

Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Química

Umar Lucio Esper Mucelli Rezende

**DESENVOLVIMENTO DE ALGORITMOS GENÉTICOS ACOPLADOS À
TÉCNICAS DE MACHINE LEARNING PARA OTIMIZAÇÃO DE
GEOMETRIAS DE CLUSTERS ATÔMICOS E/OU MOLECULARES COM
ÊNFASE EM METODOLOGIAS *AB INITIO***

Belo Horizonte
2024

UFMG/ICEX/DQ. 1.609
T.734

Umar Lucio Esper Mucelli Rezende

**DESENVOLVIMENTO DE ALGORITMOS GENÉTICOS ACOPLADOS À
TÉCNICAS DE MACHINE LEARNING PARA OTIMIZAÇÃO DE
GEOMETRIAS DE CLUSTERS ATÔMICOS E/OU MOLECULARES COM
ÊNFASE EM METODOLOGIAS *AB INITIO***

Tese apresentada ao Departamento de Química do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais como requisito parcial para obtenção do grau de Doutor em Ciências - Química.

Orientador: Jadson Cláudio Belchior

Belo Horizonte
2024

Ficha Catalográfica

R467d
2024
T

Rezende, Umar Lucio Esper Mucelli.
Desenvolvimento de algoritmos genéticos acoplados à técnicas de machine learning para otimização de geometrias de clusters atômicos e/ou moleculares com ênfase em metodologias ab initio [manuscrito] / Umar Lucio Esper Mucelli Rezende. 2024.
131 f. : il., gráfs., tabs.

Orientador: Jadson Cláudio Belchior.

Tese (doutorado) – Universidade Federal de Minas Gerais – Departamento de Química.
Bibliografia: f. 100-118.
Apêndices: f. 119-131.

1. Físico-química – Teses. 2. Algoritmos genéticos – Teses. 3. Funcionais de densidade – Teses. 4. Aprendizado do computador – Teses. 5. Redes neurais (Computação) – Teses. 6. Química quântica – Teses. I. Belchior, Jadson Cláudio, Orientador. II. Título.

CDU 043



UNIVERSIDADE FEDERAL DE MINAS GERAIS



"Desenvolvimento de Algoritmos Genéticos Acoplados à Técnicas de Machine Learning para Otimização de Geometrias de Clusters Atômicos e/ou Moleculares com Ênfase em Metodologias *ab initio*"

Umar Lucio Esper Mucelli Rezende

Tese aprovada pela banca examinadora constituída pelos Professores:

Prof. Jadson Cláudio Belchior - Orientador
UFMG

Prof. Frederico Vasconcellos Prudente
UFBA

Prof. Maicon Pierre Lourenço
UFES

Prof. Willian Ricardo Rocha

UFMG

Prof. Guilherme Ferreira de Lima
UFMG

Belo Horizonte, 03 de junho de 2024.



Documento assinado eletronicamente por **Guilherme Ferreira de Lima, Professor do Magistério Superior**, em 03/06/2024, às 18:06, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Jadson Claudio Belchior, Professor do Magistério Superior**, em 05/06/2024, às 09:56, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Frederico Vasconcellos Prudente, Usuário Externo**, em 05/06/2024, às 20:06, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Maicon Pierre Lourenço, Usuário Externo**, em 07/06/2024, às 09:27, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Willian Ricardo Rocha, Professor do Magistério Superior**, em 12/06/2024, às 09:45, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3277078** e o código CRC **1E61185A**.

Dedico esta, bem como todas as minhas demais conquistas, aos meus amados pais Luciana Esper Barão e Umar Lucio Mucelli Rezende por sempre estarem comigo em todos os momentos.

Agradecimentos

Agradeço aos meus pais por pelo apoio incondicional.

Agradeço ao meu orientador, Jadson Cláudio Belchior, por todos os conselhos, pela paciência e ajuda ao longo da minha jornada acadêmica.

Aos meus amigos por sempre estarem comigo em qualquer situação.

À UFMG e ao departamento de química por toda infraestrutura, e seriedade em suas atividades.

Aos órgão de fomento, em especial, a CAPES pelo apoio financeiro para realização deste trabalho de pesquisa.

“Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito. Não sou o que deveria ser, mas Graças a Deus, não sou o que era antes.”
(Martin Luther King)

Resumo

Clusters são agregados de partículas que podem abranger propriedades bem diferentes, dependendo de seu tamanho e composição, entre outras características, podendo se assemelhar a átomos ou moléculas individuais, até propriedades observadas no limite do *bulk* do material. Devido a essa diferenciação de propriedades, clusters podem ser considerados como uma nova classe de materiais, atraindo amplo interesse dentro das diversas áreas científicas. Para estudar os clusters, é muito importante encontrar ou determinar a forma, ou a conformação, que este assumiria na natureza, dentro das condições que ele está sendo estudado. Para determinar a forma destes clusters, algoritmos genéticos têm sido implementados e desenvolvidos ao longo das últimas décadas e se mostraram ferramentas essenciais, embora ainda longe de serem perfeitas, para resolver este tipo de problema. Neste trabalho, um novo algoritmo genético, o NQGA ("*New Quantum Genetic Algorithm*"), com um conjunto de novos operadores genéticos, os quais são propostos para ser uma ferramenta mais eficiente em localizar o mínimo global da superfície de energia potencial para clusters atômicos e moleculares. Destacam-se os operadores Predador de sócia (DGP) e o Predição por *Machine Learning* (MLP) como os principais responsáveis por reduzir drasticamente o número de visitas à superfície de energia quântica e possibilitar otimizações de estruturas utilizando menos recursos computacionais. O NQGA é capaz de realizar otimização de estrutura utilizando cálculo de energia clássico (potenciais interatômicos parametrizados) ou diretamente por métodos *ab initio*, se aproveitando do acoplamento com pacote quântico GAMESS-US e ORCA. O NQGA foi validado para o método clássico utilizando clusters de cobre, ouro e nanoliga de cobre-ouro e ouro-prata para diferentes casos e em todos apresentou resultado coerente com o encontrado na literatura. Pequenos clusters de Lítio foram estudados utilizando a metodologia CCSD(T) para testar a capacidade do NQGA de ser executado em métodos *ab initio* de alto nível. O NQGA se mostrou capaz de encontrar a energia mínima de clusters moleculares, como no caso do cluster $(H_2O)_{11}$ e obteve resultados melhores para o mínimo global do cluster Mg_6H_4 .

Palavras-chaves: clusters; *ab initio*; algoritmo genético; operadores genéticos; otimização; machine learning.

Abstract

Clusters are small aggregates of particles that can exhibit vastly different properties depending on their size, composition, and other characteristics ranging from resembling individual atoms or molecules to properties observed in the bulk limit of the material. Due to this differentiation in properties, clusters can be considered as a new class of materials, attracting significant interest within the scientific field. To study clusters, it is crucial to find or determine the geometry, or conformation, that they would assume in nature under the conditions they are being studied. To determine the geometry of these clusters, genetic algorithms have been implemented and developed over the past decades, proving to be essential tools, although still far from perfect in solving this type of problem.

In this work, a new genetic algorithm, the NQGA (New Quantum Genetic Algorithm), along with a set of new genetic operators, is proposed to be a more efficient tool in locating the global minimum of potential energy surface for atomic and molecular clusters. The doppelgänger predator (DGP) and Machine Learning Prediction (MLP) operators stand out as the main contributors to drastically reduce the number of samples required from the quantum energy surface, enabling more efficient structure optimizations. The NQGA is capable of performing structure optimization using classical energy calculations (parametrized interatomic potentials) or directly through *ab initio* methods, by being coupled with the well developed quantum packages GAMESS-US and ORCA.

The NQGA methodology was validated for the classical method using clusters of copper, gold, and copper-gold and gold-silver nanoalloys for different cases, and in all cases, it yielded coherent results with those found in the literature. Small clusters of lithium were studied using the CCSD(T) methodology to test the NQGA's ability to operate in high-level *ab initio* methods. The NQGA demonstrated great capability to find the geometry of minimum energy of molecular clusters, as in the case of the $(H_2O)_{11}$ cluster, and achieved improved results for the global minimum of the Mg_6H_4 cluster.

Keywords: clusters; *ab initio*; genetic algorithm; genetic operators; geometry optimization; machine learning.

