observável  $A$  pode ser definido como

$$a_P(\mathcal{A}_1, \dots, \mathcal{A}_P) = \frac{1}{P} \sum_{k=1}^P a(\mathcal{A}_k). \quad (1.168)$$

O valor esperado  $\langle \hat{A} \rangle$  pode ser aproximado, para um valor finito de  $P$ , como a

média do estimador sobre a distribuição de probabilidade, logo

$$\langle A \rangle \approx \langle a_P(\mathcal{A}_1, \dots, \mathcal{A}_P) \rangle_f, \quad (1.169)$$

onde o índice  $f$  indica a média sobre a função de probabilidade. No limite em que  $P \rightarrow \infty$ ,  $\langle A \rangle_P$  converge para o valor exato  $\langle A \rangle$ .

Vamos considerar o caso da energia média. Sabemos que

$$E = \langle \hat{H} \rangle = \left\langle \frac{\hat{p}^2}{2m} + \hat{U}(\hat{x}) \right\rangle. \quad (1.170)$$

Apesar da dependência tanto das coordenadas quanto dos momentos, a energia média do sistema pode ser obtida via relação termodinâmica, ou seja

$$E = -\frac{\partial}{\partial \beta} \ln \mathcal{Z}(L, T). \quad (1.171)$$

Utilizando-se a expressão da função de partição para caminhos cíclicos, temos o que precisamos para calcular o valor de  $E$ . Derivando a função de partição com relação a  $\beta$ , obtemos a seguinte equação

$$E = \lim_{P \rightarrow \infty} \langle \epsilon_P(\mathcal{A}_1, \dots, \mathcal{A}_P) \rangle_f, \quad (1.172)$$

onde  $\epsilon_P$  é o estimador termodinâmico da energia

$$\epsilon_P(\mathcal{A}_1, \dots, \mathcal{A}_P) = \frac{P}{2\beta} - \sum_{k=1}^P \frac{m\omega_P^2}{2} (\mathcal{A}_{k+1} - \mathcal{A}_k)^2 + \frac{1}{P} \sum_{k=1}^P U(\mathcal{A}_k). \quad (1.173)$$

### 1.3.2 Dinâmica Molecular com Integrais de Caminho (PIMD)

Após uma breve discussão sobre as integrais de caminho, é preciso aplicá-las nos cálculos de dinâmica molecular. É importante ressaltar que, neste contexto, a dinâmica molecular serve apenas como um método para amostrar a distribuição canônica quântica.

Consideremos novamente o problema unidimensional cuja função de partição é apresentada na equação 1.166. Esta equação pode ser manipulada a fim de que se torne similar à função de partição clássica de uma cadeia de polímero cíclico, na qual cada partícula se move em um potencial externo  $U(x)/P$ . Logo, o isomorfismo clássico nos permite escrever

$$\mathcal{Z}(L, T) = \lim_{P \rightarrow \infty} \left( \frac{mP}{2\pi\beta\hbar^2} \right)^{P/2} \int_{D(L)} d\mathcal{A}_1 \cdots d\mathcal{A}_P \times \exp \left\{ -\beta \sum_{k=1}^P \left[ \frac{p_i^2}{2m'} + \frac{1}{2} m\omega_P^2 (\mathcal{A}_{k+1} - \mathcal{A}_k)^2 + \frac{1}{P} U(\mathcal{A}_k) \right] \right\} \Big|_{\mathcal{A}_{P+1}=\mathcal{A}_1}, \quad (1.174)$$

onde  $\omega_P = \sqrt{P}/(\beta\hbar)$  é conhecida como a frequência do polímero, relacionada à constante de mola que acopla as contas vizinhas.

Essa similaridade fez com que Chandler e Wolynes [77] explorassem a analogia entre a função de partição quântica (aproximada, para um  $P$  finito) e a clássica. Devido à sua estrutura em anel, os  $P$  pontos do polímero são frequentemente chamados de contas (*beads*). Representamos esta estrutura do polímero cíclico na Figura 1.6.

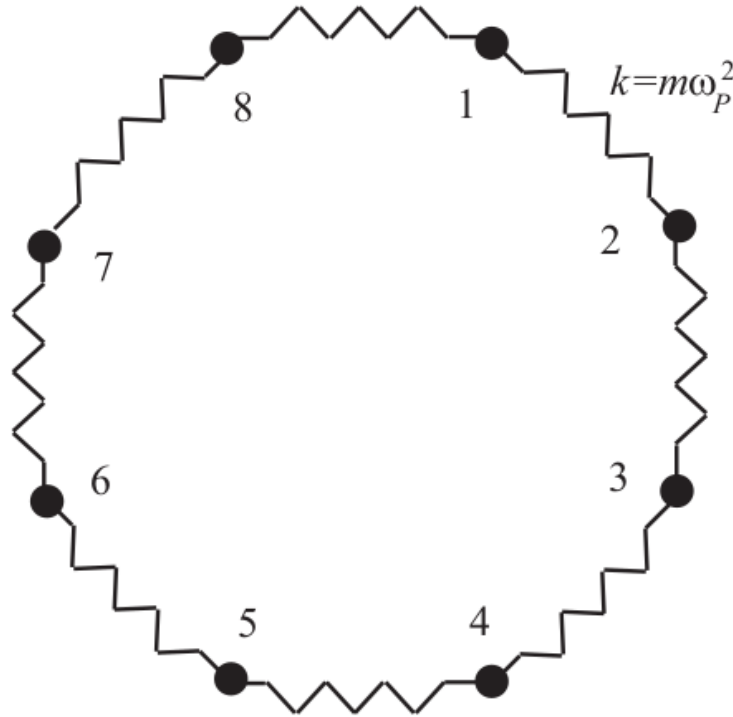


Figura 1.6: Representação do isomorfismo entre uma partícula quântica e um polímero cíclico clássico com constante de mola  $k = m\omega_P^2$ . Adaptado de Tuckerman [2].

Deste modo, pode-se utilizar o seguinte Hamiltoniano clássico efetivo para

gerar as equações de movimento

$$\mathcal{H}_{eff}(x, p) = \sum_{k=1}^P \left[ \frac{p_k^2}{2m'_k} + \frac{1}{2} m \omega_P^2 (x_{k+1} - x_k)^2 + \frac{1}{P} U(x_k) \right] \Big|_{x_{P+1}=x_1}, \quad (1.175)$$

onde, por conveniência, a notação das posições foi alterada de  $\mathcal{A}_k$  para  $x_k$ .

Com este Hamiltoniano, temos as seguintes equações de movimento

$$\dot{x}_k = \frac{p_k}{m'_k}, \quad \dot{p}_k = -m \omega_P^2 (2x_k - x_{k+1} - x_{k-1}) - \frac{1}{P} \frac{\partial U}{\partial x_k}. \quad (1.176)$$

Na forma atual, o problema ainda é complicado de se resolver numericamente devido ao forte acoplamento do termo harmônico. Porém, ao se desacoplar este termo por meio de uma transformação de coordenadas, pode-se otimizar a dinâmica. A transformação para as coordenadas de *staging* é uma forma de se desacoplar o termo harmônico

$$\begin{aligned} u_1 &= x_1, \\ u_k &= x_k - \frac{(k-1)x_{k+1} + x_1}{k}, \quad k = 2, \dots, P. \end{aligned} \quad (1.177)$$

A seguir, propõe-se a formulação da dinâmica molecular em termos das variáveis de *staging* ( $u_k$ ). Ao expressar o termo de acoplamento harmônico nessas variáveis, obtém-se

$$\sum_{k=1}^P (x_{k+1} - x_k)^2 = \sum_{k=2}^P \frac{k}{k-1} u_k^2, \quad (1.178)$$

uma equação na qual os termos harmônicos estão desacoplados. Uma vez que o Jacobiano da transformação é igual a 1, pode-se reescrever a função de partição da seguinte forma

$$\begin{aligned} \mathcal{Z}_P(L, T) &= \int dp_1 \cdots dp_P \int_{D(L)} du_1 \cdots du_P \\ &\times \exp \left\{ -\beta \sum_{k=1}^P \left[ \frac{p_k^2}{2m'_k} + \frac{1}{2} m_k \omega_P^2 u_k^2 + \frac{1}{P} U(x_k(u)) \right] \right\}. \end{aligned} \quad (1.179)$$

Os parâmetros  $m_k$  são definidos da seguinte forma

$$\begin{aligned} m_1 &= 0, \\ m_k &= \frac{k}{k-1} m, \quad k = 2, \dots, P. \end{aligned} \quad (1.180)$$

Além disso, temos que  $m'_1 = m$  e  $m'_k = m_k$ . De forma análoga ao feito anteriormente, podemos escrever o seguinte Hamiltoniano a ser resolvido para as equações de movimento

$$\mathcal{H}_{eff}(u, p) = \sum_{k=1}^P \left[ \frac{p_k^2}{2m'_k} + \frac{1}{2} m_k \omega_P^2 u_k^2 + \frac{1}{P} U(x_k(u)) \right]. \quad (1.181)$$

Para finalizar esta seção, explicita-se a expressão para a função de partição de um sistema com  $N$  partículas em 3 dimensões

$$\begin{aligned} \mathcal{Z}_P(N, V, T) &= \mathcal{Z}_{ideal}^{(P)} \int \prod_{i=1}^N (d\mathbf{r}_i^{(1)} \cdots d\mathbf{r}_i^{(P)}) \\ &\times \exp \left\{ -\beta \left[ \sum_{k=1}^P \sum_{i=1}^N \frac{1}{2} m_i \omega_P^2 (\mathbf{r}_i^{(k+1)} - \mathbf{r}_i^{(k)})^2 + \sum_{k=1}^P \frac{1}{P} U(\mathbf{r}_1^{(k)}, \dots, \mathbf{r}_N^{(k)}) \right] \right\}. \end{aligned} \quad (1.182)$$

Um ponto importante a ressaltar é o fato de que o potencial físico  $U$  atua apenas entre partículas que possuem o mesmo índice de conta ( $k$ ). A implementação de um algoritmo de dinâmica molecular com integral de caminho para  $N$  partículas em 3 dimensões segue a mesma metodologia aplicada a uma única partícula em uma dimensão. Com isto, temos o seguinte Hamiltoniano efetivo para o sistema

$$\mathcal{H}_{eff} = \sum_{k=1}^P \left[ \sum_{i=1}^N \frac{(\mathbf{p}_i^{(k)})^2}{2m'_i} + \sum_{i=1}^N \frac{1}{2} m_i \omega_P^2 (\mathbf{r}_i^{(k+1)} - \mathbf{r}_i^{(k)})^2 + \frac{1}{P} U(\mathbf{r}_1^{(k)}, \dots, \mathbf{r}_N^{(k)}) \right]. \quad (1.183)$$

É crucial destacar que, se as equações de movimento estiverem acopladas a cadeias de Nosé-Hoover, é fundamental seguir o protocolo de conectar cada componente de cada variável de *staging* ou modo normal ao seu próprio termostato, totalizando  $3NP$  termostatos.

À primeira vista, isso pode parecer excessivo, pois adiciona  $3NMP$  graus de liberdade extras ao sistema (onde  $M$  é o comprimento de cada cadeia de Nosé-Hoover). No entanto, o custo computacional associado a  $3NMP$  termostatos é pequeno quando comparado ao cálculo das forças.

Outro ponto a se destacar é que uma partícula  $i$  interage com uma partícula  $j$  apenas se ambas pertencerem à mesma conta ( $k$ ). Isto faz com que o cálculo das forças físicas em cada conta seja independente das outras contas.

### 1.3.3 Equações de Langevin

Como visto na seção anterior, é possível, por meio das integrais de caminho, aproximar um sistema quântico com um sistema clássico de partículas conectadas por molas. Para tal metodologia, no limite em que o número de contas  $P$  tende ao infinito, recupera-se o resultado quântico exato.

Contudo, devido ao custo computacional, realiza-se um truncamento no número de contas. Como exemplo, é necessário utilizar ao menos  $P = 32$  para convergir propriedades estruturais de sistemas que contenham ligações covalentes O-H em temperatura ambiente [68].

Devido à limitação computacional, são necessários métodos que permitam diminuir o número de contas (réplicas) de modo a facilitar a aplicação das integrais de caminho em cálculos de dinâmica molecular.

Existem alguns métodos eficientes para a aplicação das integrais de caminho. Neste trabalho, vamos focar em métodos que utilizam as equações de Langevin generalizadas (GLE), baseando-se na intuição de que flutuações quânticas podem ser efetivamente imitadas por dinâmicas que não obedecem ao teorema clássico de flutuação-dissipação. Durante os anos 1980, alguns autores propuseram uma abordagem quântica para as equações de Langevin [138, 139].

Recentemente, vários grupos propuseram o uso de uma equação de Langevin generalizada [132, 133] para garantir que cada modo normal do sistema amostra a energia média correta de um oscilador harmônico quântico,  $\langle E(\omega) \rangle = (\hbar\omega/2) \coth(\beta\hbar\omega/2)$ , simulando o efeito da quantização dos graus de liberdade nucleares. Havendo acoplamentos anarmônicos, o fluxo de energia entre os modos de alta e baixa frequência resulta em desvios significativos em relação à distribuição de energia desejada.

Uma maneira de corrigir este problema é combinar as integrais de caminho com a equação de Langevin generalizada (PI+GLE). Apesar dessa boa descrição, o uso de um termostato GLE em uma simulação clássica (sem réplicas) é suficiente apenas para acelerar a convergência de observáveis estruturais, mas não para convergir outros estimadores (como a energia cinética) que dependem da distribuição quântica completa.

É possível estender o conceito das GLEs para acelerar a convergência das simulações PIMD, especialmente para baixos valores de  $P$ . Para tal, apresenta-se a seguir o método PIGLET [3].

Primeiramente, vamos apresentar um Hamiltoniano genérico relativo ao problema com integrais de caminho

$$\mathcal{H}_{eff}(q, p) = \sum_{k=1}^P \left( \frac{p_k^2}{2m'_k} + \frac{1}{2}m\omega_P^2(q_{k+1} - q_k)^2 + \frac{1}{P}U(q_k) \right). \quad (1.184)$$

De maneira geral, para um número  $P$  de réplicas, o valor esperado do potencial  $U$  é estimado como

$$\langle U \rangle = \left\langle \frac{1}{P} \sum_{k=1}^P U(q_k) \right\rangle. \quad (1.185)$$

A energia cinética quântica pode ser dada, utilizando o estimador do virial da centróide [140, 141], da seguinte forma

$$\langle T \rangle = \frac{3N}{2\beta} + \frac{1}{2P} \sum_{k=1}^P \langle (\mathbf{q}_k - \tilde{\mathbf{q}}) \cdot \nabla U(\mathbf{q}_k) \rangle, \quad (1.186)$$

onde  $\tilde{\mathbf{q}} = (1/P) \sum_k \mathbf{q}_k$  é a centróide, ou seja, a posição média da cadeia de polímero.

Para contornar os problemas de convergência da PIMD, utiliza-se a técnica de PI+GLE, na qual cada modo normal do polímero está sujeito a uma instância independente do mesmo termostato GLE, projetado de forma que o valor esperado  $\langle U \rangle$  seja exato para qualquer número de réplicas. Com isto, pode-se reduzir o número de réplicas em um fator de até 4 em comparação com a PIMD convencional.

Contudo, o método PI+GLE por si só não acelera a convergência da energia cinética quântica. A ideia central do método PIGLET é também convergir  $\langle T \rangle$  de forma rápida, assim como  $\langle U \rangle$ . Deste modo, podemos analisar as equações 1.185 e 1.186 para um oscilador harmônico unidimensional com frequência  $\omega$ . Logo

$$\begin{aligned} \langle T \rangle &= \frac{1}{2\beta} + \frac{1}{2P} \omega^2 \sum_{k=1}^P \langle q_k^2 \rangle - \frac{1}{2} \omega^2 \langle \tilde{q}^2 \rangle = \\ &= \langle U \rangle + \frac{1}{2\beta} - \frac{1}{2} \omega^2 \langle \tilde{q}^2 \rangle. \end{aligned} \quad (1.187)$$

A partir desta equação, percebe-se que, para o oscilador harmônico, a diferença entre a energia cinética e potencial depende apenas da flutuação da centróide. A

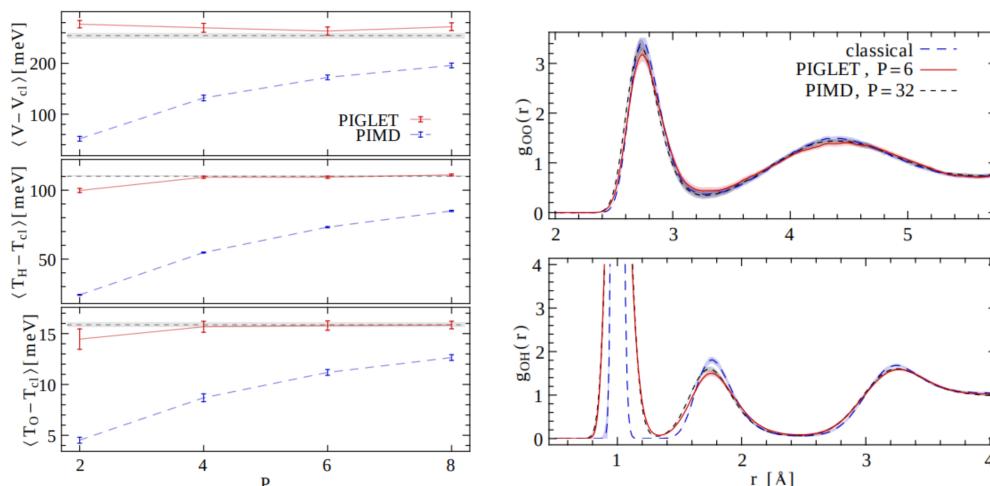


Figura 1.7: Comparação da convergência de quantidades físicas para PIMD e PIGLET em função do número de contas ( $P$ ). Retirado da Referência [3].

estatística quântica surge da distribuição das réplicas em torno da centróide. Para garantir a amostragem correta, o método PIGLET aplica um termostato clássico à centróide e termostatos PI+GLE aos modos internos do polímero. A diminuição no número de réplicas devido à aplicação do método PIGLET pode ser verificada na Figura 1.7.

É de suma importância ressaltar que os termostatos GLE são parametrizados de acordo com o número de réplicas, a temperatura e as frequências características do sistema. Uma ferramenta para se obterem os parâmetros para cada caso está disponível na Ref. [135], e o método pode ser utilizado em diversos pacotes de simulação.

## 1.4 Potenciais de *Machine Learning*

O cérebro humano é formado por uma rede de neurônios que estão acoplados a receptores e efetores. Os receptores são chamados de dendritos e os efetores de axônios. Os dendritos coletam sinais de muitos outros neurônios em uma área limitada; um corpo celular ou soma integra os sinais coletados e gera um sinal de resposta; e um axônio ramificado distribui essa resposta por meio de contatos com as árvores dendríticas de muitos outros neurônios [142]. Deste modo, a ideia das redes neurais é simular o comportamento do cérebro, onde existe uma camada que recebe os dados a serem processados (chamada de camada de entrada) que se comunicam com camadas ocultas, onde os dados serão processados. A

quantidade de camadas define se ela será ou não uma rede profunda, e por fim há uma camada de saída, onde o resultado de interesse é gerado. Dentro de cada camada é necessário um número de neurônios que recebem uma certa informação, processam e geram um dado a ser compartilhado com outros neurônios [143].

Anteriormente, na seção 1.2.1, foi apresentado o conceito de campos de força. Estes campos de força podem ser obtidos de diversas formas, porém os mais comuns são parametrizados para condições específicas.

Outro ponto a se ressaltar é a alta precisão dos resultados provenientes de cálculos *ab initio*. Entretanto, apesar de sua qualidade, o custo computacional é extremamente elevado. Para contornar o alto custo computacional, o uso de aprendizado de máquina — em inglês, *machine learning* (ML) — surge como uma possibilidade para unir a escalabilidade da MD com a acurácia dos cálculos *ab initio*.

Desde o fim da década de 2000, técnicas de ML vêm sendo empregadas para modelar a superfície de energia potencial (PES) de sistemas a partir de dados de DFT [44, 45, 11]. A partir disto, pode-se utilizar essa modelagem da PES para gerar um campo de força com precisão comparável à dos cálculos DFT.

Behler e Parrinello [44] desenvolveram uma representação geral de superfícies de energia potencial utilizando redes neurais feed-forward (NN), introduzindo características cruciais que permitiram o uso geral de redes neurais como a PES. Eles representaram a energia total do sistema como uma soma de energias atômicas, com uma rede neural diferente para cada espécie atômica; e, mais importante, introduziram as funções de simetria centradas no átomo [46] (ACSF) para descrever a geometria da configuração atômica dentro de um certo raio de corte, isto é, para representar o ambiente químico local que envolve um dado átomo.

Desta forma, em 2018, Zhang *et al.* [97] propuseram a abordagem *Deep Potential Molecular Dynamics*, na qual o modelo de rede neural mantém todas as simetrias naturais presentes no problema. Além disso, sua construção segue princípios fundamentais, não contendo elementos arbitrários além da própria estrutura da rede.

Em sequência, as redes neurais de mensagem passada (MPNN), um tipo de rede neural de grafos (GNN), que parametrizam um mapeamento de um grafo rotulado para um espaço-alvo, que pode ser um grafo ou um espaço vetorial foram desenvolvidas [5]. O grafo é incorporado em um espaço euclidiano 3D, no qual cada nó representa um átomo, e arestas conectam os nós (átomos) correspondentes que estejam dentro de uma certa distância entre si.

Existem outras abordagens (que não serão utilizadas neste trabalho) para obter os potenciais através de aprendizado de máquina, tal qual o *Gaussian Approximation Potential* (GAP) que utiliza um método bayesiano, o que permite a formulação de conhecimento prévio sobre o sistema atômico e suas interações na forma de hiperparâmetros, bem como a estimativa de incerteza. Isto é vantajoso devido ao ajuste do modelo constitui um problema de otimização convexa, equivalente à solução de um sistema linear, de modo que muitos problemas associados à minimização da função de perda em redes neurais não ocorrem [144].

Diversos grupos desenvolveram *softwares* capazes de gerar campos de força com a precisão de simulações DFT por meio de ML [93, 5, 57]. Neste trabalho, o foco será em dois *softwares* com abordagens distintas: o DeePMD-kit, que utiliza redes neurais profundas (DNN), e o MACE, um código que utiliza redes neurais de grafos baseadas na passagem de mensagens (MPNN) para modelar a PES.

### 1.4.1 DeePMD-kit

Inicialmente, apresentamos a estrutura de treinamento do DeePMD-kit [93]. O código possui partes escritas em C++ e Python. Além disso, até a segunda versão [93], era construído sobre o TensorFlow [145], porém, na versão mais recente [98], passou a suportar novos *backends* — como PyTorch [146], JAX [147] e PaddlePaddle [148] —, permitindo que os modelos gerados sejam utilizados por diversos *softwares*, como o LAMMPS [149] e o i-PI [150]. Nesta seção, será discutida a formulação matemática por trás do treinamento da rede neural.

Vamos definir um sistema que tenha  $N$  átomos, cujas posições são representadas por  $\mathbf{R}_1, \dots, \mathbf{R}_N$ . A energia total do sistema pode ser dada em função de todas as posições, ou seja,  $E = E(\mathbf{R}_1, \dots, \mathbf{R}_N)$ . Para este método, podemos considerar também que a energia total pode ser decomposta como a soma das energias atômicas, logo

$$E = \sum_{i=1}^N E_i. \quad (1.188)$$

A energia de cada átomo,  $E_i$ , pode ser determinada por sua vizinhança, ou seja, pela sua posição e pela interação com os demais átomos no interior de um raio de corte  $r_c(i)$

$$E_i = E_{s(i)}(\mathbf{R}_i, \{\mathbf{R}_j | j \in r_c(i)\}), \quad (1.189)$$

onde  $s(i)$  denota a espécie química do átomo  $i$ .

A ideia mais direta para modelar a energia atômica por meio de uma DNN é

treinar a rede com os dados de entrada sendo simplesmente as posições de cada átomo e seus vizinhos contidos em  $r_c$ .

No entanto, as simetrias de translação, rotação e permutação inerentes à PES não são garantidas por essa abordagem simples. Por isso, é necessário um pré-processamento adequado das posições atômicas, que mapeie o ambiente químico atômico por meio de descritores.

A construção deste descritor no método DeePMD inicia-se com a definição de um sistema de coordenadas centrado no átomo  $i$ . A coordenada relativa  $\mathbf{R}_{ij}$  pode ser reescrita da seguinte forma

$$\mathbf{R}_{ij} = x_{ij}^0 \mathbf{e}_x^0 + y_{ij}^0 \mathbf{e}_y^0 + z_{ij}^0 \mathbf{e}_z^0. \quad (1.190)$$

Neste caso, preserva-se a simetria translacional do problema. Para garantir a preservação da simetria rotacional, constrói-se um referencial local para cada átomo e registram-se as coordenadas dos vizinhos nesse referencial. Primeiramente, dois átomos, indexados como  $a(i)$  e  $b(i)$ , são selecionados dentre os vizinhos. Em seguida, o referencial local do átomo  $i$  é construído por

$$\begin{aligned} \mathbf{e}_{i1} &= \mathbf{e}(\mathbf{R}_{ia(i)}), \\ \mathbf{e}_{i2} &= \mathbf{e}(\mathbf{R}_{ib(i)} - (\mathbf{R}_{ib(i)} \cdot \mathbf{e}_{i1}) \mathbf{e}_{i1}), \\ \mathbf{e}_{i3} &= \mathbf{e}_{i1} \times \mathbf{e}_{i2}, \end{aligned} \quad (1.191)$$

onde  $\mathbf{e}(\mathbf{R})$  representa o versor de  $\mathbf{R}$ . Sendo assim, as coordenadas locais podem ser obtidas a partir das coordenadas globais por meio de uma rotação

$$(x_{ij}, y_{ij}, z_{ij})^T = \mathcal{R}(R_{ia(i)}, R_{ib(i)})^T (x_{ij}^0, y_{ij}^0, z_{ij}^0)^T, \quad (1.192)$$

onde a matriz de rotação é  $\mathcal{R} = [\mathbf{e}_{i1}, \mathbf{e}_{i2}, \mathbf{e}_{i3}]$ .

A informação descrita para o átomo  $i$  por meio do átomo  $j$  pode ser escrita de forma completa (utilizando a informação angular e radial) ou de forma simplificada (utilizando apenas a informação radial)

$$\mathcal{D}_{ij}^\alpha = \begin{cases} \left\{ \frac{s(R_{ij})}{R_{ij}}, \frac{s(R_{ij})x_{ij}}{R_{ij}}, \frac{s(R_{ij})y_{ij}}{R_{ij}}, \frac{s(R_{ij})z_{ij}}{R_{ij}} \right\}, & \text{inf. radial e angular} \\ \left\{ \frac{s(R_{ij})}{R_{ij}} \right\}, & \text{inf. radial} \end{cases} \quad (1.193)$$

onde  $s(r)$  é uma função de suavização que vai a zero em  $r_c$ .

A informação completa será utilizada quando  $\alpha > 0$ . Para  $\alpha = 0$ , apenas

a informação radial será utilizada para o treino da rede neural. As interações covalentes, como as de ligação e de ângulos diédricos, são bem descritas com a informação angular e radial dos primeiros e segundos vizinhos do átomo de interesse. Enquanto isso, interações de longo alcance podem ser capturadas apenas com a informação radial de vizinhos mais distantes.

Para melhor eficiência, organizam-se os átomos  $j$  de acordo com sua espécie atômica e, para cada espécie, organiza-se de acordo com o inverso da distância entre  $i$  e  $j$ , ou seja,  $1/R_{ij}$ . Dessa forma, a simetria permutacional é naturalmente preservada. Seguindo os procedimentos mencionados acima, construímos o mapeamento das posições atômicas para os descritores, que é denotado por

$$\mathcal{D}_i = \mathcal{D}_i(\mathbf{R}_i, \{\mathbf{R}_j | j \in r_c(i)\}). \quad (1.194)$$

Os descritores  $\mathcal{D}_i$  preservam as simetrias de translação, rotação e de permutação e são utilizados pela DNN para o cálculo da energia atômica. Assim, os descritores mapeiam a rede neural para a energia atômica da seguinte forma

$$E_{s(i)} = \mathcal{N}_{s(i)}(\mathcal{D}_i). \quad (1.195)$$

Temos, então, o caso de uma rede *feedforward* na qual os dados fluem da camada de entrada  $\mathcal{D}_i$ , passando por múltiplas camadas ocultas totalmente conectadas, até a camada de saída, que corresponde à energia atômica. Logo, para uma rede com  $N_h$  camadas ocultas, temos o seguinte mapeamento

$$\mathcal{N}_{s(i)}(\mathcal{D}_i) = \mathcal{L}_{s(i)}^{\text{out}} \circ \mathcal{L}_{s(i)}^{N_h} \circ \mathcal{L}_{s(i)}^{N_h-1} \circ \dots \circ \mathcal{L}_{s(i)}^1(\mathcal{D}_i), \quad (1.196)$$

onde  $\circ$  simboliza a composição de funções.

O mapeamento apresentado por  $\mathcal{L}_{s(i)}^p$  ocorre da camada  $p-1$  para a camada  $p$ , o que é interpretado como a composição de uma transformação linear com uma função não linear, o que se chama de função de ativação

$$\mathbf{d}_i^p = \mathcal{L}_{s(i)}^p(\mathbf{d}_i^{p-1}) = \varphi(\mathbf{W}_{s(i)}^p \mathbf{d}_i^{p-1} + \mathbf{b}_{s(i)}^p), \quad (1.197)$$

onde  $\mathbf{d}_i^p$  denota o vetor de neurônios na camada  $p$ ,  $\mathbf{W}_{s(i)}^p$  é a matriz de pesos com dimensão  $M_p \times M_{p-1}$  (onde  $M_p$  é o número de neurônios), e o vetor  $\mathbf{b}_{s(i)}^p$  é o viés (*bias*) da transformação linear.

Logo, a energia atômica pode ser dada por meio do mapeamento de saída  $\mathcal{L}_{s(i)}^{\text{out}}$ ,

então

$$E_{s(i)} = \mathcal{L}_{s(i)}^{\text{out}}(\mathbf{d}_i^{N_h}) = \mathbf{W}_{s(i)}^{\text{out}} \mathbf{d}_i^{N_h} + b_{s(i)}^{\text{out}}. \quad (1.198)$$

As forças e o tensor de virial podem ser obtidos analiticamente por meio da diferenciação da energia total em relação às posições atômicas e ao tensor de deformação, respectivamente.

Os parâmetros  $\{\mathbf{W}_{s(i)}^p, \mathbf{b}_{s(i)}^p\}$  são determinados através da minimização da função de perda (*loss function*). Esta função é definida como uma soma ponderada dos erros quadráticos médios (RMSE) da energia, das forças e do virial

$$L(p_e, p_f, p_\xi) = p_e \Delta E^2 + \frac{p_f}{3N} \sum_i \|\Delta \mathbf{F}_i\|^2 + \frac{p_\xi}{9} \|\Delta \Xi\|^2, \quad (1.199)$$

onde  $\Delta$  indica a diferença entre o valor de referência e o previsto pela rede, e os pré-fatores  $p_e, p_f, p_\xi$  são hiperparâmetros ajustáveis durante o treinamento. Em seu trabalho, Wang et al. [4] definem a seguinte abordagem para cada pré-fator

$$p_l(t) = p_l^{\text{limit}} \left[ 1 - \frac{r_l(t)}{r_l^0} \right] + p_l^{\text{start}} \left[ \frac{r_l(t)}{r_l^0} \right], \quad (1.200)$$

onde  $r_l^0$  e  $r_l(t)$  correspondem à taxa de aprendizagem inicial e à taxa de aprendizagem no passo  $t$ . Além disso, o pré-fator inicia com o valor  $p_l^{\text{start}}$  e encerra com o valor  $p_l^{\text{limit}}$ . A taxa de aprendizagem adotada possui um decaimento exponencial

$$r_l(t) = r_l^0 \times d_r^{t/d_s}, \quad (1.201)$$

onde  $d_r$  é a taxa de decaimento e  $d_s$  é o número de passos de decaimento.

Isto influencia diretamente nos RSME tanto da *loss function* quanto das grandezas físicas, como energias e forças. Para o caso na água líquida apresentado por Wang et al. [4], definiu-se os pesos iniciais para a energia e força sendo, respectivamente, 0,02 e 1000 e os pesos finais como 8 e 1. A análise do RSME para a energia e força para este treinamento pode ser observado na Figura 1.8. Neste caso, ao aumentar o número de passos, verifica-se uma diminuição no RSME para ambas as grandezas. Além disso, podemos observar que a taxa de decaimento do erro é maior nos primeiros passos de treino devido ao maior valor da taxa de decaimento, isto ocorre também pelo fato de que nos primeiros passos, o modelo não reconhece a totalidade dos passos, precisando de uma taxa de conhecimento superior comparado ao final do treinamento.

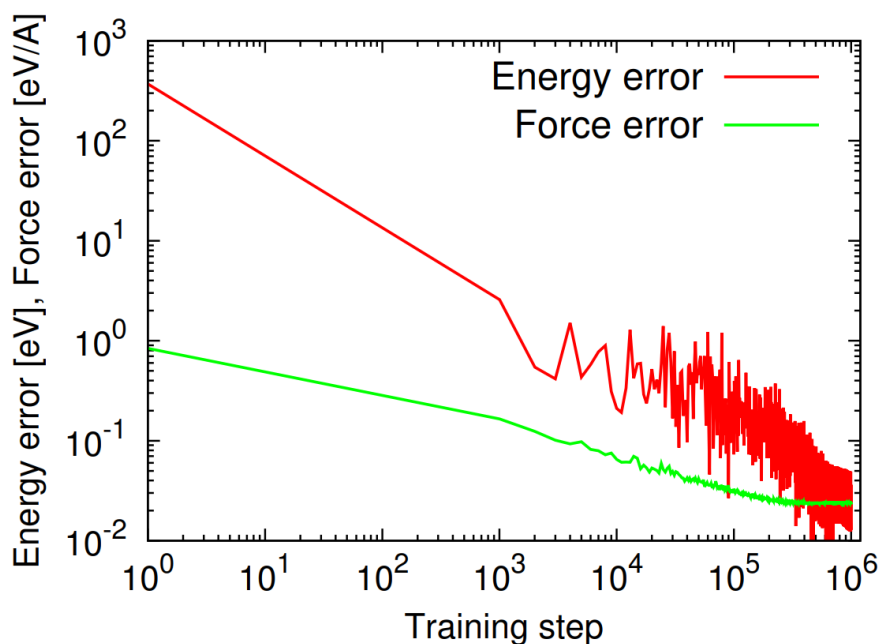


Figura 1.8: RSME da energia e treinamento de uma rede neural para água líquida utilizando pesos iniciais para a energia e força sendo, respectivamente, 0,02 e 1000 e os pesos finais como 8 e 1. Retirado da Referência [4].

### 1.4.2 MACE

Além da abordagem DNN, discutimos a construção do MACE [5]. Como citado anteriormente, o MACE é uma rede neural de passagem de mensagens. Para estes modelos, temos um conceito simples para o aprendizado da rede.

O estado do  $i$ -ésimo nó do grafo na camada  $t$  pode ser representado da seguinte forma

$$\sigma_i^{(t)} = (\mathbf{r}_i, z_i, \mathbf{h}_i^{(t)}), \quad (1.202)$$

onde  $\mathbf{r}_i$  representa a posição,  $z_i$  o respectivo elemento químico e  $\mathbf{h}_i^{(t)}$  são as características aprendidas do nó.

A rede, então, consiste em três etapas: construção da mensagem, atualização da informação e leitura (*readout*). A construção da mensagem  $\mathbf{m}_i^{(t)}$  é feita para cada nó em função de cada vizinho contido no interior do raio de corte  $r_c(i)$

$$\mathbf{m}_i^{(t)} = \bigoplus_{j \in r_c(i)} M_t(\sigma_i^{(t)}, \sigma_j^{(t)}), \quad (1.203)$$

onde  $M_t$  é a função de mensagem aprendível e  $\bigoplus_{j \in r_c(i)}$  é uma operação de *pooling* aprendível e invariante a permutações sobre os vizinhos do átomo  $i$  (por exemplo,

uma soma).

Uma vez obtida a mensagem, na etapa de atualização, ela é utilizada para gerar uma nova representação para o nó, ou seja

$$\mathbf{h}_i^{(t+1)} = U_t(\sigma_i^{(t)}, \mathbf{m}_i^{(t)}), \quad (1.204)$$

com  $U_t$  sendo a função de atualização aprendível.

Após a criação e atualização das mensagens ao longo de  $T$  camadas, a função de leitura  $\mathcal{R}_t$  é utilizada para se encontrar o observável de interesse — neste caso, a energia do átomo  $i$

$$E_i = \sum_{t=1}^T \mathcal{R}_t(\sigma_i^{(t)}). \quad (1.205)$$

Com essa simples construção, podem-se obter redes precisas [5, 57]. Apesar de as MPNNs equivariantes apresentarem melhores resultados em comparação com modelos baseados em descritores locais, seu custo computacional pode ser elevado [56]. Como alternativa, o Multi-ACE (Atomic Cluster Expansion) [151] surge como um *framework* unificado, parametrizando a mensagem  $m_i^{(t)}$  por meio de uma expansão em uma base de funções polinomiais completas e de alta ordem de correlação.

Desta forma, a inovação do MACE reside em uma construção eficiente para a mensagem passada. Expandindo-se as mensagens em uma hierarquia de ordens de correlação, tem-se

$$\mathbf{m}_i^{(t)} = \sum_j \mathbf{u}_1(\sigma_i^{(t)}; \sigma_j^{(t)}) + \sum_{j_1, j_2} \mathbf{u}_2(\sigma_i^{(t)}; \sigma_{j_1}^{(t)}, \sigma_{j_2}^{(t)}) + \cdots + \sum_{j_1, \dots, j_\nu} \mathbf{u}_\nu(\sigma_i^{(t)}; \sigma_{j_1}^{(t)}, \dots, \sigma_{j_\nu}^{(t)}), \quad (1.206)$$

onde as funções  $\mathbf{u}$  são aprendíveis, as somas são realizadas sobre os vizinhos de  $i$ , e  $\nu$  é um hiperparâmetro que corresponde à ordem máxima de correlação.

A construção desta mensagem inicia-se pela representação das interações de dois corpos utilizando uma base radial aprendível, harmônicos esféricos e uma incorporação (*embedding*) aprendível das características dos nós. De forma simplificada, esta construção pode ser dada pela função  $A_{i,klm}^{(t)}$

$$\mathbf{A}_{i,klm}^{(t)} = \sum_{j \in r_c(i)} R_{kl}(r_{ji}) Y_l^m(\hat{\mathbf{r}}_{ji}) \mathbf{W}_{kz_j}^{(t)}, \quad (1.207)$$

onde a matriz  $\mathbf{W}$  contém os pesos aprendíveis que misturam as características do nó vizinho.

A operação fundamental do MACE é a construção eficiente de características de ordens superiores a partir das representações de dois corpos  $\mathbf{A}_i^{(t)}$ . Isso é realizado primeiro através da formação de produtos tensoriais, seguida por uma etapa de simetrização

$$\mathbf{B}_{i,\eta\nu kLM}^{(t)} = \sum_{l_1 m_1, \dots, l_\nu m_\nu} C_{\eta\nu, l_1 m_1 \dots}^{LM} \left( \prod_{\tilde{\zeta}=1}^{\nu} \sum_{\tilde{k}} w_{\tilde{k}\tilde{l}\tilde{\zeta}}^{(t)} \mathbf{A}_{i,\tilde{k}\tilde{l}\tilde{\zeta} m_{\tilde{\zeta}}}^{(t)} \right), \quad (1.208)$$

onde  $C_{\dots}^{LM}$  são os coeficientes generalizados de Clebsch-Gordan, garantindo que  $\mathbf{B}_{i,\eta\nu kLM}^{(t)}$  seja  $L$ -equivariante, com pesos aprendíveis  $w_{\tilde{k}\tilde{l}\tilde{\zeta}}^{(t)}$  e  $\nu$  sendo a ordem de correlação.

Logo, de forma geral, a mensagem pode ser escrita como a seguinte expansão linear

$$\mathbf{m}_{i,kLM}^{(t)} = \sum_{\nu} \sum_{\eta\nu} \mathbf{W}_{z_i, kL, \eta\nu}^{(t)} \mathbf{B}_{i,\eta\nu kLM}^{(t)} \quad (1.209)$$

onde a matriz de pesos  $\mathbf{W}$  também depende da espécie atômica do nó central ( $z_i$ ) e da simetria  $L$ .

Sob condições moderadas, as características de ordem superior  $\mathbf{B}_{i,\eta\nu kLM}^{(t)}$  podem ser interpretadas como uma base completa de interações de muitos-corpos, as quais podem ser computadas a um custo comparável ao de interações par-a-par.

Finalmente, podemos escrever as expressões das funções de atualização e leitura. Para a atualização das informações, temos a seguinte relação

$$\mathbf{h}_{i,kLM}^{(t+1)} = U_{kL}^{(t)}(\sigma_i^{(t)}, \mathbf{m}_i^{(t)}) = \sum_{\tilde{k}} \mathbf{W}_{kL, \tilde{k}}^{(t)} \mathbf{m}_{i, \tilde{k}LM}^{(t)} + \sum_{\tilde{k}} \mathbf{W}_{z_i, kL, \tilde{k}}^{(t)} \mathbf{h}_{i, \tilde{k}LM}^{(t)}. \quad (1.210)$$

Como já apresentado, o passo de leitura é dado através de uma soma sobre as camadas  $t$ . A energia atômica pode ser escrita como

$$E_i = E_i^{(0)} + \dots + E_i^{(T)}, \quad (1.211)$$

onde

$$E_i^{(t)} = \begin{cases} \sum_{\tilde{k}} \mathbf{W}_{\text{readout}, \tilde{k}} \cdot \mathbf{h}_{i, \tilde{k}00}^{(t)} & \text{se } t < T \\ \text{MLP}_{\text{readout}}^{(t)} \left( \left\{ \mathbf{h}_{i, \tilde{k}00}^{(t)} \right\}_k \right), & \text{se } t = T, \end{cases} \quad (1.212)$$

sendo MLP um Perceptron Multicamadas.

Temos, assim, que a leitura depende apenas das características invariantes ( $L = 0$ )  $\mathbf{h}_{i, \tilde{k}00}^{(t)}$  para garantir que as contribuições de energia por sítio também

sejam invariantes. Para preservar a linearidade em relação à ordem de correlação, utilizam-se funções de leitura lineares em todas as camadas, exceto na última, na qual se emprega um MLP.

Nos modelos GNN, as características internas transformam-se de forma específica sob a ação de um determinado grupo de simetria, o que permite reduzir o número de configurações necessárias no processo de treinamento. Além disso, esses modelos apresentam uma diminuição significativa nos erros associados a forças e energias quando comparados a modelos baseados em DNN [5].

Em treinos utilizando o MACE é possível alterar a arquitetura da rede para melhoria dos resultados. Durante o processo de treinamento de um modelo com configurações da molécula de aspirina, no artigo introdutório do código (Ref. [5]), os autores mostraram o erro absoluto médio (MAE) das forças e mostram que aumentar o hiperparâmetro correspondente a correlação,  $\nu$ , e a equivariância, ou seja, aumentar o respeito da rede às simetrias do problema, diminuem o MAE, consequentemente melhorando a precisão do modelo. Isto pode ser observado no gráfico presente na Ref. [5] na Figura 1.9.

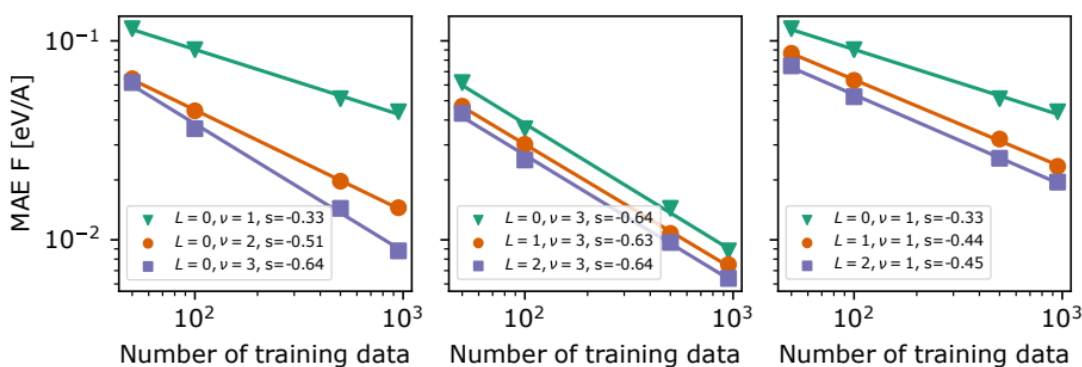


Figura 1.9: MAE da força em função do número de configurações incluídas no treino para a molécula aspirina. Esquerda: mantém a equivariância como  $L = 0$  e varia a correlação. Centro: Mantém a correlação com  $\nu = 3$  e varia a equivariância. Direita: Mantém a correlação com  $\nu = 1$  e varia a equivariância. Para todos os casos,  $s$  representa a inclinação da reta. Retirado da Referência [5].

## 1.5 Integração Termodinâmica

Para o cálculo da diferença de energia livre, pode-se supor um sistema que transita de um estado  $\mathcal{A}$  (descrito pelo potencial  $U_{\mathcal{A}}$ ) para um estado  $\mathcal{B}$  (descrito pelo potencial  $U_{\mathcal{B}}$ ). Esta mudança de potenciais e estados não é, em geral, simples

ou direta. Para isto, pode-se definir um caminho reversível no qual o sistema transita lentamente do estado  $\mathcal{A}$  para o estado  $\mathcal{B}$  [2].

Para que se possa ter um caminho adiabático, é necessário que o sistema esteja relaxado para todos os pontos deste caminho. Para se mover entre os estados neste caminho, introduz-se uma nova variável,  $\lambda$ , que age como uma parametrização externa. Deste modo, pode-se definir a energia potencial do sistema a partir dos potenciais  $U_{\mathcal{A}}$  e  $U_{\mathcal{B}}$ , além do caminho parametrizado por  $\lambda$ , do seguinte modo

$$U(\mathbf{r}, \lambda) \equiv f(\lambda)U_{\mathcal{A}}(\mathbf{r}) + g(\lambda)U_{\mathcal{B}}(\mathbf{r}), \quad (1.213)$$

onde as funções  $f(\lambda)$  e  $g(\lambda)$  definem o caminho adiabático a ser percorrido. Nestes casos, apesar de arbitrárias, elas devem satisfazer as seguintes condições:

- $f(0) = 1$ ;
- $f(1) = 0$ ;
- $g(0) = 0$ ;
- $g(1) = 1$ .

Deste modo, temos que  $U(\mathbf{r}, 0) = U_{\mathcal{A}}$  e  $U(\mathbf{r}, 1) = U_{\mathcal{B}}$ . A maneira mais simples de se definir esta parametrização é dada por  $f(\lambda) = 1 - \lambda$  e  $g(\lambda) = \lambda$ .

Para se encontrar a diferença de energia livre ao longo do caminho definido pelo potencial da equação 1.213, podemos escrever a função de partição para este problema

$$\mathcal{Z}(N, V, T, \lambda) = C_N \int d^N \mathbf{p} d^N \mathbf{r} \exp \left\{ -\beta \left[ \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} + U(\mathbf{r}, \lambda) \right] \right\}. \quad (1.214)$$

A partir da função de partição do sistema, pode-se, de modo simples, encontrar a energia livre de Helmholtz

$$A(N, V, T, \lambda) = -k_B T \ln \mathcal{Z}(N, V, T, \lambda). \quad (1.215)$$

Derivando-se a expressão da energia livre em função de  $\lambda$ , obtém-se

$$\frac{\partial A}{\partial \lambda} = -\frac{k_B T}{\mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial \lambda}. \quad (1.216)$$

Usando o resultado da equação 1.214, temos uma expressão para o segundo termo da equação 1.216

$$\frac{\partial \mathcal{Z}}{\partial \lambda} = -\beta \int d^N \mathbf{p} d^N \mathbf{r} \left( \frac{\partial U}{\partial \lambda} \right) e^{-\beta H(\mathbf{p}, \mathbf{r}, \lambda)}. \quad (1.217)$$

Combinando as equações 1.216 e 1.217, chega-se a uma expressão final para a derivada da energia livre em função de  $\lambda$

$$\frac{\partial A}{\partial \lambda} = \left\langle \frac{\partial U}{\partial \lambda} \right\rangle_{\lambda}. \quad (1.218)$$

A diferença de energia livre entre o estado  $\mathcal{A}$  e o estado  $\mathcal{B}$  é dada por

$$\Delta A_{\mathcal{A} \rightarrow \mathcal{B}} = A_{\mathcal{B}} - A_{\mathcal{A}} = \int_0^1 \left\langle \frac{\partial U(\mathbf{r}, \lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda. \quad (1.219)$$

Neste caso, faz-se a média sobre o ensemble para um  $\lambda$  fixo. Voltemos para as expressões previamente definidas para  $f(\lambda)$  e  $g(\lambda)$ . Temos que o potencial em função de  $\lambda$  pode ser escrito da seguinte forma

$$U(\mathbf{r}, \lambda) = (1 - \lambda)U_{\mathcal{A}}(\mathbf{r}) + \lambda U_{\mathcal{B}}(\mathbf{r}). \quad (1.220)$$

Aplicando a derivada, chega-se à expressão final para a variação de energia livre

$$\Delta A_{\mathcal{A} \rightarrow \mathcal{B}} = \int_0^1 \langle U_{\mathcal{B}} - U_{\mathcal{A}} \rangle_{\lambda} d\lambda. \quad (1.221)$$

A este método dá-se o nome de integração termodinâmica, que pode ser empregada para diversos cálculos de variação da energia livre.

### 1.5.1 Energia de Ligação

Uma possível aplicação é o cálculo da energia livre de ligação entre dois sistemas. Na Figura 1.10, mostramos o esquema de passos para o cálculo da energia de ligação. Fica claro, observando a figura, que  $\Delta A_{\text{lig}}$  pode ser obtido por meio de um ciclo termodinâmico. O estado não ligado corresponde ao ponto  $\lambda = 0$ , enquanto o estado ligado corresponde ao ponto  $\lambda = 1$ .

A implementação deste algoritmo é relativamente simples, uma vez que há apenas a necessidade de se calcular a diferença de energia para cada valor de  $\lambda$  e, então, integrar numericamente. Entretanto, há a necessidade de se fazer uma boa

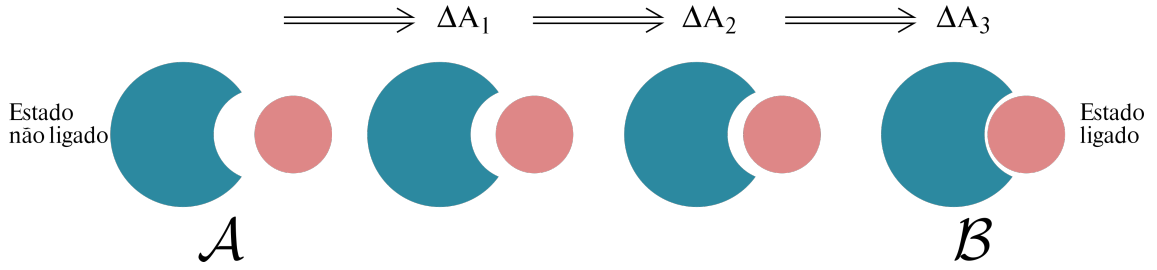


Figura 1.10: Esquema para o ciclo termodinâmico utilizado no cálculo da energia livre de ligação entre dois sistemas.

escolha dos valores de  $\lambda$  para a integração numérica, especialmente em regiões onde a derivada  $\langle \partial U / \partial \lambda \rangle$  varia rapidamente.

Com isto, a expressão para a energia de ligação entre dois sistemas clássicos, A e B, pode ser escrita da seguinte maneira

$$\Delta A_{\text{lig}} = \Delta A_{\text{complexo}} - (\Delta A_A + \Delta A_B). \quad (1.222)$$

Ao se tratar de sistemas quânticos (como os apresentados na seção 1.3), pode-se utilizar a mesma metodologia empregada para sistemas clássicos. Contudo, isso demanda um alto custo computacional, não sendo prático. Desta forma, pode-se utilizar a abordagem feita por Rossi *et al.* [6], na qual se aproveita a integração termodinâmica no caso clássico e, a partir deste, obtêm-se as energias livres dos estados quânticos por meio de uma correção. Um esquema na Figura 1.11 nos apresenta o caminho feito por esta abordagem.

Com isto, a expressão apresentada na equação 1.222 deve ser reescrita, a fim de incluir os novos termos de correção quântica,  $\delta^{q-c}$

$$\Delta A_{\text{lig}}^{\text{quant}} = \Delta A_{\text{lig}}^{\text{class}} + \delta_{\text{complexo}}^{q-c} - (\delta_A^{q-c} + \delta_B^{q-c}), \quad (1.223)$$

onde os subscritos representam o sistema complexado e os sistemas isolados.

A forma mais simples de se transitar de um sistema quântico para um sistema clássico é modificar as massas atômicas, incluindo um fator multiplicativo,  $m_i \rightarrow \mu m_i$ . Ao se tender a massa de um átomo ao infinito, recupera-se o limite clássico. Isto pode ser demonstrado a partir da variância típica entre *beads*, onde

$$\langle (q_j - q_{j+})^2 \rangle \sim \frac{k_B T}{k_s}, \quad (1.224)$$

com  $k_s = m(P/(\beta\hbar))^2$ . Deste modo

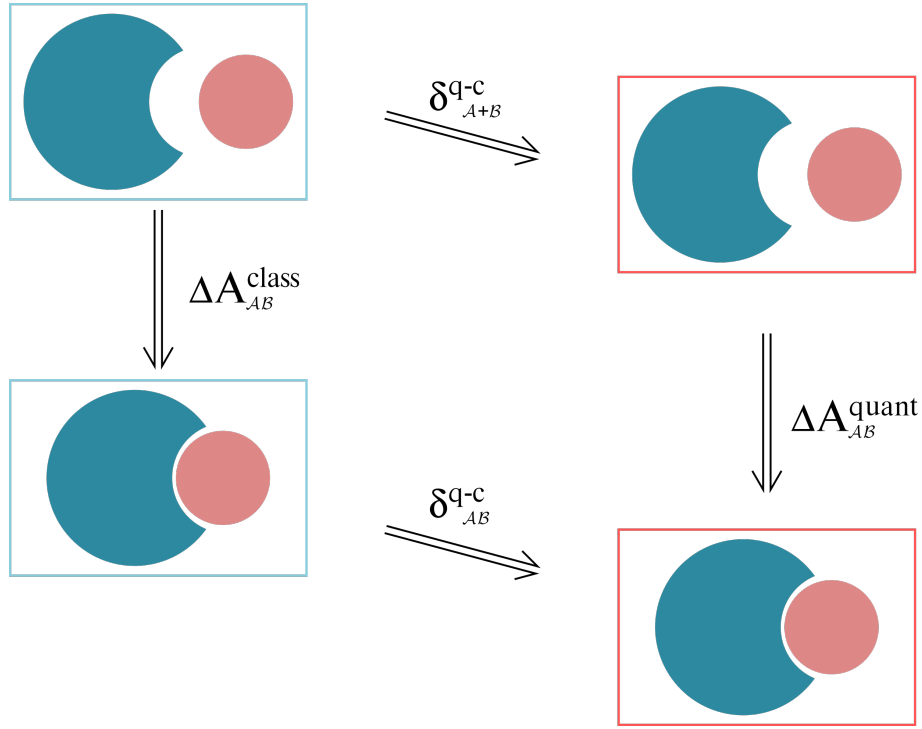


Figura 1.11: Esquema para o cálculo da energia de ligação em simulações quânticas, utilizado por Rossi *et al.* [6] O contorno azul representa as simulações clássicas e o contorno vermelho, as simulações quânticas.

$$\langle (q_j - q_{j+})^2 \rangle \propto \frac{1}{m}, \quad (1.225)$$

onde no limite  $m \rightarrow 0$ , todas as *beads* convergem para um único ponto.

Sendo assim, temos a seguinte equação para encontrar os termos  $\delta$

$$\delta^{q-c} = A^{\text{quant}} - A^{\text{class}} = \int_m^\infty d\mu \left\langle \frac{\partial \mathcal{H}}{\partial \mu} \right\rangle = \int_m^\infty \frac{d\mu}{\mu} \langle T(\mu) - T^{\text{class}} \rangle, \quad (1.226)$$

onde  $\mu$  representa as massas atômicas rescaladas,  $T(\mu)$  é a energia cinética do sistema quântico para uma determinada massa e  $T^{\text{class}}$  é a energia cinética do sistema clássico considerando as massas reais. Como, numericamente, não é factível se utilizar uma massa infinita, pode-se reescrever a integral em função de um fator de escala de massa,  $g = \sqrt{m_0/\mu}$ . Com a mudança de variável, a expressão pode ser reescrita da seguinte forma

$$\delta^{q-c} = \int_0^1 \frac{2}{g} \left( \langle T(m_0/g^2) \rangle - \langle T^{\text{class}} \rangle \right) dg, \quad (1.227)$$

onde  $m_0$  é a massa do sistema clássico.

# Capítulo 2

## Testes de Potenciais

A construção de uma rede neural envolve diversas etapas, principalmente a elaboração do *dataset* e a escolha dos parâmetros. Portanto, antes de calcular qualquer propriedade de um sistema, é necessário realizar testes relacionados à convergência de parâmetros e à qualidade dos dados que compõem o treino da rede. Devido ao fato de ser uma área de interesse recente, ainda não há um conjunto de parâmetros padrão para o treinamento dos modelos.

Neste capítulo, testamos hipóteses que julgamos essenciais para a construção de uma metodologia sólida, visando à obtenção de uma rede neural precisa com o menor custo computacional possível, uma vez que, em nosso contexto, isso representa um gargalo. Investigamos a qualidade da rede para *datasets* construídos com distintos potenciais de correlação e troca, analisando os seguintes parâmetros: o número de passos de treino, o número de moléculas de água, a extensão do raio de corte e o número de configurações para o treino. Além disso, discutimos a importância da inclusão de dados de gelo para o cálculo das propriedades da fase sólida. O procedimento para a construção dos *datasets* está descrito na referência [42]. Apenas a análise referente ao número de configurações utiliza o *software* MACE, enquanto as demais são feitas a partir de modelos treinados com o DeePMD-kit.

Para a construção do *dataset*, Torres *et al.* [42] fizeram simulações com a inclusão de NQE com o campo clássico q-TIP4P/F [152] em diferentes temperaturas (300 e 600 K) e densidades (0,88, 1,0 e 1,2 g/cm<sup>3</sup>), gerando 18.000 configurações selecionadas que capturam diferentes regiões do espaço de fases. Essas geometrias foram posteriormente avaliadas em nível *ab initio* usando DFT com diferentes funcionais. O procedimento assegura que a rede neural não fique restrita a uma única região do espaço de fases, evitando sobreajuste e garantindo maior generalização.

Para as análises referentes ao número de passos de treino, tamanho do raio de corte, número de moléculas de água e inclusão de configurações de gelo, foi utilizada a versão para GPU do *software* DeePMD-kit v2.1.3 [4, 93, 153] para a geração dos potenciais de rede neural profunda (DNN). Empregou-se o descritor

*Deep Potential-Smooth Edition*, no qual todas as coordenadas relativas (distâncias e ângulos) são consideradas na construção do descritor [154].

A arquitetura da rede foi composta por quatro camadas com 320, 160, 32 e 16 neurônios, respectivamente, utilizando como função de ativação a tangente hiperbólica. A função de perda (*loss function*), definida como a soma dos erros quadráticos médios das energias e forças, foi minimizada pelo otimizador Adam. Os parâmetros adotados para a função de perda estão listados na Tabela 2.1.

Parâmetro	Valor
start_pref_e	0.02
limit_pref_e	8.00
start_pref_f	1500
limit_pref_f	1.00

Tabela 2.1: Parâmetros da função de perda (*loss function*) empregados no treinamento.

Para a divisão dos dados, 90% das configurações foram destinadas ao treinamento e 10% à validação, totalizando 17.100 estruturas de água utilizadas durante o treinamento. Os cálculos foram realizados em um único processador AMD Epyc 7662 com uma GPU NVIDIA Tesla A100. Após o término, os modelos foram comprimidos com o objetivo de otimizar o desempenho computacional [155]. A descrição do treinamento dos modelos utilizados para a análise referente ao número de configurações está presente na respectiva seção.

## 2.1 Número de Passos de Treino

O primeiro parâmetro a ser analisado é o número de passos de treino. Para esta análise, utilizamos dados obtidos com dois diferentes funcionais de correlação e troca: optB88 [30] e vdW-cx [31]. Para cada rede, 90% dos dados foram utilizados para o treino efetivo do modelo, enquanto os 10% restantes foram usados para testar a precisão da rede.

Para avaliar o número de passos necessários, foram treinadas sete redes distintas no intervalo entre  $1 \times 10^6$  e  $7 \times 10^6$  passos, ou seja, atualizamos os pesos da rede após a rede aprender uma configuração, uma vez que o *batch size* é igual a 1. Para o treino, utilizamos o DeePMD-kit com aceleração em GPU [4, 93, 153] e o descritor *Deep Potential-Smooth Edition* [154], com um raio de corte de 6.0 Å e raio de suavização de 0.5 Å.

A primeira forma de avaliar a qualidade da rede é calcular o RMSE do modelo para as forças e energias em relação aos dados de teste. O RSME é dada por

$$RSME = \sqrt{\frac{1}{N} \sum_i (A_i^{DFT} - A_i^{NN})^2}, \quad (2.1)$$

onde  $A_i$  representa o observável de interesse para cada configuração. Adicionalmente, devemos observar se, em uma dinâmica longa, a energia total do sistema mantém-se em equilíbrio durante todo processo de produção, uma vez que um dos principais interesses em potenciais baseados em redes neurais é a capacidade de simular sistemas por longos períodos.

Nesta etapa, realizamos uma *deep potential molecular dynamics* (DP-MD) para cada modelo em uma caixa contendo 64 moléculas de água, utilizando o código LAMMPS [40]. O processo de termalização foi dividido em duas etapas: primeiramente, uma simulação NVT de 30 ps com o termostato *Canonical sampling through velocity rescaling* (CSVR) [156] a 300 K; em seguida, uma dinâmica NPT durante 150 ps na mesma temperatura e com pressão de 1 bar, utilizando o termostato de Nosé-Hoover [129, 130]. Com o sistema termalizado, fizemos uma simulação de 2 ns no mesmo ponto do diagrama de fases para avaliar a evolução da energia durante o tempo de simulação. Para todas as dinâmicas, utilizamos um *timestep* de 0.25 fs, e os tempos de relaxação foram mantidos em 25 fs para a temperatura e 250 fs para a pressão.

Primeiramente, apresentamos na Figura 2.1 a comparação entre as energias e forças obtidas com a rede e com cálculos DFT para ambos os funcionais, com redes treinadas durante  $1 \times 10^6$  passos. Como observado, o gráfico apresenta um comportamento linear, o que indica que a rede possui boa precisão quando comparada ao resultado obtido com DFT para as mesmas configurações. A seguir, na Tabela 2.2, apresentamos os erros de energia e força para cada modelo, considerando os diferentes números de passos de treino.

Para obter um modelo preciso, definimos que os erros devem estar abaixo de 0.55 meV/átomo para energia e 50 meV/Å para forças. Como visto, os modelos treinados com  $1 \times 10^6$  passos são os únicos que possuem uma discrepância superior ao limite considerado para a energia. Desse modo, a análise do RMSE é o primeiro passo para avaliar a qualidade de um modelo.

Dando continuidade, após uma longa dinâmica de 2 ns, verificamos que os modelos treinados com  $5 \times 10^6$  passos ou mais concluíram as simulações com sucesso (exceto o modelo treinado com o funcional vdW-cx e  $7 \times 10^6$  passos),

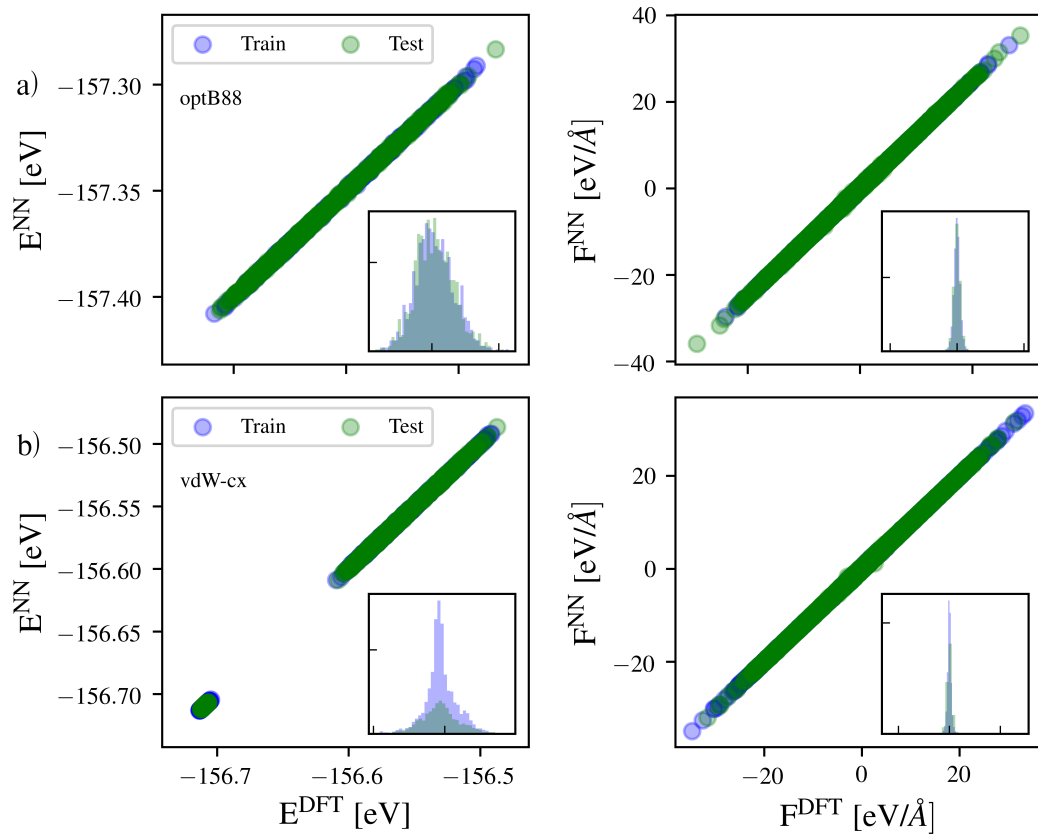


Figura 2.1: Comparação das energias (painel esquerdo) e forças (painel direito) obtidas com a rede neural treinada com  $1 \times 10^6$  passos e com DFT para os funcionais (a) optB88 e (b) vdW-cx.

Número de passos	RMSE Energia (meV/átomo)		RMSE Forças (meV/Å)	
	optB88	vdW-cx	optB88	vdW-cx
$1 \times 10^6$	0.58	0.59	48.7	50.3
$2 \times 10^6$	0.55	0.54	47.1	47.2
$3 \times 10^6$	0.51	0.54	45.4	47.5
$4 \times 10^6$	0.53	0.54	44.5	47.3
$5 \times 10^6$	0.49	0.49	44.1	44.0
$6 \times 10^6$	0.48	0.47	43.9	42.8
$7 \times 10^6$	0.48	0.51	42.8	44.9

Tabela 2.2: RMSE para energias e forças dos modelos em comparação com dados DFT para cada número de passos de treino.

mantendo a energia em equilíbrio durante todo o período analisado. Na Tabela 2.3, apresentamos o resultado das simulações para cada modelo. Verificamos que modelos treinados por poucos passos não são capazes de inferir uma boa

representação para a energia, fazendo com que ela não estabilize, ou mantendo mais de um patamar de estabilização para o mesmo sistema. Na Figura 2.2, exibimos os gráficos da energia do sistema para um modelo que falha e outro que obtém êxito em conservar a energia.

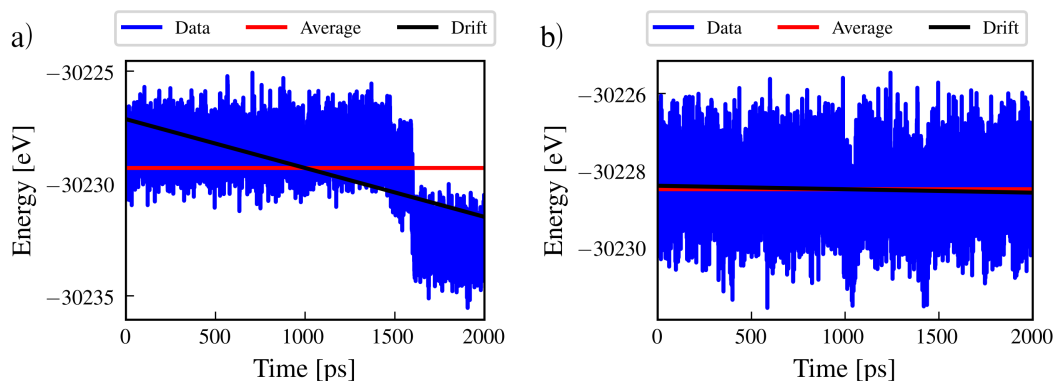


Figura 2.2: Energia total em função do tempo para modelos treinados com (a)  $1 \times 10^6$  e (b)  $5 \times 10^6$  passos. Dados obtidos para modelos com o funcional optB88.

Número de passos	vdW-cx	optB88
$1 \times 10^6$	×	×
$2 \times 10^6$	×	×
$3 \times 10^6$	✓	✓
$4 \times 10^6$	×	×
$5 \times 10^6$	✓	✓
$6 \times 10^6$	✓	✓
$7 \times 10^6$	×	✓

Tabela 2.3: Tabela de sucesso de dinâmicas de 2 ns para cada modelo treinado com diferentes números de passos. O marcador verde (✓) indica que a energia foi conservada durante a simulação, enquanto o marcador vermelho (×) indica o oposto.

Com base nesse resultado, apesar de os modelos treinados com mais de  $5 \times 10^6$  passos concluírem a simulação mantendo a energia total em equilíbrio, sugerimos treinar os modelos por  $7 \times 10^6$  passos para garantir maior robustez.

## 2.2 Raio de Corte

Dando sequência à análise dos parâmetros, passamos a observar a dependência da qualidade do modelo em relação ao raio de corte, onde é coletada a informação

dos átomos no interior da região definida por este raio em relação ao átomo centrado na origem desta região. Considerando que o parâmetro de rede de uma caixa com 64 moléculas de água é 12,43 Å, expandir o raio de corte para valores superiores a 6.5 Å torna-se redundante, porém não é possível afirmar o mesmo para um sistema com 256 moléculas.

Para essa análise, utilizamos o processo descrito por Torres *et al.* [42] para gerar 18.000 configurações de uma caixa contendo 256 moléculas de água. A rede neural foi gerada com o funcional PBE [65]. Os dados de energia e força DFT foram coletados com o *software* SIESTA [157], utilizando uma base otimizada *double- $\zeta$*  polarizada (DZP) [158] e um pseudopotencial de norma conservada [114], com um *mesh-cutoff* de 500 Ry.

Com os dados dos cálculos DFT, treinamos redes neurais com raios de corte entre 5 Å e 8 Å, utilizando o DeePMD-kit com aceleração em GPU [4, 93, 153], o descritor *Deep Potential-Smooth Edition* [154] e um raio de suavização de 0.5 Å. Conforme definido anteriormente, os modelos foram treinados por  $7 \times 10^6$  passos. Para avaliar a qualidade dos modelos, calculamos seus erros em relação aos resultados de DFT. Para obter as propriedades estruturais, realizamos uma DP-MD com o mesmo protocolo anterior, utilizando o LAMMPS [40], mas com a etapa de produção estendida para 500 ps.

Após o treino de cada rede, calculamos o RMSE de cada modelo em relação aos dados DFT. Os resultados são apresentados na Figura 2.3 e na Tabela 2.4. Observamos que os modelos com raio de corte de 5.5 Å e 6.5 Å não apresentam erros de energia abaixo do estipulado. Por outro lado, apenas o modelo com raio de corte de 5.0 Å não atendeu ao critério de RMSE para as forças.

Raio de corte (Å)	RMSE Energia (meV/átomo)	RMSE Forças (meV/Å)
5.0	0.30	52.6
5.5	0.73	50.0
6.0	0.54	48.8
6.5	0.64	47.7
7.0	0.26	45.9
7.5	0.26	45.8
8.0	0.27	45.8

Tabela 2.4: RMSE para energias e forças dos modelos em comparação com dados DFT para cada raio de corte.

Apesar de alguns modelos apresentarem erros acima do limite, prosseguimos com todos para investigar a precisão das propriedades estruturais. Durante as

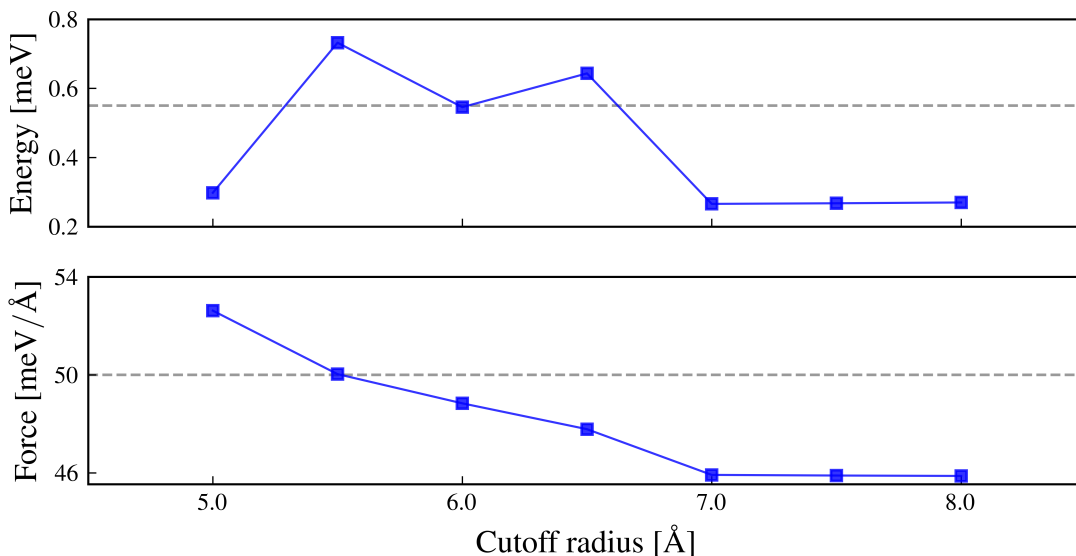


Figura 2.3: RMSE para cada modelo treinado com diferentes raios de corte.

dinâmicas, verificamos que o modelo com menor raio de corte (5.0 Å) não atinge o equilíbrio, nem durante a termalização nem nos 500 ps de produção. Já os modelos com raios de corte de 7.0 Å e 8.0 Å levam um longo tempo, acima de 200 ps para equilibrar a energia, ou seja, durante a etapa de produção. Desse modo, esses três modelos foram desconsiderados.

Para os modelos que tiveram êxito, verificamos se, durante a produção, eles reproduzem as seguintes propriedades: distância O-H intramolecular, ângulo H-O-H, densidade e a função de distribuição radial ( $g(r)$ ) para todos os pares. Para medir a precisão da  $g(r)$ , utilizamos a equação proposta por Schran e colaboradores [159]:

$$\epsilon_{ij} = 1 - \frac{\int_0^{+\infty} |g_{ij}^{\text{EXP}}(r) - g_{ij}^{\text{DP}}(r)| dr}{\int_0^{+\infty} g_{ij}^{\text{EXP}}(r) dr + \int_0^{+\infty} g_{ij}^{\text{DP}}(r) dr}, \quad (2.2)$$

onde o subscrito  $ij$  representa os pares atômicos (O-O, O-H e H-H) e os superscritos  $DP$  e  $EXP$  representam os dados da rede neural e do experimento de Soper [10], respectivamente.

As comparações entre os resultados de cada modelo e os dados experimentais estão na Tabela 2.5. É possível notar que mesmo os modelos cujos erros não satisfazem as condições impostas reproduzem resultados similares para as propriedades estruturais. Outro ponto de interesse é a discrepância entre a densidade da água obtida com o nível de teoria PBE e os dados experimentais, resultado que será

discutido em seções futuras. Na Figura 2.4, apresentamos a evolução temporal da densidade de equilíbrio e as distribuições do comprimento da ligação covalente e do ângulo H-O-H para o modelo com raio de corte de 6.0 Å.

	Raio de Corte (Å)			Exp.
	5.5	6.0	7.5	
$d_{OH}$ (Å)	0.99	0.99	0.99	0.99 [160]
$\theta_{HOH}$ (°)	105	105	105	104.47 [161]
$\rho$ (g/cm <sup>3</sup> )	0.890	0.889	0.890	0.997 [27]
$\epsilon_{OO}$ (%)	91.2	90.5	90.5	-
$\epsilon_{OH}$ (%)	83.7	83.5	83.1	-
$\epsilon_{HH}$ (%)	93.7	93.5	93.2	-

Tabela 2.5: Comparação do comprimento da ligação covalente, ângulo H-O-H, densidade de equilíbrio e precisão da função de distribuição radial para todos os pares para os modelos treinados com diferentes raios de corte em relação aos dados experimentais. O desvio padrão do comprimento de ligação é  $\pm 0.03$  Å, para o ângulo é  $\pm 5^\circ$  e para a densidade a incerteza é de  $\pm 0.008$  g/cm<sup>3</sup>. O valor experimental de  $\theta_{HOH}$  refere-se à fase gasosa em condições ambientes.

Comparando os modelos, observa-se que não há diferença significativa entre eles, exceto nos erros. Assim, qualquer escolha seria viável. Contudo, como definimos critérios de erro, apenas os modelos com raio de corte de 6.0 Å e 7.5 Å são considerados satisfatórios. Para a escolha final, devemos considerar também o custo computacional. Verificamos que o modelo com maior raio de corte (8.0 Å) levou 43% mais tempo para treinar em comparação com o de menor raio de corte (5.0 Å). Isso se deve ao maior número de vizinhos incluídos na análise de cada átomo. Portanto, utilizar o modelo com raio de corte de 6.0 Å atende às nossas expectativas de precisão e eficiência.

## 2.3 Número de Moléculas de Água

Uma vez que o raio de corte de 6.0 Å se mostrou adequado, podemos analisar se o número de moléculas de água em cada configuração interfere na qualidade do modelo. Para isso, utilizando o mesmo protocolo da rede com 256 moléculas e os dados de Torres *et al.* [42], treinamos uma rede com configurações de 64 moléculas no nível PBE. Para diferenciá-las, chamaremos a rede com 64 moléculas de PBE-64 e a com 256 moléculas de PBE-256. Para ambas as redes, foram utilizadas 18.000

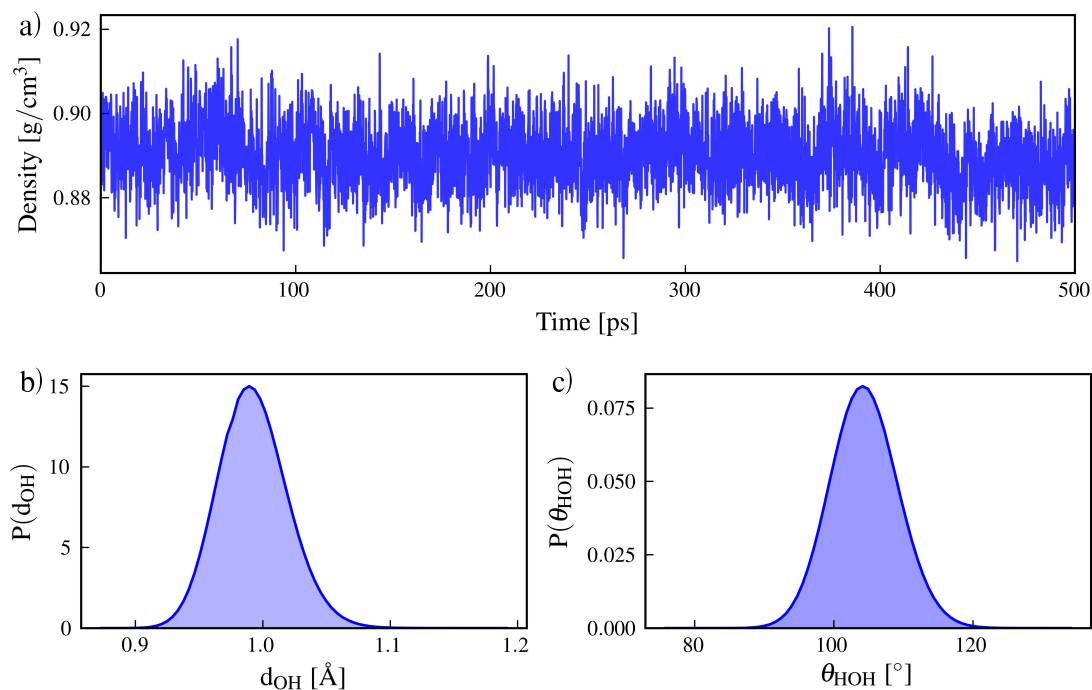


Figura 2.4: (a) Densidade de equilíbrio e distribuições (b) do comprimento da ligação covalente e (c) do ângulo H-O-H para a rede treinada com raio de corte de 6.0 Å.

configurações para o treinamento, sendo 90% para o conjunto de treino e o restante para o conjunto de teste.

Com as duas redes, realizamos simulações para caixas com 64 e 256 moléculas. Um dos objetivos das redes neurais é escalar as interações de um sistema pequeno para reproduzi-las em sistemas consideravelmente maiores. Portanto, usando o mesmo protocolo de dinâmica molecular da seção anterior, comparamos as densidades de equilíbrio das duas redes para ambos os sistemas, além de comparar a  $g(r)$  da rede PBE-64 com a referência da rede PBE-256 para a caixa com 256 moléculas. Na Tabela 2.6, apresentamos a densidade de equilíbrio para os dois modelos em diferentes volumes (número de moléculas). Na Figura 2.5, é mostrada a evolução temporal da densidade. Verificamos que ambos os modelos possuem precisão similar na densidade de equilíbrio, indicando que modelos com menos dados de treino podem obter resultados semelhantes aos que usam um número maior de moléculas.

Além disso, para ambos os modelos, as RDFs são praticamente idênticas, com precisão acima de 95% para a simulação com 64 moléculas e cerca de 99% de fidelidade entre as RDFs para a caixa com 256 moléculas. Dessa forma, um modelo

Modelo	Densidade ( $\text{g}/\text{cm}^3$ )	
	64 moléculas	256 moléculas
PBE-64	$0.90 \pm 0.02$	$0.896 \pm 0.008$
PBE-256	$0.89 \pm 0.02$	$0.889 \pm 0.008$
Experimental [10]	0,917	

Tabela 2.6: Densidade de equilíbrio para os modelos PBE-64 e PBE-256 para diferentes números de moléculas de água na caixa de simulação.

treinado com poucas moléculas pode capturar as propriedades estruturais de um sistema maior, o que permite reduzir o custo computacional tanto dos cálculos *ab initio* quanto do processo de treinamento.