Lista de Figuras

Figura 2.1 - Fluxograma de funcionamento de um algoritmo genético tradicional.	30
Figura 4.1 - Exemplo de rede neural artificial com uma entrada, uma camada oculta e uma saída.	52
Figura 4.2 - Exemplo de rede neural atomística multidimensional artificial. . . .	53
Figura 5.1 - Estruturas dos complexos estudados.	57
Figura 5.2 - Conversão de átomos e clusters em objetos programáveis.	59
Figura 5.3 - Fluxograma de funcionamento do NQGA. Linhas tracejadas representam caminhos opcionais. DGP e MLP são representativos dos operadores predador de sócias (<i>Doppelgänger Predator</i>) e predição por <i>Machine learning</i> (<i>Machine Learning Predictor</i>), respectivamente.	60
Figura 5.4 - Fluxograma de funcionamento simplificado do NQGA utilizando duas etapas, incluindo métodos de avaliação de energia clássicos e quântico.	61
Figura 5.5 - Método de geração aleatório por volume esférico representado no espaço 2D.	62
Figura 5.6 - Aplicação do operador mutação <i>flip</i> no cluster A resultando no cluster B.	65
Figura 5.7 - Aplicação do operador mutação rotacional no cluster A resultando no cluster B.	65
Figura 5.8 - Cluster genérico A sofrendo uma mutação e se tornando o cluster B.	69
Figura 6.1 - Novo mínimo global para o $Cu_{12}Au_7$	77
Figura 6.2 - Progresso do NQGA ao otimizar o cluster $Cu_{12}Au_7$	78
Figura 6.3 - Os três mínimos locais de menor energia encontrados para o cluster $Cu_{12}Au_7$	78
Figura 6.4 - Estruturas otimizadas para os clusters Li_n (para n=3-7) e suas respectivas energias (a.u.), obtidas utilizando DFT/PBE com a base aug-cc-pVTZ.	81
Figura 6.5 - Estruturas de mínimo obtidas pelo NQGA para os clusters Li_8^A , Li_8^B e suas respectivas energias (a.u.), obtidas utilizando DFT/PBE com a base aug-cc-pVTZ.	81
Figura 6.6 - Tempo necessário para obter um cálculo de energia pontual para clusters de lítio. Dados obtidos utilizando o pacote quântico ORCA em um processador de 24 núcleos Intel Xeon E5-2650.	84

Figura 6.7 - Distâncias de ligação e energia (em a.u.) para o Li_2 e LiH , obtidas no nível de teoria DLPNO-CCSD(T)/aug-CC-PVQZ.	85
Figura 6.8 - Estruturas otimizadas dos clusters Li_3 , Li_4 , Li_2H , Li_3H e suas energias (em a.u.), obtidas no nível de teoria DLPNO-CCSD(T)/aug-CC-PVQZ.	86
Figura 6.9 - Estrutura otimizada do $(H_2O)_5$ e sua energia (em a.u.), obtidas no nível de teoria MP2/aug-cc-pVDZ.	87
Figura 6.10 - Estrutura otimizada do $(H_2O)_{11}$ e sua energia (em a.u.), obtidas no nível de teoria MP2/aug-cc-pVDZ.	89
Figura 6.11 - Estrutura otimizada do Mg_6 , obtida pelo NQGA no nível de teoria MP2/aug-cc-pVDZ.	90
Figura 6.12 - Estruturas de mínimo obtidas pelo NQGA (A) para o cluster Mg_6H_4 comparado com o mínimo reportado na literatura (B) e suas energias respectivas (em a.u.), obtidas no nível de teoria DFT/aug-cc-pVDZ com funcional de troca e correlação B3PW91 e correção de dispersão de Grimme's (DFT-D3).	90
Figura 6.13 - Erro quadrático médio (RMSE) do potencial gerado pelo operador MLP durante a otimização do cluster Mg_6H_4	92
Figura 6.14 - Mínimo global do cluster Mg_6H_4 utilizando o método HF-3c para otimização global.	93
Figura 6.15 - Mínimo global do cluster $Mg_6(H_2)$, em A obtido pelo NQGA e B obtido na literatura e suas respectivas energias, em a.u. utilizando o nível de teoria DFT/aug-cc-pVDZ com funcional de troca e correlação B3PW91 e correção de dispersão de Grimme's (DFT-D3).	94
Figura 6.16 - Varredura da distância da molécula de hidrogênio ao centro do cluster Mg_6 em função da energia total do cluster Mg_6H_2	95
Figura 6.17 - Comparativo da energia de adsorção do $Mg_6(H_2)_{x/2}$ com Mg_6H_x utilizando de 1 a 10 pares de hidrogênio.	96
Figura 6.18 - Estruturas dos clusters $Mg_6(H_2)_6$ (A) e Mg_6H_{12} (B).	96

Lista de Tabelas

Tabela 3.1 - Parâmetros para o potencial de Gupta.	44
Tabela 6.1 - Testes preliminares feitos considerando Cu_nAu_m ($n + m \leq X$ para $X = 14, 19, 38, \text{ e } 55$). O potencial de Gupta foi utilizado como função de energia [3.33], e a Tabela 2 compara alguns clusters puros e ligas com trabalhos da literatura.	76
Tabela 6.2 - Testes preliminares feitos considerando Au_nAg_n ($n = 10, 20, 30, 40, 50, 60, 70, \text{ e } 75$). O potencial de Gupta foi utilizado como função de energia [3.33], e a Tabela 2 compara alguns clusters puros e ligas com trabalhos da literatura.	76
Tabela 6.3 - O efeito do operador MLP no número de visitas a SEP quântica para os clusters de lítio Li_n ($n=3-8$).	82

Sumário

1	INTRODUÇÃO	15
2	O ALGORITMO GENÉTICO	25
2.1	Introdução	25
2.2	O algoritmo canônico	26
2.3	Operadores genéticos	27
3	QUÍMICA COMPUTACIONAL	32
3.1	Introdução	32
3.2	Método de Hartree-Fock	34
3.3	Teoria do funcional da densidade (DFT)	38
3.4	Potenciais interatômicos	43
4	APRENDIZADO DE MÁQUINA	46
4.1	Introdução	46
4.2	Aplicação na química	48
4.3	Modelo <i>kernel-based</i>	49
4.4	Modelo de redes neurais artificiais	51
5	A METODOLOGIA <i>New Quantum Genetic Algorithm</i>(NQGA)	56
5.1	Introdução	56
5.2	Geração da população inicial	60
5.3	Operadores tradicionais	63
5.3.1	Imigração	63
5.3.2	Mutação	64
5.3.3	Cruzamento	66
5.4	Operadores adicionais	66
5.4.1	Predador de sócia (DGP)	67
5.4.2	Predador	69
5.4.3	Apex	70
5.4.4	Simetrizador	71
5.4.5	Predição por <i>Machine learning</i> (MLP)	71
5.4.6	Avaliação de energia	73
6	RESULTADOS E DISCUSSÃO	74
6.1	Introdução	74

6.2	Clusters de ligas cobre-ouro e ouro-prata	75
6.3	Clusters de lítio	80
6.4	Clusters de água	86
6.5	Clusters atômicos e moleculares de magnésio-hidrogênio	89
7	CONCLUSÃO	97
7.1	Perspectivas futuras	98
	REFERÊNCIAS	100
A	PARÂMETROS DE ENTRADA DO NQGA	119
A.1	Cobre, Ouro, liga Cobre-Ouro	119
A.2	Liga Ouro-Prata	120
A.3	Lítio (DFT)	122
A.4	Lítio (DLPNO-CCSD(T))	123
A.5	Água	125
A.6	Magnésio-Hidrogênio	126
A.7	MLP	128
B	PUBLICAÇÕES	131

Capítulo 1

Introdução

Existem várias abordagens para descrever a agregação de elementos individuais, cada uma variando conforme o comportamento coletivo desses elementos. Quando um conjunto de indivíduos exibe comportamento semelhante entre si, mas distinto quando comparado ao seu comportamento individual, nos referimos a esse conjunto como um *cluster*. O conceito de *cluster* é amplamente empregado em diversas disciplinas: na astronomia, para designar grupos de corpos celestes unidos pela gravidade; na estatística, para identificar conjuntos de dados semelhantes em algum aspecto; na ciência da computação, para descrever grupos de unidades computacionais trabalhando em conjunto; e na química, para definir aglomerados de átomos ou moléculas. Embora sua tradução mais adequada para o português seja “aglomerados”, o termo “cluster” tem sido mantido em uso, portanto, um cluster pode ser definido como um grupo ou aglomerado de unidades, sejam elas computadores, átomos ou moléculas, que estão fortemente interligados, de modo que, em muitos aspectos, podem ser considerados como um único sistema coeso. Essa interconexão robusta entre as unidades resulta em propriedades e comportamentos distintos que diferem das unidades individuais isoladas. Na química, o estudo dos clusters é essencial para compreender as características e propriedades emergentes que surgem do coletivo, proporcionando *insights* valiosos no entendimento da estrutura, propriedades e comportamento de diversos materiais. Qualquer material presente no dia a dia pode, de certa forma ser considerado um aglomerado de átomos ou moléculas, e o que diferencia eles do que chamamos de clusters químicos é, de maneira genérica, a quantidade de partículas. Para ser mais exato, o que diferencia um cluster de um material em escala macroscópica é a dependência da suas propriedades em relação ao número de elementos que compõe este sistema. São considerados clusters sistemas contendo desde espécies individuais até sistemas observados no limite do *bulk* do material (1). O conceito de *bulk* em um material refere-se ao ponto em que o sistema atinge um tamanho suficientemente grande para que suas propriedades se tornem bem definidas e não variem significativamente com o aumento do tamanho(2). Em outras palavras, o *bulk* representa um regime onde as propriedades do material são consideradas intensivas, ou seja, independentes da escala do sistema. Estruturas em escala nanométrica, conhecidas

como nanoestruturas, têm despertado grande interesse entre pesquisadores devido às suas propriedades frequentemente distintas daquelas observadas no material *bulk* (3). Essas propriedades muitas vezes exibem uma forte dependência do tamanho das estruturas, podendo ser caracterizadas como extensivas, ou seja, dependentes de parâmetros como massa, volume e número de partículas. (4) Um simples exemplo é o metal paládio que no seu estado sólido não apresenta características magnéticas; porém em alguns clusters apresenta momento magnético diferente de zero (5).