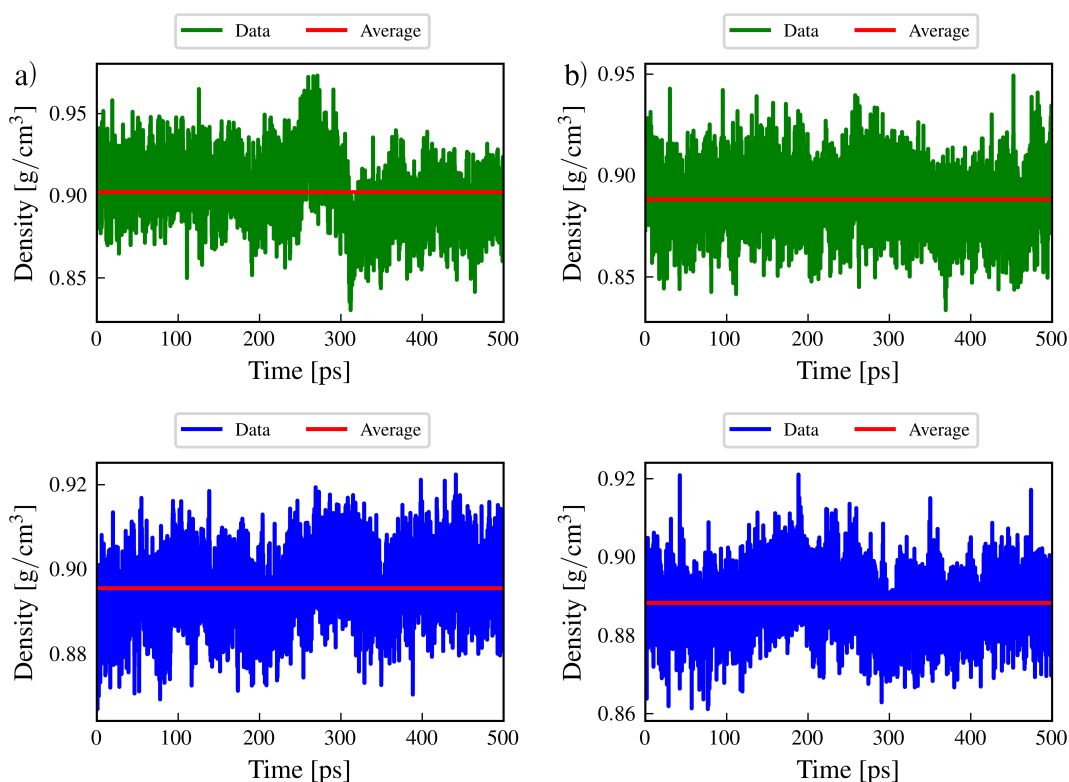


Figura 2.5: Evolução temporal da densidade para os sistemas com (a) 64 e (b) 256 moléculas na caixa de simulação. As curvas em verde representam o modelo PBE-64 e, em azul, o modelo PBE-256.

## 2.4 Inclusão de Configurações de Gelo

Até o momento, observamos o comportamento de um modelo treinado apenas com dados de água líquida. Surge, então, a seguinte questão: uma rede treinada com dados de água líquida reproduz apenas resultados para essa fase ou para qualquer sistema formado por moléculas de H<sub>2</sub>O?

Inicialmente, utilizamos o *dataset* construído com dados do funcional SCAN [34] por Torres *et al.* [42]. A rede foi treinada com o mesmo protocolo anterior, e o modelo foi denominado SCAN-W. De forma similar, treinamos dois outros modelos incluindo 1.000 configurações de gelo no *dataset*. No primeiro (SCAN-WI), as configurações foram geradas por dinâmica molecular a volume constante. No segundo (SCAN-WI2), as configurações foram obtidas por dinâmica a pressão constante. Em ambos os casos, o modelo SCAN-W foi usado na DP-MD com temperatura de 220 K. Para o modelo SCAN-WI, as configurações possuem densidade próxima ao valor experimental para essa temperatura (0.92 g/cm<sup>3</sup>), enquanto para o SCAN-WI2, a pressão foi mantida em 1 bar. Forças e energias foram obtidas com o *software* VASP [162], com um *cutoff* de energia de 118 Ry e o método PAW (*Projected Augmented Wave*) [163].

Após o treinamento, realizamos uma simulação de dinâmica molecular NPT para cada modelo em um sistema de água com 64 moléculas (300 K, 1 bar) e em uma estrutura de gelo hexagonal com 96 moléculas (220 K, 1 bar). Em todas as simulações, os sistemas foram equilibrados em duas etapas: mais de 30 ps no ensemble NVT (termostato CVSR [156]) e 150 ps no ensemble NPT (termostato de Nose-Hoover [129, 130]), seguidos por uma etapa de produção de 500 ps, com *timestep* de 0.25 fs e coleta de dados a cada 0.1 ps. As simulações foram realizadas com o código LAMMPS [40].

Na Figura 2.6, apresentamos o erro dos três modelos para configurações de água, gelo e para o conjunto total de dados usado no treino do SCAN-WI2. Como esperado, o modelo SCAN-W não atinge o critério de RMSE para a energia do gelo, o que demonstra a necessidade de incluir configurações dessa fase. No entanto, quando incluímos configurações de gelo com densidade fixa (SCAN-WI), o modelo apresenta erros maiores e não se mostra adequado para obter as propriedades de nenhuma das fases.

A seguir, discutimos as propriedades da água e do gelo. O modelo SCAN-WI não completou a simulação para nenhum dos sistemas, e a energia total não foi conservada. A Tabela 2.7 mostra que a inclusão de configurações de gelo no

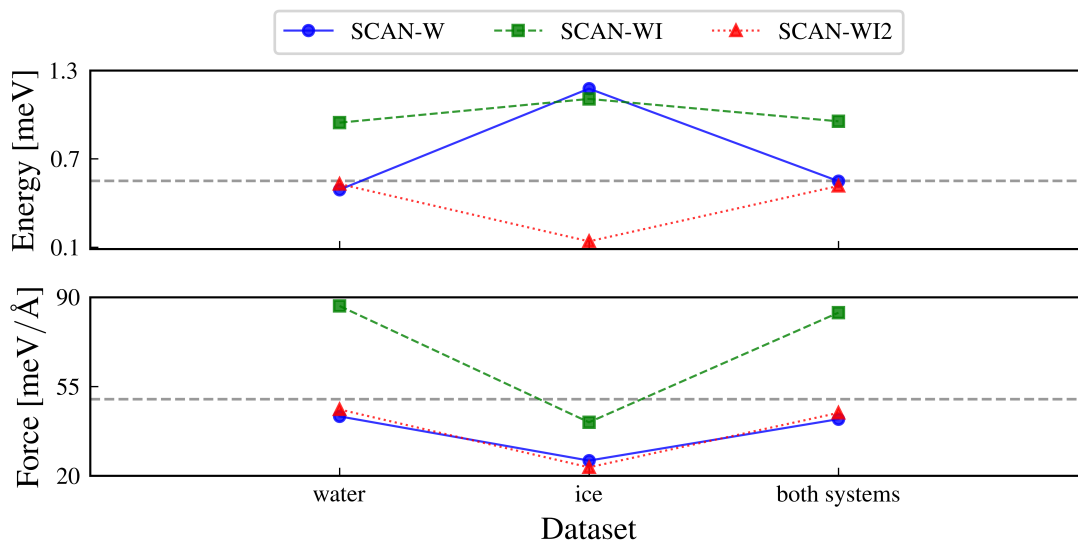


Figura 2.6: RMSE para as energias (painel superior) e forças (painel inferior) dos modelos SCAN-W, SCAN-WI e SCAN-WI2 para os dados de treino de água e gelo.

modelo SCAN-WI2 manteve a qualidade da simulação de água. Além disso, embora as densidades do gelo sejam próximas, o SCAN-WI2 descreve a energia total do sistema de gelo com maior precisão que o SCAN-W. Portanto, a inclusão de configurações diversas traz benefícios para a construção do modelo.

Propriedade	SCAN-W		SCAN-WI2		Experimental	
	Água	Gelo	Água	Gelo	Água	Gelo
$d_{OH}$ (Å)	0,98	0,99	0,98	0,98	0,99 [160]	1,008 [164]
$\theta_{HOH}$ (°)	105	106	105	106	104,47 [161]	104,52 [164]
$\rho$ (g/cm <sup>3</sup> )	1,01	0,964	1,00	0,974	0,999 [10]	0,917 [10]

Tabela 2.7: Comprimento da ligação  $d_{OH}$ , ângulo  $\theta_{HOH}$  e densidade de equilíbrio  $\rho$  para os modelos SCAN-W e SCAN-WI2. O desvio padrão para as simulações de água é  $\pm 0,03$  Å (ligação),  $\pm 5^\circ$  (ângulo) e  $\pm 0,008$  g/cm<sup>3</sup> (densidade). Para o gelo, os desvios são  $\pm 0,02$  Å (ligação),  $\pm 4^\circ$  (ângulo) e  $\pm 0,008$  g/cm<sup>3</sup> (densidade).

## 2.5 Número de Configurações

O tamanho do *dataset* também é um fator relevante. Recentemente, demonstrou-se que modelos treinados com MACE [5, 55] com menor número de configurações reproduziram resultados de *datasets* maiores [165]. Para analisar a quantidade de

configurações necessárias, treinamos três redes com dados de Torres *et al.* [42] usando o funcional SCAN [34]. Estudos anteriores, utilizando o DeePMD-kit, mostraram que, com o mesmo *dataset*, 10% dos dados são suficientes para reproduzir propriedades estruturais, enquanto 30% são necessários para as propriedades dinâmicas [66].

A escolha das configurações inclusas no *dataset* é de suma importância. Por exemplo, a escolha das configurações utilizadas neste trabalho e feita por Torres *et al.* [42] foi feita extraíndo as configurações de uma simulação longa, totalizando 60.000 configurações totais. Em seguida, foi utilizada a métrica de Jensen–Shannon para selecionar as 18.000 configurações finais. Deste modo, é possível abranger diferentes pontos do espaço de fase.

Utilizamos o código MACE [5, 55] para treinar modelos com 540, 1.080 e 1.620 configurações, ao longo de 500 épocas, com a opção SWA ativada após 250 épocas. Similarmente aos outros modelos, selecionamos um raio de corte de 6.0 Å. Para cada *dataset*, 5% das configurações foram usadas para validação. Os demais parâmetros foram mantidos de acordo com a sugestão presente na documentação do *software* [5].

Para avaliar o impacto do número de configurações, utilizamos o *software* i-Pi [150] para realizar simulações DP-MD. A temperatura foi controlada com o termostato SVR [156] e a pressão com o barostato isotrópico de Bussi-Zykova-Parrinello [166]. O sistema foi simulado a 300 K e 1 bar por 2 ns, sendo os primeiros 100 ps para termalização. O *timestep* foi de 0.5 fs, com parâmetros de amortecimento de 40 fs (temperatura) e 200 fs (pressão). Para as propriedades dinâmicas, realizamos cinco simulações distintas para obter a média.

Para finalizar a discussão, apresentamos na Tabela 2.8 os erros de cada modelo treinado com MACE.

RMSE	N° de configurações		
	540	1080	1620
Energia (meV/átomo)	0.2	0.2	0.2
Força (meV/Å)	13.5	10.8	10.1

Tabela 2.8: RMSE para energia e força dos modelos treinados com diferentes números de configurações.

Os modelos treinados com MACE apresentam valores de RMSE menores que os do DeePMD-kit, mesmo com menor número de configurações. A redução no

tamanho do *dataset* não resultou em um modelo ruim, e todas as simulações foram concluídas com um pequeno *drift* na densidade ao longo do tempo de simulação, como visto na Figura 2.7. A Tabela 2.9 mostra que a densidade de equilíbrio é similar para todos os modelos MACE e apresenta pequeno desvio em relação ao modelo treinado com 18.000 configurações (SCAN-W, Tabela 2.7). Vale destacar que a simulação DP-MD com MACE é computacionalmente mais custosa que a com DeePMD-kit.

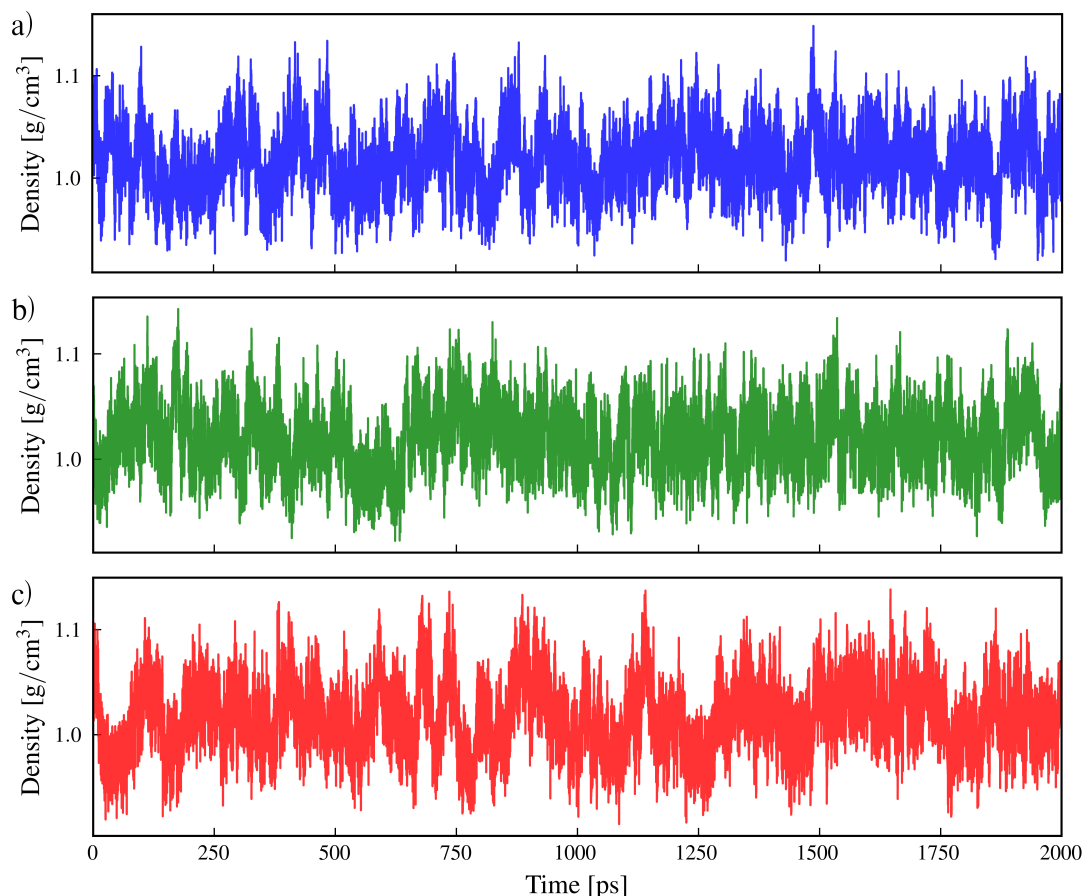


Figura 2.7: Densidade de equilíbrio para os modelos treinados com (a) 540, (b) 1080 e (c) 1620 configurações.

A redução no número de configurações não afeta as propriedades estruturais, como a densidade, conforme demonstrado por Gomes-Filho *et al.* [66]. Entretanto, conforme o mesmo estudo, a convergência da difusão só foi observada com 30% do *dataset*. Na Figura 2.8, apresentamos a evolução temporal do coeficiente de difusão para cada modelo. Para calcular o coeficiente de difusão foi utilizada a Equação 2.3:

Propriedades	Nº de configurações			SCAN [42]
	540	1080	1620	
$\rho$ (g/cm <sup>3</sup> )	1.02	1.02	1.02	1.02
D (Å <sup>2</sup> /ps)	0.047	0.045	0.045	0.040

Tabela 2.9: Densidade de equilíbrio e coeficiente de difusão para os modelos MACE treinados com 540, 1080 e 1620 configurações. O desvio padrão é da ordem de  $\pm 0.03$  g/cm<sup>3</sup> para a densidade e  $\pm 0.004$  Å<sup>2</sup>/ps para a difusão.

$$D = \lim_{t \rightarrow \infty} \frac{1}{6t} \langle |r(t) - r(0)|^2 \rangle. \quad (2.3)$$

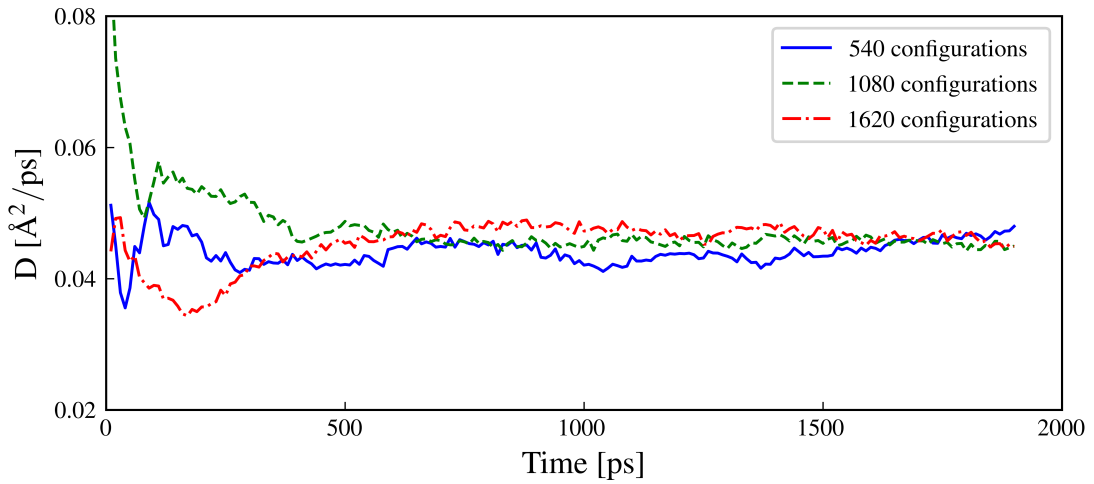


Figura 2.8: Evolução temporal do coeficiente de difusão para cada modelo, com diferentes números de configurações no *dataset*.

Como mostrado na Tabela 2.9, os modelos treinados com 1080 e 1620 configurações apresentam convergência no coeficiente de difusão, e os resultados estão de acordo com o esperado, com ambas as curvas tendendo ao mesmo limite em simulações longas. É evidente que o valor experimental para a difusão permanece distante do obtido. Cabe ressaltar que a inclusão de efeitos quânticos nucleares melhora a precisão desse valor [42]. Por fim, na Figura 2.9, apresentamos o coeficiente de difusão e o deslocamento quadrático médio para todas as simulações.

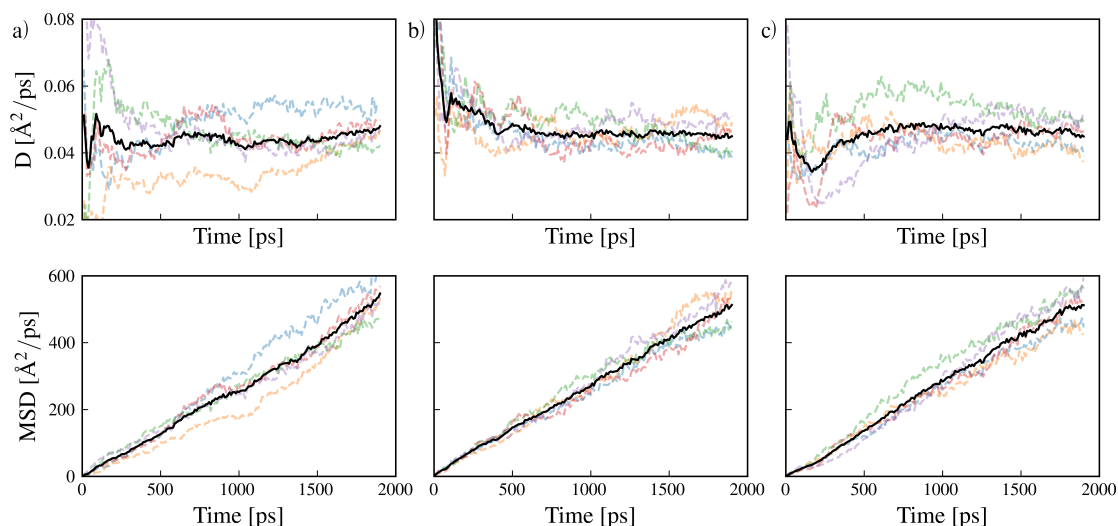


Figura 2.9: Evolução temporal do coeficiente de difusão e do deslocamento quadrático médio para todas as simulações dos modelos treinados com (a) 540, (b) 1080 e (c) 1620 configurações. As linhas tracejadas coloridas representam as cinco diferentes simulações e a linha sólida em preto representa a média sobre as simulações.

## 2.6 Considerações Finais

Em resumo, apresentamos uma visão geral dos principais parâmetros usados em scripts de treinamento para sistemas de água. Primeiro, é importante destacar que o RMSE é apenas um indicador inicial da qualidade de um modelo DP; mesmo modelos com erros relativamente baixos de energia e força podem não ser necessariamente precisos. Em geral, aumentar o número de passos de treinamento reduz o RMSE e, para sistemas de água, recomenda-se aproximadamente  $5 \times 10^6$  passos para obter um modelo confiável, capaz de reproduzir as propriedades estruturais da água. A extensão do raio de corte não traz benefícios significativos: acima de  $6,5 \text{ \AA}$ , a energia do sistema já não é bem reproduzida e, devido ao maior número de vizinhos dentro do raio de corte, as simulações tornam-se mais lentas. Além disso, um sistema de 64 moléculas de água é a escolha mais eficiente para a construção do conjunto de dados, já que modelos treinados com 256 moléculas apresentam essencialmente a mesma precisão, mas com um custo computacional *ab initio* muito maior.

Também observamos que um conjunto de dados construído apenas a partir de configurações de água líquida não captura corretamente a energia de fases de gelo, tornando essencial a inclusão de configurações sólidas para o estudo de

propriedades do gelo. A adição de configurações isovolumétricas não melhora o desempenho dos modelos, enquanto a inclusão de amostras extraídas de simulações NPT garante melhor transferibilidade. Por fim, comparamos DeePMD-kit com os modelos MACE. Um modelo treinado com MACE pode reproduzir as propriedades estruturais da água usando apenas 3% do conjunto de dados e, com 6%, também reproduzir o coeficiente de autodifusão relatado em estudos anteriores. No entanto, simulações DP-MD com MACE são mais lentas do que com DeePMD-kit. Por outro lado, como os cálculos *ab initio* no nível SCAN são altamente custosos, o MACE surge como uma alternativa atraente para gerar modelos com precisão de funcionais híbridos a um custo menor.

## Capítulo 3

# Efeitos Quânticos Nucleares em Gelo Hexagonal

O gelo hexagonal, também conhecido como gelo  $I_h$ , constitui virtualmente todo o gelo encontrado na Terra e exibe propriedades essenciais para o controle da vida e do clima no planeta [16]. Além disso, possui importância em diversas áreas, como criobiologia, química atmosférica e ciência de materiais [13]. A estrutura do gelo  $I_h$  é tetragonal e foi primeiramente proposta por Pauling [21], com base em resultados experimentais [167, 168, 19]. Apesar de sua estrutura química simples, observamos um comportamento anômalo em sua densidade [12], além de algumas propriedades que ainda não são compreendidas em sua totalidade, em particular aquelas em que as ligações de hidrogênio são preponderantes [17]. Na Figura 3.1, apresentamos a estrutura do gelo  $I_h$ .

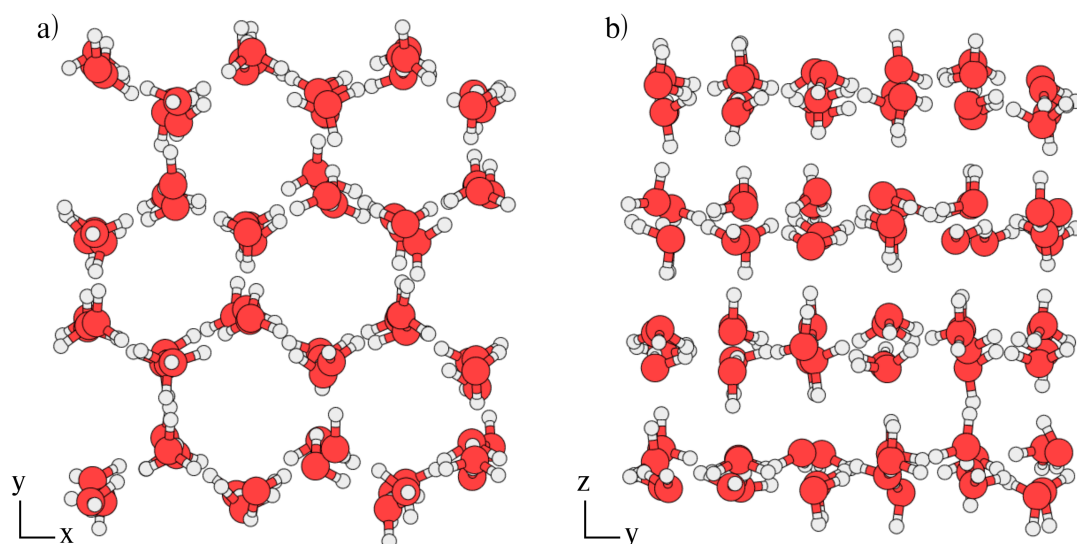


Figura 3.1: Vista (a) superior e (b) lateral da estrutura do gelo  $I_h$ . As esferas vermelhas representam os átomos de oxigênio e, as cinzas, os átomos de hidrogênio.

Embora alguns modelos clássicos possam produzir bons resultados para fa-

ses do gelo [36, 37], como o TIP4P/Ice [39], sua precisão e transferibilidade são frequentemente limitadas, apresentando desafios para reproduzir com exatidão dados experimentais em diversas condições com um único conjunto de parâmetros. Por outro lado, simulações *ab initio* são precisas (no limite da aproximação do potencial de correlação e troca), porém extremamente custosas e restritas a sistemas com poucos átomos. Para sistemas aquosos, os resultados ainda são extremamente dependentes dos funcionais de correlação e troca empregados [28].

Os potenciais de *machine learning*, em particular os campos de força baseados em redes neurais profundas [4, 93, 97, 169], oferecem uma forma eficiente de representar a superfície de energia potencial e, portanto, são atraentes para a realização de simulações de dinâmica molecular, pois integram a precisão de cálculos *ab initio* com uma eficiência computacional comparável à da MD clássica [62, 170]. Recentemente, os potenciais de *machine learning* foram utilizados para investigar algumas propriedades do gelo  $I_h$  [35, 67] e do gelo superiônico [15]. Em outro estudo, com qualidade equivalente ao funcional SCAN, a temperatura de fusão do gelo  $I_h$  foi obtida sendo de aproximadamente 310 K [35].

A natureza quântica dos núcleos desempenha um papel importante em diversos sistemas, especialmente quando estão envolvidos átomos leves. O hidrogênio, por exemplo, manifesta comportamento quântico mesmo em temperaturas ambientes [68, 69, 70]. Em sistemas aquosos, já foi demonstrado que os efeitos nucleares quânticos (NQE) estão associados à melhoria na descrição das propriedades estruturais de diferentes fases do gelo [73, 75, 76].

Cheng *et al.* investigaram as propriedades termodinâmicas da água líquida com DFT utilizando funcionais híbridos (revPBE0-D3), incorporando explicitamente os NQE. Seus resultados mostraram que esses efeitos aumentam em cerca de 1% a densidade da água líquida, bem como das fases de gelo hexagonal e cúbico, em comparação com simulações clássicas [78]. Resultados análogos foram encontrados para a água líquida ao empregar os funcionais meta-GGA SCAN [53] e SCAN0 [79]. Contudo, a combinação de AIMD e PIMD é computacionalmente dispendiosa, restringindo-se normalmente a escalas temporais curtas. Em contraste, quando se utilizam potenciais clássicos, a inclusão de NQE pode levar a tendências opostas [80].

Além de afetarem a densidade de equilíbrio, os NQE também exercem influência sobre a rede de ligações de hidrogênio. Simulações de PIMD baseadas em DFT mostram que há uma densidade distinta de zero para valores positivos da coordenada de transferência de prótons, o que evidencia o compartilhamento de

prótons e a aproximação entre moléculas de água, resultando em um alargamento da distribuição do momento de dipolo [81].

Deste modo, neste capítulo, apresentaremos as propriedades estruturais e vibracionais do gelo  $I_h$  obtidas a partir de potenciais de aprendizado de máquina (*machine learning*) com diferentes níveis de teoria, a fim de observar a influência dos funcionais de correlação e troca. Traremos uma discussão sobre a construção do *dataset*, o treinamento dos modelos utilizados e os métodos computacionais das dinâmicas moleculares. Adicionalmente, obteremos as propriedades considerando os efeitos quânticos nucleares (NQE) para quantificar a influência da deslocalização dos átomos de hidrogênio nas propriedades do gelo  $I_h$ .

### 3.1 Métodos Computacionais

Nossa investigação seguiu um fluxo de trabalho (*workflow*) estruturado em cinco etapas principais. Na primeira, treinamos modelos preliminares para vários funcionais de correlação e troca: PBE [65], vdW-cx [31], optB88-vdW [30] e SCAN [34]. O conjunto de dados DFT foi obtido por Torres *et al.* [42]. O PBE é um dos funcionais mais utilizados, apesar de sua conhecida tendência a superestruturar a água líquida [171]. Os funcionais vdW-cx e optB88-vdW incorporam interações de van der Waals (vdW), essenciais para refinar as propriedades estruturais e dinâmicas da água [29, 11]. Já o SCAN, pertencente à família de funcionais meta-GGA, foi incluído devido ao seu sucesso recente na modelagem precisa da água líquida [12].

Na segunda etapa, realizamos simulações *Deep Potential Molecular Dynamics* (DP-MD) no ensemble isotérmico-isobárico (NPT) com os modelos preliminares para gerar configurações de gelo, que foram então incorporadas ao conjunto de dados de treinamento. A terceira etapa envolveu o cálculo de energias e forças DFT para cada configuração de gelo selecionada, utilizando cada um dos funcionais. Na quarta etapa, retreinamos e validamos cada modelo, incorporando as novas configurações de gelo ao conjunto de treinamento.

Finalmente, na quinta etapa, os modelos refinados foram utilizados para realizar simulações DP-MD e *Deep Potential Path Integral Molecular Dynamics* (DP-PIMD) no ensemble NPT. Todas as simulações foram realizadas utilizando uma caixa com 96 moléculas e orientações aleatórias das ligações de hidrogênio, garantindo momento de dipolo elétrico líquido nulo na célula de simulação. A seguir, detalharemos cada etapa. De modo resumido, a Figura 3.2 ilustra o fluxo de trabalho

seguido neste capítulo.

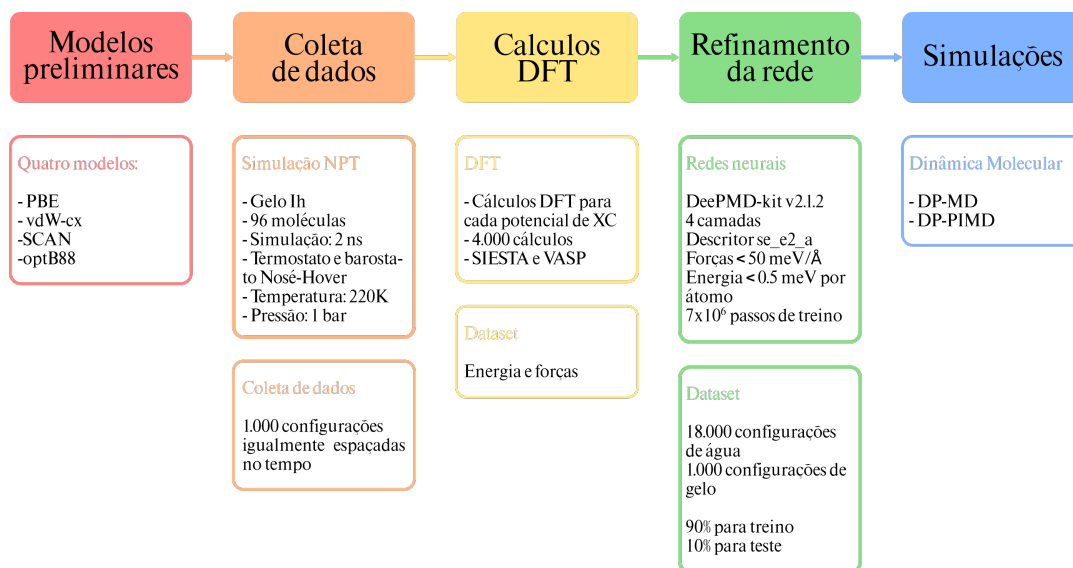


Figura 3.2: Estrutura do fluxo de trabalho (*workflow*) para a obtenção das propriedades do gelo  $I_h$ .

### 3.1.1 Treinamento das Redes Neurais

Os modelos preliminares foram desenvolvidos a partir de um *dataset* que abrange uma ampla variedade de geometrias da água líquida. Essas configurações foram extraídas de *snapshots* não correlacionados de longas simulações de dinâmica molecular (MD) clássica, que amostraram diversas condições do sistema, incluindo diferentes densidades e temperaturas. Para cada funcional, o *dataset* inicial consistiu em energias e forças calculadas via DFT para 18.000 configurações obtidas através de cálculos de PIMD, compostas por uma caixa de 64 moléculas de água. Informações detalhadas sobre o *dataset* inicial podem ser encontradas em Torres *et al.* [42].

Em seguida, realizamos simulações DP-MD no ensemble NPT para cada modelo preliminar, utilizando uma estrutura de gelo  $I_h$  composta por 96 moléculas de água, a  $T = 220$  K e  $P = 1$  bar, com uma etapa de produção de 2 ns e um passo de tempo de 0.25 fs no código LAMMPS [40]. Para essas simulações, foram utilizados o termostato encadeado de Nosé-Hoover [129, 130] e um barostato com três cadeias. De cada simulação, 1.000 configurações foram extraídas em intervalos de tempo uniformes e adicionadas ao *dataset* inicial. Essa melhoria do modelo pode ser interpretada como uma abordagem de aprendizado por transferência

(*transfer learning*) [172].

Realizamos, então, cálculos DFT de ponto único (*single-point*) para cada configuração de gelo utilizando o código SIESTA [157] para os funcionais PBE, vdW-cx e optB88-vdW. Esses cálculos empregaram pseudopotenciais de norma conservada [114], um *mesh cutoff* de 500 Ry e uma base *double- $\zeta$*  otimizada com orbitais de polarização para elétrons de valência [158]. Já os cálculos com o funcional SCAN foram realizados no VASP [162], com um *cutoff* de energia de 118 Ry e o método *Projected Augmented Wave* (PAW) [163].

Utilizamos a versão para GPU do *software* DeePMD-kit v2.1.3 [4, 93, 153] para gerar os potenciais de rede neural profunda (DNN). O descritor *Deep Potential-Smooth Edition* foi empregado, no qual todas as coordenadas relativas (distâncias e ângulos) são usadas para construir o descritor [154], com um raio de corte de 6.0 Å e um raio de suavização de 0.5 Å. Neste raio, o número máximo de vizinhos selecionados é de 46 átomos de oxigênio e 92 de hidrogênio.

A arquitetura da rede consiste em quatro camadas com 320, 160, 32 e 16 neurônios, respectivamente, e a função de ativação selecionada foi a tangente hiperbólica. A função de perda (*loss function*), composta pelos erros quadráticos médios nas energias e forças, foi minimizada usando o método de gradiente descendente estocástico de Adam ao longo de  $7 \times 10^6$  passos de treino. Esse procedimento foi repetido para cada funcional, produzindo um modelo distinto para cada um. Os parâmetros da função de perda estão presentes na Tabela 3.1.

Parâmetro	Valor
start_pref_e	0.02
limit_pref_e	8.00
start_pref_f	1500
limit_pref_f	1.00

Tabela 3.1: Parâmetros da função de perda (*loss function*) utilizados no processo de treinamento.

Selecionamos 90% das configurações como conjunto de treinamento e 10% como conjunto de teste, correspondendo a 17.100 estruturas de água e 900 de gelo. Para o treinamento, utilizamos um único processador AMD Epyc 7662 com uma GPU NVIDIA Tesla A100. Cada potencial levou 22 horas para completar as etapas de treinamento. Após a conclusão, o modelo foi comprimido para melhorar o desempenho computacional [155].

Para avaliar a qualidade dos modelos, realizamos duas validações:

- Validação *a priori*: os erros nas previsões de energia e força dos modelos foram comparados aos valores de referência DFT usando o conjunto de teste.
- Validação *a posteriori*: examinamos se as coordenadas geradas pelas simulações DP-MD produzem estruturas com erros aceitáveis para energias e forças.

Para a validação *a priori*, dividimos o conjunto de teste em dois subconjuntos: um com configurações de água e outro com configurações de gelo. Em seguida, testamos o modelo para os dois subconjuntos e para o conjunto total. O valor do Erro Quadrático Médio (RMSE) para energia e força de cada modelo e conjunto está apresentado na Tabela 3.2. A comparação entre os resultados obtidos com os cálculos DFT e os modelos está na Figura 3.3.

Modelo	RMSE Energia (meV/átomo)			RMSE Forças (meV/Å)		
	Água	Gelo	Água+Gelo	Água	Gelo	Água+Gelo
PBE	0.50	0.27	0.49	44.1	39.5	43.7
vdW-cx	0.58	0.19	0.52	49.5	22.9	44.6
SCAN	0.56	0.19	0.56	48.8	34.5	46.9
optB88	0.51	0.20	0.50	45.2	39.5	44.9

Tabela 3.2: RMSE para cada modelo treinado para os subconjuntos de teste: água, gelo e a combinação de ambos (Água+Gelo).

Podemos verificar que, para cada funcional, obtivemos erros pequenos no conjunto de teste, abaixo de 50 meV/Å (forças) e 0.55 meV/átomo (energia), dentro da faixa aceitável para um modelo confiável [169]. Além disso, os erros para cada funcional são próximos, indicando que as estruturas utilizadas são representativas do espaço de fase. Outro ponto a destacar é que o erro do conjunto de dados de água domina o erro total da amostra, devido à maior quantidade de informações nesse subconjunto. Contudo, vimos que os dados de gelo possuem um erro menor, tanto para a energia quanto para as forças, o que justifica o uso de um número menor de configurações de gelo no *dataset* final.

Para a validação *a posteriori*, geramos 200 configurações de gelo com nosso modelo para cada funcional e replicamos os cálculos DFT com o mesmo protocolo utilizado na construção do *dataset*. Deste modo, apresentamos na Figura 3.4 a comparação entre as energias e forças para este novo conjunto de dados e, na Tabela 3.3, mostramos os erros correspondentes para cada funcional. É possível verificar que os modelos geram estruturas que estão de acordo com os resultados

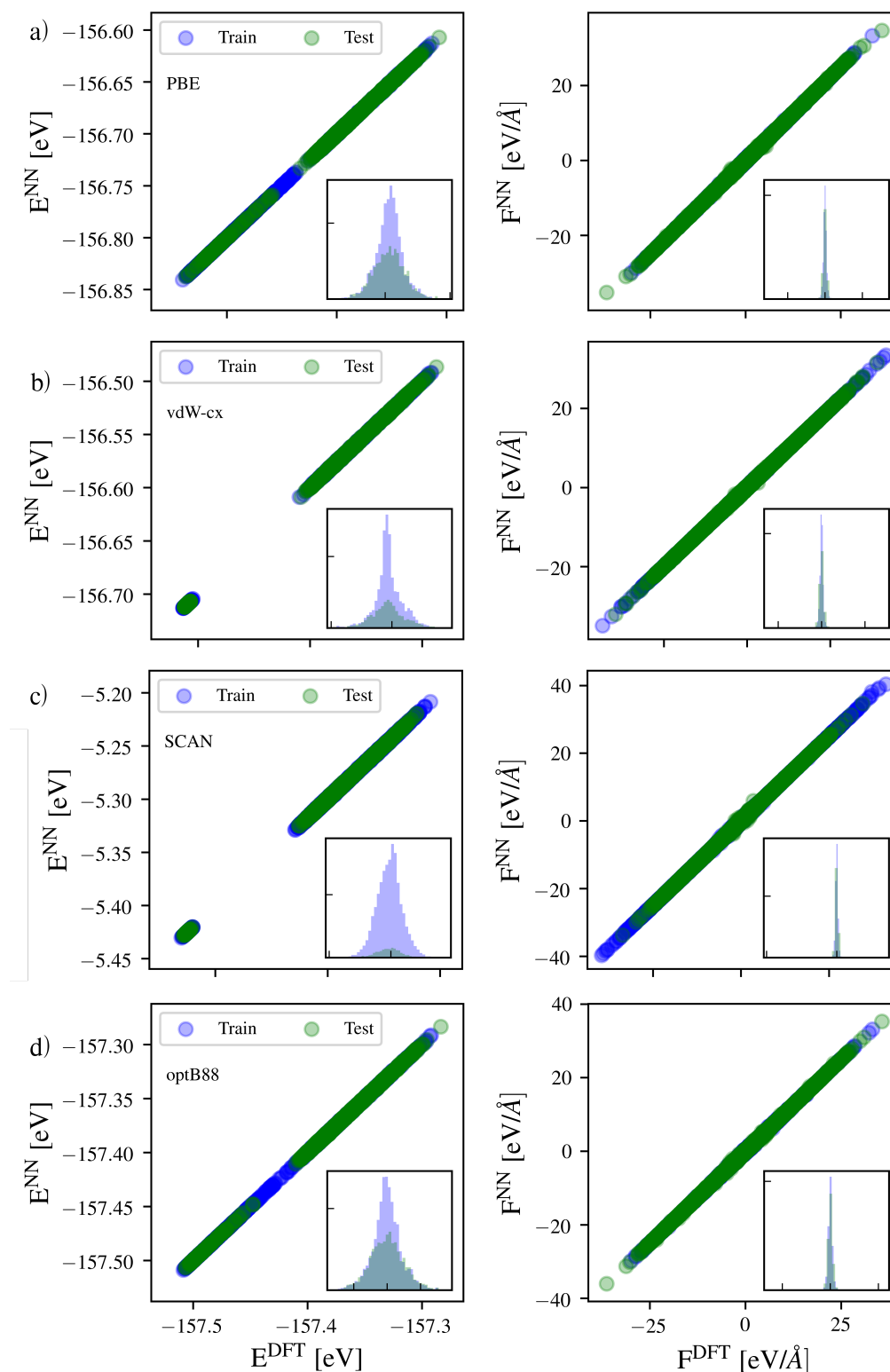


Figura 3.3: Comparação das energias e forças entre os cálculos DFT e a rede neural para os funcionais: (a) PBE, (b) vdW-cx, (c) SCAN e (d) optB88. Os *insets* representam a distribuição das diferenças entre os valores obtidos via DFT e os modelos.