O tamanho reduzido para sistemas de clusters acarreta em uma razão superfície/volume elevada e, portanto, apresentando efeitos de energia de superfície, podendo levar à formação de novos materiais e estruturas não cristalinas que não seria possível com o material *bulk* (1, 4). Outro fator importante é a composição destas nanoestruturas. Alterações na composição ou a simples adição de um novo elemento em sua estrutura pode levar a propriedades físicas e químicas completamente diferentes de suas contrapartes separadas (6). Clusters são compostos por diferentes átomos e podem ser classificados como um novo tipo de material, e como existem inúmeras combinações de tamanho e composição diferentes, este é um campo muito atrativo para pesquisas experimentais e teóricas com resultado notáveis (4, 7–10). Existe um grande apelo científico por clusters metálicos e nanoligas, especialmente envolvendo metais de transição devido ao seu potencial de aplicações em catálise, microeletrônica e nanotecnologia (11). O interesse por tais sistemas vai além de entender e determinar suas estruturas, sendo o foco entender como estas estruturas crescem e como suas propriedades se comportam perante este crescimento.

Há várias formas diferentes de descrever as ligações entre os átomos de um cluster, seja usando potenciais atomísticos empíricos, teoria do funcional de densidade (DFT) ou outros métodos quânticos. É importante encontrar, para um dado número de átomos, qual é o arranjo de átomos ou moléculas responsável pela estrutura de menor energia potencial, o mínimo global (GM) na hipersuperfície de energia potencial. É importante determinar a estrutura do clusters de mínimo global, pois estes são os candidatos mais prováveis de serem encontrados/observados caso seja feito um estudo experimental (12). Encontrar este GM é uma tarefa árdua, pois a demanda computacional (também usado na literatura custo computacional) para otimizar um sistema que possui N átomos e, conseqüentemente, $3N-6$ graus de liberdade cresce exponencialmente (13, 14). Além disso, é importante considerar o custo computacional associado à avaliação da energia potencial do sistema por método quântico em cada iteração. Essa demanda computacional aumenta significativamente em sistemas com dispersão, polarização e interações complexas dos elétrons de valência, especialmente em metais de transição. Portanto, é compreensível por que há uma grande quantidade de estudos dedicados a discutir e propor metodologias mais eficientes para determinar o mínimo global.

Um tratamento mecânico-quântico completo demanda tempo e computação de alto desempenho, em contrapartida, produz bons resultados para o estudo teórico de pequenos

clusters. Muitas vezes utiliza-se um potencial empírico para estabelecer um bom ponto de partida para o método quântico. Isso ocorre porque os métodos quânticos aplicados a metais de transição são bem sensíveis à condição inicial. Além disso, esses potenciais são apenas expressões analíticas parametrizadas. São funções matemáticas com embasamento físico extremamente rápidas de serem calculadas. Não existem muitos potenciais empíricos parametrizados para metais de transição, mas por exemplo, o potencial de *Finnis Sinclair* (FS) (15,16) foi usado por James *et al.* (17) para o estudo de clusters de Mo_N e Fe_N , para $N = 3-100$, com resultados muito bons, sendo possível determinar os números mágicos, tipos preferenciais de estruturas e ainda uma comparação com métodos semi-empíricos (17). Outro potencial muito utilizado para trabalhar sistemas metálicos é o de Gupta (18) utilizado em sistemas de cobre e ouro (19,20), ouro e prata (21) e vários outros (22–26). Independente do método usado para avaliar a energia dos clusters, sabe-se que o número de mínimos locais presentes na hipersuperfície de energia potencial aumenta drasticamente à medida que o cluster aumenta de tamanho. Muitas metodologias foram desenvolvidas para buscar/determinar o mínimo global, porém a existência de múltiplos mínimos locais ocultando a presença de um mínimo mais profundo faz com que muitos desses métodos falhem, incentivando o desenvolvimento de novos métodos mais eficientes.

Algoritmos de otimização global baseado na avaliação do gradiente, por exemplo o BFGS (27–30) falham em identificar o GM pois devido aos mínimos profundos da superfície de energia potencial (SEP) o algoritmo fica estagnado (ou "preso") em mínimos locais profundos. Uma abordagem eficaz para utilizar o método BFGS na busca pelo mínimo global de um cluster é iniciar a otimização com uma geometria inicial próxima da expectativa do mínimo global. Isso reduz as chances do algoritmo ficar estagnado em um mínimo local. Esta é uma técnica muito empregada para estudo da estrutura eletrônica e propriedades de cluster utilizando pacotes quânticos como Orca (31, 32) e GAMESS-US (33) que utiliza o BFGS para otimizar a geometria do cluster a ser estudado. Para evitar erros provindos de estruturas iniciais, que convergiriam para mínimos locais, vários métodos, em maioria métodos heurísticos capazes de otimização global, têm sido estudados e empregados na otimização de clusters atômicos e moleculares de diversos tamanhos.

Os algoritmos de otimização heurística são técnicas que visam encontrar soluções aproximadas para problemas de otimização. Ao contrário dos métodos exatos, que buscam encontrar a solução ótima, os algoritmos heurísticos buscam soluções satisfatórias em tempo viável. Eles são frequentemente empregados quando encontrar a solução ótima é computacionalmente inviável devido à complexidade do problema ou à falta de uma solução exata conhecida. Em contraste, os métodos não heurísticos, como os algoritmos de programação linear ou inteira, buscam encontrar a solução ótima por meio de procedimentos matemáticos precisos e determinísticos.

Existem vários tipos diferentes de algoritmos de otimização heurísticos, cada um com

suas próprias características e aplicações. Alguns exemplos comuns incluem o *basin-hopping* (BH)(34, 35), colônia de abelhas artificiais (ABC)(36, 37), enxame de partículas(38, 39), Monte-Carlo (MC)(40) e algoritmo genético (GA)(41, 42).

O algoritmo genético pertence a uma classe de algoritmos evolucionários e é inspirado na teoria da evolução de Darwin, e utiliza técnicas de seleção natural para encontrar soluções aproximadas para um problema. Ele funciona criando uma população de soluções candidatas e utilizando técnicas (operadores genéticos) de "cruzamento" e "mutação" para gerar novas soluções a partir das existentes. As soluções são avaliadas de acordo com uma função de aptidão e as melhores são selecionadas para serem utilizadas como base para a próxima geração de soluções. Este processo é repetido até que uma solução satisfatória seja encontrada ou um critério de parada seja atingido.

Existem dois tipos principais de operadores genéticos: o cruzamento (ou "*crossover*") e a mutação. O cruzamento consiste em combinar duas soluções existentes para gerar uma nova solução. Isso é feito escolhendo aleatoriamente uma "posição de corte" e combinando os elementos das duas soluções à esquerda e à direita da posição de corte. A mutação é o processo de alterar aleatoriamente uma ou mais partes de uma solução existente. Isso é feito escolhendo aleatoriamente uma ou mais posições da solução e alterando o valor nessa posição (43).

Adaptando os operadores genéticos para trabalhar com clusters e ajustando a função de aptidão para representar a energia do cluster de acordo com o método escolhido, clássico ou quântico, pode-se utilizar o algoritmo genético (GA) para otimizar a superfície de energia potencial (SEP) de um determinado cluster. Essa técnica tem se mostrado bastante adequada como opção para abordar problemas de minimização em SEP, apresentando vantagens como: dispensa da necessidade de calcular o gradiente na SEP, baixo custo computacional e, por ser um método estocástico, maior chance de identificar o mínimo global, ao invés de ficar preso em mínimos locais. Os operadores genéticos desempenham um papel crucial na eficácia dos algoritmos genéticos para essa tarefa considerando o custo computacional de cada amostragem feita na SEP. Utilizar algoritmos genéticos associados a potenciais empíricos tem sido uma técnica comum para estabelecer um ponto de partida para posterior refinamento. Isso é mencionado em referências como (44, 45). No entanto, é mais raro encontrar associação direta a métodos quânticos *ab initio*. Isso envolve determinar, em princípio, o mínimo global diretamente na superfície de energia quântica. Cruzamento e mutação são os operadores mais comuns em algoritmos genéticos porém diversos outros operadores são propostos e testados para atingir um determinado objetivo e melhorar a eficiência do GA. Apesar de ser uma tarefa árdua, novos algoritmos genéticos estão sendo desenvolvidos e um dos principais focos de melhoria em um GA é manter sua população o mais diversa possível para garantir a exploração apropriada da SEP.

Guimarães et al.(46) introduziram dois novos operadores no algoritmo: o "aniquilador"

e o “história”. O operador “aniquilador” é responsável por eliminar toda a população, efetivamente reiniciando o algoritmo. Por outro lado, o operador “história” mantém armazenados os melhores indivíduos gerados ao longo das iterações. Esses indivíduos são então usados para comparação, visando a obtenção do melhor cluster possível, idealmente correspondente ao mínimo global. Este GA foi utilizado para diversos sistemas atômicos e moleculares e obteve elevado grau de sucesso (19, 47, 48).

Até o momento os algoritmos genéticos propostos utilizavam a função de aptidão, também conhecida como função custo, como uma medida da energia potencial clássica para distinguir entre clusters de melhor qualidade (menor energia) e clusters de qualidade inferior (maior energia). Os clusters otimizados, com base em potenciais clássicos, frequentemente serviam como ponto de partida, mas nem sempre representavam mínimos globais na superfície de energia potencial quântica (SEPQ). Vale ressaltar que a função de aptidão pode assumir diferentes formas, podendo ser contínua ou discreta. Davis, Horswell e Johnston(49) adotaram uma abordagem diferente ao utilizar a metodologia DFT (Teoria do Funcional da Densidade) com ondas planas calculadas pelo pacote quântico *Quantum Espresso* como função de aptidão em seu algoritmo genético BCGA (*Birmingham Cluster Genetic Algorithm*), visando otimizar clusters de Irídio e Platina.

O QGA (*Quantum Genetic Algorithm*) foi nossa primeira proposta de algoritmo genético que incorpora um método quântico, utilizando funções de base gaussianas. Essa escolha se deve às propriedades específicas das funções de base gaussianas, como sua localização, escalabilidade e capacidade eficiente de representação. Essas características as tornam ideais para a modelagem de sistemas quânticos, especialmente nanoclusters. O QGA utiliza apenas operadores clássicos, cruzamento e mutação, acoplado com o pacote quântico GAMESS-US para avaliar a energia dos clusters. Uma das vantagens deste pacote quântico é sua capacidade de utilizar diferentes níveis de teoria, possibilitando do QGA em diversos sistemas utilizando o nível de teoria que melhor reproduz dados experimentais disponíveis.(50). Uma das desvantagens é que devido a simplicidade de seus operadores, o QGA se torna computacionalmente proibitivo para sistemas maiores. Esta não é uma desvantagem exclusiva do QGA, mas sim uma limitação dos métodos de minimização globais, que obrigatoriamente, aumenta o número de visitas a SEP.

Os cálculos de estrutura eletrônica utilizando teorias quânticas requerem considerável demanda computacional. Entretanto, dentre as várias abordagens disponíveis neste campo, o método DFT se destaca por sua impressionante relação custo-benefício. Esta abordagem oferece um resultado preciso sem as taxas computacionais exorbitantes em comparação com outros métodos. Portanto, não é surpresa que o método DFT seja uma das técnicas mais amplamente utilizadas no estudo de estrutura eletrônica. A DFT é uma metodologia utilizada em química e na física para estudar a estrutura eletrônica de átomos, moléculas e até mesmo para o estudo de sistemas no estado sólido. Embora exista uma polêmica em relação se a DFT deve ou não ser considerada um método *ab initio*, não devemos deixar

de considerar que é uma metodologia que já foi utilizada com muito sucesso em inúmeros trabalhos e é ideal para sistemas grandes ou de alta complexidade (51).

Para que um GA quântico seja viável para sistemas maiores, é necessário propor operadores que faça o menor número de amostragens possível na SE PQ e ainda mantendo uma boa variabilidade da população. Silva, Silva e Belchior(52) investigaram diversos operadores genéticos disponíveis em sistemas de 26 e 55 átomos, utilizando o potencial interatômico de Lennard-Jones. Com isso, foi possível concluir que certos operadores, por exemplo o operador torção, não falham para o sistema estudado e utilizam um número médio de iterações bem pequeno, enquanto outros como a mutação dinâmica falha 100% das vezes. Para melhorar ainda mais esta metodologia, Vilhelmsen e Hammer(53) mostraram a importância de detectar e eliminar clusters com estrutura similar presentes na população para evitar que haja convergência da geometria para estruturas já estudadas durante a etapa relaxamento local. Lazauskas, Sokol e Woodley(54) estudaram um método para eliminar estruturas que têm uma alta probabilidade de falhar durante a etapa de minimização local, reduzindo o gasto de recursos computacionais e melhorando a eficiência geral do algoritmo. O fator limitante dos algoritmos genéticos quânticos é justamente as visitas na superfície quântica ou a modelagem da SE PQ. Uma das ferramentas que pode ajudar com essa modelagem são as redes neurais artificiais (55, 56).

As redes neurais artificiais (RNA) são modelos computacionais inspirados na estrutura e função do sistema nervoso humano. Elas são capazes de aprender a partir de dados, e são comumente utilizadas para tarefas de classificação, reconhecimento de padrões e predição (57). RNA é uma área ativa de pesquisa e desenvolvimento, e tem sido usada para uma variedade de aplicações, incluindo processamento de linguagem natural, visão computacional e reconhecimento de fala. RNA têm sido amplamente utilizadas em química computacional, devido à sua capacidade de aprender padrões complexos a partir de dados(58). Algumas das aplicações comuns incluem a previsão de propriedades químicas(59), a descoberta de novos compostos(60), a otimização de processos químicos(61) e a análise de dados experimentais(62).

As redes neurais artificiais têm sido utilizadas como uma ferramenta para reproduzir superfícies de energia potencial quântica. Essas superfícies são usadas para descrever como a energia potencial de um sistema químico varia com a configuração geométrica dos seus átomos. Elas são fundamentais para a compreensão de reações químicas e podem ser usadas para prever como os átomos se movem durante uma reação. Os métodos tradicionais para calcular a SE PQ são baseados em métodos quânticos, como a teoria do funcional da densidade (DFT), que são computacionalmente intensivos e podem ser limitados para sistemas de maior complexidade. As redes neurais artificiais podem ser usadas para reproduzir essas superfícies de forma mais rápida e eficiente(63, 64). Existem várias abordagens para usar redes neurais para reproduzir a SE PQ. Um exemplo é usar uma rede neural para aprender a mapear configurações geométricas de átomos para

energias potenciais correspondentes, a partir de dados gerados por métodos quânticos tradicionais(65). Outra abordagem é usar uma rede neural para aprender a prever a configuração geométrica ótima de um sistema químico a partir de sua energia potencial(66). Algumas abordagens também combinam as redes neurais com outros métodos, como o método de Monte Carlo, para aprimorar a precisão da previsão da SEPQ(67). É importante ressaltar que, a reprodução de superfície de energia potencial quântica utilizando RNA ainda é um campo em desenvolvimento, e ainda existem desafios a serem superados, como a falta de dados disponíveis, a generalidade e a interpretabilidade dos resultados.

Este trabalho visa condensar os avanços feitos ao longo dos anos em algoritmos genéticos junto a novas ideias para propor um novo GA capaz de fazer otimização de clusters utilizando métodos quânticos. Um GA quântico robusto deve ser capaz de otimizar clusters utilizando o menor número possível de visitas a superfície de energia potencial quântica para ser computacionalmente viável, e por isso deve conter operadores capazes de discernir dentre todos os clusters gerados ao longo das iterações quais devem passar por tratamento quântico e quais devem ser descartados. Outra propriedade desejável é que seja possível utilizar diferentes níveis de teoria para abordagem teórica e que o algoritmo comporte a maior diversidade possível de sistemas como atômico ou molecular, neutro ou carregado, rígido ou flexível e misturas, conhecidas como *alloys*.

Torna-se particularmente importante, a partir das informações apresentadas até aqui, o interesse em descrever clusters cada vez com a maior exatidão e rigor possíveis para que se possa fazer o estudo adequado de suas propriedades eletrônicas e elucidar possíveis aplicações práticas dos clusters estudados. Afinal, a dependência das propriedades dos clusters em relação a seu tamanho já foi extensivamente discutida. Devemos reforçar que um tratamento teórico de alto nível para sistemas grandes, e ou complexos, se torna rapidamente proibitivo mesmo com a metodologia proposta neste trabalho, porém com avanços nos algoritmos e na capacidade de processamento dos computadores será possível estudar sistemas cada vez maiores e mais complexos.

O objetivo deste trabalho é desenvolver uma metodologia de otimização geométrica de clusters e nanoclusters utilizando o conceito de algoritmo genético nomeado NQGA (*New Quantum Genetic Algorithm*). O maior diferencial do NQGA será o foco em minimizações de energia na superfície de energia potencial quântica de maneira direta. Isso será realizado com o auxílio de pacotes quânticos como Orca e GAMESS-US. Essa abordagem é diferente da implementação anterior em nosso grupo de pesquisa, onde o QGA (ref. (68)) foi exclusivamente implementado no programa GAMESS-US, enfrentando, portanto, as limitações inerentes a este programa. Em particular, a implementação também no Orca permitirá maior flexibilidade em otimizações na superfície quântica, tais como, um algoritmo melhorado para cálculo de integrais denominado “SHARK” e um controle granular para trabalhar sistemas nos quais efeitos relativísticos devem ser considerados. A construção desta metodologia levará em consideração vários avanços feitos nos GAs e

operadores genéticos para reduzir o número de visitas a SEPQ, reiterando que estas visitas representam a maior barreira neste tipo de algoritmo devido ao alto custo computacional de uma metodologia quântica. Para contribuir com a etapa de modelar e minimizar a SEPQ irá contar com um operador genético baseado em RNA. Finalmente esta metodologia será testada em sistemas previamente estudados na literatura para validar sua eficácia, assim verificando a usabilidade desta metodologia para estudar propriedades da estrutura eletrônica de novos sistemas. Portanto, este trabalho irá focar nos seguintes pontos:

- Desenvolvimento da metodologia do NQGA;
 - Inicialização eficiente uma otimização global;
 - Comparação de estruturas químicas similares;
 - Balanço entre número de visitas a SEPQ e a exatidão do algoritmo;
 - Geração de novas estruturas candidatas a mínimo global;
- Validação da metodologia em sistemas previamente disponíveis na literatura.

Nosso grupo de pesquisa tem se empenhado vigorosamente no desenvolvimento de metodologias sofisticadas aplicadas a sistemas complexos, com um enfoque particular em estudos de sistemas atômicos e moleculares para cálculos de otimização estrutural. Existem vários trabalhos na literatura que se concentram em algoritmos genéticos para minimização de funções de energia potencial (campos de força), e mais recentemente, soluções na superfície de energia potencial quântica, conforme discutido anteriormente neste capítulo. Esses esforços mostram o interesse e a necessidade de se desenvolver algoritmos mais eficientes que terão contribuição significativa para o avanço do campo da química computacional.