DFT, atingindo, assim, um dos objetivos: replicar cálculos para grandes sistemas por um longo tempo de simulação com a mesma qualidade de cálculos *ab initio*.

Modelo	RMSE Energia (meV/átomo)	RMSE Forças (meV/Å)
PBE	0.25	38.9
vdW-cx	0.34	46.6
SCAN	0.13	23.7
optB88-vdW	0.20	39.6

Tabela 3.3: RMSE para cada modelo treinado para as configurações de gelo com os dados obtidos para o método de validação *a posteriori*.

### 3.1.2 Dinâmica Molecular: Propriedades Estruturais

Após verificada a precisão dos modelos, partimos para a obtenção das propriedades do gelo  $I_h$ . Realizamos duas simulações independentes de dinâmica molecular NPT: uma incluindo NQEs (DP-PIMD) e outra sem (DP-MD). Essas simulações foram executadas em uma estrutura de gelo  $I_h$  protonicamente desordenada, cujo dipolo líquido é nulo, composta por 96 moléculas de água. Em todas as simulações, os sistemas foram inicialmente equilibrados por 100 ps, seguidos por 500 ps de produção com um passo de tempo de 0.25 fs, e os dados foram coletados a cada 0.05 ps. As simulações foram conduzidas a  $T = 220$  K e  $P = 1$  bar, permitindo comparações diretas com resultados experimentais [10].

Todas as simulações utilizaram o código i-PI [41]. Empregamos o termostato da Equação de Langevin Generalizada (GLE) [133] para DP-MD, com os parâmetros extraídos do GLE4MD [135]. Para as simulações DP-PIMD, discutiremos a seguir a escolha do termostato adequado. Nosso modelo de rede neural fornece uma forma rápida e eficiente de avaliar um grande número de estruturas a uma fração do custo computacional dos cálculos de DFT. Isso também oferece um meio de testar a convergência de parâmetros — algo que demandaria muito mais tempo em comparação com simulações de DFT. Contudo, ao aumentar o número de moléculas ou alterar a desordem das ligações de hidrogênio faz com que as propriedades estruturais variem por uma fração de por cento, não sendo necessária esta análise.

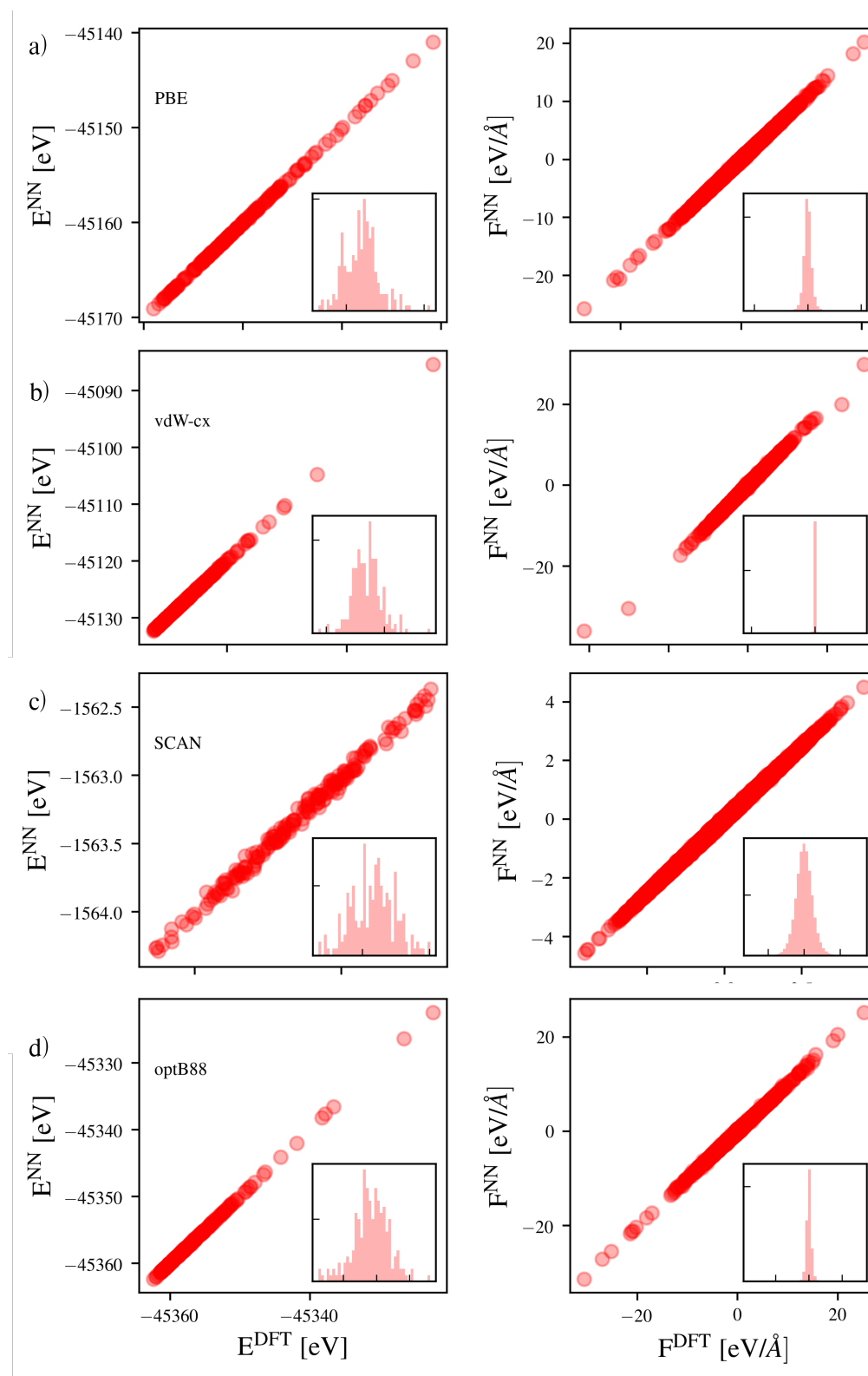


Figura 3.4: Comparação das energias e forças entre os cálculos DFT e a rede neural utilizando o método de validação *a posteriori* para os funcionais: (a) PBE, (b) vdW-cx, (c) SCAN e (d) optB88. Os *insets* representam a distribuição das diferenças entre os valores obtidos via DFT e os modelos.

### 3.1.3 Dinâmica Molecular: Propriedades Vibracionais

As propriedades vibracionais foram obtidas de duas maneiras distintas. Para os cálculos que não consideraram os NQEs, utilizamos o ensemble NVE. Para os cálculos que consideram os NQEs, utilizamos a abordagem da Dinâmica Molecular de Polímeros em Anel Termostatizada (TRPMD) [8, 9].

O espectro de frequência foi obtido utilizando a transformada de Fourier da função de autocorrelação das velocidades,  $Z(t)$ , com a seguinte equação:

$$Z(t) = \sum_i \frac{\langle v_i(t) \cdot v_i(0) \rangle}{\langle v_i(0) \cdot v_i(0) \rangle}, \quad (3.1)$$

onde  $v_i(t)$  representa a velocidade do  $i$ -ésimo átomo no instante  $t$ .

Para os cálculos com NQEs, partimos da configuração final obtida via DP-PIMD e utilizamos o método TRPMD em uma simulação de 10 ps, com as velocidades coletadas a cada 1 fs. Para a análise de convergência, testamos diferentes números de réplicas (*beads*): 32 e 48.

Como visto na Figura 3.5(a), verificamos que os cálculos com 32 *beads* estão convergidos. Em seguida, extraímos 10 configurações igualmente espaçadas no tempo da simulação DP-PIMD e repetimos o processo para cada uma delas, a fim de obter uma amostragem mais ampla. Verificamos na Figura 3.5(b) que a média sobre as diferentes condições iniciais resulta no mesmo espectro de uma única simulação de 100 ps. Para o cálculo do espectro vibracional sem NQEs, utilizaremos este mesmo procedimento de 10 simulações curtas, com as condições iniciais obtidas das simulações DP-MD.

### 3.1.4 Termostato das Simulações DP-PIMD

Como discutido na Seção 1.3, as simulações que incluem NQEs necessitam de um número suficientemente grande de *beads*. Para tal, verificamos a quantidade necessária para a convergência das propriedades estruturais. Inicialmente, utilizando o termostato PILE-L [173], calculamos a densidade de equilíbrio e a função de distribuição radial para cada par atômico para 48 e 64 *beads*. Apresentamos na Figura 3.6 as propriedades estruturais para os diferentes números de *beads*. Verificamos que utilizar 48 *beads* para cada cálculo atinge o mesmo resultado que um número superior de réplicas. Com isso, o custo computacional pode ser reduzido sem perda na qualidade dos resultados.

Outra solução para diminuir o custo computacional seria utilizar o termostato

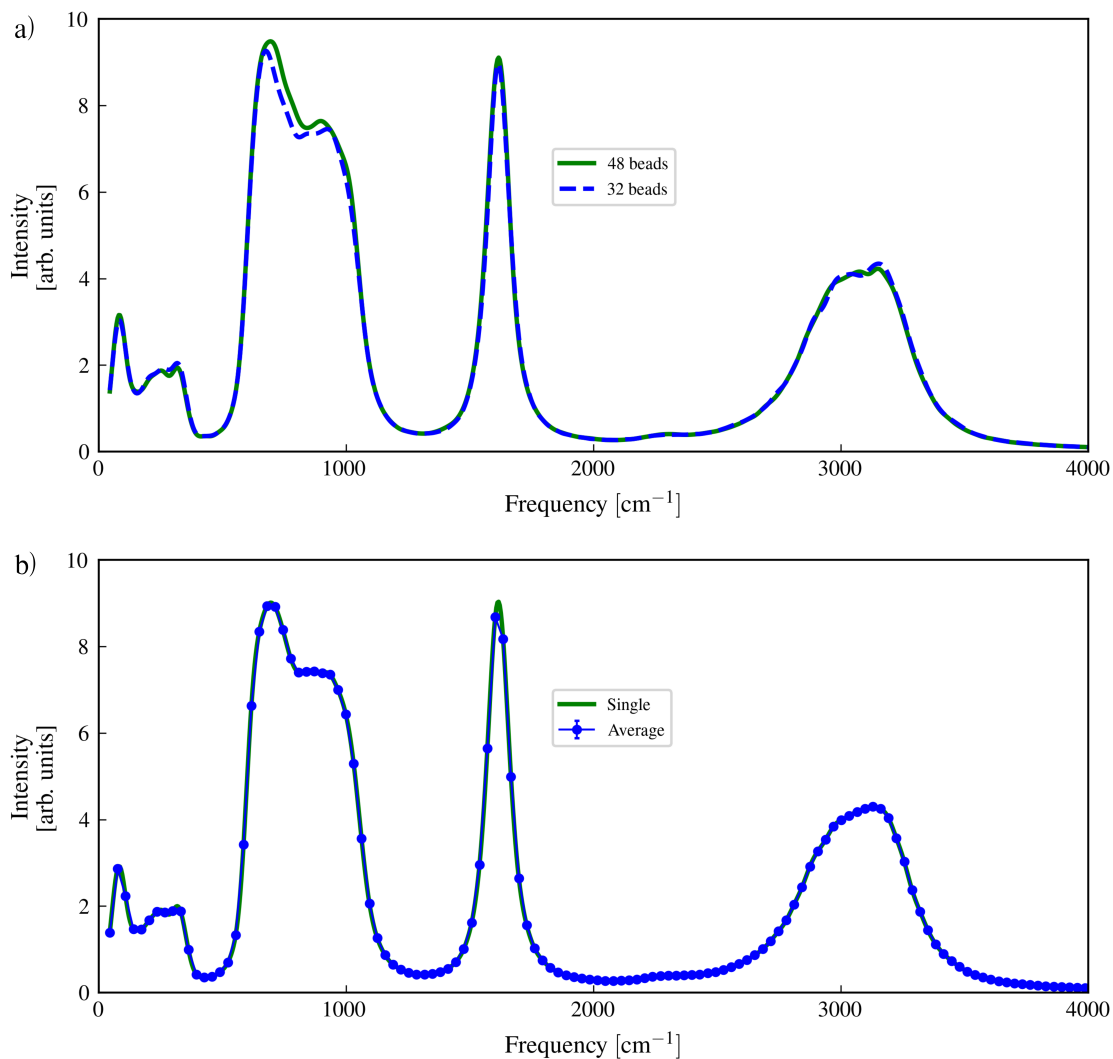


Figura 3.5: Espectro vibracional para (a) teste de convergência do número de *beads* e (b) comparação entre uma simulação longa e a média sobre 10 simulações curtas.

PIGLET [3, 174]. Comparamos, então, na Figura 3.7, as mesmas propriedades observadas anteriormente para um cálculo utilizando o termostato PILE-L com 48 *beads* e outro com o termostato PIGLET com 8 *beads*.

Como as propriedades estruturais obtidas são essencialmente idênticas, prosseguiremos com o termostato PIGLET, diminuindo o custo computacional e facilitando a análise dos resultados. Para as simulações DP-PIMD, utilizaremos os demais parâmetros estabelecidos para as simulações DP-MD. Os parâmetros do termostato PIGLET foram extraídos do *site* GLE4MD [135].

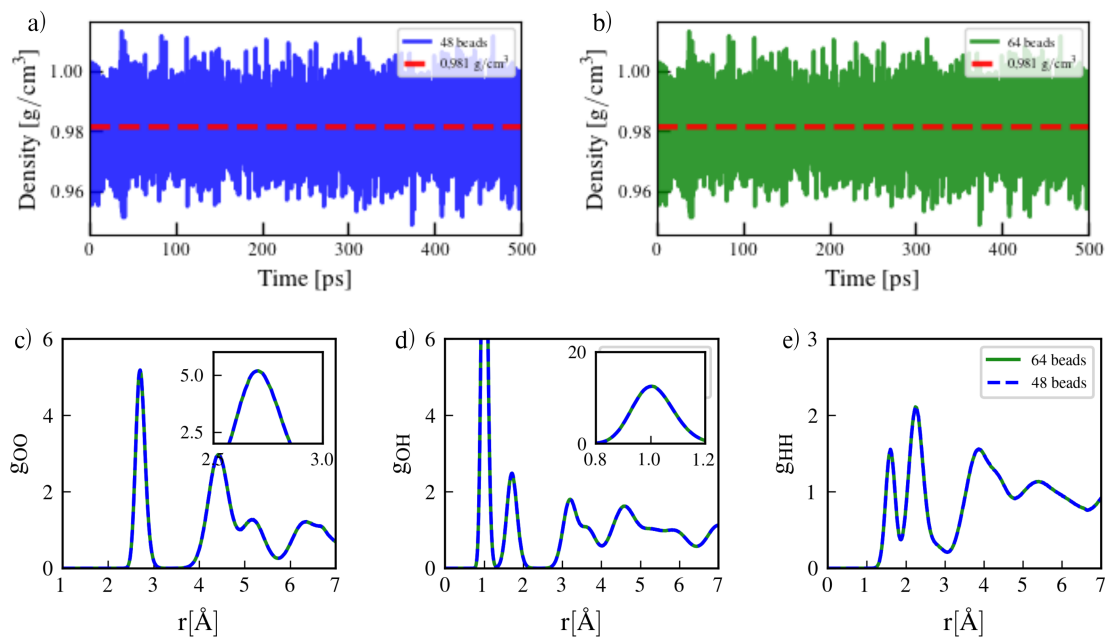


Figura 3.6: Comparação das propriedades estruturais para cálculos com 48 e 64 beads. Densidade de equilíbrio para (a) 48 e (b) 64 beads e (c-e) função de distribuição radial.

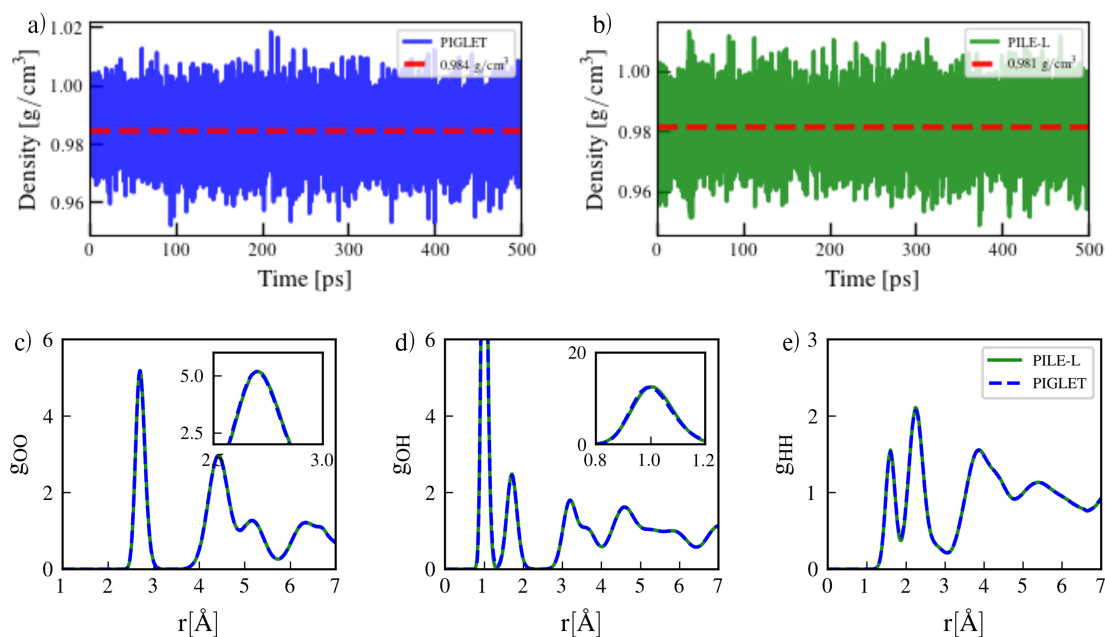


Figura 3.7: Comparação das propriedades estruturais para cálculos com os termostatos PILE-L e PIGLET. Densidade de equilíbrio para (a) PIGLET e (b) PILE-L e (c-e) função de distribuição radial.

## 3.2 Propriedades Estruturais

Primeiramente, apresentamos a densidade de equilíbrio para todos os modelos, com e sem NQEs. A Figura 3.8 reporta a evolução temporal da densidade durante o período de produção de dados. As densidades de equilíbrio do gelo  $I_h$  para cada funcional estão listadas na Tabela 3.4. Observamos que, para os funcionais optB88-vdw, vdW-cx e SCAN, a densidade obtida via DP-MD apresenta um desvio significativo em relação aos valores experimentais, excedendo-os em mais de 5%. No entanto, para o PBE, a densidade está mais próxima dos dados experimentais. Como discutiremos, este é provavelmente um efeito espúrio. Sabe-se que o PBE tende a superestruturar a água líquida, e é provável que esteja se beneficiando dessa propriedade neste caso.

Modelo	Densidade ( $\text{g}/\text{cm}^3$ )	
	DP-MD	DP-PIMD
PBE	0.937 (2.18%)	0.961 (4.80%)
vdW-cx	0.981 (6.98%)	1.013 (10.47%)
SCAN	0.974 (6.22%)	1.001 (9.16%)
optB88	0.964 (5.51%)	0.984 (7.31%)
Exp. [10]	0.917	

Tabela 3.4: Densidade de equilíbrio para cada modelo nas simulações DP-MD e DP-PIMD. Os valores entre parênteses denotam os desvios percentuais em relação aos dados experimentais [10]. A incerteza das densidades é da ordem de  $\pm 0.009 \text{ g}/\text{cm}^3$ .

Vale destacar que as densidades obtidas aqui são semelhantes às calculadas via AIMD para PBE ( $0.936 \text{ g}/\text{cm}^3$ ) e SCAN ( $0.964 \text{ g}/\text{cm}^3$ ) por período mais curtos, conforme observado por Chen *et al.* [12]. Em comparação com trabalhos anteriores [42], verificamos que o PBE prevê o gelo mais denso que a água, falhando em reproduzir a conhecida anomalia de densidade da água [12]. Este comportamento anômalo é corretamente capturado pelos demais funcionais.

A inclusão dos NQEs aumenta a densidade de equilíbrio para todos os modelos. Este aumento de densidade (contração de volume) também foi observado em diferentes resultados baseados em DFT para tempos de simulação mais curtos [78, 53, 79, 73]. Em nossos cálculos, o aumento de densidade varia entre 1% e 3%, dependendo do modelo (Tabela 3.4).

Para investigar o aumento na densidade, calculamos as funções de distribuição

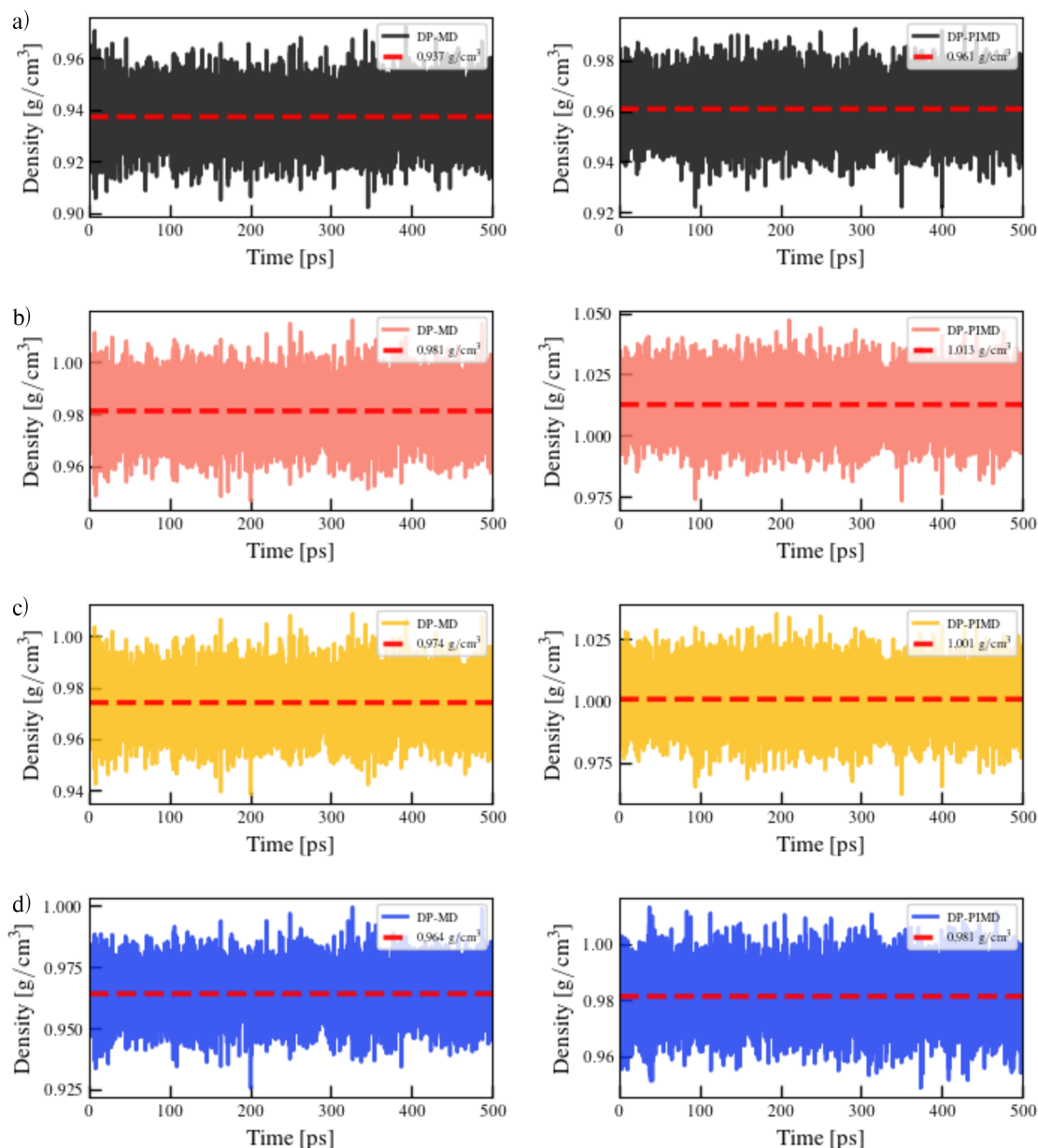


Figura 3.8: Evolução temporal da densidade para os modelos: (a) PBE, (b) vdW-cx, (c) SCAN e (d) optB88. As simulações DP-MD estão à esquerda e as simulações DP-PIMD, à direita.

radial (RDFs) para os pares oxigênio-oxigênio ( $g_{OO}$ ), oxigênio-hidrogênio ( $g_{OH}$ ) e hidrogênio-hidrogênio ( $g_{HH}$ ), conforme mostrado na Figura 3.9. De modo geral, as RDFs O-O apresentam uma distância O-O mais curta em comparação com os dados experimentais. A inclusão dos NQEs não resulta em mudanças significativas nas curvas O-O, com exceção da posição do primeiro pico, como mostrado no detalhe da primeira coluna, que se desloca um pouco para a esquerda em linha

com um aumento de densidade.

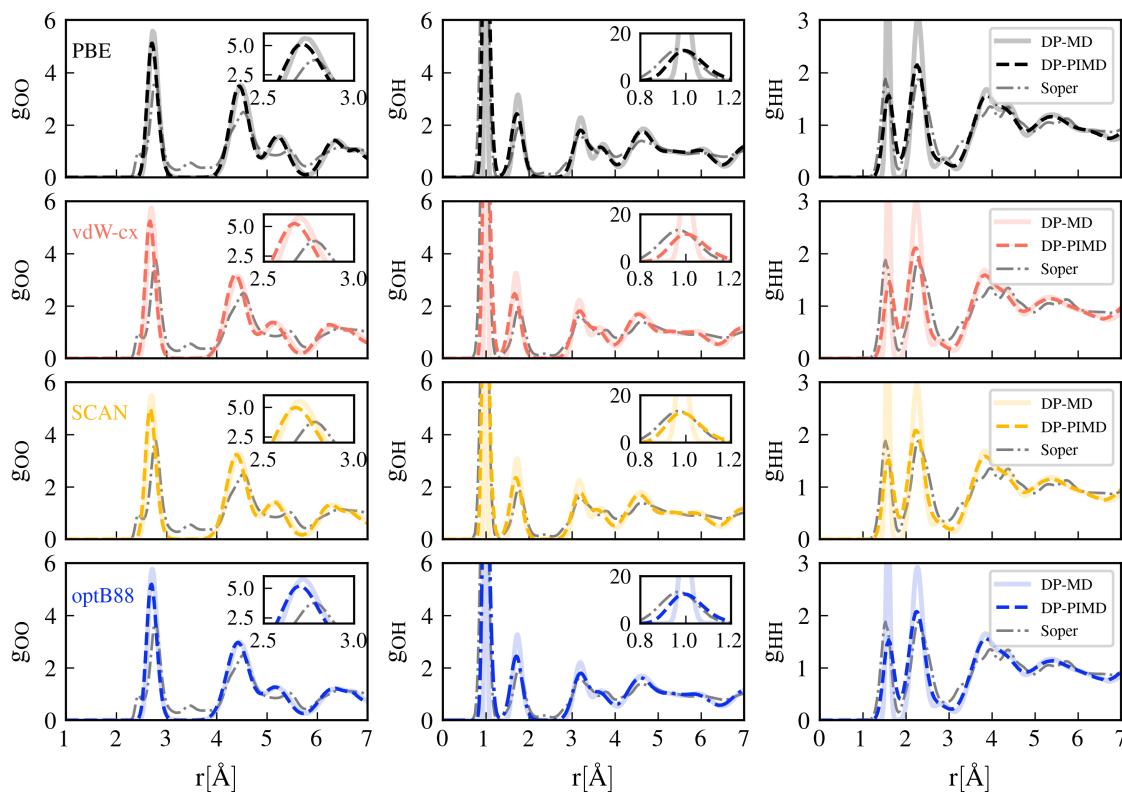


Figura 3.9: Função de distribuição radial para cada funcional. Os funcionais estão descritos no painel esquerdo. O painel esquerdo apresenta a  $g_{OO}$ , o central a  $g_{OH}$  e o direito a  $g_{HH}$ .

Os dados experimentais para as RDFs O-H e H-H desviam-se significativamente daqueles obtidos com simulações DP-MD, que apresentam maior estruturação. No entanto, a inclusão de NQEs melhora substancialmente as previsões estruturais dessas RDFs. Observamos uma suavização das curvas, e a intensidade dos máximos e mínimos mostra boa concordância com os resultados experimentais [10]. Adicionalmente, é interessante notar que essa melhoria parece tornar as curvas menos sensíveis à escolha do funcional, como visto na Figura 3.10.

Para quantificar a melhora das RDFs, utilizamos a Equação 2.2. Os valores finais estão presentes na Tabela 3.5. Observa-se que os NQEs afetam principalmente a interação O-H, com um aumento na precisão de aproximadamente 10%. Esta melhoria está correlacionada com um efeito de suavização no primeiro pico da RDF O-H. Similarmente, os NQEs levam a um melhor ajuste na RDF H-H. É importante destacar que, na região do terceiro pico de  $g_{OH}$ , observamos uma melhoria significativa ao incluir NQEs. Isso indica que as flutuações quânticas

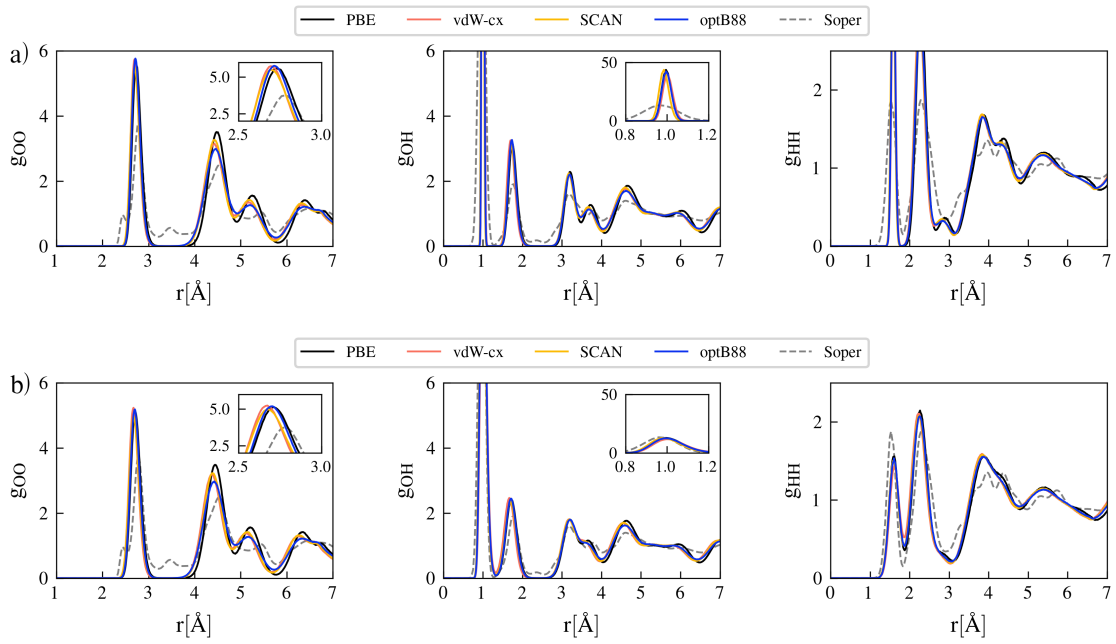


Figura 3.10: Comparação de cada RDF para todos os funcionais (a) sem e (b) com a inclusão dos NQEs.

aumentam a entropia configuracional das estruturas de gelo, alargando as distribuições e aproximando-as das observações experimentais, fenômeno também reportado para água líquida [53].

Modelo	DP-MD			DP-PIMD		
	$\epsilon_{OO}$ (%)	$\epsilon_{OH}$ (%)	$\epsilon_{HH}$ (%)	$\epsilon_{OO}$ (%)	$\epsilon_{OH}$ (%)	$\epsilon_{HH}$ (%)
PBE	81.6	77.7	86.9	81.1	87.5	89.9
vdW-cx	81.1	76.5	86.2	78.3	84.5	87.6
SCAN	80.9	78.2	86.5	79.3	87.9	88.5
optB88	83.7	78.4	87.3	82.5	88.0	89.6

Tabela 3.5: Desempenho percentual de cada modelo na reprodução das RDFs experimentais, calculado com a Equação 2.2.

De modo geral, a incorporação de NQEs aumentou a precisão das RDFs para todos os funcionais. A única exceção ocorre na RDF O-O, onde a precisão diminuiu ligeiramente, pois os NQEs deslocam as posições dos dois primeiros picos para distâncias menores, refletindo a mudança na densidade. Para entender a mudança na densidade associada aos NQEs, apresentamos na Figura 3.11(a) a distribuição da coordenada de transferência de próton,  $\nu = d_{OH} - d_{O...H'}$ , que serve como indicador de eventos de transferência de próton [53]. Também apresentamos

distribuições para distâncias de ligações intra e intermoleculares. É notável que os NQEs alargam os picos e mostram densidade não nula para valores positivos de  $\nu$ , similar ao observado em água líquida [94, 53]. Mais importante, o máximo da distribuição,  $\nu_{max}$ , exibe um ligeiro deslocamento para maiores distâncias (Figura 3.11(a) e Tabela 3.6). Esta observação está ligada ao fortalecimento das interações intermoleculares, uma vez que este pico é sensível à rede de ligações de hidrogênio. Este efeito sutil provavelmente resulta da competição entre várias interações atômicas.

Modelo	$d_{OH}$ (Å)		$d_{O...H'}$ (Å)		$d_{O...O'}$ (Å)	
	DP-MD	DP-PIMD	DP-MD	DP-PIMD	DP-MD	DP-PIMD
PBE	0.995	1.001	1.748	1.719	2.736	2.710
vdW-cx	1.003	1.011	1.706	1.677	2.694	2.671
SCAN	0.985	0.987	1.728	1.697	2.698	2.678
optB88	0.999	1.003	1.733	1.715	2.710	2.699
Exp.	1.008 [164]		1.750 [175]		2.759 [164]	

Tabela 3.6: Distâncias interatômicas ( $d_{OH}$ ,  $d_{O...H'}$  e  $d_{O...O'}$ ) para as simulações DP-MD e DP-PIMD. Os valores correspondem ao comprimento de ligação mais provável, apresentado nas Figuras 3.11(b-d).

As Figuras 3.11(b-d) mostram as distribuições para a distância intramolecular O-H ( $d_{OH}$ ), a distância intermolecular O-H ( $d_{O...H'}$ ) e a distância intermolecular O-O ( $d_{O...O'}$ ). A Tabela 3.6 lista as posições de pico correspondentes. Observamos que as distâncias médias  $d_{OH}$  de ambas as simulações são comparáveis, embora a distribuição DP-PIMD seja mais ampla e levemente assimétrica à direita, sugerindo um enfraquecimento da ligação covalente.

Ao mesmo tempo, as distribuições de  $d_{O...H'}$  e  $d_{O...O'}$  também são mais amplas, mas assimétricas à esquerda (Figuras 3.11(c-d)). Isso está de acordo com as expectativas, pois os NQEs alongam as ligações covalentes e promovem o compartilhamento de prótons na água líquida [73, 1]. Li *et al.* [7] argumentaram que a ligação de hidrogênio na água líquida é relativamente fraca e, portanto, tende a se tornar ainda mais fraca com a inclusão de NQEs. Para o gelo, entretanto, observamos o oposto: o valor médio para  $d_{O...H'}$  e  $d_{O...O'}$  é significativamente menor, levando à conclusão de que a ligação de hidrogênio é forte no gelo e se fortalece quando os NQEs são considerados.

Li *et al.* [7] observaram teoricamente a relação entre a projeção das ligações O-H na direção O-O e a distância O-O, conforme mostrado no detalhe da Figura

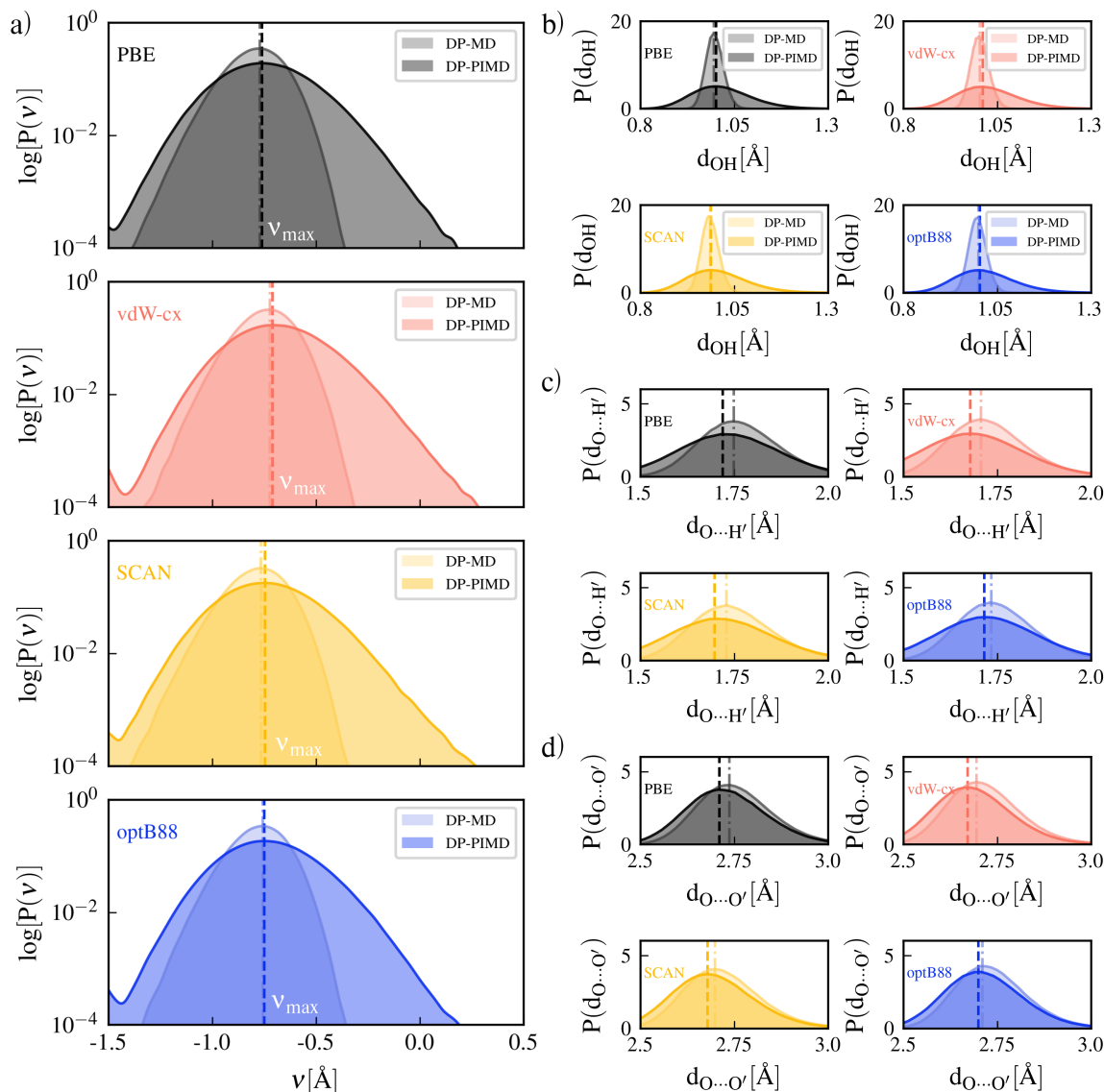


Figura 3.11: Distribuição da (a) coordenada de transferência de próton  $v$  e dos comprimentos de ligação (b) intramolecular O-H ( $d_{OH}$ ), (c) intermolecular O-H ( $d_{O...H'}$ ) e (d) intermolecular O-O ( $d_{O...O'}$ ) obtidos via DP-MD e DP-PIMD. A linha tracejada (tracejada e pontilhada) refere-se ao valor mais provável para as simulações DP-PIMD (DP-MD).