É amplamente aceito que a eficácia na obtenção do mínimo global - seja de um conjunto de partículas descritas por uma função matemática ou de um conjunto de dados específico - é fundamentalmente um conceito teórico. Embora seja possível alcançar esse mínimo, na maioria dos casos, as técnicas de minimização e otimização são continuamente aperfeiçoadas. Com o tempo, essas melhorias permitem determinações cada vez mais precisas do mínimo global, aproximando-se cada vez mais do ideal teórico. Nossos estudos seguem essa tendência. Em 2015, propusemos nosso primeiro algoritmo quântico (QGA) para o estudo de clusters atômicos e os testes demonstraram alta eficiência para os sistemas de sódio-potássio(68). Em um estudo subsequente, conseguimos aplicar com êxito a técnica aos clusters de nitrogênio(69). Em um trabalho mais recente, expandimos nossa abordagem para incluir clusters bimetálicos de sódio-lítio(50).

Nosso objetivo na presente proposta de estudo é criar um algoritmo genético inovador e, para aprimorá-lo, uma das nossas estratégias é integrá-lo com a nova metodologia de *Machine Learning*, que consiste numa técnica de inteligência artificial de aprendizagem

inteligente e que tem demonstrado ser altamente eficiente em diversas áreas(61–64). Entretanto, o problema de otimização na superfície quântica, resolvendo a energia através da equação de Schrödinger independente do tempo, é de extrema importância e merece atenção especial. Por esta razão, este trabalho concentra-se particularmente nessa área (otimização de geometrias na SEP quântica), visando contribuir para o avanço do entendimento e desenvolvimento de métodos eficazes para esta classe de problemas. Tendo essas premissas como foco principal do desenvolvimento do trabalho, preparamos um procedimento de desenvolvimento eficiente que garante a versatilidade e eficiência do método. Devido à grande complexidade deste tipo de problema de otimização, especialmente em sua integração com as metodologias da mecânica quântica e o desenvolvimento de operadores genéticos dentro de técnicas de inteligência artificial (IA), optou-se inicialmente pelo desenvolvimento em uma aplicação específica. Esta aplicação é exclusivamente focada em uma molécula com um certo grau de complexidade. O objetivo principal é conduzir testes preliminares dos operadores genéticos, bem como otimizar e acoplar diversas metodologias de mecânica quântica e algoritmos genéticos. Além disso, a nova proposta inclui a incorporação de outra técnica de inteligência artificial, redes neurais artificiais, para mapear a superfície quântica de clusters de átomos e moléculas.

Propusemos um estudo utilizando a metodologia proposta neste trabalho para investigar a influência de ligantes de bi-piridina dissubstituídos com grupos doadores e receptores de elétrons nas propriedades estruturais e eletrônicas dos complexos de ibuprofenato de Eu^{+3} . Estes complexos possuem várias características desejáveis para um trabalho preliminar explorando a aplicabilidade e funcionalidade dos operadores genéticos que foram propostos para o NQGA, tais como: quantidade total de átomos grande o suficiente para representar vários outros sistemas, ao mesmo tempo que não requer tempo computacional excessivo para determinação da estrutura eletrônica, quantidade fixa de átomos e relevância científica. Além disso, é possível testar de forma eficiente os operadores genéticos a serem aplicados na metodologia, bem como o desenvolvimento dos novos operadores acoplados às redes neurais. Com base nos resultados de aplicação deste operadores no complexo de ibuprofenato de Eu^{+3} , é possível selecionar quais operadores são, a princípio, mais eficientes, e portanto, tornando possível expandir esses resultados para o NQGA na forma de técnicas generalizadas que constituem a classe de operadores genéticos. Com base nos dados e resultados obtidos, foi possível desenvolver o algoritmo proposto com foco nos testes preliminares e no desenvolvimento da metodologia que incorpora técnicas sofisticadas de aprendizado de máquina, destacando a interseção entre a ciência dos dados e a química computacional. Prevê-se que a implementação completa do NQGA permitirá a reavaliação e exploração mais precisa e acessível de trabalhos semelhantes aos que serão apresentados nos próximos capítulos, bem como outros de natureza semelhante. Isso permitirá um foco primordial na análise dos resultados, eliminando um tempo considerável dedicado à elaboração das estruturas e pré-processamento necessário para a execução dos cálculos

quânticos, uma vez que todo o trabalho de formação do complexo, interação e obtenção da estrutura eletrônica será facilitado pelo NQGA.

Este trabalho representa um passo significativo em nossa busca contínua para aprimorar a eficiência e a aplicabilidade dos algoritmos genéticos na otimização estrutural de sistemas atômicos e moleculares complexos. Acreditamos que nossos esforços contribuirão significativamente para o campo e esperamos continuar a expandir e aprimorar nossas metodologias no futuro. Nos capítulos seguintes serão apresentados o detalhamento do desenvolvimento teórico da parte de algoritmos genéticos e os correspondentes operadores genéticos, o escopo teórico das metodologias da mecânica quântica, o detalhamento das técnicas de IA primariamente via *Machine Learning* e finalmente a construção do novo NQGA com o detalhamento de todo o modelo desenvolvido. Finalmente algumas aplicações de clusters atômicos, tais como lítio, magnésio, ligas cobre-ouro, ouro-prata, lítio-hidrogênio, assim como aplicações moleculares como clusters de água e magnésio-hidrogênio serão desenvolvidas para os testes do NQGA. Posteriormente uma proposta para os trabalhos futuros, focados neste tipo de estudo, será apresentada e finalmente as principais conclusões.

Capítulo 2

O algoritmo genético

2.1 Introdução

Algoritmo genético (GA) é um modelo computacional, inspirado em processos biológicos, com modelo originário nas raízes da ciência evolucionária. Geralmente, os GA's representam um procedimento de busca e otimização em um determinado espaço, porém, diferente das técnicas tradicionais, ele não avalia e melhora apenas uma solução, e sim um conjunto de soluções, ou hipóteses, chamado população. O GA produzirá uma hipótese descendente desta população usando mecanismos de mutação e recombinação das melhores hipóteses disponíveis na população, logo, a cada iteração a população atual será substituída pelos descendentes das soluções mais adequadas.

Matematicamente, os GA's são otimizadores de funções que codificam potenciais soluções em um formato que pode ser alterado sem perder informações essenciais de cada solução. As propriedades interessantes advindas do uso de GA's são algoritmos extremamente robustos com relação ao conjunto de dados de entrada, eficientes em sua execução e são de fácil programabilidade (70). As principais características que diferem um GA de outros métodos de otimização são:

- Possibilidade de manipulação direta do código de otimização;
- O resultado é uma população de soluções e não uma solução única;
- A busca é feita via amostragem, ou as vezes chamada, busca cega;
- O algoritmo usa operadores estocásticos e não regras determinísticas;

A propriedade mais importante dos algoritmos genéticos é sua robustez, isto porque, de certa forma, o GA simula o "algoritmo" presente na natureza que seleciona os indivíduos mais bem adaptados. Isso significa que é possível encontrar uma ou mais soluções mesmo que o conjunto de dados não facilite a obtenção destas soluções. Os GA's são geralmente utilizados quando se trabalha com funções não deriváveis ou funções com múltiplos mínimos

locais. Apesar de todas as vantagens mencionadas até agora, a grande desvantagem do GA é que muitas vezes ele deve ser cuidadosamente desenvolvido para tratar um tipo de problema, onde uma escolha de operadores genéticos desfavorável pode causar problemas nos resultados e, por isso, é importante entender muito bem o problema o qual deseja-se aplicar o GA (71).

2.2 O algoritmo canônico

A apresentação do algoritmo genético canônico é feita no formato de conjunto numérico binário e emprega técnicas simples de processamento de informação como, por exemplo, cópia, inserção, substituição de partes numéricas e geração aleatória de números. Geralmente, a população inicial é gerada aleatoriamente. Cada indivíduo da população é avaliado por uma função de aptidão, ou função custo, que servirá como parâmetro para ordenar, conforme uma hierarquia, a população, que no caso de clusters químicos é sua energia no atual estado. Esse ranqueamento além de definir as melhores hipóteses da população, também serve para aumentar a chance de reprodução dessas hipóteses.

A definição canônica para um GA é dado por f_i/\bar{f} , onde f_i é o valor de aptidão associado a cada indivíduo i , e \bar{f} é o valor custo médio considerando todos os indivíduos da população

$$\bar{f} = \frac{1}{n} \sum_{i=1}^n f_i \quad (2.1)$$

Após avaliar a aptidão de cada indivíduo, a população será submetida a operadores genéticos para gerar novos indivíduos. Operadores genéticos são técnicas utilizadas em algoritmos genéticos para modelar o processo de evolução biológica. Eles são utilizados para manipular a representação de indivíduos em uma população, a fim de produzir novos indivíduos com características desejadas. Os operadores mais comuns são a seleção, o cruzamento, e a mutação (72):

- Seleção: É o processo de escolher indivíduos para a próxima geração com base em sua aptidão (*fitness*), que é uma medida de quão bem eles se adaptam ao ambiente. Indivíduos com maior aptidão são mais propensos a ser selecionados para a próxima geração por meio da aplicação de outros operadores. Este operador preserva os indivíduos mais adaptados e favorece a propagação de seus genes.
- Cruzamento/Crossover: O operador de cruzamento, reprodução ou recombinação (em inglês *crossover*), é um dos principais operadores genéticos utilizados em algoritmos genéticos. Ele é usado para combinar as características de dois indivíduos diferentes para produzir um novo indivíduo, de forma a passar características desejáveis para a próxima geração. É importante notar que a taxa de cruzamento também é um fator

crítico, uma taxa muito baixa pode resultar em uma população estagnada, enquanto uma taxa muito alta pode resultar em uma perda de características desejadas.

- **Mutação:** Este operador também é um dos principais operadores genéticos utilizados em algoritmos genéticos. Ele é usado para introduzir variações aleatórias nas características de um indivíduo, a fim de explorar novos espaços de busca e evitar que a população se torne estagnada em um único ponto ótimo. A mutação é aplicada de forma aleatória e é uma técnica que é usada para introduzir novas variações genéticas na população, o que ajuda a evitar o problema de otimização local. A taxa de mutação é um fator importante, pois uma taxa muito alta pode resultar em uma perda de características desejadas, enquanto uma taxa muito baixa pode resultar em uma população estagnada.

É importante notar que a mutação é um processo estocástico, e sua aplicação deve ser criteriosa, pois ela pode afetar a estabilidade do algoritmo genético. Assim, é necessário encontrar um equilíbrio entre a taxa de mutação e a taxa de cruzamento, para evitar que o algoritmo se torne instável e que a população evolua para uma solução sub ótima(73). Múltiplas operações de cruzamento entre os indivíduos com maior aptidão tendem a criar uma população cada vez mais homogênea, e a mutação contribui para que novas combinações sejam feitas para aumentar a probabilidade de que soluções melhores sejam criadas.

2.3 Operadores genéticos

A implementação das classes de operadores genéticos difere bastante quando aplicada a sistemas químicos, bem diferentes em relação aos sistemas matemáticos(42). Para os sistemas matemáticos, as representações e operações são em base binária, enquanto para os sistemas químicos são utilizados sistemas de coordenadas, geralmente cartesianos, o que torna essa aplicação um caso bem mais desafiador (42). Estes são alguns dos operadores de cruzamento disponíveis na literatura:

- **Cortar e emendar(C&S)(42):** Átomos do cluster k são separados em dois grupos por um plano gerado de maneira aleatória. Outro plano aleatório é definido para dividir o cluster l . Os grupos gerados pela separação de k e l devem conter o mesmo número de átomos. Então, os grupos com números de átomos equivalentes serão trocados entre os clusters k e l para gerar um novo cluster m com mesmo número de átomos dos progenitores;
- **Cortar e emendar generalizado(genC&S)(74):** Esta versão do cortar e emendar é menos tendenciosa, pois é baseada apenas nas distâncias euclidianas entre os átomos, ao invés de planos de corte. Iniciando um cruzamento entre os cluster k e l , começa-se

selecionando um átomo aleatório, nomeado CP, do cluster k , e posteriormente um número $X \in [1, N - 2]$, onde N é o número de átomos que compõe o cluster k . Um novo cluster m é gerado a partir da cópia do átomo CP e dos X átomos mais próximos, ainda usando o cluster k como referência. Do cluster l , são selecionados $N - (X + 1)$ átomos para completar m . Será dada preferência para os átomos mais próximos de CP no espaço 3D. Átomos muito próximos (e.g. distancia menor que δ) de partículas já presentes em m são desconsiderados. No caso de não haver átomos o suficiente em l para completar m , então átomos aleatórios serão gerados;

- Cortar e emendar esférico(C&SE)(75): Operador análogo ao C&S, porém, ao invés de utilizar planos para separar os átomos, são utilizadas esferas. Uma esfera é definida para separar o cluster k em dois grupos de átomos, um na parte interior da esfera e o outro na parte exterior. A mesma esfera é gerada para o cluster l . As partes equivalentes serão trocadas para gerar o novo cluster m ;
- Cruzamento aritmético(70): Sejam $x^{(k)}$, $x^{(l)}$ e $x^{(m)}$ vetores de coordenadas cartesianas dos indivíduos k , l e m respectivamente. Este operador gera o novo individuo m utilizando a seguinte equação: $x^{(m)} = 0.5(x^{(k)} + x^{(l)})$;
- Cruzamento de dois pontos(76): As coordenadas dos átomos dos clusters selecionados para cruzamento, k e l , são colocadas em um vetor unidimensional. Dois números inteiros são gerados: $s_1 = [1, (3N - 1)]$ e $s_2 = [(s_1 + 1), 3N]$. As coordenadas do cluster k que estiverem entre s_1 e s_2 são substituídas pelas coordenadas do cluster l que estão no mesmo intervalo;
- Cruzamento uniforme(76): Para gerar o vetor de coordenadas cartesianas do novo individuo gerado $x^{(m)}$, cada coordenada $x_i^{(m)}$ tem uma probabilidade especifica de vir de cada um de seus progenitores k e l . E.g. $x_i^{(m)}$ tem α chance de vir do cluster k $x_i^{(k)}$ e β chance de vir do cluster l $x_i^{(l)}$;

Os próximos operadores são do tipo mutação, portanto atuam em clusters de forma singular, não requerendo um par de clusters como no caso do cruzamento:

- Operador angular(77): Este operador age em um intervalo de 1-5% do total de número de átomos do cluster escolhidos de forma aleatória. Cada átomo selecionado é movido aleatoriamente na superfície de uma esfera de raio R_i (representado pela distancia geométrica entre o átomo e o centro do cluster);
- Operador de deslocamento cartesiano(78): Este operador age em um intervalo de 1 a N átomos escolhidos aleatoriamente, onde N é o número total de átomos do cluster. Os átomos selecionados são modificados de acordo com a equação $x_i^{(m)} = x_i^{(k)} + Sr_{min}([-1, +1]\hat{i} + [-1, +1]\hat{j} + [-1, +1]\hat{k})$ onde $x_i^{(m)}$ são as novas

coordenadas do átomo i , $x_i^{(k)}$ são as antigas coordenadas do átomo i , S é um parâmetro arbitrário, $\hat{i}, \hat{j}, \hat{k}$ são vetores unitários e r_{min} é a distancia entre o átomo i e o átomo mais próximo.

- Mutação dinâmica(76): Este operador age em todos os átomos do cluster selecionado e modifica de acordo com a equação: $r_i^{(m)} = [(1 - \delta), (1 + \delta)]r_i^{(k)}$ onde $r_i^{(m)}$ são as novas coordenadas do átomo i , $r_i^{(k)}$ são as antigas coordenadas do átomo i e δ é um parâmetro arbitrário;
- Modificação do interior(79): Este operador move um átomo em direção ao centro do cluster. Sendo R a distancia entre o átomo selecionado e o centro do cluster. O átomo sera movido para uma posição aleatória na superfície de uma esfera de raio $[0.01, 0.1]R$, centrada no centro geométrico do cluster;
- Operador de superfície angular(80): Este operador move um átomo para a superfície do cluster. Um átomo aleatório é selecionado. Sendo R_{max} a distancia entre o centro geométrico do cluster e o átomos mais distante deste centro. O átomo selecionado é, então, movido para uma posição aleatória na superfície de uma esfera com raio R_{max} centrada no centro geométrico do cluster;
- Operador de torção(76, 78): Um plano aleatório é definido para separar o cluster selecionado em duas partes, não necessariamente do mesmo tamanho. Uma das porções é rotacionada ao longo de um eixo formado pela normal do plano de corte.

Como a definição de um operador de mutação é simplesmente a alteração de alguma característica geométrica do cluster, há uma variedade bem maior de operadores disponíveis na literatura além dos que foram exemplificados. Muitos outros operadores foram propostos na literatura (70, 81, 82) e, como explicado anteriormente, devem ser escolhidos com cuidado considerando as particularidades do tipo de função a ser otimizada. A partir das informações apresentadas, podemos ilustrar o funcionamento de um GA genérico para clusters, como:

Pode-se notar que é um algoritmo simples, com relação ao número de decisões, com funcionamento iterativo linear em que a natureza heurística beneficia da aleatoriedade presente nas etapas “População inicial”, “Seleção”, “Cruzamento” e “Mutação”.

É importante ressaltar que, dentro do algoritmo, a avaliação de energia é um passo independente. Isso significa que é possível utilizar qualquer função ou método que correlacione coordenadas cartesianas com a energia disponível para realizar essa avaliação. Dessa forma, a flexibilidade na escolha da função de energia permite adaptar o algoritmo para diferentes sistemas químicos e níveis de teoria.