**3.12.** Em sua análise, com o aumento da projeção da ligação O-H, as distâncias O-O diminuem. O comportamento oposto foi reportado quando a projeção foi reduzida. Como os NQEs tendem a enfraquecer as ligações fracas e fortalecer as fortes, todos os funcionais fortalecem as ligações de hidrogênio e reproduzem a diminuição das distâncias O-O, conforme visto na Figura 3.12. Os dados de distância (Tabela 3.6) confirmam os achados de Ceriotti *et al.* [94], de que mudanças

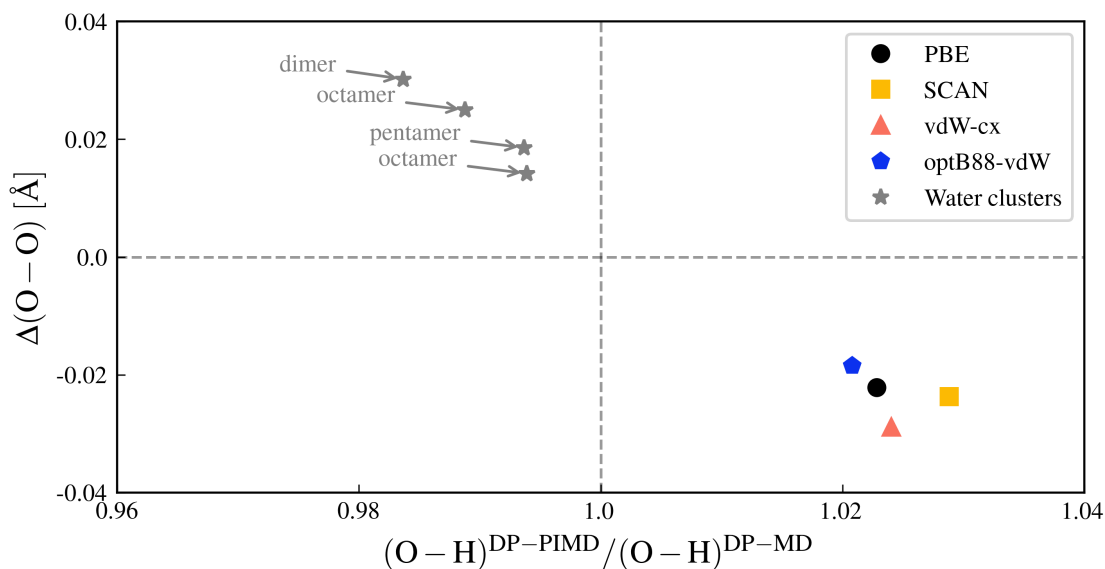


Figura 3.12: Relação entre a projeção das ligações O-H na direção O-O e a distância O-O. Os valores para *clusters* de água foram obtidos teoricamente por Li *et al.* [7]

na densidade alteram principalmente a distância média O-O. Comparando os funcionais, vdW-cx apresenta a menor  $d_{O...O'}$  e a maior densidade em ambos os tipos de simulação, enquanto o PBE apresenta o comportamento oposto.

De modo geral, estes resultados sugerem que a força das ligações de hidrogênio aproxima os átomos de oxigênio, aumentando a densidade, apesar do pequeno aumento na ligação covalente e dos maiores ângulos OHO. Isso é independente da escolha do funcional. Mais importante, leva-nos à conclusão de que todos os funcionais tendem a superligar as moléculas de água no contexto do gelo  $I_h$ . Considerar a natureza quântica do átomo de hidrogênio apenas torna esta superligação mais pronunciada.

### 3.3 Propriedades Dinâmicas

Finalmente, mostramos na Figura 3.13 os espectros vibracionais do gelo  $I_h$ . A inclusão de NQEs resulta em uma diminuição da frequência de deformação angular (*bending*) das moléculas em todos os modelos. Porém, a mudança mais proeminente aparece na faixa de frequência de estiramento (*stretching*) (2500-3500  $\text{cm}^{-1}$ ). Este deslocamento para o vermelho (*redshift*) pode ser atribuído ao fato de que estes são os modos mais afetados por efeitos anarmônicos.

Apesar da pequena mudança no valor mais provável para o comprimento

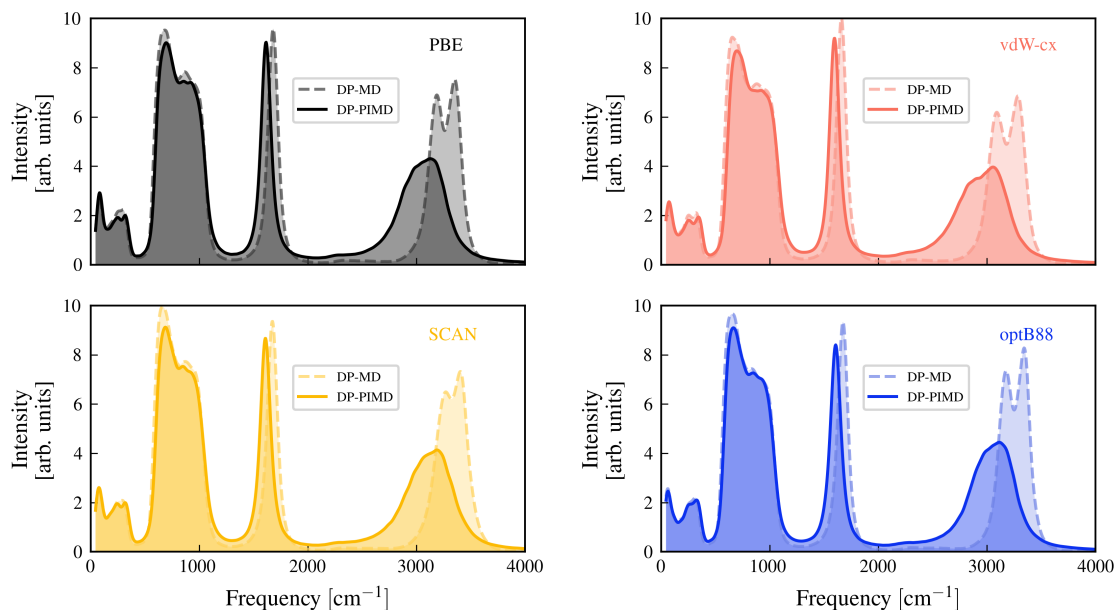


Figura 3.13: Espectros vibracionais para o gelo  $I_h$  a 220 K obtidos com o método TRPMD [8, 9] (linhas contínuas). Para comparação, também mostramos os espectros obtidos via DP-MD no ensemble NVE (linhas tracejadas).

da ligação, a inclusão de NQEs leva a maiores populações em comprimentos de ligação maiores. Isso, por sua vez, leva a um alargamento dos modos e a um consequente *redshift*. O deslocamento está na faixa de 218-250  $\text{cm}^{-1}$ , sendo o optB88 o funcional com maior deslocamento (na sequência: optB88, vdW-cx, PBE e SCAN). Notamos que funcionais com interações de van der Waals apresentam os maiores deslocamentos, implicando ligações de hidrogênio mais fortes.

Os NQEs também levam a um alargamento de cerca de 380  $\text{cm}^{-1}$  na faixa de frequência de estiramento em comparação com os resultados DP-MD para todos os funcionais. Mais uma vez, a escolha do funcional determina a largura da banda, já que o vdW-cx apresenta a maior largura para ambas as simulações, enquanto o optB88 tem a menor.

### 3.4 Considerações Finais

Em resumo, realizamos simulações de dinâmica molecular de estruturas de gelo utilizando um potencial profundo treinado com diferentes funcionais de troca-correlação. A inclusão dos efeitos quânticos nucleares leva a uma melhora geral na descrição das propriedades estruturais do gelo em comparação com experimentos. De fato, as funções de distribuição radial obtidas com DP-PIMD foram, em geral,

mais precisas do que as obtidas com DP-MD, devido a uma melhor descrição da distribuição O–H. A distribuição O–O não apresentou mudanças significativas com a inclusão dos efeitos quânticos nucleares para grandes distâncias; entretanto, para distâncias abaixo de 4 Å, a posição do pico desloca-se ligeiramente para distâncias menores.

Observamos que a densidade, para todos os funcionais de troca-correlação, aumentou quando os efeitos quânticos nucleares foram considerados. Ao inspecionar as diferentes ligações no sistema, identificamos que a ligação intramolecular O–H enfraquece ligeiramente, enquanto as ligações de hidrogênio tornam-se mais fortes, fazendo com que as distâncias O–O diminuam, levando a densidades maiores, mais distantes dos valores experimentais. Isso nos leva a concluir que os funcionais do tipo GGA e o SCAN tendem a superligar as moléculas de água no gelo, resultando em uma forte rede de H-bonds, que é ainda mais reforçada pela inclusão de interações de van der Waals e efeitos quânticos nucleares.

# Capítulo 4

## Defeitos em Gelo

Após estudar a influência das interações interatômicas na densidade do gelo e as implicações da inclusão dos efeitos quânticos nucleares, partimos a observar os defeitos pontuais na estrutura cristalina do gelo. Em relação a defeitos pontuais, podemos observar duas categorias distintas: os defeitos puramente eletrônicos, como os defeitos de Bjerrum e iônico, que apenas alteram a configuração das ligações de hidrogênio [84] e os defeitos que correspondem a inclusão ou falta de moléculas de água no sistema, no caso os defeitos intersticiais e defeitos de vacância. Neste capítulo, focaremos neste último tipo de defeito. A Figura 4.1 mostra a estrutura para os defeitos de vacância, e duas configurações para os defeitos intersticiais: Tu e Bc.

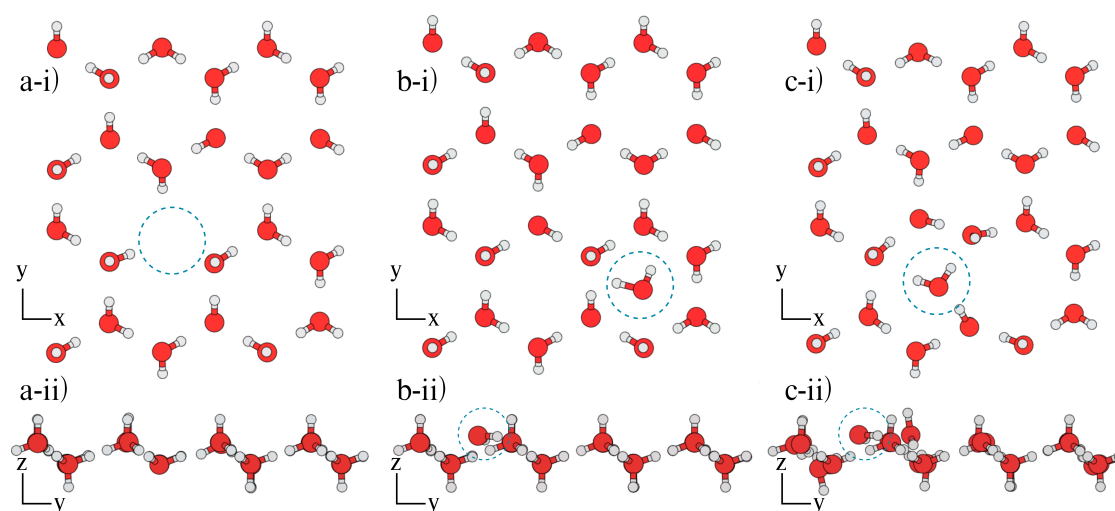


Figura 4.1: Geometria dos defeitos: (a) vacância, (b) Tu e (c) Bc. Os círculos pontilhados em azul indicam a posição dos defeitos na estrutura do gelo  $I_h$ . As estruturas são apresentadas nos planos (i)  $xy$  e (ii)  $yz$ .

Do ponto de vista experimental, há a indicação que os defeitos de vacância são dominantes em relação aos defeitos intersticiais para temperaturas abaixo de  $-50^\circ\text{C}$  e os defeitos intersticiais apresentando maior ocorrência para temperaturas mais altas [85]. Contudo, os resultados experimentais obtidos para a vacância por

meio da técnica de aniquilação de pósitrons [88, 87] apresentam valores para a energia de formação que divergem dos valores teóricos [85, 86, 84].

Em contrapartida, os defeitos intersticiais podem ocupar principalmente duas posições na estrutura do gelo  $I_h$ : *trigonal capped* (Tc) e *trigonal uncapped* (Tu) [16]. Além destas possibilidades, cálculos de dinâmica molecular realizados por Ikeda-Fukazawa [91] identificaram uma posição intermediária de uma molécula de água entre dois átomos de oxigênio, resultante da difusão da molécula intersticial. Essa configuração recebe o nome de defeito *bond center* (Bc). As análises para o defeito Tc foram suprimidas devido ao alto valor encontrado para sua energia de formação [85].

De acordo com trabalhos experimentais e teóricos [89, 90, 85], em torno de 230 K, verifica-se uma inversão na concentração relativa dos defeitos: para temperaturas superiores, os defeitos intersticiais dominam, enquanto para temperaturas inferiores, os defeitos de vacância são predominantes. Entretanto, a divergência entre os valores teóricos e experimentais para a energia de formação da vacância, somados a dependência dos resultados devido à escolha do funcional de XC [27], faz com que esta inversão possa ser questionada.

Além disso, a difusão de moléculas de água no gelo ocorre como uma unidade, ou seja, a molécula como um todo migra de um sítio para outro. No caso da difusão do defeito intersticial, a barreira de migração é de 0,16 eV e com coeficiente de difusão sendo  $1,8 \times 10^{-6} \text{ m}^2\text{s}^{-1}$  [92]. Esse processo pode ocorrer na direção paralela ou perpendicular ao eixo  $c$  (Figura 5): no primeiro caso, a molécula se move dentro do mesmo hexágono, enquanto no segundo desloca-se entre hexágonos, expulsando uma molécula da rede para outro interstício. Embora Goto *et al.* [92] apontem maior frequência de salto na direção paralela, simulações de dinâmica molecular indicam que o coeficiente de difusão perpendicular é superior, mesmo apresentando maior barreira de migração [91].

Conforme visto no capítulo anterior, as redes neurais foram capazes de obter bons resultados para as propriedades estruturais do gelo  $I_h$ . Neste capítulo, será apresentada uma análise das energias envolvidas nos processos de formação e difusão dos defeitos, tais como a energia de ligação das moléculas intersticiais e de uma molécula no *bulk* de gelo, e a barreira de difusão dos defeitos. Além disto, será trazida uma análise da energia de formação destes defeitos e a influência da adição de termos não locais no funcional de troca e correlação sobre o ordenamento energético dos defeitos de interesse. Para expandir as análises, utilizaremos potenciais obtidos por meio de *machine learning*, evidenciando a abordagem computacional

para as simulações e o treinamento dos modelos. Por fim, será obtida a energia livre de ligação da molécula de água no contexto dos defeitos, considerando os efeitos de temperatura.

## 4.1 Métodos Computacionais

Para a análise das energias referentes aos defeitos, seguimos com duas abordagens distintas: a primeira com respeito aos cálculos DFT e a segunda segue o estudo utilizando redes neurais.

Os cálculos DFT foram realizados como referência (*benchmark*) para as redes neurais. Uma vez que a principal referência para a análise das energias de formação de cada defeito utiliza o funcional PW91 [176], iniciamos a nossa análise com o mesmo funcional e também com o funcional PBE [65]. Em seguida, passamos a analisar a influência dos termos não locais para o cálculo da energia de formação dos defeitos. Para isso, utilizamos o funcional SCAN [34] e também uma correção de van der Waals [33] nos cálculos com o funcional PBE. Para estes cálculos, a estrutura foi minimizada para cada funcional em função de diferentes parâmetros de rede.

Em relação à abordagem com as redes neurais, utilizamos os modelos gerados no Capítulo 3, nos níveis PBE e SCAN, como ponto de partida para melhorar a precisão dos modelos. Com as redes, geramos configurações por meio de simulações DP-MD para cada defeito, além de configurações de monômero e dímero. Na etapa seguinte, calculamos as forças e a energia de cada configuração utilizando cálculos DFT e com estes dados, refinamos os modelos treinados com configurações de água e gelo tanto com o código DeePMD-kit [4, 93, 153] quanto com o código MACE [5, 55].

Com os novos modelos treinados, replicamos os cálculos de minimização utilizando as redes para se compararmos os resultados com os provenientes de cálculos *ab initio*. Além disso, encontramos os modos normais para todos os defeitos para se determinarem as energias livres de formação. Finalizamos essa análise com a barreira de migração dos defeitos dentro da estrutura cristalina e também com a energia livre de ligação, considerando os efeitos de temperatura, obtida via integração termodinâmica.

### 4.1.1 Treinamento das Redes Neurais

Com o intuito de melhorar a precisão da rede, utilizamos as redes PBE e SCAN apresentadas no Capítulo 3 para gerar configurações para os diferentes defeitos, além do monômero e dímero de água. Todas as configurações de defeito foram extraídas de uma simulação DP-MD de 500 ps à temperatura de 220 K e volume constante (o volume médio obtido para o modelo da estrutura sem defeitos). As configurações para o dímero e o monômero foram extraídas de uma DP-PIMD com a mesma duração e temperatura, porém em uma caixa cúbica com aresta de 20 Å.

Todas as estruturas foram extraídas em intervalos de tempo uniformes. Para cada defeito, selecionamos 250 configurações e, para o monômero e o dímero, selecionamos 125 configurações cada. No total, 1.000 novas configurações foram incluídas em nosso conjunto de dados. Para cada configuração, obtivemos as forças e a energia por meio de cálculos DFT com os mesmos parâmetros apresentados na Seção 3.1.1.

Após a inclusão das energias e forças em nosso banco de dados, treinamos duas redes neurais em diferentes códigos. Primeiramente, utilizamos o DeePMD-kit com os mesmos parâmetros (com exceção do número de passos, reduzido para  $3 \times 10^6$ ) utilizados para o treinamento dos modelos apresentados no Capítulo 3 com as 20.000 configurações (sendo 90% destinadas ao treino e o restante ao teste). Utilizamos o modelo previamente treinado (sem os dados de defeitos e moléculas isoladas) como ponto de partida (*transfer learning*). Com a finalidade de melhorar a precisão da rede para as energias do monômero e do dímero, alteramos o peso correspondente para cada subconjunto: 30% para as configurações de água, 20% para as configurações de gelo  $I_h$  e cada defeito, além de 25% para as configurações de monômero e 25% para as de dímero.

Na Tabela 4.1, apresentamos o RMSE da rede treinada com dados PBE para cada subconjunto de dados. É possível perceber que a rede consegue capturar com fidelidade as forças e energias das configurações de água e com alta precisão as de gelo  $I_h$ , defeitos e monômero.

Para o segundo modelo, utilizamos apenas 18% dos dados gerados por Torres *et al.* [42], somados às 1.000 configurações de gelo  $I_h$  e às 1.000 configurações explícitas anteriormente. O treinamento deste segundo modelo, utilizando dados no nível SCAN, foi feito com o MACE. A quantidade de configurações para cada fase está presente na Tabela 4.2. A diminuição do número de configurações deve-se

RMSE	Subconjuntos						
	Água	Gelo	Tu	Vacância	Bc	Monômero	Dímero
Energia (meV/átomo)	0,55	0,24	0,11	0,12	0,13	0,07	1,25
Forças (meV/Å)	50,2	50,8	25,4	24,9	25,4	3,1	15,3

Tabela 4.1: RMSE para as energias (em meV/átomo) e forças (em meV/Å) de cada subconjunto para o modelo treinado com dados PBE no DeePMD-kit.

ao exposto no Capítulo 2 e, devido a isto, mantivemos os mesmos parâmetros apresentados anteriormente.

Subconjunto	Número de Configurações
Água	3240
Gelo	1000
Monômero	125
Dímero	125
Tu	250
Bc	250
Vacância	250

Tabela 4.2: Número de configurações utilizadas para cada subconjunto para o treinamento da rede utilizando o MACE.

Para se confirmar a precisão da rede, mostramos na Tabela 4.3 os RMSEs para força e energia do modelo para os subconjuntos de treinamento e de validação. É importante notar que, neste modelo, não separamos os erros por subconjuntos geométricos. Isto deve-se ao fato de as redes treinadas com MACE já apresentarem valores de RMSE bem abaixo do limiar de qualidade para a amostra total (na qual os erros nas configurações de água são mais relevantes devido ao volume de dados), não havendo necessidade de se explicitar que o modelo apresenta RMSEs ainda menores para as configurações de interesse (gelo  $I_h$ , defeitos e monômero).

Subconjunto	RMSE Energia (meV/átomo)	RMSE Forças (meV/Å)
Treinamento	0,3	3,4
Validação	0,2	3,6

Tabela 4.3: RMSE para as energias (em meV/átomo) e forças (em meV/Å) de cada subconjunto para o modelo treinado com dados SCAN no MACE.

### 4.1.2 Análise Energética

A análise energética dos defeitos pontuais envolverá as energias de ligação e de formação (e suas respectivas energias livres), além da investigação das energias envolvidas no processo de mobilidade dos defeitos no interior da estrutura do gelo.

O cálculo das energias de formação e de ligação compartilha a mesma ideia, modificando apenas o estado de referência (potencial químico) da molécula adicionada (ou removida) da estrutura cristalina do gelo  $I_h$ . Para o cálculo da energia de formação, o potencial químico corresponde a uma molécula de água no *bulk* do gelo, ou seja, representa a energia por molécula na estrutura do gelo  $I_h$ . Deste modo,  $\mu_{\text{form}} = E_{\text{gelo}}/n_{\text{gelo}}$ , e a equação para a energia de formação é reescrita da seguinte forma:

$$E_{\text{form}} = E_{\text{defeito}} - n_{\text{defeito}} \cdot \mu_{\text{form}}, \quad (4.1)$$

onde  $n_{\text{defeito}}$  corresponde ao número de moléculas do sistema com defeito.

Em relação ao contexto da energia de ligação, a molécula de referência está isolada no vácuo. Deste modo,  $\mu_{\text{bind}} = E_{\text{mol}}$ , e a energia de ligação é:

$$E_{\text{bind}} = E_{\text{completo}} - (E_{\text{dissociado}} + E_{\text{mol}}), \quad (4.2)$$

onde  $E_{\text{completo}}$  corresponde a energia total do sistema com maior número de moléculas, ou seja, para o defeito intersticial, consideramos o sistema com defeito e para o defeito de vacância, consideramos o gelo perfeito como o sistema completo.

Para finalizar esta análise, introduzimos as contribuições vibracionais para se encontrarem as energias livres de formação. Os detalhes para os cálculos dos modos normais serão apresentados a seguir, porém, a expressão geral para a energia livre é da seguinte forma:

$$F_{\text{form}} = E_{\text{form}} + \Delta U^{\text{vib}} - T\Delta S_{\text{vib}}, \quad (4.3)$$

onde  $\Delta U^{\text{vib}}$  refere-se à variação da energia vibracional e  $T\Delta S_{\text{vib}}$  corresponde à contribuição da entropia vibracional.

De modo geral, a expressão para a energia vibracional é dada por:

$$U^{\text{vib}} = \sum_i \frac{1}{2} h\nu_i + \frac{h\nu}{e^{h\nu/k_B T} - 1}, \quad (4.4)$$

onde  $\nu_i$  corresponde às frequências dos modos normais.

A variação da energia vibracional pode ser escrita da seguinte forma:

$$\Delta U^{vib} = U_{\text{completo}}^{vib} - n_{\text{completo}} \frac{U_{\text{gelo}}^{vib}}{n_{\text{gelo}}}. \quad (4.5)$$

A expressão para a entropia vibracional é dada pela seguinte equação:

$$S_{\text{vib}} = k_B \sum_i \left[ \frac{h\nu_i/k_B T}{e^{h\nu_i/k_B T} - 1} - \ln \left( 1 - e^{-h\nu_i/k_B T} \right) \right]. \quad (4.6)$$

A lógica para  $\Delta S_{\text{vib}}$  segue a mesma da equação 4.5. Em todos os casos, desprezamos os primeiros seis modos, correspondentes à translação e rotação do sistema como um todo.

Primeiramente, será apresentada uma análise realizada a partir de cálculos DFT. Utilizamos os funcionais PW91, SCAN e PBE com correção vdW para minimizar as forças das estruturas com o código VASP [162]. Nestes cálculos, utilizamos um *cutoff* de energia de 118 Ry e pseudopotenciais PAW. Com o funcional PBE, também utilizamos o VASP com os mesmos parâmetros. De modo a se investigar a influência da base escolhida, para minimizações com o código SIESTA [157], realizamos cálculos com duas bases diferentes: a primeira *double- $\zeta$*  e a segunda *triple- $\zeta$* . Em ambas as minimizações, foram mantidos o *mesh cutoff* de 500 Ry e pseudopotenciais de norma conservada [114]. Como feito por de Koning *et al.* [85], utilizamos os parâmetros de rede experimentais para as seguintes temperaturas: 10 K, 205 K e 265 K. Todos os valores de parâmetro de rede foram retirados a partir da Referência [16].

Após a obtenção das energias de ligação e formação a partir de cálculos de primeiros princípios, replicamos a análise utilizando as redes neurais treinadas com os potenciais DeePMD e MACE. Em todos os processos necessários para a minimização das estruturas utilizando as redes neurais, foi seguido o mesmo protocolo. O *software* LAMMPS [40] foi escolhido para se aplicar o algoritmo de gradiente conjugado com critério de convergência para a força. Como posteriormente foram calculados os modos vibracionais, utilizamos um limiar de  $10^{-8}$  eV/Å para a força. Dado o ganho computacional, utilizamos os parâmetros de rede experimentais de 18 temperaturas diferentes, todos extraídos da Referência [16].

O cálculo das energias livres inclui as contribuições vibracionais obtidas a partir

dos modos normais das estruturas de interesse. Para o cálculo dos modos normais, utilizamos as estruturas relaxadas para gerar deslocamentos de 0.01 Å para todos os átomos em todas as direções, ou seja, geramos  $6N$  novas configurações, onde  $N$  corresponde ao número de átomos da configuração de interesse. A partir disto, calculamos a energia e as forças destas novas configurações utilizando o LAMMPS [40] e, com o auxílio do phonopy [177, 178], encontramos a matriz hessiana e os modos normais apenas para o modelo treinado com o MACE.

Por fim, a análise energética encerra com as barreiras envolvidas na mobilidade dos defeitos. Resultados experimentais [92] evidenciam a difusão de moléculas intersticiais na estrutura do gelo  $I_h$  na direção paralela ao eixo cristalográfico  $c$ . Além disso, cálculos de dinâmica molecular [91] confirmam a mobilidade destes defeitos também na direção perpendicular ao mesmo eixo, com a molécula intersticial não se restringindo a um único tipo de sítio. Para se calcular a taxa de difusão,  $D$ , deve-se utilizar a equação:

$$D = \alpha \lambda^2 \Gamma, \quad \text{com} \quad \Gamma = \nu_0 e^{-E_m/k_B T}, \quad (4.7)$$

onde  $\alpha$  é um fator geométrico,  $\lambda$  representa a distância entre os saltos,  $\Gamma$  é a frequência de saltos,  $E_m$  é a barreira de migração e  $\nu_0$  é a frequência de tentativa.

Na Figura 4.2 retratamos os possíveis caminhos para a difusão dos defeitos. Como podemos observar, os defeitos intersticiais apresentam 12 possibilidades de difusão na direção perpendicular ao eixo  $c$ , uma vez que a molécula intersticial pode expulsar seis moléculas e após ser expulsa da rede cristalina, a molécula pode migrar para dois sítios possíveis. Na direção paralela ao eixo  $c$  a molécula possui duas possibilidades de difusão. Enquanto para a vacância, três moléculas podem ocupar o espaço vazio na rede cristalina.

O valor de  $\nu_0$  pode ser obtido através da Teoria Harmônica do Estado de Transição (HTST) [179]:

$$\nu_0 = \frac{\prod_i^{3N} \nu_i^R}{\prod_i^{3N-1} \nu_i^S}, \quad (4.8)$$

onde  $\nu^R$  e  $\nu^S$  representam os modos normais das estruturas de mínima energia e do estado de transição (ponto de sela), respectivamente. A diferença no número de frequências utilizadas deve-se ao fato de que o estado de transição possui um único modo vibracional imaginário.

Para se obter a energia de migração e os modos normais, utilizamos o pacote ASE [180]. A energia de migração foi calculada a partir do método CI-NEB (*Clim-*

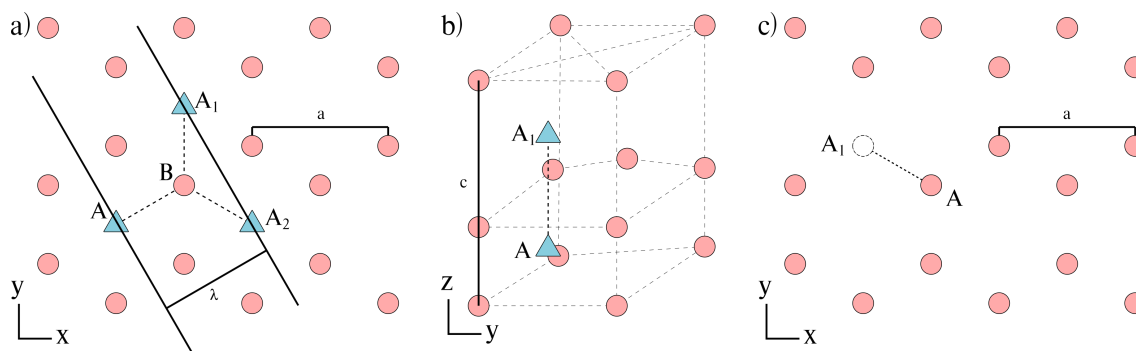


Figura 4.2: Análise geométrica dos saltos de defeitos. Os círculos representam as posições dos oxigênios da rede cristalina e os triângulos, os oxigênios intersticiais. Vistas sobre os planos (a)  $xy$  e (b)  $yz$  para o defeito intersticial e (c)  $xy$  para o defeito de vacância.

*bing Image Nudged Elastic Band*) [181] com as estruturas inicial e final previamente minimizadas (critério de força de  $0.01 \text{ eV}/\text{\AA}$ ). Após a relaxação, interpolamos 8 imagens, conectadas entre si por molas. A minimização de energia entre as imagens foi realizada até que a força máxima fosse inferior a  $0.05 \text{ eV}/\text{\AA}$ . Para os modos normais, o ASE repete o processo de deslocamentos finitos já descrito para o cálculo da matriz hessiana.

### 4.1.3 Efeitos de Temperatura na Dinâmica dos Defeitos

Até o momento, os efeitos de temperatura ficaram restritos ao tamanho do parâmetro de rede. A fim de se incluir os efeitos de temperatura de forma explícita, realizamos dois cálculos distintos. Primeiramente, realizamos cálculos de dinâmica molecular utilizando o potencial treinado com o DeePMD no nível SCAN, apenas para se verificar a mobilidade dos defeitos em função da temperatura. Para tal, essas simulações foram executadas para os defeitos Tu e de vacância. Em todas as simulações, os sistemas foram inicialmente equilibrados por 100 ps, seguidos por 2 ns de produção com um passo de tempo de 0.25 fs, e os dados foram coletados a cada 0.05 ps. As simulações foram conduzidas para as temperaturas de 10 K, 100 K e 265 K a volume constante, com os parâmetros de rede correspondentes ao valor experimental para cada temperatura.

Em seguida, estudamos as energias livres de ligação para os defeitos nas mesmas temperaturas. Nestas simulações, utilizamos o método de integração termodinâmica, apresentado no Capítulo 2. Todas as simulações utilizaram o código i-PI [41]. Os sistemas foram inicialmente equilibrados por 100 ps, seguidos

por 500 ps de produção com um passo de tempo de 0.25 fs, e os dados foram coletados a cada 0.05 ps. Para que não haja a difusão dos defeitos, incluímos um potencial harmônico para confinar a molécula intersticial. O processo consiste em duas etapas: na primeira, o potencial de confinamento é gradualmente desligado com a molécula isolada ( $\lambda_1 : 0 \rightarrow 1$ ). Na segunda, a molécula livre é gradualmente acoplada ao sistema com defeito ( $\lambda_2 : 0 \rightarrow 1$ ). Deste modo, a energia livre de ligação é dada pela seguinte expressão:

$$\Delta F_{\text{bind}} = \int_0^1 d\lambda_1 \langle U_{\text{harm}} \rangle_{\lambda_1} + \int_0^1 d\lambda_2 \langle U_{\text{completo}} - (U_{\text{dissociado}} + U_{\text{mol}} + U_{\text{harm}}) \rangle_{\lambda_2} \quad (4.9)$$

De forma complementar, a equação 4.3 pode ser adaptada para o cálculo da energia livre de ligação, na qual os efeitos de temperatura estão inclusos apenas implicitamente, por meio da entropia vibracional e do parâmetro de rede. Com isto, temos a seguinte expressão:

$$F_{\text{bind}} = E_{\text{bind}} + (U_{\text{completo}}^{\text{vib}} - U_{\text{dissociado}}^{\text{vib}} - U_{\text{mol}}^{\text{vib}}) - T(S_{\text{vib}}^{\text{completo}} - S_{\text{vib}}^{\text{dissociado}} - S_{\text{vib}}^{\text{mol}}). \quad (4.10)$$

## 4.2 Análise Energética

### 4.2.1 Análise Energética via DFT

A análise da energia de formação inicia-se com os resultados obtidos via DFT para os funcionais PW91 e PBE. Devido à similaridade dos funcionais, espera-se que os resultados obtidos sejam similares. A Figura 4.3 apresenta a energia de formação para ambos os funcionais. Neste momento, a energia de formação para o PBE foi calculada com a base *double- $\zeta$*  com o *software* SIESTA.

Verificamos que, apesar de as energias de formação do defeito Bc serem comparáveis entre os funcionais, o PBE apresenta um grande desvio para as energias de formação do defeito Tu. De modo a se analisar se os cálculos feitos com o SIESTA são suscetíveis à base utilizada, incluímos as energias de formação para uma base mais completa (usando o SIESTA) e para um cálculo com ondas planas. A comparação entre as diferentes abordagens é mostrada na Figura 4.4.

O funcional PBE apresenta resultados comparáveis ao PW91, sendo um importante ponto de partida, uma vez que os resultados para o funcional PW91 estão consolidados na literatura [85]. A escolha da base nos cálculos realizados com o

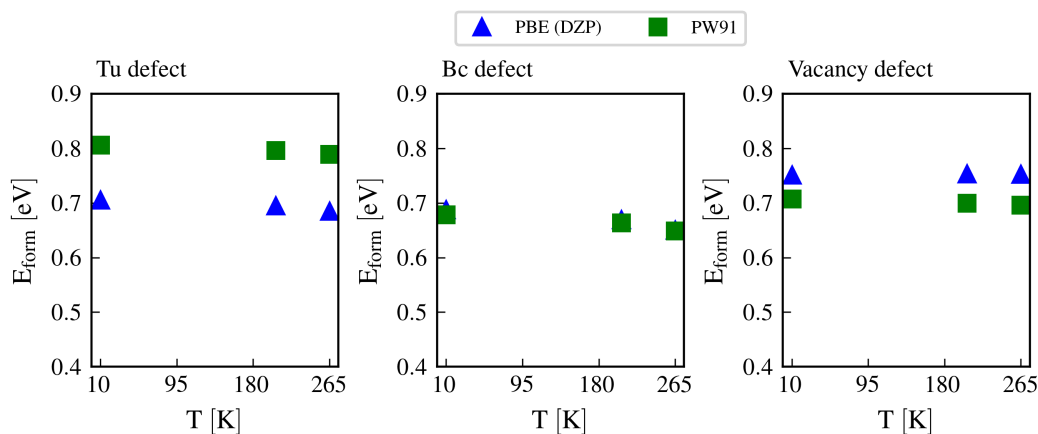


Figura 4.3: Energia de formação para cada defeito para os funcionais PW91 e PBE com base *double- $\zeta$* .

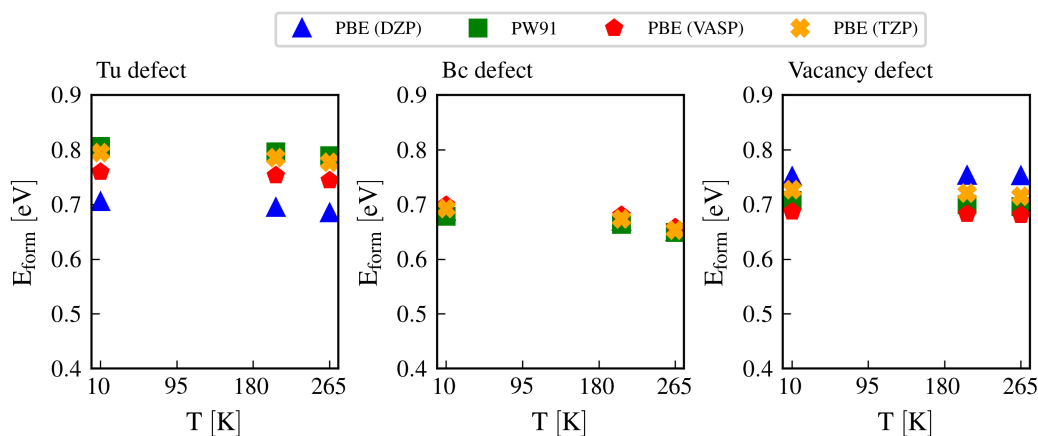


Figura 4.4: Energia de formação para cada defeito para o funcional PBE com diferentes abordagens, em comparação com o funcional PW91.

SIESTA influencia a qualidade dos resultados para a energia. Contudo, no Capítulo 3, verificamos que o funcional PBE não é uma boa escolha para o estudo de estruturas de gelo  $I_h$ . Portanto, a análise foi ampliada, incluindo efeitos não locais com o funcional SCAN e também interações de van der Waals nos cálculos com o funcional PBE. Na Figura 4.5, apresentamos a comparação entre as energias de formação. Observamos que a inclusão de termos não locais influencia as energias de formação principalmente dos defeitos intersticiais, alterando o ordenamento energético tanto em relação ao apresentado por este trabalho para o funcional PBE, quanto ao ordenamento mostrado por de Koning *et al.* [85].

Para finalizar as análises, compilamos na Tabela 4.4 os valores de energia de formação em função da temperatura para cada funcional estudado. Os dados

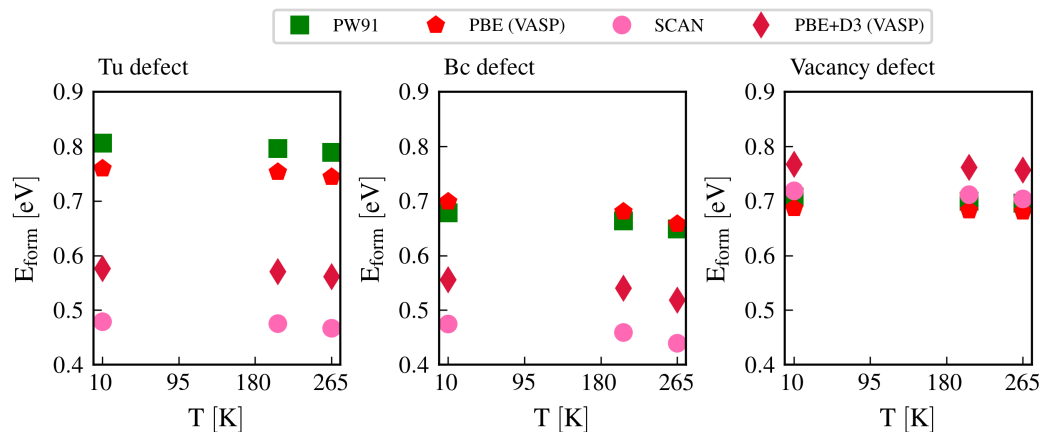


Figura 4.5: Energia de formação comparando funcionais locais com funcionais que incluem termos não locais.

referentes aos cálculos PBE foram obtidos com o código VASP.

Funcional	Defeito	Energia de Formação (meV)		
		10 K	205 K	265 K
PW91	Tu	806,2	796,6	789,8
	Bc	678,7	663,8	649,2
	Vacância	707,4	700,1	696,5
PBE	Tu	705,3	694,9	684,8
	Bc	688,1	669,9	657,0
	Vacância	751,6	753,5	752,6
PBE+D3	Tu	576,7	571,0	561,9
	Bc	556,2	540,2	518,8
	Vacância	767,5	761,8	757,3
SCAN	Tu	479,0	475,2	467,2
	Bc	474,5	459,3	439,6
	Vacância	718,7	712,1	703,9

Tabela 4.4: Energia de formação para os funcionais PW91, PBE, PBE+D3 e SCAN.

Finalizando esta análise, calculamos a energia de ligação da molécula de água no contexto de cada defeito. Vale ressaltar que, para a vacância, é considerada a energia de ligação de uma molécula na estrutura cristalina do gelo, enquanto para os defeitos intersticiais, o valor apresentado refere-se à energia de ligação da molécula extra na estrutura cristalina. Os valores são apresentados na Tabela 4.5. Verificamos que as energias de ligação da molécula na estrutura cristalina do gelo

são comparáveis para os dois funcionais, além de apresentarem uma diminuição em módulo quando se aumenta o parâmetro de rede da estrutura. Por outro lado, para os defeitos intersticiais, a ligação se torna mais forte ao se expandir o volume. Contudo, para o funcional PBE, observamos que as moléculas intersticiais são fracamente ligadas, com a energia de ligação sendo uma ordem de magnitude menor que as energias encontradas para o funcional SCAN.

Funcional	Defeito	Energia de Ligação (meV)		
		10 K	205 K	265 K
PBE	Tu	-6,7	-16,3	-25,3
	Bc	-23,8	-41,2	-59,4
	Vacância	-1463,6	-1464,6	-1462,7
SCAN	Tu	-198,6	-201,0	-252,2
	Bc	-218,8	-232,7	-284,5
	Vacância	-1399,7	-1390,5	-1344,9

Tabela 4.5: Energia de ligação para os funcionais PBE e SCAN.

## 4.2.2 Análise Energética via Redes Neurais

Com as novas redes neurais treinadas a partir dos modelos apresentados no Capítulo 3, minimizamos as estruturas para 18 parâmetros de rede diferentes. Considerando os dados DFT como referência, medimos a discrepância entre as energias de formação obtidas com os potenciais de ML e os resultados de referência DFT nos níveis PBE e SCAN. Como as redes neurais foram treinadas com dados obtidos com a base *double- $\zeta$*  no SIESTA, os resultados dos modelos foram comparados com essa abordagem DFT. Nesta etapa, utilizamos apenas os modelos treinados com o DeePMD-kit.