Como o algoritmo genético (GA) é um método heurístico, não existe convergência garantida. Portanto, é fundamental estabelecer um critério de parada para finalizar o

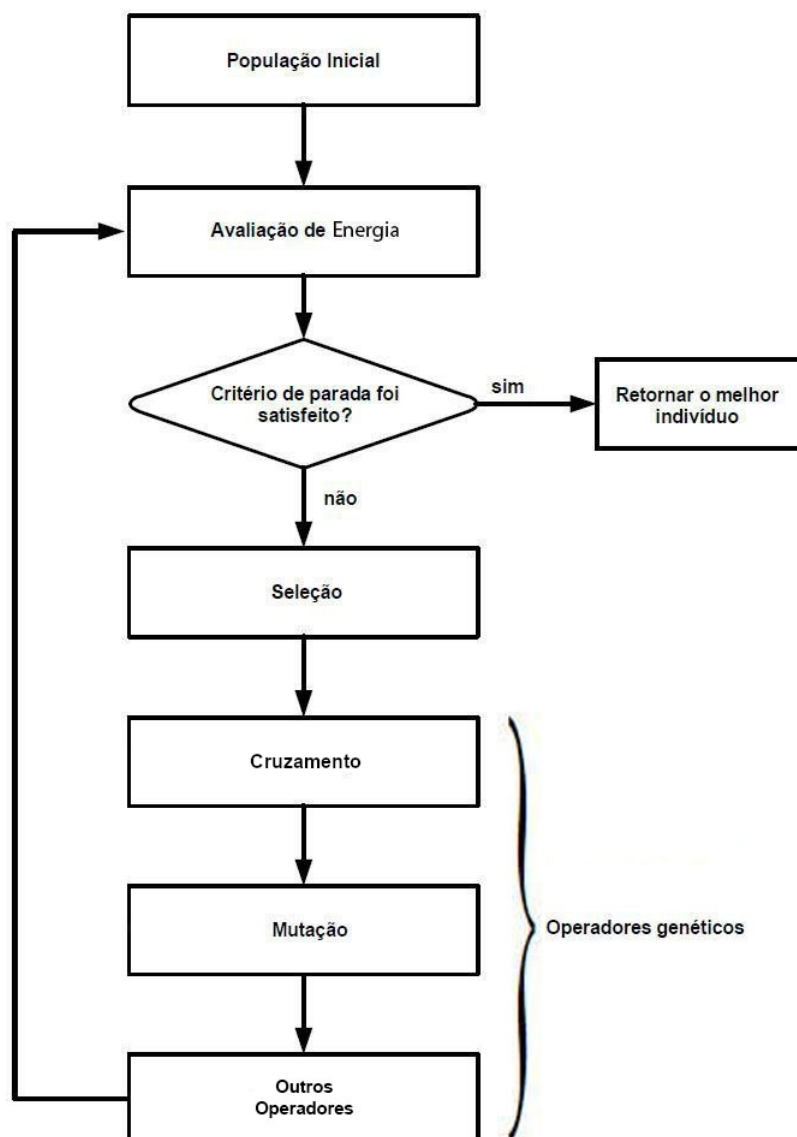


Figura 2.1: Fluxograma de funcionamento de um algoritmo genético tradicional.

algoritmo. No estudo de clusters químicos, um critério de parada comum é quando o indivíduo mais bem adaptado (ou seja, o cluster com menor energia) permanece como o melhor indivíduo por um número pré-definido de iterações. Essa abordagem ajuda a evitar a execução desnecessária de iterações adicionais quando a melhoria na solução é inexistente ou mínima(71). Devido à natureza heurística do GA e à falta de um parâmetro convergente, juntamente com o fato de que as avaliações de energia são pontuais na superfície de energia potencial (SEP), surge um problema interessante: a dificuldade em determinar o mínimo exato dos vales na SEP. Para contornar esse problema, foi proposta uma arquitetura híbrida de algoritmo genético, que consiste na combinação do GA para busca global na SEP e de um algoritmo convergente, como o BFGS, para realizar minimizações locais, aumentando consideravelmente a eficiência dos algoritmos genéticos híbridos em relação aos não híbridos (42).

Destaca-se que todos os operadores genéticos, métodos de avaliação de energia e critérios de parada são detalhados e justificados no **Capítulo 5**. Neste capítulo, será apresentada a metodologia desenvolvida neste trabalho, oferecendo uma visão mais aprofundada sobre como esses elementos são aplicados e configurados para atender aos objetivos específicos dos sistemas estudados.

Capítulo 3

Química computacional

3.1 Introdução

Sistemas físicos com dimensões próximas ou abaixo da escala atômica são tratados por uma teoria física chamada mecânica quântica, originada no início do século XX, quando cientistas como Max Planck e Albert Einstein introduziram a ideia de que a energia é quantizada, ou seja, não é emitida ou absorvida de forma contínua, mas sim em pequenos pacotes discretos chamados de “quanta”. Essa teoria revolucionária mudou completamente a forma como compreendemos o mundo microscópico da matéria(83).

Embora a mecânica quântica seja um ramo fundamental da física, ela também é a base teórica de vários campos da química, e a química quântica é um ramo da química que se baseia nos princípios e nas leis da mecânica quântica para entender e descrever o comportamento dos átomos e das moléculas. A química quântica fornece uma estrutura matemática rigorosa para descrever as propriedades dos sistemas químicos em um nível fundamental. Ela utiliza conceitos como a função de onda, que descreve o comportamento probabilístico das partículas subatômicas, e o princípio da superposição, que afirma que uma partícula pode estar em vários estados simultaneamente. Esses conceitos permitem uma descrição mais precisa e detalhada dos átomos e das moléculas, incluindo a determinação de suas estruturas eletrônicas, as energias dos seus estados eletrônicos e as propriedades químicas resultantes(84). O pilar mais básico da mecânica quântica é a equação de Schrödinger dependente do tempo:

$$\hat{\mathcal{H}}\Psi = i\hbar\frac{\partial\Psi}{\partial t} \quad (3.1)$$

Na equação 3.1, $\hat{\mathcal{H}}$ é um operador responsável por descrever a soma das parcelas de energia cinética e potencial do sistema.

$$\hat{\mathcal{H}} = -\frac{\hbar^2}{2m}\nabla^2 + V \quad (3.2)$$

No formato do hamiltoniano apresentado em 3.2 o primeiro termo está associado à energia cinética eletrônica e o segundo termo à energia potencial. Observa-se que não há um termo específico para descrever a energia cinética do núcleo. Isso ocorre porque a maioria dos hamiltonianos é construída utilizando a aproximação de Bohr-Oppenheimer. Em resumo, essa aproximação considera que, dado que a massa do núcleo é muito superior à dos elétrons, o movimento do núcleo em relação aos elétrons será muito mais lento e, portanto, pode ser desprezado. Em termos mais técnicos, os movimentos podem ser considerados desacoplados(85). Isso quer dizer que para uma dada configuração nuclear tem-se a solução para a estrutura eletrônica do sistema. Se o termo de energia potencial for independente do tempo, a equação 3.1 pode ser simplificada para a equação de Schrödinger independente do tempo.

$$\hat{H}\Psi = E\Psi \quad (3.3)$$

A equação de Schrödinger passa a ser uma equação de autovalor correspondente a energia total do sistema, representado por E na equação 3.3. Para tentarmos extrair uma propriedade de um determinado sistema, é necessário obter a função de onda Ψ do sistema que por sua vez dependerá da determinação do operador Hamiltoniano adequado ao problema. Embora essas equações sejam bastante difundidas e amplamente utilizadas, nota-se que só é possível obter uma solução exata para os casos mais simples, cujo Hamiltoniano pode ser escrito como:

$$\hat{H} = \sum_{j=1}^Z \left\{ -\frac{\hbar}{2m} \nabla_j^2 - \left(\frac{1}{4\pi\epsilon_0} \right) \frac{Ze^2}{r_j} \right\} + \frac{1}{2} \left(\frac{1}{4\pi\epsilon_0} \right) \sum_{j \neq k}^Z \frac{e^2}{|\vec{r}_j - \vec{r}_k|} \quad (3.4)$$

A equação 3.4 se refere à um hamiltoniano para um átomo neutro de número atômico Z , com carga nuclear Ze , rodeado por Z elétrons com carga $-e$. O primeiro termo está relacionado à energia cinética eletrônica somado à energia potencial eletrostática entre elétron e núcleo. O segundo termo da equação trata a relação elétron-elétron, mais especificamente, o potencial de repulsão entre o elétron e seus vizinhos. O termo referente a repulsão eletrônica é o termo que impossibilita a resolução analítica da equação de Schrödinger, por isso a solução exata só é possível para átomos que possuam apenas um elétron, ou como são chamados, átomos hidrogenóides. Uma opção seria assumir que a contribuição da repulsão elétron-elétron seria algo próximo de nula, assim o hamiltoniano sempre poderia ser tratado como o de um átomo hidrogenóide. Porém, para um simples átomo de hélio contendo dois elétrons, desconsiderar o termo de repulsão eletrônica gera um erro de 40% (86). Com isso podemos afirmar que a repulsão eletrônica tem contribuição pertinente e não deve ser ignorada. Se não podemos ignorar o termo de repulsão eletrônica, mas também não é possível resolver a equação com ele, será necessário desenvolver métodos e aproximações que tornem possível resolver a equação de Schrödinger usando hamiltoniano do tipo apresentado na equação 3.4. Desde que esse problema foi descoberto nos primórdios da teórica quântica,

tem-se tentado encontrar soluções para resolver problemas multieletrônicos, e um dos artifícios usado até hoje é o método aproximativo de Hartree-Fock.