As Figuras 4.6 apresentam as energias de formação obtidas para diferentes parâmetros de rede utilizando os modelos PBE e SCAN, respectivamente, treinados com o DeePMD. Conforme apresentado, os modelos treinados com DeePMD-kit não capturam com fidelidade as energias de formação. A diferença máxima entre as energias de formação obtidas com DFT e com as redes neurais é de 40 meV para o defeito Bc. Para os demais defeitos, temos uma qualidade razoável nos resultados obtidos. Os valores exatos das diferenças são apresentados na Tabela 4.6.

Apesar da inclusão de estruturas de gelo  $I_h$  com defeitos no treinamento, os

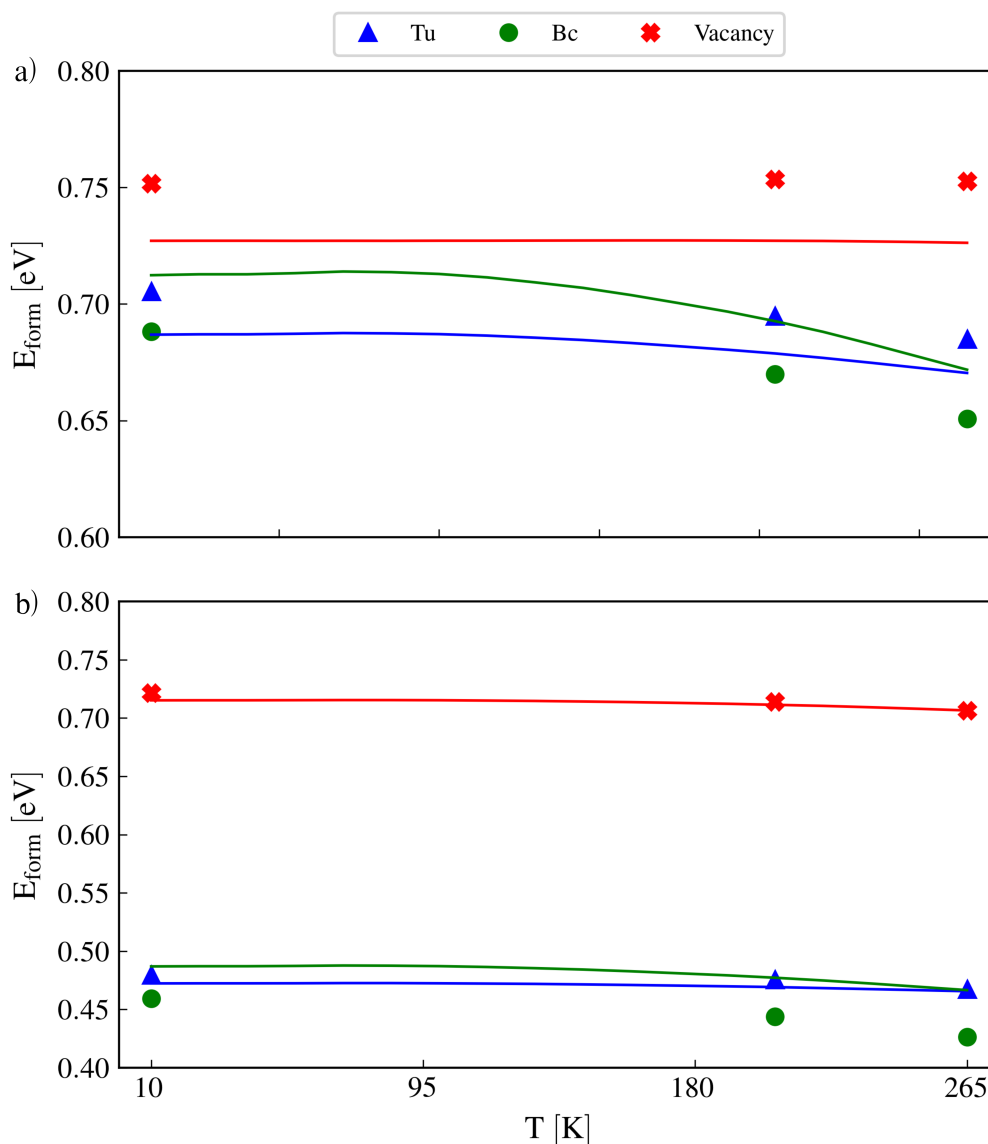


Figura 4.6: Energia de formação dos defeitos para os potenciais treinados com o DeePMD-kit no nível de precisão (a) PBE e (b) SCAN.

modelos não foram suficientes para atingir a precisão desejada. A rede treinada com qualidade PBE apresenta um largo desvio para todas as estruturas, enquanto a rede SCAN falha em representar a energia do defeito Bc.

Na Figura 4.7, apresentamos a comparação entre as energias de formação obtidas com o modelo treinado com os dados SCAN no *software* MACE e os dados de referência DFT. É notável que o modelo treinado utilizando o MACE apresenta baixos desvios na energia de formação, como mostrado na Tabela 4.7. Os erros correspondem a valores em torno de 1% da energia de formação obtida pelos

Funcional	T (K)	$\Delta E_{form}$ (meV)		
		Tu	Bc	Vacância
PBE	10	18,47	24,18	24,51
	205	16,07	22,75	26,36
	265	14,42	22,11	26,39
SCAN	10	7,16	27,57	6,45
	205	6,36	33,36	2,61
	265	1,78	40,05	0,06

Tabela 4.6: Diferença entre as energias de formação ( $\Delta E_{form} = E_{form}^{NN} - E_{form}^{DFT}$ ) para todos os defeitos e parâmetros de rede para os modelos treinados com os funcionais PBE e SCAN utilizando o código DeePMD.

cálculos de primeiros princípios. Deste modo, esta rede será a utilizada no restante deste trabalho.

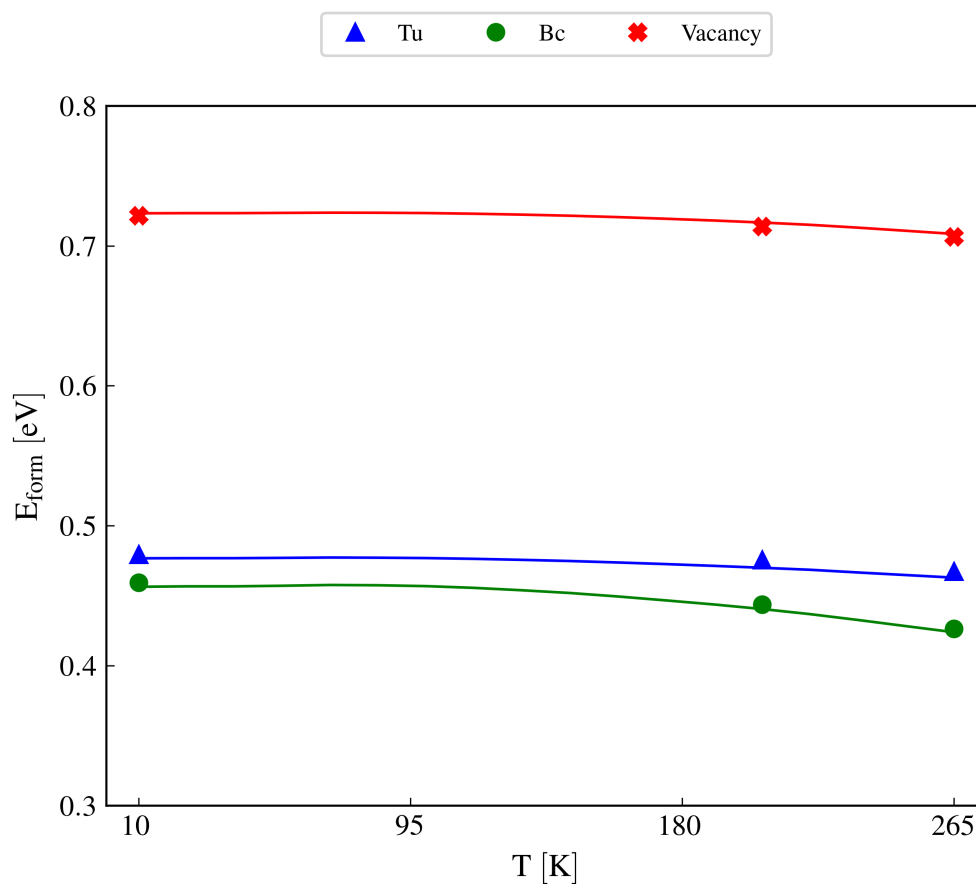


Figura 4.7: Energia de formação dos defeitos para o potencial treinado com o MACE no nível de precisão SCAN.

Funcional	T (K)	$\Delta E_{form}$ (meV)		
		Tu	Bc	Vacância
SCAN	10	2,75	2,94	1,60
	205	5,55	3,34	2,69
	265	4,46	2,59	2,06

Tabela 4.7: Diferença entre as energias de formação para todos os defeitos e parâmetros de rede para o modelo treinado com o funcional SCAN utilizando o MACE.

A energia de ligação obtida para cada defeito por este modelo é mostrada na Figura 4.8. Para o defeito de vacância, a energia de ligação refere-se à energia de ligação de uma molécula na estrutura cristalina do gelo. As diferenças de energia de ligação para os defeitos são da ordem de 5 meV para as temperaturas  $T = 10$  K e  $T = 205$  K. Entretanto, para o maior volume, as energias de ligação obtidas com a rede possuem desvios em torno de 35-40 meV. Apesar disto, verificamos um comportamento similar ao encontrado nos cálculos DFT, uma vez que as energias de ligação diminuem (em módulo) para a vacância com o aumento do volume, e o inverso ocorre para os defeitos intersticiais.

Para se computar a energia livre de formação, incluímos as contribuições de energia vibracional e de entropia vibracional. A energia livre de formação para cada defeito e temperatura é apresentada na Figura 4.9.

Como esperado, os efeitos entrópicos vibracionais têm maior relevância em altas temperaturas. A variação da energia vibracional é basicamente constante para todas as temperaturas, em torno de 60 meV para os defeitos Tu e vacância e 30 meV para o defeito Bc. Já a contribuição entrópica ( $T\Delta S$ ) inicia com valor nulo e, para as temperaturas mais altas, a contribuição é de 63 meV para o defeito Tu, 22 meV para o defeito Bc e 85 meV para a vacância. A contribuição entrópica para todos os parâmetros de rede está presente na Figura 4.10.

A partir da energia livre de formação, podemos encontrar a concentração de defeitos em função da temperatura utilizando a seguinte equação:

$$c = zN_s e^{-F_{form}/k_B T}, \quad (4.11)$$

onde  $N_s$  corresponde ao número de sítios por molécula de água em que o defeito pode ocorrer e  $z$  representa o número de possíveis orientações que a molécula intersticial tem em um dado sítio. Para o defeito Tu, temos que  $N_s = 1/2$  e  $z = 40$ ,

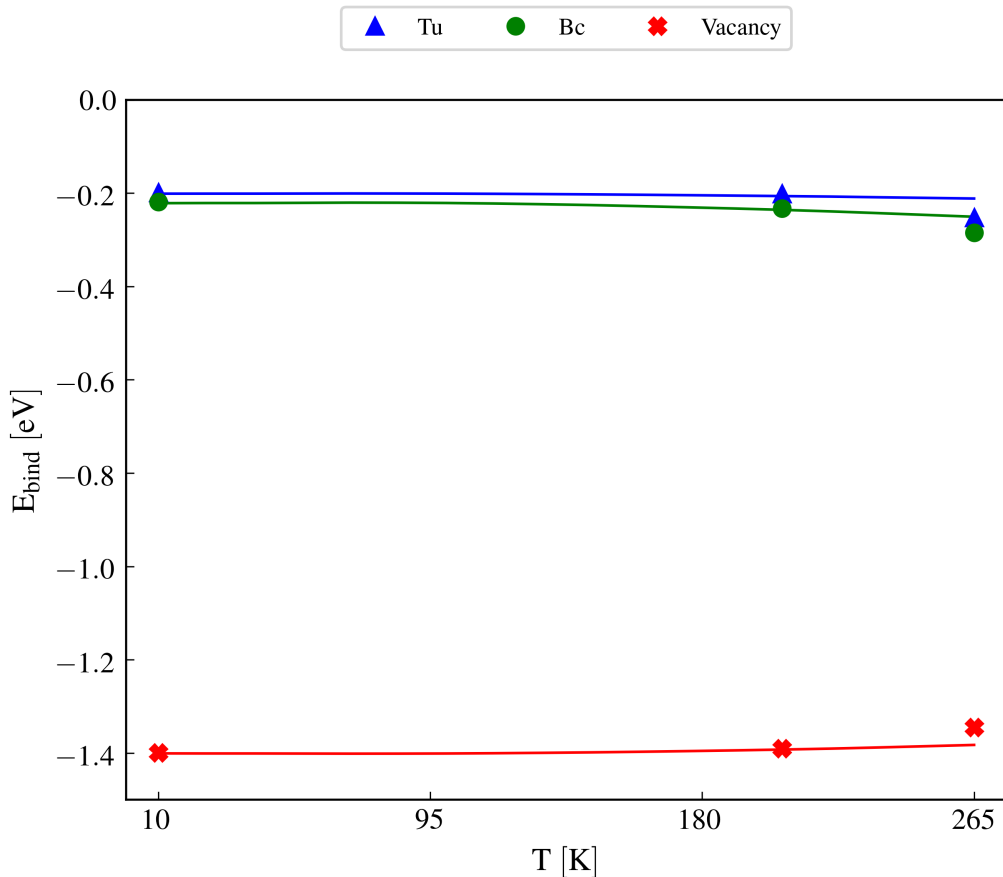


Figura 4.8: Energia de ligação dos defeitos para o potencial treinado com o MACE no nível de precisão SCAN.

enquanto para o defeito Bc, temos  $N_s = 2$  e  $z = 4$ . Para a vacância, ambos os parâmetros são iguais a 1. Todos os valores foram retirados da Referência [85].

Na Figura 4.11 apresentamos a concentração dos defeitos em função da temperatura. Diferentemente do encontrado por de Koning *et al.* [85], verificamos que não há uma transição entre a concentração de defeitos, com uma ordem totalmente estabelecida, sendo o defeito Tu o mais provável de ocorrer, e a vacância, o que possui a menor concentração. A diferença entre nossos resultados e os dados experimentais [88, 87, 89, 90] pode ser uma evidência de que a técnica experimental pode não ser precisa para a determinação da energia de formação da vacância. Outra possibilidade deve-se ao fato do defeito intersticial ser mais móvel.

Considerando a mobilidade dos defeitos, é possível medir a difusão destes na estrutura cristalina do gelo. Para tal, vamos analisar três diferentes mobilidades: a da molécula intersticial nas direções perpendicular e paralela ao eixo  $c$ , e a mobilidade da vacância. A representação gráfica das mobilidades foi apresentada

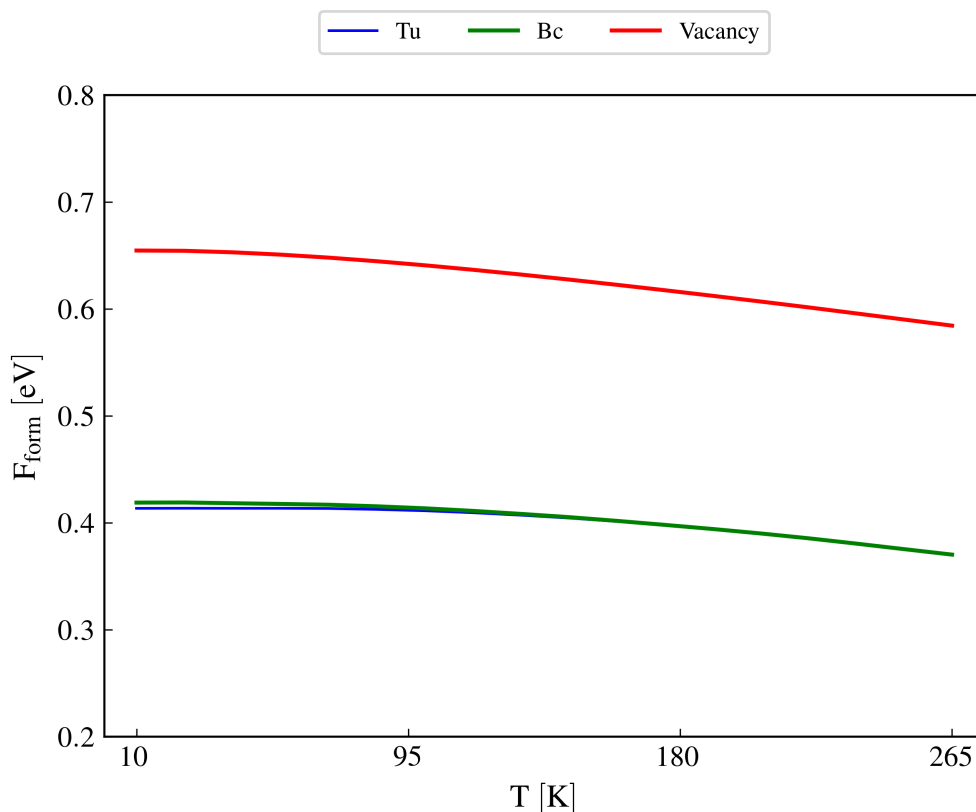


Figura 4.9: Energia livre de formação dos defeitos para o potencial treinado com o MACE no nível de precisão SCAN.

anteriormente na Figura 4.2. O cálculo da difusão dos defeitos depende de algumas etapas anteriores. O primeiro passo é calcular a barreira de migração para as diferentes mobilidades. Por definição, a energia de barreira de migração,  $E_m$ , é dada por:

$$E_m = E_S - E_R, \quad (4.12)$$

onde  $E_S$  corresponde à energia da estrutura do ponto de sela e  $E_R$  à energia da estrutura de menor energia.

Inicialmente, na Figura 4.12a) apresentamos o perfil de energia para todos os pontos do caminho de reação para o defeito de vacância, para todos os parâmetros de rede de interesse. Neste gráfico, é possível observar que a barreira de migração aumenta conforme se aumenta o parâmetro de rede.

De forma complementar, as imagens relaxadas durante o caminho de migração são apresentadas na Figura 4.12b). Devido à desordem protônica, a molécula que troca de sítio rotaciona, de modo a ocupar o espaço vazio e manter a orientação das ligações de hidrogênio. Para os defeitos intersticiais, temos duas diferentes

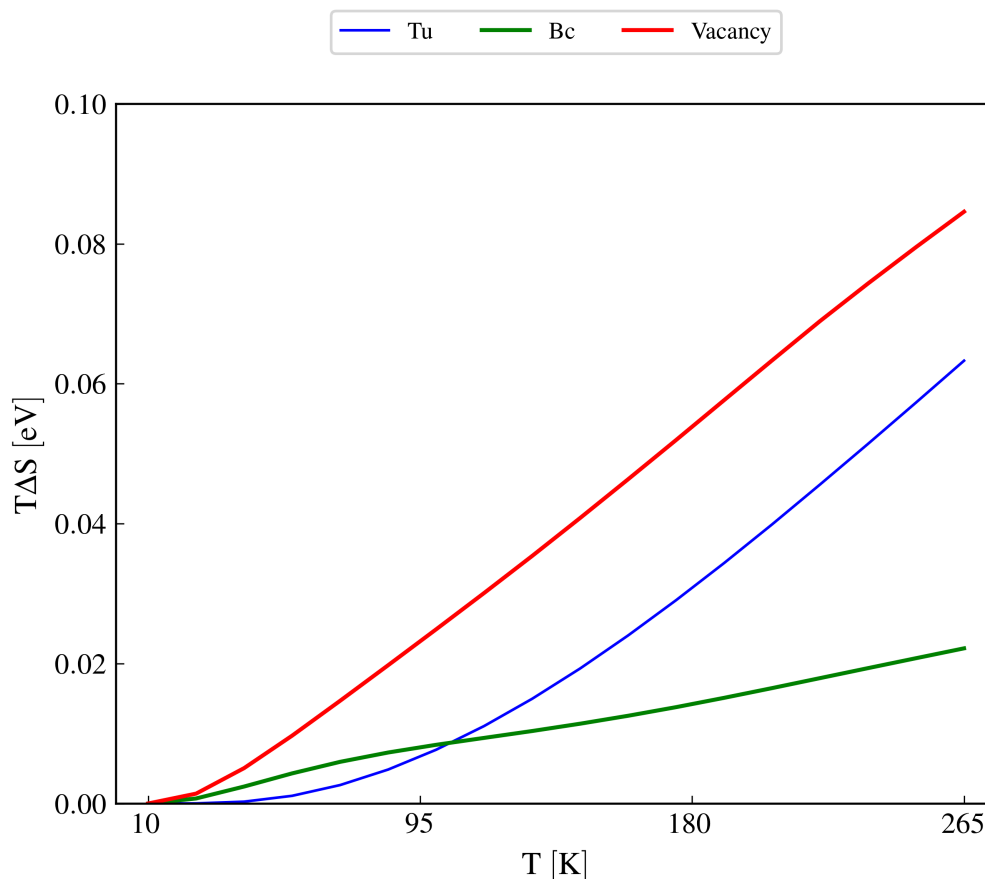


Figura 4.10: Contribuição vibraional ( $T\Delta S$ ) na energia livre de formação dos defeitos para o potencial treinado com o MACE no nível de precisão SCAN.

possibilidades, nas direções perpendicular e paralela ao eixo  $c$ . Primeiramente, na Figura 4.13a), apresentamos o perfil da barreira de migração para a difusão do defeito na direção perpendicular ao eixo  $c$ . De forma oposta ao verificado para a vacância, a barreira diminui ao se aumentar o parâmetro de rede. Assim, é maior a facilidade de o defeito se difundir em uma estrutura com maior volume. Além disso, a partir de resultados experimentais [92], encontramos que a barreira de migração dos defeitos intersticiais é de 0.16 eV, valor próximo ao encontrado teoricamente neste trabalho.

O caminho da mobilidade dos defeitos intersticiais na direção perpendicular ao eixo  $c$  é apresentado na Figura 4.13b). A difusão do defeito intersticial nesta direção ocorre com a molécula intersticial "expulsando" uma molécula de água da estrutura cristalina e ocupando seu lugar. Logo, a molécula expulsa passa a ocupar o sítio intersticial. Este mecanismo é distinto do apresentado por Goto *et. al* [92], onde os autores propõem que a molécula intersticial se difunde abrindo um

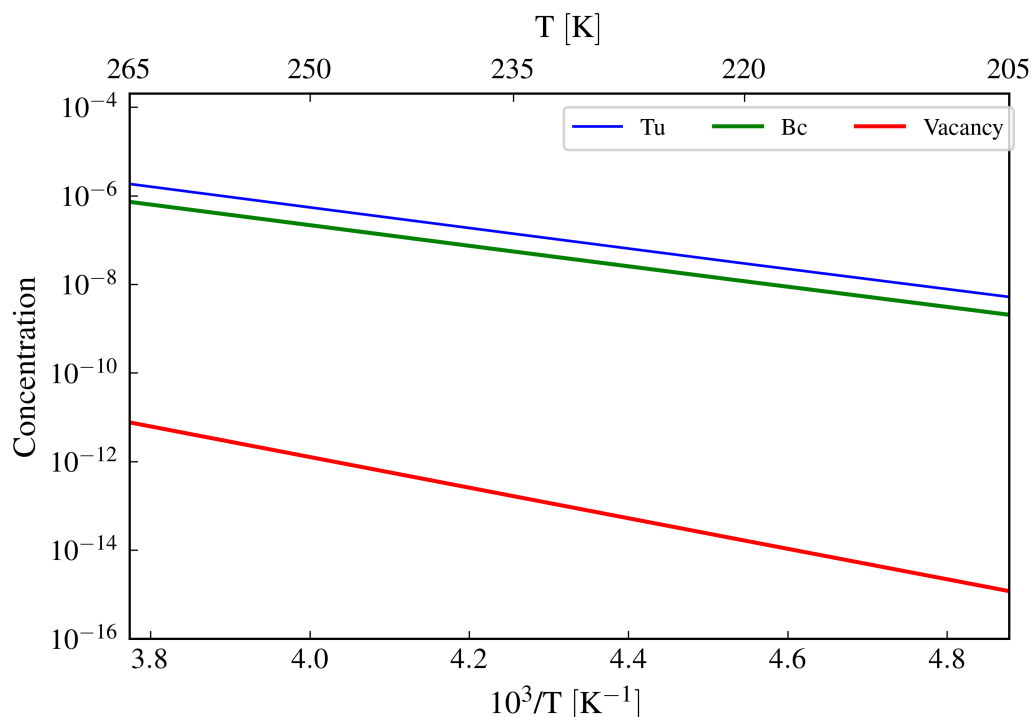


Figura 4.11: Concentração dos defeitos em função da temperatura, calculada com o potencial treinado com MACE no nível de precisão SCAN.

espaço entre outras duas moléculas. Outro ponto a se destacar é o aparecimento da estrutura referente ao defeito Bc na terceira imagem, evidenciando que o defeito Bc corresponde a um estado de transição para a difusão da molécula intersticial.

Contudo, esta não é a única direção em que o defeito pode se difundir. Na Figura 4.14a) apresenta o perfil de energia para as imagens referentes à mobilidade do defeito na direção paralela ao eixo *c*. Novamente, verificamos que a barreira de migração diminui com o aumento do parâmetro de rede, ou seja, essa característica está relacionada diretamente com o defeito, não havendo relação com a direção de difusão. O caminho pode ser visto na Figura 4.14b).

Finalizando a discussão sobre a barreira de migração, apresentamos na Figura 4.15 as energias de barreiras de migração, além da energia de barreira com a contribuição da energia vibracional para as três mobilidades analisadas e previamente discutidas. Fica evidente o fato da barreira de migração dos defeitos intersticiais diminuir ao expandir o volume e a barreira diminuir para os defeitos de vacância. Outro ponto a se destacar é o aumento da energia de barreira para os defeitos intersticiais quando adicionamos a contribuição vibracional para baixas temperaturas. Isto ocorre devido a energia de ponto zero ( $\hbar\nu/2$ ) ser maior para

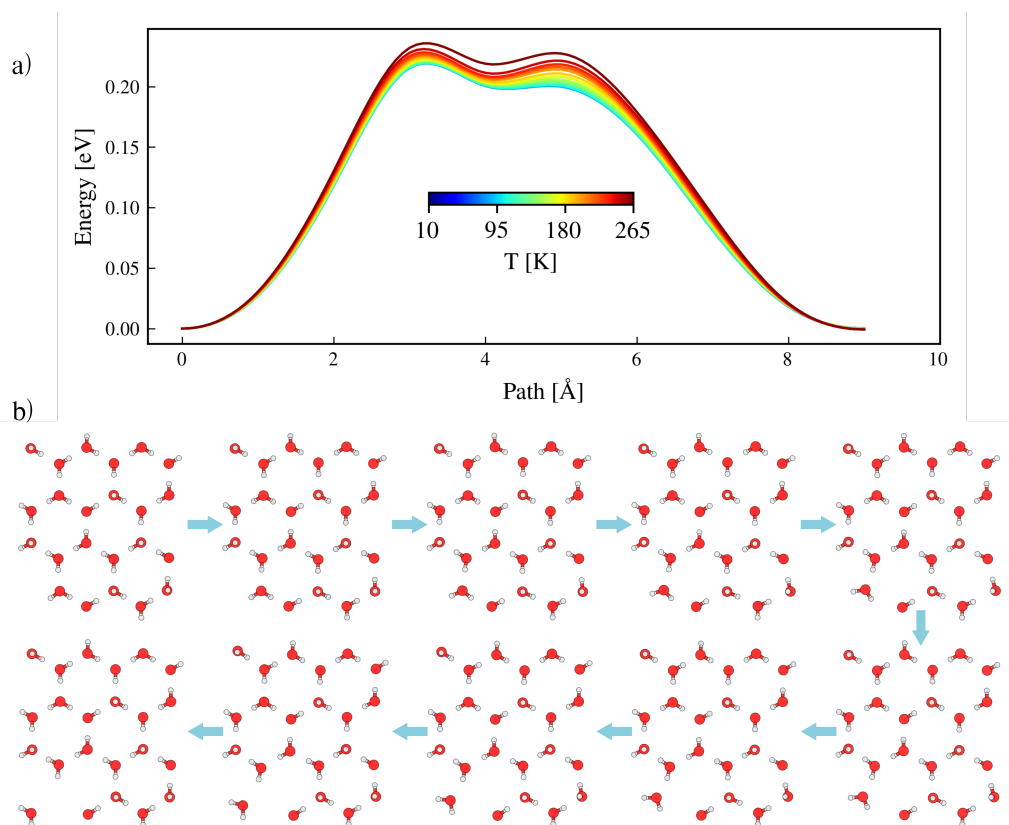


Figura 4.12: (a) Perfil de energia para as imagens ao longo do caminho de migração e (b) caminho de mobilidade para o defeito de vacância.

o ponto de sela que para o estado de mínimo para todas as temperaturas para a difusão paralela a direção  $c$  e para temperaturas abaixo de 200 K para a difusão perpendicular ao eixo  $c$ . Contudo, quando o termo relativo a temperatura é relevante, notamos uma diminuição da energia de barreira, como esperado.

Passando para a análise da frequência de tentativa de migração do defeito, calculamos esta quantidade de acordo com a equação 4.8. Na Figura 4.16, apresentamos a frequência de tentativa de migração do defeito para todos os parâmetros de rede. Verificamos que o defeito de vacância apresenta uma maior frequência de tentativa, com valores superiores aos da difusão do defeito intersticial em duas ordens de grandeza. Outro ponto a ser considerado é a diferença de regimes apresentada para os defeitos intersticiais. Enquanto a frequência de tentativa para a vacância apresenta um crescimento consistente ao se aumentar o parâmetro de rede, os defeitos intersticiais possuem dois regimes distintos para a mobilidade perpendicular ao eixo  $c$ : o primeiro, constante para baixas temperaturas; e o segundo, para temperaturas acima de 190 K, no qual a frequência de tentativa quadruplica, então, apresenta um comportamento decrescente.

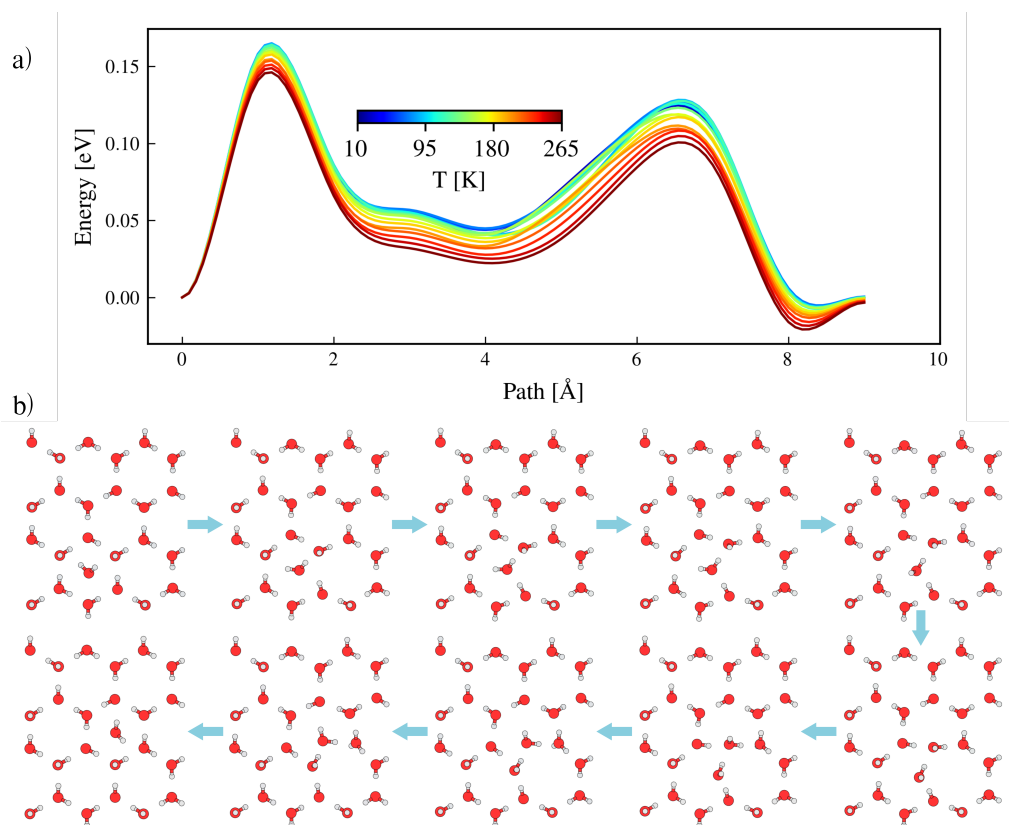


Figura 4.13: (a) Perfil de energia para as imagens ao longo do caminho de migração e (b) caminho de mobilidade para o defeito intersticial na direção perpendicular ao eixo  $c$ .

Com a frequência de tentativa e a energia de barreira, os coeficientes de difusão para ambos os defeitos são apresentados na Figura 4.17. Apesar da grande diferença entre os valores das frequências de tentativa, o coeficiente de difusão para cada defeito possui valores comparáveis.

Os defeitos de vacância, apesar de possuírem a maior frequência de tentativa, também possuem a maior barreira de migração. Deste modo, a difusão da vacância ocorre com menor frequência, em comparação com o defeito  $T_u$  na direção perpendicular ao eixo  $c$ . Similarmente, para a difusão dos defeitos intersticiais na direção paralela ao eixo  $c$ , apesar da menor barreira de migração, o defeito apresenta a menor difusão em comparação com os demais, devido à sua baixa frequência de tentativa.

Dados experimentais [92] mostram que a taxa de salto dos defeitos intersticiais paralelamente à direção  $c$ , com temperatura próxima ao ponto de fusão, é da ordem de  $\Gamma_{\parallel} \approx 1,4 \times 10^{10}$  Hz. Os autores do trabalho experimental afirmam que  $\Gamma_{\perp}$  deve ser menor que  $\Gamma_{\parallel}$ ; contudo, para altas temperaturas, encontramos

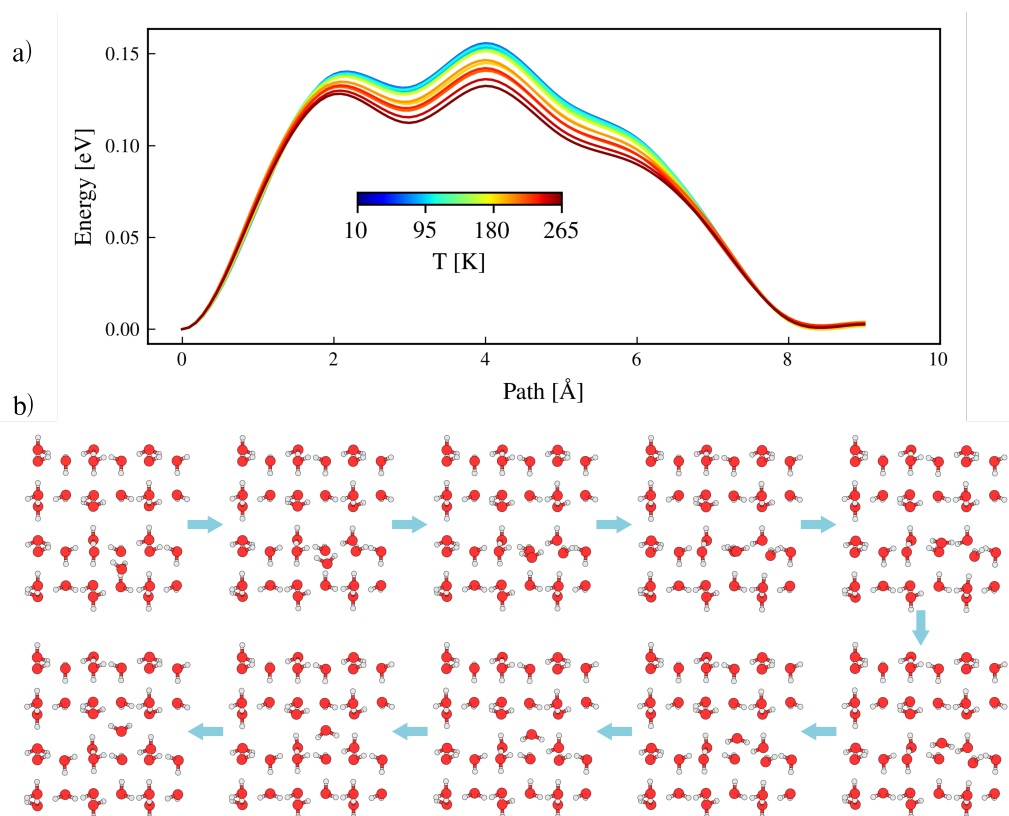


Figura 4.14: (a) Perfil de energia para as imagens ao longo do caminho de migração e (b) caminho de mobilidade para o defeito intersticial na direção paralela ao eixo  $c$ .

o comportamento oposto, com a difusão perpendicular ao eixo  $c$  acontecendo com maior frequência. Apesar desta afirmação, no trabalho experimental, não há uma explicação contundente sobre a medição ser feita apenas para a difusão paralela ao eixo  $c$ . Além disto, cálculos de dinâmica molecular com potencial clássico [91] indicam que a difusão é maior na direção perpendicular ao eixo  $c$ . É possível que a explicação dada a época surgiu de um equívoco em relação ao mecanismo de difusão, onde acreditava-se que a molécula intersticial passaria por um espaço estreito entre duas moléculas, o que seria energeticamente desfavorável [92]. Entretanto, apresentamos um mecanismo de difusão diferente, que possui uma barreira de migração próxima a difusão paralela ao eixo  $c$ , porém com uma frequência de tentativa superior, o que leva a um maior valor de difusão do defeito.

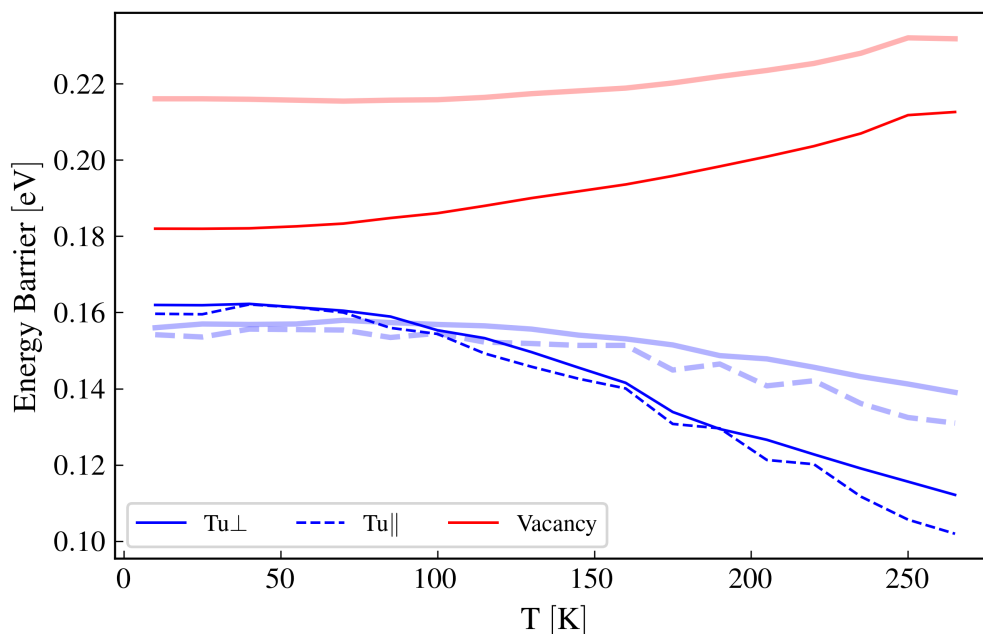


Figura 4.15: Barreira de migração para o defeito Tu (nas direções paralela e perpendicular ao eixo  $c$ ) e para o defeito de vacância. As linhas com transparência representam a energia de barreira, enquanto as sem transparência representam a energia de barreira com a adição da contribuição vibracional, enquanto a energia de barreira é

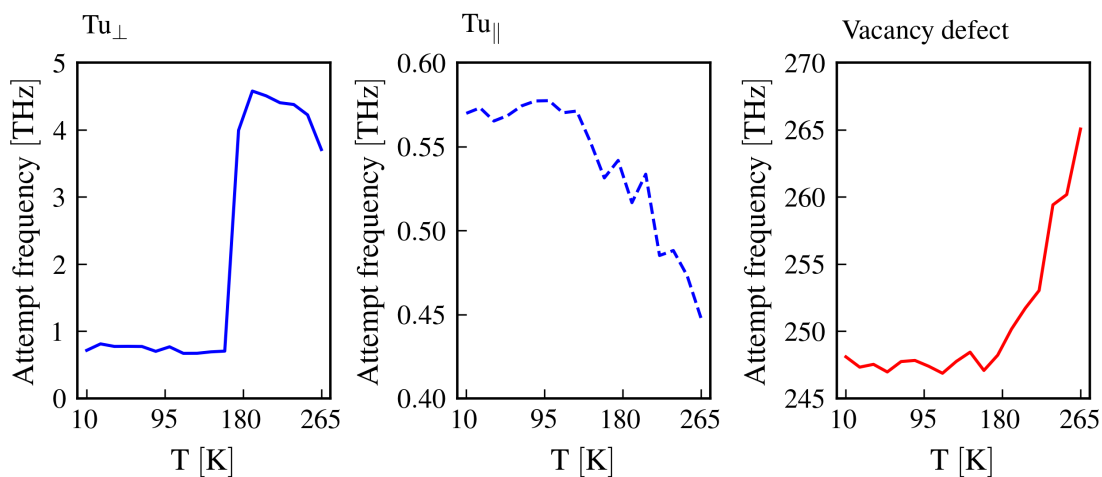


Figura 4.16: Frequência de tentativa de migração dos defeitos intersticiais e de vacância.

### 4.3 Efeitos de Temperatura na Dinâmica dos Defeitos

Até este momento, não foram discutidos os efeitos diretos da temperatura na mobilidade dos defeitos. A temperatura esteve envolvida até então apenas

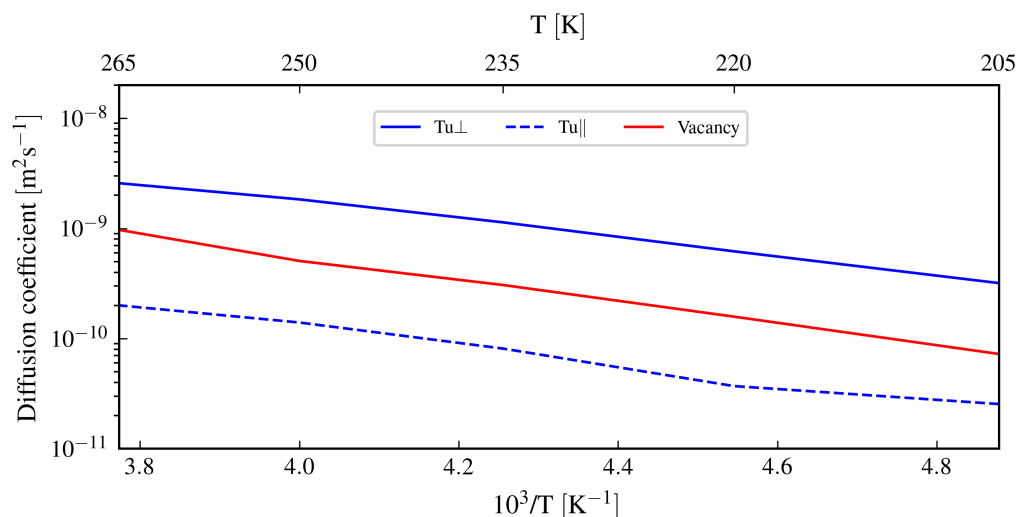


Figura 4.17: Coeficiente de difusão dos defeitos intersticiais e de vacância em função do parâmetro de rede.

de forma implícita, por meio dos parâmetros de rede da estrutura cristalina. Como mostrado no Capítulo 1, a integração termodinâmica nos permite calcular a energia livre de um sistema. Para os defeitos, podemos utilizar esta teoria para se encontrarem as energias livres de ligação. Inicialmente, podemos utilizar a mesma abordagem da energia livre de formação (baseada na aproximação harmônica) para se obter a energia livre de ligação para todos os parâmetros de rede, sem incluir explicitamente os efeitos dinâmicos da temperatura. Este resultado pode ser observado na Figura 4.18.

Os resultados obtidos a partir da estrutura relaxada nos mostram um aumento (em módulo) na energia livre de ligação para os defeitos intersticiais e o comportamento oposto para o defeito de vacância. Dando seguimento, passamos a calcular a energia livre de ligação considerando explicitamente os efeitos da temperatura por meio de MD. O número de pontos escolhidos para  $\lambda$  deve ser analisado caso a caso, com a finalidade de se obter uma curva suave para a integração. Para estes cálculos, definimos três temperaturas (10 K, 100 K e 265 K) para se observar a tendência da energia livre de ligação. Devido ao fato do defeito Bc ser apenas um estado de transição, seguiremos apenas com a análise sobre os defeitos Tu e de vacância.

Iniciamos a análise com o confinamento da molécula livre no espaço por meio de um potencial harmônico. As curvas da integração termodinâmica são apresentadas na Figura 4.19.

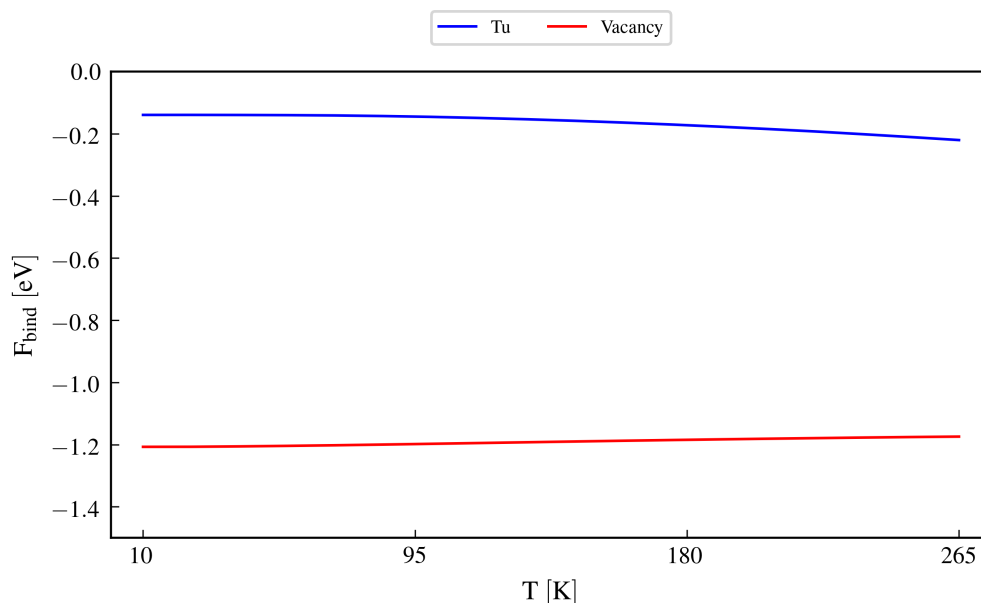


Figura 4.18: Energia livre de ligação para o defeito intersticial Tu e de vacância calculados a partir da aproximação harmônica.

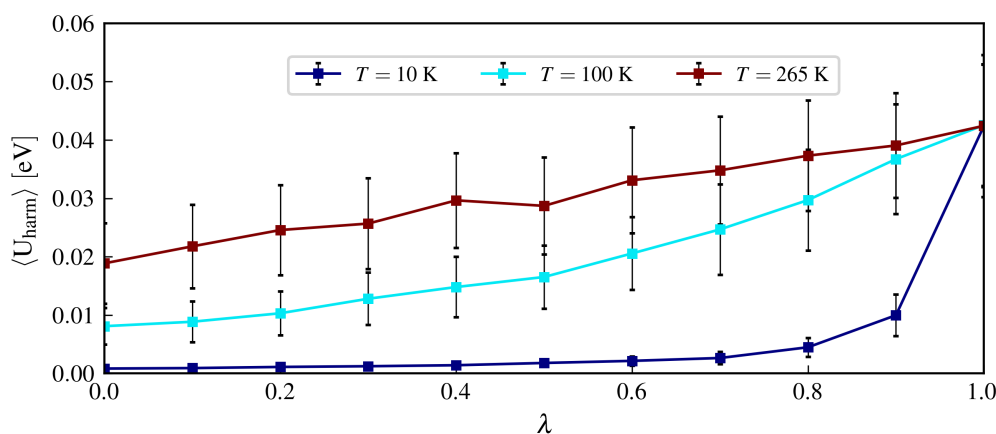


Figura 4.19: Derivada da energia livre em relação a  $\lambda$  para o potencial de confinamento em função do parâmetro  $\lambda$ , para a molécula isolada.

O valor obtido através da integração das curvas está presente na Tabela 4.8. Denotamos esta quantidade como  $I_1$ . Verificamos que, como esperado, ao se aumentar a temperatura, o valor da integral do potencial de confinamento aumenta.

Em sequência, passamos a calcular a segunda parte da equação 4.9. Nesta etapa, a interação entre a molécula isolada e a estrutura de interesse é gradualmente "ligada", enquanto o confinamento da molécula é desligado. Na Figura 4.20, apresentamos as curvas de integração termodinâmica para todos os defeitos

Temperatura	10 K	100 K	265 K
$I_1$ (eV)	$0,0042 \pm 0,0003$	$0,019 \pm 0,002$	$0,030 \pm 0,002$

Tabela 4.8: Valores obtidos para a integral do potencial de confinamento ( $I_1$ ) para a molécula isolada.

analisados. A integral sob cada curva é dada pela seguinte equação:

$$I_2 = \int_0^1 d\lambda_2 \langle U_{\text{completo}} - (U_{\text{dissociado}} + U_{\text{mol}} + U_{\text{harm}}) \rangle_{\lambda_2}. \quad (4.13)$$

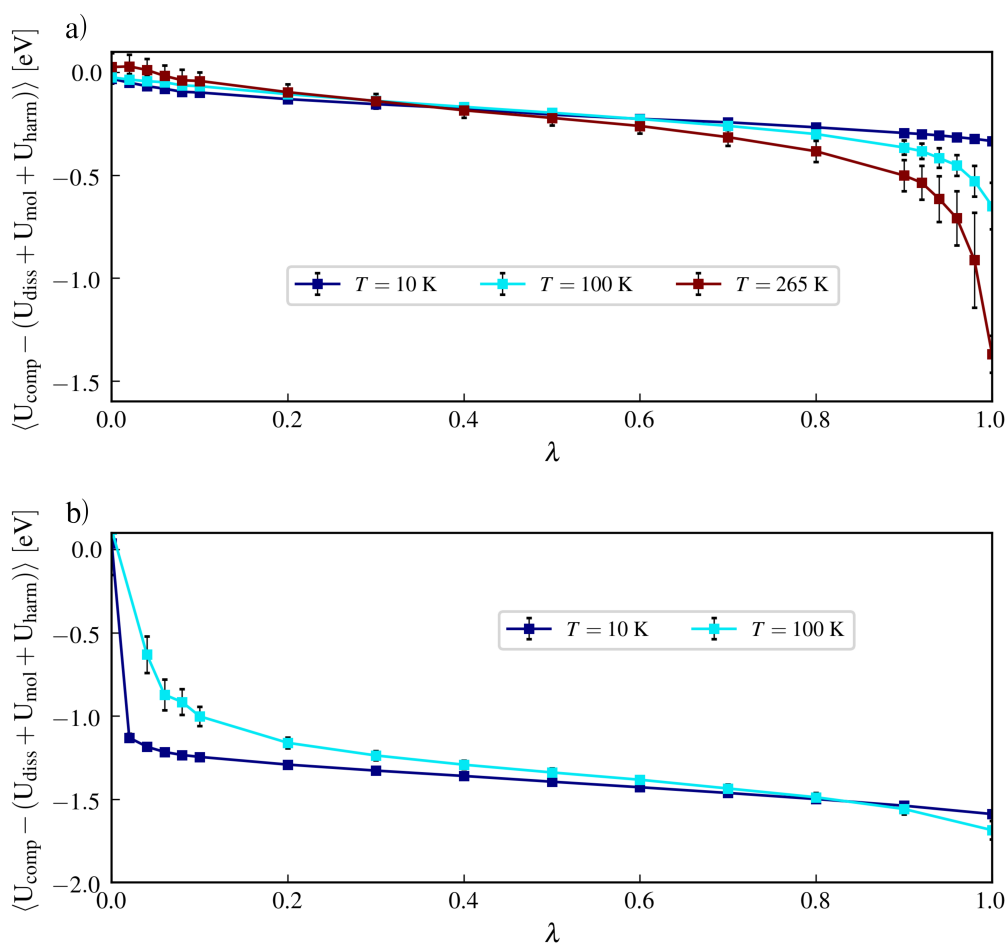


Figura 4.20: Derivada da energia livre em relação a  $\lambda$  para a diferença de energia entre os sistemas em função do parâmetro  $\lambda$ , para (a) o defeito intersticial  $T_u$  e (b) defeito de vacância.

É possível observar o mesmo comportamento obtido para a energia livre de ligação sem se considerarem os efeitos dinâmicos da temperatura. Quanto maior

a temperatura, menor (em módulo) é a energia de ligação para uma molécula na rede cristalina do gelo  $I_h$  (vacância). Enquanto para os defeitos intersticiais, a ligação se torna mais forte. Os valores para  $I_2$  são apresentados na Tabela 4.9.

Defeito	$I_2$ (eV)		
	10K	105K	265K
Tu	$-0,1893 \pm 0,0007$	$-0,210 \pm 0,001$	$-0,261 \pm 0,003$
Vacância	$-1,383 \pm 0,003$	$-1,29 \pm 0,01$	

Tabela 4.9: Valores obtidos para a integral  $I_2$  (interação molécula-sistema) para cada defeito.

Finalmente, com os valores de  $I_1$  e  $I_2$ , podemos encontrar a energia livre de ligação de cada defeito para cada temperatura. Na Figura 4.21, comparamos a energia livre de ligação obtida pelos dois métodos apresentados. Verificamos que o método realizado por meio da aproximação harmônica (minimização) apresenta a mesma tendência do resultado obtido através da integração termodinâmica (MD). Contudo, a inclusão dos efeitos de temperatura aumenta, em módulo, a energia livre de ligação para ambos os casos, comparado apenas com a aproximação harmônica. Notamos que a aproximação harmônica apresenta baixos desvios para as temperaturas de 100 K e 265 K para ambos os defeitos, com desvios na faixa de 4%-7% comparado aos resultados obtidos via integração termodinâmica. Contudo, para a temperatura de 10 K o desvio é alto, sendo 24% para o defeito Tu e 13% para a vacância.

Dada a geometria dos defeitos, podemos estimar o valor da energia livre por ligação de hidrogênio com os efeitos de temperatura. No caso do defeito Tu, a molécula intersticial forma apenas duas ligações de hidrogênio, enquanto para o defeito de vacância, formam-se quatro. Logo, como era esperado, a energia por ligação de hidrogênio para uma molécula na estrutura cristalina do gelo  $I_h$  (caso da vacância) possui o maior valor (em módulo), em comparação com o defeito intersticial, uma vez que a molécula intersticial pode ter qualquer orientação na estrutura.

## 4.4 Considerações Finais

A partir do estudo da mobilidade dos defeitos pontuais na estrutura cristalina do gelo  $I_h$  constatamos que a inclusão de termos não locais no funcional de

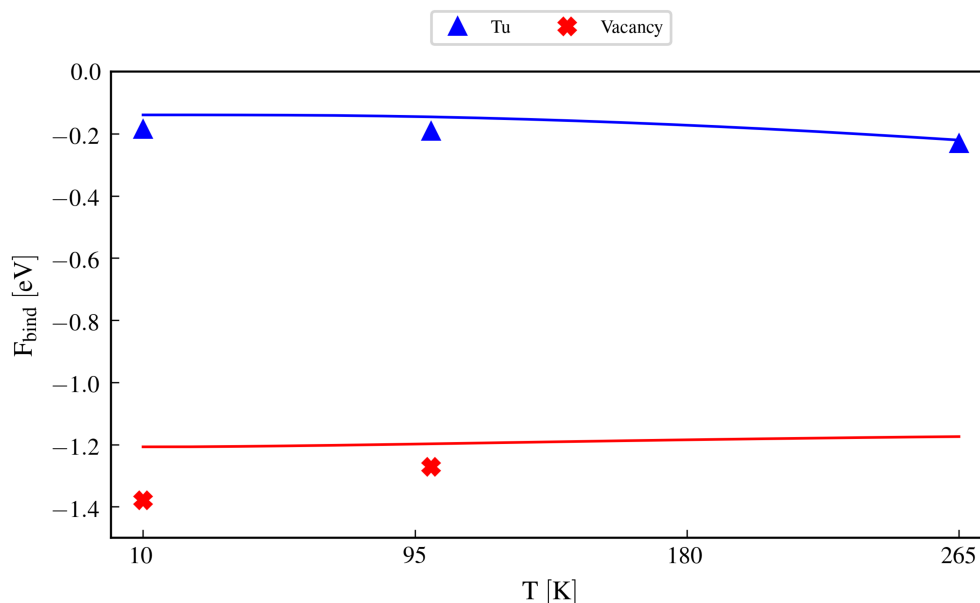


Figura 4.21: Comparação da energia livre dos defeitos intersticiais Tu e de vacância entre os métodos de aproximação harmônica e integração termodinâmica. A linha contínua representa os cálculos obtidos com a aproximação harmônica, enquanto os pontos representam os resultados obtidos por meio da integração termodinâmica.

correlação e troca altera o ordenamento da energia de formação, uma vez que os defeitos intersticiais, antes menos energeticamente favoráveis, se tornam os mais favoráveis. Porém, essa adição não resulta em uma mudança significativa na energia de formação do defeito de vacância. Além disso, o defeito de vacância apresenta uma barreira de migração mais elevada que o defeito Tu. Um aspecto particularmente interessante é a forma como essas barreiras respondem à variação do parâmetro de rede: à medida que a rede é expandida, a barreira associada ao defeito de vacância aumenta, enquanto a do defeito Tu tende a diminuir. No caso do defeito Tu, as duas direções possíveis de mobilidade exibem o mesmo comportamento de redução da barreira, sendo a migração paralela ao eixo  $c$  a energeticamente mais favorável.

Embora o defeito de vacância possua uma barreira mais alta, ele também apresenta uma frequência de tentativa de migração superior, o que resulta em um coeficiente de difusão mais elevado. Por outro lado, o defeito Tu, devido à sua geometria e à baixa energia de migração, difunde-se mais facilmente na direção perpendicular ao eixo  $c$ . O mecanismo inicialmente proposto para a difusão desse defeito considerava que a molécula localizada no interstício se deslocava

como uma unidade única dentro da rede cristalina, atravessando os espaços disponíveis. No entanto, nossas simulações indicam um mecanismo distinto: a molécula intersticial empurra uma molécula pertencente à rede, ocupa seu lugar, e esta, por sua vez, migra para outro interstício. Assim, o processo de difusão não ocorre pelo movimento de uma única molécula, mas sim pela propagação do defeito através da substituição sucessiva entre diferentes moléculas de água.

Adicionalmente, calculamos a energia livre de ligação dos defeitos  $T_u$  e de vacância, considerando os efeitos da temperatura por meio da integração termodinâmica. Verificamos que, para o defeito de vacância, o módulo da energia livre de ligação diminui com o aumento da temperatura, refletindo o enfraquecimento das interações entre as moléculas de água na rede. Em contraste, o defeito intersticial apresenta o comportamento oposto, com o aumento do módulo da energia livre de ligação conforme a temperatura se eleva.

# Capítulo 5

## Considerações Finais

Nesta tese abordamos o estudo de propriedades de água e gelo a partir de potenciais obtidos através de *machine learning*. Para a obtenção de um bom potencial, treinamos diferentes redes com diversos parâmetros a fim de comparar a discrepância entre dados obtidos pela rede com dados de referência, obtidos através da Teoria do Funcional da Densidade. Mostramos que para sistemas de água, uma rede treinada durante  $7 \times 10^6$  passos obtém sucesso em conservar a energia total do sistema em longas simulações, enquanto para a rede treinada durante  $1 \times 10^6$  passos observamos que a energia total do sistema estabiliza em dois patameres distintos durante o tempo total da simulação. Por outro lado, uma rede com um número reduzido de moléculas, além do modelo aprender a informação de moléculas dentro de um raio menor, como um sistema com 64 moléculas de água e coletando a informação de moléculas dentro do raio de corte de  $6,0\text{\AA}$ , consegue reproduzir as propriedades estruturais de um sistema com 256 moléculas de água, ou seja, os potenciais de *machine learning* são capazes de transferir a informação de um sistema pequeno e escalar para um sistema com um número grande de moléculas.

Por outro lado, uma rede que contém apenas a informação de configurações de água não possui boa acurácia quando aplicada para o estudo de gelo  $I_h$ , principalmente para prever as forças entre as espécies atômicas. Além disso, a inclusão de dados de gelo com volume fixo não resultam em melhoria do modelo, uma vez que os erros para força e energia preditas para este modelo, tanto para dados de água e gelo, estão acima do considerável bom neste estudo. Estas análises foram feitas utilizando o código DeePMD [4]. Os modelos gerados a partir deste código apresentam baixos erros e bons resultados para propriedades estruturais. Contudo, apresentamos uma alternativa as redes DNN, o código MACE [5], que utiliza redes neurais de grafos. Um modelo treinado com o código DeePMD já mostrou eficiente para calculos estruturais e dinâmicos com 30% dos dados utilizados nesta tese. Porém, um modelo treinado com o código MACE necessita apenas de 6% destes mesmos dados para obter os mesmos resultados obtidos com

o modelo treinado com o *dataset* completo. Isto é vantajoso para a diminuição do custo computacional para o treinamento das redes.

Após a revisão sobre a construção de uma rede neural precisa, partimos a estudar os efeitos quânticos nucleares na estrutura do gelo e sua influência nas interações interatômicas. Primeiramente, observamos que a inclusão de termos não locais faz com que a densidade do gelo aumente, comparado com o funcional PBE, sem a inclusão destes termos. Contudo, o funcional PBE viola o comportamento anômalo para água, uma vez que para este funcional, obtemos uma densidade menor da água comparada a densidade do gelo  $I_h$ . Desta forma, as ligações de hidrogênio são superestimadas para os funcionais optB88, vdW-cx e SCAN, fazendo com que as moléculas de água se aproximem, tornando o gelo mais denso. Esses resultados indicam que os funcionais do tipo GGA, assim como o SCAN, tendem a superestimar a ligação entre as moléculas de água no gelo, produzindo uma rede de ligações de hidrogênio mais forte, a qual é ainda intensificada pela inclusão das interações de van der Waals e dos efeitos quânticos nucleares.

Ao incluir os NQE verificamos um aumento na densidade, o que nos comprova que a inclusão de tais efeitos fortalece as ligações de hidrogênio, apesar de enfraquecer as ligações covalentes. Como resultado direto, observamos o surgimento de valores positivos da coordenada  $\nu$ . De modo oposto, as funções de distribuição radial obtidas com DP-PIMD mostraram-se, em geral, mais precisas do que aquelas calculadas com DP-MD, sobretudo pela melhor descrição da distribuição O–H. No caso da distribuição O–O, não observamos mudanças relevantes na presença dos efeitos quânticos nucleares em longas distâncias; entretanto, para valores abaixo de 4Å, notamos um leve deslocamento do pico para menores distâncias.

Em relação ao estudo referente aos defeitos pontuais na estrutura cristalina do gelo  $I_h$ , observamos que os funcionais PBE e PW91 apresentam o mesmo ordenamento da energia de formação, onde o defeito Tu é menos favorável energeticamente de ser formar. Apesar disto, os valores ainda distam do valor experimental. Entretanto, verificamos que para os funcionais SCAN e PBD+D3, o ordenamento se altera, com os defeitos intersticiais apresentando energias de formação mais baixas e com valores próximos ao resultado experimental. Porém, mesmo com esta nova abordagem, a energia de formação do defeito de vacância não apresenta mudança significativa, o que nos permite supor que as técnicas experimentais podem ser defasadas, não possibilitando a medida com precisão. Outro ponto a destacar é a diferença na energia de ligação da molécula intersticial entre os funcionais SCAN e PBE. Para o funcional SCAN, verificamos que a energia de

ligação é baixa, com a molécula sendo fracamente ligada. Enquanto isso, para o funcional SCAN, a energia de ligação é 30 vezes superior ao funcional PBE.

A adição dos modelos treinados com o MACE nos possibilitou a melhoria na precisão dos resultados. Com um número reduzido de configurações no treino a energia de formação da rede neural de grafos foi inferior a 6 meV comparado ao resultado DFT, enquanto para a rede neural profunda, as diferenças de energia chegaram a 40 meV para o defeito Bc. Neste ponto, apesar de um custo computacional mais alto, a rede treinada pelo código MACE [55, 5] é mais eficiente para uma análise energética. Com os resultados de energia de formação, foi possível calcular a concentração dos defeitos na estrutura cristalina do gelo  $I_h$ . Diferentemente do observado na literatura, em nossa análise, não se observa a inversão de concentração dos defeitos, sendo os defeitos intersticiais presentes em maior concentração.

A mobilidade dos defeitos também é importante de ser discutida. Verificamos que o defeito de vacância possui uma maior barreira para se propagar comparado ao defeito Tu. Outro ponto relevante a ser discutido é o aumento desta barreira para o defeito de vacância ao passo em que o parâmetro de rede é expandido e o comportamento inverso para os defeitos Tu. As duas possíveis direções de mobilidade do defeito Tu apresentam o mesmo perfil de diminuição, tendo a direção paralela ao eixo  $c$  com menor barreira. Apesar da maior barreira, o defeito de vacância é o defeito que apresenta maior frequência de tentativa de migração de um sítio a outro, aumentando, assim, o coeficiente de difusão. Devido a sua geometria e o baixo valor da energia de migração, o defeito Tu apresenta maior coeficiente de difusão na direção perpendicular ao eixo  $c$ . O mecanismo da difusão do defeito Tu na direção perpendicular era proposta como a molécula que está presente no interstício se movesse como uma unidade no interior da rede cristalina, difundindo-se por espaços estreitos. Contudo, observamos que a difusão desta molécula é feita a partir de outro mecanismo, onde a molécula intersticial expulsa uma molécula da rede cristalina, ocupa sua posição e a molécula expulsa escolhe outro interstício. Sendo assim, não temos uma única molécula que se difunde, mas sim, o defeito que se propaga através de diferentes moléculas de água.

Por fim, calculamos a energia livre de ligação para os defeitos Tu e vacância com efeitos de temperatura por meio da integração termodinâmica. Verificamos que a energia livre de ligação do defeito de vacância diminui em módulo ao aumentar a temperatura, enfraquecendo a ligação entre as moléculas de água na rede cristalina. Para o defeito intersticial, observamos o comportamento oposto, com a energia

livre de ligação aumentando em módulo com o aumento da temperatura.

Com esta tese foi possível responder todas as questões feitas na Introdução, possibilitando o estudo de propriedades de água e gelo  $I_h$  através de potenciais obtidos através de *machine learning*. Esta possibilidade é importante principalmente para o estudo da inclusão de NQE, uma vez que cálculos AIMD, que já são altamente custosos, demandam um longo tempo computacional para um tempo curto de simulação e para cálculos de barreira de migração, o que nos permite a inclusão de um grande número de réplicas para melhorar a análise do perfil de energia.

# Apêndice A

## Unidades Atômicas

Para o primeiro nível de simplificação da equação de muitos corpos, Equação 1.13, nos apoiaremos nas unidades de medida. Algumas grandezas presentes na equação já são conhecidas do ponto de vista experimental, tais como as massas do elétron e do próton, a constante de Planck reduzida, a carga do elétron e também a permissividade no vácuo. Estas grandezas são constantes e independentes em relação ao material.

Vamos imaginar o sistema mais simples possível: o átomo de hidrogênio. Nele, temos apenas um elétron e um núcleo. No estado fundamental, o sistema possui um raio  $a_0$ . Desta forma, podemos escrever a energia de interação de acordo com a equação 1.3 em módulo:

$$E_{Ha} = \frac{e^2}{4\pi\epsilon_0 a_0}. \quad (\text{A.1})$$

A equação A.1 traz a unidade "Ha", que corresponde a Hartree, ou seja, a energia de interação entre um par de elétrons ou de núcleos é da mesma ordem de  $E_{Ha}$ . Porém, ainda não há uma forma de simplificar a equação 1.13. Quando Bohr propôs seu modelo atômico para o hidrogênio, obteve um resultado a partir do momento angular do estado fundamental, uma vez que a trajetória do elétron é quantizada. Assim:

$$m_e v a_0 = \hbar, \quad (\text{A.2})$$

com  $v$  sendo a velocidade do elétron. Por estar em um referencial não inercial, o elétron sofre uma força centrífuga, que se iguala em módulo à força de interação devido ao núcleo. Portanto:

$$m_e \frac{v^2}{a_0} = \frac{e^2}{4\pi\epsilon_0 a_0^2}. \quad (\text{A.3})$$

Ao observar as equações A.1 a A.3, verificamos que a energia cinética do elétron é  $E_{Ha}$ . Sendo assim, ao observar a equação 1.13, notamos que os termos referentes às interações Coulombianas também apresentam uma relação com  $E_{Ha}$ . Logo,

por conveniência, dividem-se os dois lados da equação 1.13, reescrevendo-a da seguinte forma:

$$\left[ - \sum_{i=1}^N \frac{1}{2} a_0^2 \nabla_i^2 - \sum_{I=1}^K \frac{1}{2(M_I/m_e)} a_0^2 \nabla_I^2 + \frac{1}{2} \sum_{i \neq j} \frac{a_0}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2} \sum_{I \neq J} Z_I Z_J \frac{a_0}{|\mathbf{R}_I - \mathbf{R}_J|} - \sum_{i,I} Z_I \frac{a_0}{|\mathbf{r}_i - \mathbf{R}_I|} \right] \Psi = \frac{E_{tot}}{E_{Ha}} \Psi. \quad (\text{A.4})$$

A equação A.4 nos mostra a Equação para Muitos Corpos em função da energia de Hartree, além da unidade de medida  $a_0$  e da massa do elétron. Para se resolver isto, faz-se uma mudança de unidades, usando as conhecidas unidades atômicas de Hartree, com as seguintes equivalências:

$$1 \text{ Ha} = 4,3597 \times 10^{-18} \text{ J},$$

$$1 \text{ bohr} = 0,529177 \times 10^{-10} \text{ m},$$

$$1 \text{ a.u. de massa} = 9.10938291 \times 10^{-31} \text{ kg},$$

com "a.u." significando unidade atômica. Logo, ao se utilizar as unidades de Hartree, as constantes fundamentais assumem o valor 1 ( $\hbar = m_e = e = 4\pi\epsilon_0 = 1$ ). Agora, a equação 1.13 pode ser escrita de uma forma mais elegante:

$$\left[ - \sum_{i=1}^N \frac{\nabla_i^2}{2} - \sum_{I=1}^K \frac{\nabla_I^2}{2M_I} + \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} - \sum_{i,I} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} \right] \Psi = E_{tot} \Psi. \quad (\text{A.5})$$

Esta forma mais elegante faz com que os únicos parâmetros externos sejam  $Z_I$  e  $M_I$ , que são características do material a ser estudado. A equação A.5 é a forma mais utilizada para a modelagem a partir de primeiros princípios.

# Apêndice B

## Dinâmica molecular *ab initio*

Após a apresentação das características que norteiam a dinâmica molecular clássica, passaremos brevemente pelo método de Car-Parrinello [96] para apontar as desvantagens do método, o que justifica, em muitos casos, o uso de potenciais empíricos.

Neste método, as equações de movimento para os orbitais eletrônicos e para os núcleos são integradas simultaneamente. A lagrangiana do problema é um funcional dos orbitais de Kohn-Sham e de suas derivadas temporais. Deste modo, a equação de Euler-Lagrange para os orbitais é:

$$\frac{d}{dt} \frac{\delta \mathcal{L}}{\delta \dot{\psi}_i^*} - \frac{\delta \mathcal{L}}{\delta \psi_i^*} = 0, \quad (\text{B.1})$$

resultando em:

$$\mu_i \ddot{\psi}_i(\mathbf{r}, t) = -\hat{H}_{KS} \psi_i(\mathbf{r}, t) + \sum_j \Lambda_{ij} \psi_j(\mathbf{r}, t), \quad (\text{B.2})$$

onde  $\hat{H}_{KS}$  é o Hamiltoniano efetivo de Kohn-Sham (apresentado na equação 1.56),  $\mu_i$  é uma massa fictícia associada ao orbital  $\psi_i$ , e  $\Lambda_{ij}$  são os multiplicadores de Lagrange que garantem a ortonormalidade.

Supõe-se que os núcleos estão fixos na posição  $\mathbf{R}_I$ . Isto implica que os termos  $\Lambda_{ij}$  são os multiplicadores de Lagrange que garantem a ortonormalidade dos orbitais. A equação de movimento dos orbitais é, então, dada por:

$$\mu_i \ddot{\psi}_i = -[\hat{H}_{KS} - \epsilon_i] \psi_i. \quad (\text{B.3})$$

Aplicando-se o algoritmo de Verlet, tem-se a seguinte expressão para  $\psi_i$  em um passo de tempo  $\Delta t$ :

$$\psi_i(\Delta t) = 2\psi_i(0) - \psi_i(-\Delta t) + (\Delta t)^2 \ddot{\psi}_i(0). \quad (\text{B.4})$$

Utilizando a equação B.3 no resultado anterior, obtemos:

$$\psi_i(\Delta t) = 2\psi_i(0) - \psi_i(-\Delta t) - \frac{(\Delta t)^2}{\mu_i} [\hat{H}_{KS} - \epsilon_i] \psi_i(0). \quad (\text{B.5})$$

Pode-se expandir este conceito para as posições nucleares. Deste modo, precisa-se resolver a seguinte equação:

$$\frac{d}{dt} \frac{\delta \mathcal{L}}{\delta \dot{\mathbf{R}}_I} - \frac{\delta \mathcal{L}}{\delta \mathbf{R}_I} = 0. \quad (\text{B.6})$$

Então:

$$M_I \ddot{\mathbf{R}}_I = - \frac{\delta E}{\delta \mathbf{R}_I}. \quad (\text{B.7})$$

A resolução desta equação de movimento com o algoritmo de Verlet é dada por:

$$\mathbf{R}_I(\Delta t) = 2\mathbf{R}_I(0) - \mathbf{R}_I(-\Delta t) - \frac{(\Delta t)^2}{M_I} \left. \frac{\delta E}{\delta \mathbf{R}_I} \right|_{t=0}. \quad (\text{B.8})$$

A vantagem desta abordagem é a possibilidade de se ter uma dinâmica que resolve tanto o problema eletrônico quanto o nuclear de forma conjunta. Deve-se assegurar que a separação adiabática entre os graus de liberdade eletrônicos e nucleares seja mantida. Em contraste, a resolução do problema eletrônico a cada passo ainda possui um custo computacional elevado, o que pode tornar a simulação de sistemas grandes por longos períodos inviável. A Figura B.1 apresenta o fluxograma de uma dinâmica molecular *ab initio* (AIMD).

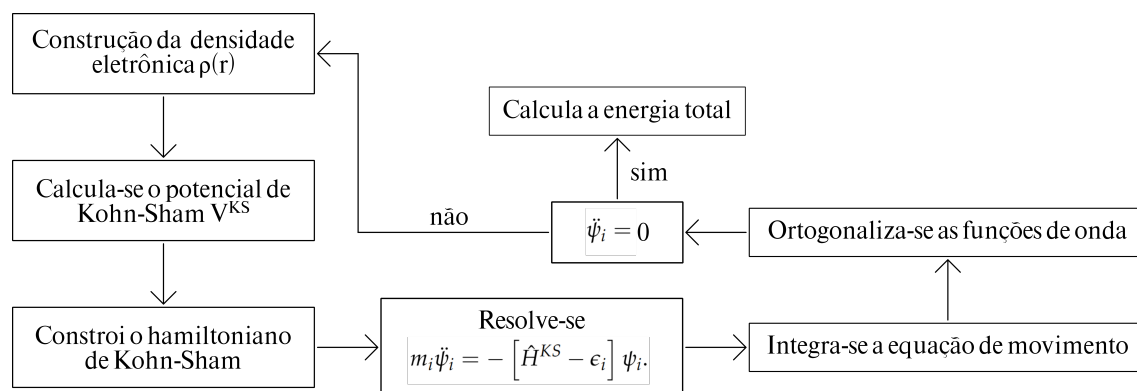


Figura B.1: Fluxograma de uma simulação de dinâmica molecular *ab initio* (AIMD).

# Apêndice C

## Hipótese ergódica

Visto que as simulações de dinâmica molecular descrevem a evolução temporal de um sistema, a relação entre a média temporal e a média estatística de uma propriedade é postulada por meio da hipótese ergódica [182]. Segundo esta relação, pode-se obter a média de uma propriedade de um sistema físico de duas maneiras equivalentes:

- A partir de um número suficientemente grande de réplicas do sistema (um *ensemble*), ou seja, diferentes configurações de um mesmo estado macroscópico; neste caso, a média é conhecida como média de ensemble (ou estatística).
- A partir de uma única simulação suficientemente longa, calculando-se a média ao longo do tempo; neste caso, a média é denominada média temporal.

Estas duas formas de se obterem as médias são importantes do ponto de vista da mecânica estatística, estabelecendo a conexão entre as descrições microscópica e macroscópica. De acordo com os princípios da mecânica estatística, os sistemas precisam estar em equilíbrio térmico para que se possa calcular seus observáveis, ou seja, os valores médios dos observáveis devem se manter constantes ao longo do tempo, e não deve haver deriva significativa na energia total do sistema (para o ensemble NVE).

Uma vez que o sistema esteja em equilíbrio, embora flutuações microscópicas ocorram, o valor médio de um observável pode ser considerado como o valor macroscópico daquela propriedade. No caso de simulações temporais, é necessário que a dinâmica seja longa o suficiente para que o sistema explore adequadamente seu espaço de fase, permitindo o uso de conceitos do limite termodinâmico, no qual pequenas flutuações são naturais na escala macroscópica. Pode-se expandir este conceito, no qual um sistema termalizado corresponde a um estado estacionário macroscópico, cujos valores médios dos observáveis não se alteram com o tempo.

Supondo que  $\bar{O}$  representa a média temporal do observável  $O$ , este pode ser calculado a partir dos valores instantâneos  $O(t_1), \dots, O(t_n)$  obtidos ao longo da

trajetória da simulação após a termalização,

$$\bar{O} = \frac{1}{n} \sum_{i=1}^n O(t_i). \quad (\text{C.1})$$

Considerando-se uma dinâmica suficientemente longa e um passo de tempo  $\Delta t \rightarrow 0$ , pode-se transformar a soma discreta em uma integral contínua no tempo

$$\bar{O} = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^{\tau} O(t) dt. \quad (\text{C.2})$$

Uma simulação longa pode ser substituída por várias simulações curtas. Escolhemos esta segunda abordagem para discutir o problema a seguir. Consideremos diferentes simulações com diferentes condições iniciais de posição e momento  $\{\mathbf{r}(t_i), \mathbf{p}(t_i)\}$ . Para cada condição inicial, definimos o observável  $O^{(i)}(t)$ . Neste caso, temos que  $\bar{O}$  pode ser escrito da seguinte forma:

$$\begin{aligned} \bar{O} = \frac{1}{M} \{ & [O^{(1)}(t=1) + \dots + O^{(1)}(t=N_1)] + [O^{(2)}(t=1) + \dots + O^{(2)}(t=N_2)] + \\ & + \dots + [O^{(n)}(t=1) + \dots + O^{(n)}(t=N_n)] \}, \end{aligned} \quad (\text{C.3})$$

onde  $M = N_1 + N_2 + \dots + N_n$ .

Em princípio, todas as  $M$  configurações microscópicas são consistentes com o mesmo estado macroscópico e, portanto, são igualmente prováveis de serem encontradas. Com isso, não há mais importância em se rotularem as configurações com relação ao tempo, e pode-se considerar que a média do observável,  $\bar{O}$ , representa a média estatística do ensemble.

Deste modo, a média estatística é dada pela seguinte equação:

$$\bar{O} = \frac{1}{M} \left[ O^{(1)} + \dots + O^{(N_1)} + \dots + O^{(N_1+N_2)} + \dots + O^{(M)} \right]. \quad (\text{C.4})$$

Agora,  $O^{(i)}$  representa o observável para cada uma das diferentes configurações, não mais para cada simulação. Com isso, pode-se usar resultados da mecânica estatística [183], na qual a média estatística pode ser obtida por meio da função densidade de probabilidade no espaço de fase,  $\rho(\{\mathbf{r}, \mathbf{p}\})$ , ou seja:

$$\langle O(\{\mathbf{r}, \mathbf{p}\}) \rangle = \frac{1}{Z} \int \rho(\{\mathbf{r}, \mathbf{p}\}) O(\{\mathbf{r}, \mathbf{p}\}) d\Gamma, \quad (\text{C.5})$$

onde  $d\Gamma = d\mathbf{r}_1 \cdots d\mathbf{r}_N d\mathbf{p}_1 \cdots d\mathbf{p}_N$  é o elemento de volume no espaço de fase, e  $Z$  é a função de partição do sistema.

Lembrando que a média temporal do observável  $O$  é denotada por  $\bar{O}$  e a média estatística por  $\langle O \rangle$ , a hipótese ergódica postula sua equivalência:

$$\bar{O}(\{\mathbf{r}(t), \mathbf{p}(t)\}) = \langle O(\{\mathbf{r}, \mathbf{p}\}) \rangle. \quad (\text{C.6})$$

# Apêndice D

## *Script* utilizado para treino das redes neurais

Abaixo, segue *scripts* utilizado para o treino da rede com o código DeePMD.

```
{
  "model": {
    "type_map": [
      "O",
      "H"
    ],
    "descriptor": {
      "type": "se_e2_a",
      "sel": [
        46,
        92
      ],
      "rcut_smth": 0.5,
      "rcut": 6.0,
      "neuron": [
        25,
        50,
        100
      ],
      "resnet_dt": false,
      "axis_neuron": 16,
      "seed": 578973,
      "activation_function": "tanh",
      "type_one_side": false,
      "precision": "default",
```

```
        "trainable": true ,
        "exclude_types": [],
        "set_davg_zero": false
    },
    "fitting_net": {
        "neuron": [
            320,
            160,
            32,
            16
        ],
        "resnet_dt": true ,
        "seed": 548948,
        "type": "ener" ,
        "numb_fparam": 0,
        "numb_aparam": 0,
        "activation_function": "tanh" ,
        "precision": "default" ,
        "trainable": true ,
        "rcond": 0.001,
        "atom_ener": []
    },
    "data_stat_nbatch": 10,
    "data_stat_protect": 0.01,
    "compress": {
        "model_file": "optB88_5b.pb" ,
        "min_nbor_dist": 0.6515838216297681,
        "table_config": [
            5,
            0.01,
            0.1,
            -1
        ],
        "type": "se_e2_a"
    }
},
```

```
"learning_rate": {
  "type": "exp",
  "decay_steps": 5000,
  "start_lr": 0.001,
  "stop_lr": 3.51e-08,
  "scale_by_worker": "linear"
},
"loss": {
  "type": "ener",
  "start_pref_e": 0.02,
  "limit_pref_e": 8,
  "start_pref_f": 1500,
  "limit_pref_f": 1,
  "start_pref_v": 0,
  "limit_pref_v": 0,
  "start_pref_ae": 0.0,
  "limit_pref_ae": 0.0,
  "start_pref_pf": 0.0,
  "limit_pref_pf": 0.0,
  "enable_atom_ener_coeff": false
},
"training": {
  "training_data": {
    "systems": [
      "data/raw_a/train",
      "data/raw_b/train",
      "data/raw_c/train",
      "data/raw_d/train",
      "data/raw_e/train",
      "data/raw_f/train",
      "data/raw_g/train"
    ],
    "batch_size": "auto",
    "set_prefix": "set",
    "auto_prob": "prob_sys_size",
    "sys_probs": null
  }
}
```

```
    },  
    "numb_steps": 7000000,  
    "seed": 877594,  
    "disp_file": "lcurve.out",  
    "disp_freq": 100,  
    "save_freq": 1000,  
    "validation_data": null,  
    "save_ckpt": "model.ckpt",  
    "disp_training": true,  
    "time_training": true,  
    "profiling": false,  
    "profiling_file": "timeline.json",  
    "enable_profiler": false,  
    "tensorboard": false,  
    "tensorboard_log_dir": "log",  
    "tensorboard_freq": 1  
  }  
}
```

Segue *script* utilizado para o treinamento da rede neural com o código MACE.

```
mace_run_train \  
--name="water_scan" \  
--train_file="train.xyz" \  
--valid_fraction=0.05 \  
--E0s="average" \  
--model="MACE" \  
--num_interactions=4 \  
--num_channels=128 \  
--max_L=1 \  
--correlation=3 \  
--r_max=6.0 \  
--forces_weight=1000 \  
--energy_weight=10 \  
--energy_key="REF_energy" \  
--forces_key="REF_forces" \  
--batch_size=2 \  
--valid_batch_size=4 \  

```

```
--max_num_epochs=500 \  
--start_swa=250 \  
--scheduler_patience=15 \  
--patience=30 \  
--eval_interval=4 \  
--ema \  
--swa \  
--error_table='PerAtomRMSE' \  
--default_dtype="float64"\  
--device=cuda \  
--seed=123 \  
--save_cpu
```

# Apêndice E

## *Script* para cálculos DFT

Abaixo, segue *script* utilizados para cálculos DFT com o código SIESTA.

```
SystemName          water
SystemLabel         water

NumberOfAtoms       288
NumberOfSpecies     2

%block ChemicalSpeciesLabel
  1  8  O.1
  2  1  H.2
%endblock ChemicalSpeciesLabel

%include water_lattice.fdf
%include ice_vectors.fdf
AtomicCoordinatesFormat Ang
%block AtomicCoordinatesAndAtomicSpecies < geom.in

%block PAO.Basis < basis_dzp.fdf

xc.functional       VDW
xc.authors          KBM

MeshCutoff          500 Ry

%block Grid.CellSampling
  0.5  0.5  0.0
  0.5  0.0  0.5
  0.0  0.5  0.5
```

```
%endblock Grid.CellSampling
```

```
SolutionMethod      Diagon
Diag.MRRR           yes
Diag.Memory         2
SCF.Mixer.Method    Pulay
SCF.Mixer.Weight    0.3
SCF.Mixer.History   7
MaxSCFIterations    200
ElectronicTemperature 300 K
DM.Tolerance        1.d-6
DM.UseSaveDM        yes

MD.TypeOfRun        CG
MD.NumCGSteps       0

LongOutput          yes
WriteCoorXmol       yes
WriteMDXmol         no
WriteMDHistory      no
```

Para a base atômica foi utilizado o seguinte arquivo:

```
O.1      3      -0.24233
n=2      0      2      E      23.36061      3.39721
         4.50769      2.64066
         1.00000      1.00000
n=2      1      2      E      2.78334      5.14253
         6.14996      2.59356
         1.00000      1.00000
n=3      2      1      E      63.98188      0.16104
         3.54403
         1.00000
H.2      2      0.46527
n=1      0      2      E      99.93138      2.59932
         4.20357      1.84463
         1.00000      1.00000
n=2      1      1      E      24.56504      2.20231
```

3.5281  
1.00000

Para os cálculos utilizando o código VASP, o seguinte *script* foi implementado:

```
PREC      = Accurate
LREAL     = False
ADDGRID   = True
#
METAGGA   = SCAN
LASPH     = True
LMIXTAU   = True
# + rVV10
# LUSE_VDW = True
# BPARAM   = 15.7
#
ALGO      = Normal
EDIFF     = 64e-7 # 1e-7 / H2O molecule
#
GGA_COMPAT = False
#
ISM EAR    = 0
SIGMA     = 0.1
#Change here for ice- set default number of bands
#Change here for ice- set default number of bands
##NBANDS  = 264 # 64 H2O
NBANDS    = 396 # 96 H2O

ENCUT     = 1600
#
ISIF      = 2
LORBIT    = 11
#
IBRION    = -1
NELMIN    = 6
#NELMDL   = -3
#
LCHARG    = False
```

```
LWAVE = True  
#  
#NSIM = 1  
NCORE = 16  
#LPLANE = False  
#LSCALAPACK = False
```

# Referências

- [1] Michele Ceriotti, Wei Fang, Peter G Kusalik, Ross H McKenzie, Angelos Michaelides, Miguel A Morales, and Thomas E Markland. Nuclear quantum effects in water and aqueous systems: Experiment, theory, and current challenges. *Chemical reviews*, 116(13):7529–7550, 2016.
- [2] Mark E Tuckerman. *Statistical mechanics: theory and molecular simulation*. Oxford university press, 2023.
- [3] Michele Ceriotti and David E Manolopoulos. Efficient first-principles calculation of the quantum kinetic energy and momentum distribution of nuclei. *Physical review letters*, 109(10):100604, 2012.
- [4] Han Wang, Linfeng Zhang, Jiequn Han, et al. Deepmd-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Computer Physics Communications*, 228:178–184, 2018.
- [5] Ilyes Batatia, David P Kovacs, Gregor Simm, Christoph Ortner, and Gábor Csányi. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields. *Advances in neural information processing systems*, 35:11423–11436, 2022.
- [6] Mariana Rossi, Wei Fang, and Angelos Michaelides. Stability of complex biomolecular structures: van der waals, hydrogen bond cooperativity, and nuclear quantum effects. *The journal of physical chemistry letters*, 6(21):4233–4238, 2015.
- [7] Xin-Zheng Li, Brent Walker, and Angelos Michaelides. Quantum nature of the hydrogen bond. *Proceedings of the National Academy of Sciences*, 108(16):6369–6373, 2011.
- [8] Mariana Rossi, Michele Ceriotti, and David E. Manolopoulos. How to remove the spurious resonances from ring polymer molecular dynamics. *The Journal of Chemical Physics*, 140(23):234116, 06 2014.
- [9] Mariana Rossi, Hanchao Liu, Francesco Paesani, Joel Bowman, and Michele Ceriotti. Communication: On the consistency of approximate quantum

- dynamics simulation methods for vibrational spectra in the condensed phase. *The Journal of Chemical Physics*, 141(18):181101, 11 2014.
- [10] AK Soper. The radial distribution functions of water and ice from 220 to 673 k and at pressures up to 400 mpa. *Chemical Physics*, 258(2-3):121–137, 2000.
- [11] Tobias Morawietz, Andreas Singraber, Christoph Dellago, and Jörg Behler. How van der waals interactions determine the unique properties of water. *Proceedings of the National Academy of Sciences*, 113(30):8368–8373, 2016.
- [12] Mohan Chen, Hsin-Yu Ko, Richard C Remsing, Marcos F Calegari Andrade, Biswajit Santra, Zhaoru Sun, Annabella Selloni, Roberto Car, Michael L Klein, John P Perdew, et al. Ab initio theory and modeling of water. *Proceedings of the National Academy of Sciences*, 114(41):10846–10851, 2017.
- [13] Daniel R Moberg, Shelby C Straight, Christopher Knight, and Francesco Paesani. Molecular origin of the vibrational structure of ice ih. *The Journal of Physical Chemistry Letters*, 8(12):2579–2583, 2017.
- [14] WA Adeagbo, A Zayak, and P Entel. Ab initio study of the structure and dynamical properties of crystalline ice. *Phase Transitions*, 78(1-3):179–196, 2005.
- [15] Filipe Matusalem, Jéssica Santos Rego, and Maurice de Koning. Plastic deformation of superionic water ices. *Proceedings of the National Academy of Sciences*, 119(45):e2203397119, 2022.
- [16] Victor F Petrenko and Robert W Whitworth. *Physics of ice*. OUP Oxford, 1999.
- [17] Jéssica Santos Rego and Maurice de Koning. Density-functional theory prediction of the elastic constants of ice ih. *The Journal of Chemical Physics*, 152(8), 2020.
- [18] DE Hare and CM Sorensen. The density of supercooled water. ii. bulk samples cooled to the homogeneous nucleation limit. *The Journal of chemical physics*, 87(8):4840–4845, 1987.
- [19] William Howard Barnes. The crystal structure of ice between 0 c. and—183 c. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 125(799):670–693, 1929.

- [20] John D Bernal and Ralph H Fowler. A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions. *The Journal of Chemical Physics*, 1(8):515–548, 1933.
- [21] Linus Pauling. The structure and entropy of ice and of other crystals with some randomness of atomic arrangement. *Journal of the American Chemical Society*, 57(12):2680–2684, 1935.
- [22] M Born and R Oppenheimer. Zur quantentheorie der molekeln. *Annalen der Physik*, 84, 1927.
- [23] Pierre Hohenberg and Walter Kohn. Inhomogeneous electron gas. *Physical review*, 136(3B):B864, 1964.
- [24] Walter Kohn and Lu Jeu Sham. Self-consistent equations including exchange and correlation effects. *Physical review*, 140(4A):A1133, 1965.
- [25] I Morrison, J-C Li, S Jenkins, SS Xantheas, and MC Payne. Ab-initio total energy studies of the static and dynamical properties of ice Ih. *The Journal of Physical Chemistry B*, 101(32):6146–6150, 1997.
- [26] Niall J English. Structural properties of liquid water and ice Ih from ab-initio molecular dynamics with a non-local correlation functional. *Energies*, 8(9):9383–9391, 2015.
- [27] Michael J Gillan, Dario Alfe, and Angelos Michaelides. Perspective: How good is dft for water? *The Journal of chemical physics*, 144(13), 2016.
- [28] Pablo Montero de Hijes, Christoph Dellago, Ryosuke Jinnouchi, and Georg Kresse. Density isobar of water and melting temperature of ice: Assessing common density functionals. *The Journal of Chemical Physics*, 161(13), 2024.
- [29] I-Chun Lin, Ari P Seitsonen, Mauricio D Coutinho-Neto, Ivano Tavernelli, and Ursula Rothlisberger. Importance of van der waals interactions in liquid water. *The Journal of Physical Chemistry B*, 113(4):1127–1131, 2009.
- [30] Jiří Klimeš, David R Bowler, and Angelos Michaelides. Chemical accuracy for the van der waals density functional. *Journal of Physics: Condensed Matter*, 22(2):022201, dec 2009.

- [31] Kristian Berland and Per Hyldgaard. Exchange functional that tests the robustness of the plasmon description of the van der waals density functional. *Phys. Rev. B*, 89:035412, Jan 2014.
- [32] Stefan Grimme. Semiempirical gga-type density functional constructed with a long-range dispersion correction. *Journal of computational chemistry*, 27(15):1787–1799, 2006.
- [33] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu. *The Journal of chemical physics*, 132(15), 2010.
- [34] Jianwei Sun, Adrienn Ruzsinszky, and John P. Perdew. Strongly constrained and appropriately normed semilocal density functional. *Physical Review Letters*, 115:036402, 2015.
- [35] Pablo M Piaggi, Athanassios Z Panagiotopoulos, Pablo G Debenedetti, and Roberto Car. Phase equilibrium of water with hexagonal and cubic ice using the scan functional. *Journal of Chemical Theory and Computation*, 17(5):3065–3077, 2021.
- [36] William L Jorgensen, Jayaraman Chandrasekhar, Jeffry D Madura, Roger W Impey, and Michael L Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics*, 79(2):926–935, 1983.
- [37] HJC Berendsen, JR Grigera, and TP Straatsma. The missing term in effective pair potentials. *Journal of Physical Chemistry*, 91(24):6269–6271, 1987.
- [38] Pekka Mark and Lennart Nilsson. Structure and dynamics of the tip3p, spc, and spc/e water models at 298 k. *The Journal of Physical Chemistry A*, 105(43):9954–9960, 2001.
- [39] JLF Abascal, E Sanz, R García Fernández, and C Vega. A potential model for the study of ices and amorphous water: Tip4p/ice. *The Journal of chemical physics*, 122(23), 2005.
- [40] Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*, 117(1):1 – 19, 1995.

- [41] Venkat Kapil, Mariana Rossi, Ondrej Marsalek, Riccardo Petraglia, Yair Litman, Thomas Spura, Bingqing Cheng, Alice Cuzzocrea, Robert H Meißner, David M Wilkins, et al. i-pi 2.0: A universal force engine for advanced molecular simulations. *Computer Physics Communications*, 236:214–223, 2019.
- [42] Alberto Torres, Luana S Pedroza, Marivi Fernandez-Serra, and Alexandre R Rocha. Using neural network force fields to ascertain the quality of ab initio simulations of liquid water. *The Journal of Physical Chemistry B*, 125(38):10772–10778, 2021.
- [43] Kurt Hornik. Approximation capabilities of multilayer feedforward networks. *Neural networks*, 4(2):251–257, 1991.
- [44] Jörg Behler and Michele Parrinello. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Physical review letters*, 98(14):146401, 2007.
- [45] Aidan P Thompson, Laura P Swiler, Christian R Trott, Stephen M Foiles, and Garritt J Tucker. Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials. *Journal of Computational Physics*, 285:316–330, 2015.
- [46] Jörg Behler. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. *The Journal of Chemical Physics*, 134(7):074106, 02 2011.
- [47] Christian Devereux, Justin S. Smith, Kate K. Davis, Kipton Barros, Roman Zubatyuk, Olexandr Isayev, and Adrian E. Roitberg. Extending the applicability of the ani deep learning molecular potential to sulfur and halogens. *Journal of Chemical Theory and Computation*, 16(7):4192–4202, 2020. PMID: 32543858.
- [48] Marco Fronzi, Sherif Abdulkader Tawfik, Mutaz Abu Ghazaleh, Olexandr Isayev, David A. Winkler, Joe Shapter, and Michael J. Ford. High throughput screening of millions of van der waals heterostructures for superlubricant applications. *Advanced Theory and Simulations*, 3:2000029, 2020.
- [49] Hao Wang and Weitao Yang. Force field for water based on neural network. *Journal of Physical Chemistry Letters*, 9:3232, 2018.

- [50] Jörg Behler. Perspective: Machine learning potentials for atomistic simulations. *Journal of Chemical Physics*, 145:170901, 2016.
- [51] Yi Yao and Yosuke Kanai. Temperature dependence of nuclear quantum effects on liquid water via artificial neural network model based on scan meta-gga functional. *Journal of Chemical Physics*, 153:044114, 2020.
- [52] Thomas E. Gartner III, Linfeng Zhang<sup>b</sup>, Pablo M. Piaggia, Roberto Cara, Athanassios Z Panagiotopoulos, and Pablo G. Debenedetti. Signatures of a liquid–liquid transition in an ab initio deep neural network model for water. *PNAS*, 117:26040, 2020.
- [53] Jianhang Xu, Chunyi Zhang, Linfeng Zhang, Mohan Chen, Biswajit Santra, and Xifan Wu. Isotope effects in molecular structures and electronic properties of liquid water via deep potential molecular dynamics based on the scan functional. *Physical Review B*, 102(21):214113, 2020.
- [54] Tran Doan Huan, Rohit Batra, James Chapman, Sridevi Krishnan, Lihua Chen, and Rampi Ramprasad. A universal strategy for the creation of machine learning-based atomistic force fields. *npj Computational Materials*, 3(1):37, 2017.
- [55] Ilyes Batatia, Simon Batzner, Dávid Péter Kovács, Albert Musaelian, Gregor NC Simm, Ralf Drautz, Christoph Ortner, Boris Kozinsky, and Gábor Csányi. The design space of e (3)-equivariant atom-centered interatomic potentials. *arXiv preprint arXiv:2205.06643*, 2022.
- [56] Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E Smidt, and Boris Kozinsky. E (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature communications*, 13(1):2453, 2022.
- [57] Albert Musaelian, Simon Batzner, Anders Johansson, Lixin Sun, Cameron J Owen, Mordechai Kornbluth, and Boris Kozinsky. Learning local equivariant representations for large-scale atomistic dynamics. *Nature Communications*, 14(1):579, 2023.
- [58] Nan Yao, Xiang Chen, Zhong-Heng Fu, and Qiang Zhang. Applying classical, ab initio, and machine-learning molecular dynamics simulations to the

- liquid electrolyte for rechargeable batteries. *Chemical Reviews*, 122(12):10970–11021, 2022.
- [59] Luigi Bonati, Daniela Polino, Cristina Pizzolitto, Pierdomenico Biasi, Rene Eckert, Stephan Reitmeier, Robert Schlögl, and Michele Parrinello. The role of dynamics in heterogeneous catalysis: Surface diffusivity and n<sub>2</sub> decomposition on fe (111). *Proceedings of the National Academy of Sciences*, 120(50):e2313023120, 2023.
- [60] Miguel de la Puente, Axel Gomez, and Damien Laage. Neural network-based sum-frequency generation spectra of pure and acidified water interfaces with air. *The Journal of Physical Chemistry Letters*, 15(11):3096–3102, 2024.
- [61] Jia-Xin Zhu and Jun Cheng. Machine learning potential for electrochemical interfaces with hybrid representation of dielectric response. *Physical Review Letters*, 135(1):018003, 2025.
- [62] Linfeng Zhang, Han Wang, Roberto Car, and Weinan E. Phase diagram of a deep potential water model. *Physical review letters*, 126(23):236001, 2021.
- [63] Thomas E Gartner III, Pablo M Piaggi, Roberto Car, Athanassios Z Panagiotopoulos, and Pablo G Debenedetti. Liquid-liquid transition in water from first principles. *Physical review letters*, 129(25):255702, 2022.
- [64] Miguel de la Puente and Damien Laage. How the acidity of water droplets and films is controlled by the air–water interface. *Journal of the American Chemical Society*, 145(46):25186–25194, 2023.
- [65] John P Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 77(18):3865, 1996.
- [66] Márcio S Gomes-Filho, Alberto Torres, Alexandre Reily Rocha, and Luana S Pedroza. Size and quality of quantum mechanical data set for training neural network force fields for liquid water. *The Journal of Physical Chemistry B*, 127(6):1422–1428, 2023.
- [67] Margaret L Berrens, Arpan Kundu, Marcos F Calegari Andrade, Tuan Anh Pham, Giulia Galli, and Davide Donadio. Nuclear quantum effects on the electronic structure of water and ice. *The Journal of Physical Chemistry Letters*, 15(26):6818–6825, 2024.

- [68] Thomas E Markland and Michele Ceriotti. Nuclear quantum effects enter the mainstream. *Nature Reviews Chemistry*, 2(3):0109, 2018.
- [69] Matti Hellström, Michele Ceriotti, and Joörg Behler. Nuclear quantum effects in sodium hydroxide solutions from neural network molecular dynamics simulations. *The Journal of Physical Chemistry B*, 122(44):10158–10171, 2018.
- [70] Mariana Rossi. Progress and challenges in ab initio simulations of quantum nuclei in weakly bonded systems. *The Journal of Chemical Physics*, 154(17):170902, 2021.
- [71] J. A. Morrone and R. Car. Nuclear quantum effects in water. *Phys. Rev. Lett.*, 101:017801, 2008.
- [72] E. M. McIntosh, K. T. Wikfeldt, J. Ellis, A. Michaelides, and W. Allison. Quantum effects in the diffusion of hydrogen on ru(0001). *J. Phys. Chem. Lett.*, 4:1565, 2013.
- [73] Bet Pamuk, Jose M Soler, R Ramírez, CP Herrero, PW Stephens, PB Allen, and M-V Fernández-Serra. Anomalous nuclear quantum effects in ice. *Physical review letters*, 108(19):193003, 2012.
- [74] Sriram Ganeshan, Rafael Ramírez, and MV Fernández-Serra. Simulation of quantum zero-point effects in water using a frequency-dependent thermostat. *Physical Review B*, 87(13):134207, 2013.
- [75] L Hernández De La Peña, MS Gulam Razul, and PG Kusalik. Quantum effects in ice ih. *The Journal of chemical physics*, 123(14):144506, 2005.
- [76] Carlos P Herrero and Rafael Ramírez. Isotope effects in ice ih: A path-integral simulation. *The Journal of chemical physics*, 134(9):094510, 2011.
- [77] David Chandler and Peter G Wolynes. Exploiting the isomorphism between quantum theory and classical statistical mechanics of polyatomic fluids. *The Journal of Chemical Physics*, 74(7):4078–4095, 1981.
- [78] Bingqing Cheng, Edgar A. Engel, Jörg Behler, Christoph Dellago, and Michele Ceriotti. Ab initio thermodynamics of liquid and solid water. *Proceedings of the National Academy of Sciences*, 116(4):1110–1115, 2019.

- [79] Chunyi Zhang, Fujie Tang, Mohan Chen, Jianhang Xu, Linfeng Zhang, Diana Y Qiu, John P Perdew, Michael L Klein, and Xifan Wu. Modeling liquid water by climbing up jacob's ladder in density functional theory facilitated by using deep neural network potentials. *The Journal of Physical Chemistry B*, 125(41):11444–11456, 2021.
- [80] Ali Eltareb, Gustavo E Lopez, and Nicolas Giovambattista. The importance of nuclear quantum effects on the thermodynamic and structural properties of low-density amorphous ice: a comparison with hexagonal ice. *The Journal of Physical Chemistry B*, 127(20):4633–4645, 2023.
- [81] Pedro Augusto Franco Pinheiro Moreira and Maurice de Koning. Nuclear quantum fluctuations in ice i h. *Physical Chemistry Chemical Physics*, 17(38):24716–24721, 2015.
- [82] Hsin-Yu Ko, Linfeng Zhang, Biswajit Santra, Han Wang, Weinan E, Robert A DiStasio Jr, and Roberto Car. Isotope effects in liquid water via deep potential molecular dynamics. *Molecular Physics*, 117(22):3269–3281, 2019.
- [83] Chunyi Zhang, Linfeng Zhang, Jianhang Xu, Fujie Tang, Biswajit Santra, and Xifan Wu. Isotope effects in x-ray absorption spectra of liquid water. *Physical Review B*, 102(11):115155, 2020.
- [84] Maurice de Koning. Crystal imperfections in ice ih. *The Journal of Chemical Physics*, 153(11), 2020.
- [85] Maurice de Koning, Alex Antonelli, Antonio JR da Silva, and Adalberto Fazzio. Structure and energetics of molecular point defects in ice i h. *Physical review letters*, 97(15):155501, 2006.
- [86] Matthew Watkins, Ding Pan, Ee Ge Wang, Angelos Michaelides, J VandeVondele, and B Slater. Large variation of vacancy formation energies in the surface of crystalline ice. *Nature materials*, 10(10):794–798, 2011.
- [87] M Eldrup, OE Mogensen, and JH Bilgram. Vacancies in hf-doped and in irradiated ice by positron annihilation techniques. *Journal of Glaciology*, 21(85):101–113, 1978.
- [88] OE Mogensen and M Eldrup. Vacancies in pure ice studied by positron annihilation techniques. *Journal of Glaciology*, 21(85):85–99, 1978.

- [89] K Goto, Takeo Hondoh, and Akira Higashi. *Experimental determinations of the concentration and the mobility of interstitials in pure ice crystals*. University of Tokyo Press, 1982.
- [90] T Hondoh, T Itoh, and A Higashi. *Behavior of point defects in ice crystals revealed by X-ray topography*. North Holland Pub., 1982.
- [91] Tomoko Ikeda-Fukazawa, Shinichiro Horikawa, Takeo Hondoh, and Katsuyuki Kawamura. Molecular dynamics studies of molecular diffusion in ice Ih. *The Journal of chemical physics*, 117(8):3886–3896, 2002.
- [92] Kumiko Goto, Takeo Hondoh, and Akira Higashi. Determination of diffusion coefficients of self-interstitials in ice with a new method of observing climb of dislocations by x-ray topography. *Japanese journal of applied physics*, 25(3R):351, 1986.
- [93] Jinzhe Zeng, Duo Zhang, Denghui Lu, Pinghui Mo, Zeyu Li, Yixiao Chen, Marián Rynik, Li’ang Huang, Ziyao Li, Shaochen Shi, et al. Deepmd-kit v2: A software package for deep potential models. *The Journal of Chemical Physics*, 159(5), 2023.
- [94] Michele Ceriotti, Jérôme Cuny, Michele Parrinello, and David E Manolopoulos. Nuclear quantum effects and hydrogen bond fluctuations in water. *Proceedings of the National Academy of Sciences*, 110(39):15591–15596, 2013.
- [95] Berni Julian Alder and Thomas Everett Wainwright. Phase transition for a hard sphere system. *The Journal of chemical physics*, 27(5):1208–1209, 1957.
- [96] Richard Car and Mark Parrinello. Unified approach for molecular dynamics and density-functional theory. *Physical review letters*, 55(22):2471, 1985.
- [97] Linfeng Zhang, Jiequn Han, Han Wang, Roberto Car, and Weinan E. Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics. *Physical review letters*, 120(14):143001, 2018.
- [98] Jinzhe Zeng, Duo Zhang, Anyang Peng, Xiangyu Zhang, Sensen He, Yan Wang, Xinzijian Liu, Hangrui Bi, Yifan Li, Chun Cai, et al. Deepmd-kit v3: A multiple-backend framework for machine learning potentials. *arXiv preprint arXiv:2502.19161*, 2025.

- [99] José David M Vianna. *Teoria Quântica de Moléculas e Sólidos: simulação computacional*. Livraria da Física, 2004.
- [100] Feliciano Giustino. *Materials modelling using density functional theory: properties and predictions*. Oxford University Press, 2014.
- [101] Richard M Martin. *Electronic structure: basic theory and practical methods*. Cambridge university press, 2020.
- [102] Vladimir Fock. Näherungsmethode zur lösung des quantenmechanischen mehrkörperproblems. *Zeitschrift für Physik*, 61:126–148, 1930.
- [103] Klaus Capelle. A bird’s-eye view of density-functional theory. *Brazilian journal of physics*, 36:1318–1343, 2006.
- [104] David M Ceperley and Berni J Alder. Ground state of the electron gas by a stochastic method. *Physical review letters*, 45(7):566, 1980.
- [105] John P Perdew and Alex Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Physical review B*, 23(10):5048, 1981.
- [106] John P Perdew and Karla Schmidt. Jacob’s ladder of density functional approximations for the exchange-correlation energy. In *AIP Conference Proceedings*, volume 577, pages 1–20. American Institute of Physics, 2001.
- [107] Federico Zahariev, Sarom Sok Leang, and Mark S Gordon. Functional derivatives of meta-generalized gradient approximation (meta-gga) type exchange-correlation density functionals. *The Journal of Chemical Physics*, 138(24), 2013.
- [108] Max Dion, Henrik Rydberg, Elsebeth Schröder, David C Langreth, and Bengt I Lundqvist. Van der waals density functional for general geometries. *Physical review letters*, 92(24):246401, 2004.
- [109] John P. Perdew and Wang Yue. Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation. *Phys. Rev. B*, 33:8800–8802, Jun 1986.
- [110] Éamonn D. Murray, Kyuho Lee, and David C. Langreth. Investigation of exchange energy density functional accuracy for interacting molecules.

- Journal of Chemical Theory and Computation*, 5(10):2754–2762, 2009. PMID: 26631788.
- [111] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38:3098–3100, Sep 1988.
- [112] James C Phillips and Leonard Kleinman. New method for calculating wave functions in crystals and molecules. *Physical Review*, 116(2):287, 1959.
- [113] GP Kerker. Non-singular atomic pseudopotentials for solid state applications. *Journal of Physics C: Solid State Physics*, 13(9):L189, 1980.
- [114] Norman Troullier and José Luís Martins. Efficient pseudopotentials for plane-wave calculations. *Physical review B*, 43(3):1993, 1991.
- [115] Peter E Blöchl. Projector augmented-wave method. *Physical review B*, 50(24):17953, 1994.
- [116] Michael P Allen and Dominic J Tildesley. *Computer simulation of liquids*. Oxford university press, 2017.
- [117] Claudio Scherer. *Métodos computacionais da física*. Editora Livraria da Física, 2005.
- [118] Pascal Friederich, Florian Häse, Jonny Proppe, and Alán Aspuru-Guzik. Machine-learned potentials for next-generation matter simulations. *Nature Materials*, 20(6):750–761, 2021.
- [119] Dylan M Anstine and Olexandr Isayev. Machine learning interatomic potentials and long-range physics. *The Journal of Physical Chemistry A*, 127(11):2417–2431, 2023.
- [120] Amir Omranpour, Pablo Montero De Hijes, Jörg Behler, and Christoph Dellago. Perspective: Atomistic simulations of water and aqueous systems with machine learning potentials. *The Journal of Chemical Physics*, 160(17), 2024.
- [121] Dennis C Rapaport. *The art of molecular dynamics simulation*. Cambridge university press, 2004.

- [122] Michael Griebel, Stephan Knapek, Gerhard Zumbusch, and Attila Caglar. *Numerische Simulation in der Moleküldynamik: Numerik, Algorithmen, Parallelisierung, Anwendungen*. Springer-Verlag, 2013.
- [123] Aneesur Rahman. Correlations in the motion of atoms in liquid argon. *Physical review*, 136(2A):A405, 1964.
- [124] Michael W Mahoney and William L Jorgensen. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *The Journal of chemical physics*, 112(20):8910–8922, 2000.
- [125] Herman JC Berendsen, James PM Postma, Wilfred F van Gunsteren, and Jan Hermans. Interaction models for water in relation to protein hydration. In *Intermolecular forces: proceedings of the fourteenth Jerusalem symposium on quantum chemistry and biochemistry held in jerusalem, israel, april 13–16, 1981*, pages 331–342. Springer, 1981.
- [126] Philippe H Hünenberger. Thermostat algorithms for molecular dynamics simulations. *Advanced computer simulation: Approaches for soft matter sciences I*, pages 105–149, 2005.
- [127] Hans C Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of chemical physics*, 72(4):2384–2393, 1980.
- [128] Herman JC Berendsen, JPM van Postma, Wilfred F Van Gunsteren, ARHJ DiNola, and Jan R Haak. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81(8):3684–3690, 1984.
- [129] Shūichi Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics*, 52(2):255–268, 1984.
- [130] William G Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Physical review A*, 31(3):1695, 1985.
- [131] Robert Zwanzig. *Nonequilibrium statistical mechanics*. Oxford university press, 2001.
- [132] Michele Ceriotti, Giovanni Bussi, and Michele Parrinello. Langevin equation with colored noise for constant-temperature molecular dynamics simulations. *Physical review letters*, 102(2):020601, 2009.

- [133] Michele Ceriotti, Giovanni Bussi, and Michele Parrinello. Colored-noise thermostats à la carte. *Journal of Chemical Theory and Computation*, 6(4):1170–1180, 2010.
- [134] Mark E Tuckerman, Dominik Marx, Michael L Klein, and Michele Parrinello. Efficient and general algorithms for path integral car-parrinello molecular dynamics. *The Journal of chemical physics*, 104(14):5579–5588, 1996.
- [135] <https://gle4md.org/>.
- [136] Richard Phillips Feynman. Space-time approach to non-relativistic quantum mechanics. *Reviews of modern physics*, 20(2):367, 1948.
- [137] RP Feynman, AR Hibbs, and George H Weiss. Quantum mechanics and path integrals, 1966.
- [138] GW Ford and M Kac. On the quantum langevin equation. *Journal of statistical physics*, 46:803–810, 1987.
- [139] George W Ford, John T Lewis, and RF943786 O’Connell. Quantum langevin equation. *Physical Review A*, 37(11):4419, 1988.
- [140] MF Herman, EJ Bruskin, and BJ Berne. On path integral monte carlo simulations. *The Journal of Chemical Physics*, 76(10):5150–5155, 1982.
- [141] David M Ceperley. Path integrals in the theory of condensed helium. *Reviews of Modern Physics*, 67(2):279, 1995.
- [142] AD Dongare, RR Kharde, Amit D Kachare, et al. Introduction to artificial neural network. *International Journal of Engineering and Innovative Technology (IJEIT)*, 2(1):189–194, 2012.
- [143] Maysam F Abbod, James WF Catto, Derek A Linkens, and Freddie C Hamdy. Application of artificial intelligence to the management of urological cancer. *The Journal of urology*, 178(4):1150–1156, 2007.
- [144] Sascha Klawohn, James P Darby, James R Kermode, Gábor Csányi, Miguel A Caro, and Albert P Bartók. Gaussian approximation potentials: Theory, software implementation and application examples. *The Journal of Chemical Physics*, 159(17), 2023.

- [145] Abadi Martín, Agarwal Ashish, Barham Paul, Brevdo Eugene, Chen Zhifeng, Citro Craig, S Corrado Greg, Davis Andy, Dean Jeffrey, Devin Matthieu, et al. Tensorflow: Large-scale machine learning on heterogeneous systems. *Software available from tensorflow.org*, 7, 2015.
- [146] A Paszke. Pytorch: An imperative style, high-performance deep learning library. *arXiv preprint arXiv:1912.01703*, 2019.
- [147] James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, et al. Jax: composable transformations of python+ numpy programs. 2018.
- [148] Yanjun Ma, Dianhai Yu, Tian Wu, and Haifeng Wang. Paddlepaddle: An open-source deep learning platform from industrial practice. *Frontiers of Data and Computing*, 1(1):105–115, 2019.
- [149] Aidan P Thompson, H Metin Aktulga, Richard Berger, Dan S Bolintineanu, W Michael Brown, Paul S Crozier, Pieter J In't Veld, Axel Kohlmeyer, Stan G Moore, Trung Dac Nguyen, et al. Lammmps-a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer physics communications*, 271:108171, 2022.
- [150] Yair Litman, Venkat Kapil, Yotam MY Feldman, Davide Tisi, Tomislav Begušić, Karen Fidanyan, Guillaume Fraux, Jacob Higer, Matthias Kellner, Tao E Li, et al. i-pi 3.0: A flexible and efficient framework for advanced atomistic simulations. *The Journal of chemical physics*, 161(6), 2024.
- [151] Ralf Drautz. Atomic cluster expansion for accurate and transferable interatomic potentials. *Physical Review B*, 99(1):014104, 2019.
- [152] Scott Habershon, Thomas E Markland, and David E Manolopoulos. Competing quantum effects in the dynamics of a flexible water model. *The journal of chemical physics*, 131(2), 2009.
- [153] Denghui Lu, Han Wang, Mohan Chen, Lin Lin, Roberto Car, E Weinan, Weile Jia, and Linfeng Zhang. 86 pflops deep potential molecular dynamics simulation of 100 million atoms with ab initio accuracy. *Computer Physics Communications*, 259:107624, 2021.

- [154] Linfeng Zhang, Jiequn Han, Han Wang, Wissam Saidi, Roberto Car, et al. End-to-end symmetry preserving inter-atomic potential energy model for finite and extended systems. *Advances in neural information processing systems*, 31, 2018.
- [155] Denghui Lu, Wanrun Jiang, Yixiao Chen, Linfeng Zhang, Weile Jia, Han Wang, and Mohan Chen. Dp compress: A model compression scheme for generating efficient deep potential models. *Journal of chemical theory and computation*, 18(9):5559–5567, 2022.
- [156] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. *The Journal of chemical physics*, 126(1), 2007.
- [157] Emilio Artacho, José María Cella, Julian D Gale, Alberto García, Javier Junquera, Richard M Martin, Pablo Ordejón, Nick Rübner Papior, Daniel Sánchez-Portal, and José M Soler. Siesta 4.1-b4. 2018.
- [158] Fabiano Corsetti, MV Fernández-Serra, José M Soler, and Emilio Artacho. Optimal finite-range atomic basis sets for liquid water and ice. *Journal of Physics: Condensed Matter*, 25(43):435504, 2013.
- [159] Christoph Schran, Fabian L. Thiemann, Patrick Rowe, Erich A. Müller, Ondrej Marsalek, and Angelos Michaelides. Machine learning potentials for complex aqueous systems made simple. *Proceedings of the National Academy of Sciences*, 118(38):e2110077118, 2021.
- [160] Anita Zeidler, Philip S Salmon, Henry E Fischer, Jörg C Neufeind, J Mike Simonson, and Thomas E Markland. Isotope effects in water as investigated by neutron diffraction and path integral molecular dynamics. *Journal of Physics: Condensed Matter*, 24(28):284126, 2012.
- [161] Felix Franks. *The physics and physical chemistry of water*, volume 1. Springer Science & Business Media, 2012.
- [162] Georg Kresse and Jürgen Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical review B*, 54(16):11169, 1996.
- [163] Georg Kresse and Daniel Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical review b*, 59(3):1758, 1999.

- [164] W. F. Kuhs and M. S. Lehmann. The structure of the ice Ih by neutron diffraction. *The Journal of Physical Chemistry*, 87(21):4312–4313, 1983.
- [165] Dávid Péter Kovács, Ilyes Batatia, Eszter Sára Arany, and Gábor Csányi. Evaluation of the mace force field architecture: From medicinal chemistry to materials science. *The Journal of Chemical Physics*, 159(4), 2023.
- [166] Giovanni Bussi, Tatyana Zykova-Timan, and Michele Parrinello. Isothermal-isobaric molecular dynamics using stochastic velocity rescaling. *The Journal of chemical physics*, 130(7), 2009.
- [167] David M Dennison. The crystal structure of ice. *Physical Review*, 17(1):20, 1921.
- [168] WH Bragg. The crystal structure of ice. *Proceedings of the Physical Society of London (1874-1925)*, 34(1):98, 1921.
- [169] Tongqi Wen, Linfeng Zhang, Han Wang, Weinan E, and David J Srolovitz. Deep potentials for materials science. *Materials Futures*, 1(2):022601, may 2022.
- [170] Jian Wang, Yonggang Zheng, Hongwu Zhang, and Hongfei Ye. Machine learning-generated tip4p-bgwt model for liquid and supercooled water. *Journal of Molecular Liquids*, 367:120459, 2022.
- [171] I-Chun Lin, Ari P Seitsonen, Ivano Tavernelli, and Ursula Rothlisberger. Structure and dynamics of liquid water from ab initio molecular dynamics comparison of blyp, pbe, and revpbe density functionals with and without van der waals corrections. *Journal of chemical theory and computation*, 8(10):3902–3910, 2012.
- [172] 5.1. transfer-learning. <https://tutorials.deepmodeling.com/en/latest/CaseStudies/Transfer-learning/Transfer-learning.html>, Accessed: Oct 15, 2023.
- [173] Michele Ceriotti, Michele Parrinello, Thomas E Markland, and David E Manolopoulos. Efficient stochastic thermostating of path integral molecular dynamics. *The Journal of chemical physics*, 133(12), 2010.
- [174] Michele Ceriotti, Giovanni Bussi, and Michele Parrinello. Nuclear quantum effects in solids using a colored-noise thermostat. *Physical review letters*, 103(3):030603, 2009.

- [175] R Brill and A Tippe. Gitterparameter von eis i bei tiefen temperaturen. *Acta Crystallographica*, 23(3):343–345, 1967.
- [176] John P Perdew and Yue Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Physical review B*, 45(23):13244, 1992.
- [177] Atsushi Togo, Laurent Chaput, Terumasa Tadano, and Isao Tanaka. Implementation strategies in phonopy and phono3py. *J. Phys. Condens. Matter*, 35(35):353001, 2023.
- [178] Atsushi Togo. First-principles phonon calculations with phonopy and phono3py. *J. Phys. Soc. Jpn.*, 92(1):012001, 2023.
- [179] George H Vineyard. Frequency factors and isotope effects in solid state rate processes. *Journal of Physics and Chemistry of Solids*, 3(1-2):121–127, 1957.
- [180] Ask Hjorth Larsen, Jens Jørgen Mortensen, Jakob Blomqvist, Ivano E Castelli, Rune Christensen, Marcin Dułak, Jesper Friis, Michael N Groves, Bjørk Hammer, Cory Hargus, Eric D Hermes, Paul C Jennings, Peter Bjerre Jensen, James Kermode, John R Kitchin, Esben Leonhard Kolsbjerg, Joseph Kubal, Kristen Kaasbjerg, Steen Lysgaard, Jón Bergmann Maronsson, Tristan Maxson, Thomas Olsen, Lars Pastewka, Andrew Peterson, Carsten Rostgaard, Jakob Schiøtz, Ole Schütt, Mikkel Strange, Kristian S Thygesen, Tejs Vegge, Lasse Vilhelmsen, Michael Walter, Zhenhua Zeng, and Karsten W Jacobsen. The atomic simulation environment—a python library for working with atoms. *Journal of Physics: Condensed Matter*, 29(27):273002, 2017.
- [181] Graeme Henkelman, Blas P Uberuaga, and Hannes Jónsson. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *The Journal of chemical physics*, 113(22):9901–9904, 2000.
- [182] Ruben Santamaria. *Molecular dynamics*. Springer, 2023.
- [183] Sílvia RA Salinas. *Introdução a física estatística vol. 09*. Edusp, 1997.