3.2 Método de Hartree-Fock

O método Hartree-Fock foi um dos primeiros métodos usado com o objetivo de determinar a estrutura eletrônica de um átomo não hidrogenóide ou de uma molécula. O objetivo do método desenvolvido por Douglas Hartree (87, 88) e Vladimir Fock (89) é resolver a equação de Schrödinger usando um hamiltoniano que considera a repulsão elétron-elétron porém, de forma aproximada, utilizando outro caminho para representar essa interação. A primeira ideia que fundamenta esse método é representada pela equação (3.5), que assume que a função de onda total do sistema é o produto das funções dos orbitais individuais para cada elétron:

$$\Psi(r_1, r_2, r_3, \dots) = \psi_1(r_1)\psi_2(r_2)\psi_3(r_3)\dots \quad (3.5)$$

Esta aproximação orbital define que a função de onda global do sistema é o produto das funções de onda eletrônicas, mas desconsidera a interação elétron-elétron. Para férmions com dois estados de spin (α) e (β), a função de onda total deve levar em consideração tanto a função de onda espacial quanto a função de spin. Os spin-orbitais ou spinors são então introduzidos, combinando as funções espaciais e de spin:

$$\Psi = \frac{1}{\sqrt{N!}} \begin{bmatrix} \chi_1(x_1) & \chi_2(x_1) & \cdots & \chi_N(x_1) \\ \chi_1(x_2) & \chi_2(x_2) & \cdots & \chi_N(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(x_N) & \chi_2(x_N) & \cdots & \chi_N(x_N) \end{bmatrix} \quad (3.6)$$

Onde (x_i) representa as coordenadas espaciais e de spin do i -ésimo elétron, e ($\chi_i(x)$) são os spin-orbitais. O determinante de Slater, representado pela equação (3.6), organiza a matriz determinante de tal forma que cada linha representa um elétron e cada coluna corresponde a um spinor, garantindo automaticamente a antissimetria requerida pelo princípio da exclusão de Pauli. Agora que foi definido como calcular a função de onda global é necessário calcular cada uma das funções de onda mono-eletrônicas. Para resolver esse problema é utilizado o princípio variacional para calcular a função de onda de forma que ela produza o menor valor de energia possível para o sistema. Em suma, minimizar a energia eletrônica como mostrado na equação 3.7. Esta equação utiliza a notação Bra-ket também conhecida como notação de Dirac (90) e será utilizada em alguns casos para simplificar visualmente a equação.

$$E_0 = \langle \psi_0^* | \hat{H} | \psi_0 \rangle \quad (3.7)$$

onde o hamiltoniano eletrônico global do sistema \hat{H} é dado por (em unidades atômicas):

$$h(i) = -\frac{1}{2}\nabla_i^2 - \sum_A \frac{Z_A}{r_{iA}} \quad (3.8)$$

$$v(i, j) = \frac{1}{r_{ij}} \quad (3.9)$$

$$\hat{H} = \sum_i h(i) + \sum_{i < j} v(i, j) \quad (3.10)$$

A equação 3.8 é o operador de um elétron que contém a energia cinética e potencial de atração do elétron i com os núcleos. Para evitar o problema da repulsão eletrônica o artifício matemático utilizado é que cada elétron vai interagir com uma nuvem eletrostática que representa a carga média dos outros elétrons. A equação 3.9 é o operador de dois elétrons, responsável por descrever o potencial médio sentido por um determinado elétron devido à presença de todos os outros elétrons do sistema. Essa é a aproximação mais grosseira do método Hartree-Fock; porém, ela permite a resolução para problemas até então complexos, por envolver muitos elétrons.

$$f(i)\chi(x_i) = \varepsilon_j\chi_j(x_i) \quad (3.11)$$

A minimização de 3.7 que garante o cálculo dos valores de spin-orbitais ótimos, levando à equação 3.11 de Hartree-Fock. A equação de Hartree-Fock é usada para calcular a energia de cada spinor através da aplicação do operador de Fock que vem da 3.10, porém para cada elétron:

$$f(i) = h(i) + v(i, j) \quad (3.12)$$

O operador de Fock 3.12 é utilizado para calcular a energia de cada spinor e a energia total do sistema será dada pelo somatório das contribuições de um e dois elétrons:

$$\langle i|\hat{h}|j\rangle = \int dx_1 \chi_i^*(x_1)\hat{h}(r_1)\chi_j(x_1) \quad (3.13)$$

$$[ij|kl] = \int dx_1 \int dx_2 \chi_i^*(x_1)\chi_j(x_1)\frac{1}{r_{12}}\chi_k^*(x_2)\chi_l(x_2) \quad (3.14)$$

$$E_{HF} = \sum_i^N \langle i|\hat{h}|i\rangle + \sum_{i > j}^N [ii|jj] - [ij|ji] \quad (3.15)$$

A equação 3.13 é usada para calcular a contribuição energética (cinética e potencial atrativa) de um elétron para seu spinor e a equação 3.14 é usada para calcular a energia de repulsão entre um elétron e o "campo" de elétrons, também conhecida como integral de

Coulomb $[ii|jj]$ ou J_{ij} . A equação 3.14 também é usada para calcular a integral de troca $[ij|ji]$, K_{ij} , que não tem um significado na interpretação clássica, pois é resultado apenas dos efeitos quânticos associados à antissimetria eletrônica. Nota-se que na equação 3.14 a aproximação imposta que a energia de repulsão eletrônica é a relação entre um elétron e a média do “campo elétrico” produzido pelos outros elétrons.

O conjunto de equações 3.11 para um determinado sistema será não linear, portanto sua solução é feita por método iterativo nomeado Campo Auto-consistente de Hartree-Fock (HF-SCF). Como na maioria dos métodos iterativos, é necessário fornecer um ponto de partida para o cálculo. Neste caso, devemos escolher uma função de onda arbitrária inicial para descrever os spinors. Usando essa função-tentativa será possível calcular as energias de todos spinors em 3.11 que por sua vez gerará uma nova função de onda melhorada e assim sucessivamente até que a diferença de energia entre a iteração passada e a presente seja pequena o suficiente.

A princípio o HF-SCF não era computacionalmente eficiente de ser resolvido até que Clemes Roothaan (91) propôs que a função de onda deve ser descrita por uma combinação linear de orbitais, ou seja, os spinors otimizados para o sistema serão obtidos através da combinação linear de um conjunto de funções, chamada de funções de base, centrada nos átomos, usualmente chamadas de orbitais atômicos(92) que pode ser expresso como:

$$\chi_i = \sum_{\mu=1}^K C_{\mu i} \tilde{\chi}_{\mu} \quad (3.16)$$

Na equação 3.16, $C_{\mu i}$ é o parâmetro linear da combinação e $\tilde{\chi}_{\mu}$ representa cada spinor. Essa definição de usar funções centradas nos átomos para descrever os orbitais atômicos foi completamente arbitrária, baseada apenas na intuição química, porém, a função de onda ainda é apenas uma função matemática imparcial aos efeitos físicos e químicos a serem estudados (93).

Usando essa nova abordagem algumas modificações serão necessárias. As equações serão trabalhadas na forma matricial uma vez que isso é computacionalmente mais eficiente e facilita a integração desse conjunto de bases à equação de Hartree-Fock. Substituindo 3.16 em 3.11 e simplificando, temos:

$$\begin{aligned} f(x_i) \sum_{\mu} C_{\mu i} \tilde{\chi}_{\mu}(x_i) &= \varepsilon_i \sum_{\mu} C_{\mu i} \tilde{\chi}_{\mu}(x_i) \\ F_{\nu\mu} &= \int dx_i \tilde{\chi}_{\nu}^*(x_i) f(x_i) \tilde{\chi}_{\mu}(x_i) \\ S_{\nu\mu} &= \int dx_i \tilde{\chi}_{\nu}^*(x_i) \tilde{\chi}_{\mu}(x_i) \\ \sum_{\mu} F_{\nu\mu} C_{\mu i} &= \varepsilon_i \sum_{\mu} S_{\nu\mu} C_{\mu i} \end{aligned} \quad (3.17)$$

A simplificação feita em 3.17 foi multiplicar a equação por $\tilde{\chi}_v^*(x_i)$ e fazer a integração. De forma bem direta, podemos transformar 3.17 em sua forma matricial:

$$FC = SC\varepsilon \quad (3.18)$$

denominada equação de Hartree-Fock-Roothaan (85). A equação 3.18 contém a matriz de Fock F onde cada elemento desta matriz é dado pela integral F (equação 3.17) em que o operador de Fock atua em um orbital atômico. A matriz S é chamada de matriz de sobreposição (*overlap*) onde cada elemento é a integral de sobreposição S (equação 3.17) em um orbital atômico. Esse valor será zero quando os orbitais forem ortogonais, pois não haverá sobreposição nenhuma, e assumirá o valor igual a unidade quando houver a sobreposição completa de orbitais, e um valor intermediário quando houver sobreposição parcial de orbitais atômicos. A matriz C é a matriz de coeficientes onde cada coluna será o conjunto de coeficientes que entram na combinação linear para cada spinor. A matriz ε é a matriz de energia, uma matriz diagonal em que cada elemento contém a energia de um spinor e o traço desta matriz é a energia total do sistema.

A equação 3.18 poderia ser tratada como equação de autovalor se a matriz de sobreposição S fosse uma matriz identidade, porém é possível trabalhar uma metodologia de mudança de bases que atenda essa condição. É importante notar que o ciclo HF-SCF ainda atende o princípio variacional, ou seja, a energia da função de onda tentativa sempre terá energia maior ou, no melhor cenário, igual a energia da função de onda real do sistema, princípio esse que, de certa forma, cria uma “garantia” de que a função obtida ao final do HF-SCF será boa. Outra vantagem da equação 3.18 é que para resolver 3.11 é necessário supor uma função de onda tentativa, mas na equação de Hartree-Fock-Roothaan 3.18 é necessário definir apenas a base e os coeficientes da combinação linear para dar início ao ciclo HF-SCF.

As funções de base são funções matemáticas usadas para descrever os orbitais atômicos, quanto maior a base mais ela se aproximará de descrever os orbitais moleculares. Dentre os possíveis modelos, as duas funções mais comuns são:

$$\chi_{sto} = Nr^a e^{-br} Y_{lm_j}(\theta, \phi) \quad (3.19)$$

$$\chi_{gto} = Nx^i y^j z^k e^{-\alpha r^2} \quad (3.20)$$

A função de Slater 3.19 é capaz de descrever a função de onda real de maneira excelente uma vez que ela mantém o mesmo esférico harmônico da função de onda real, exceto pela parte radial da função. O problema desta função está na aplicação computacional. A cúspide produzida nessa função é um desafio até para os computadores modernos o que torna sua utilização mais restrita. Por outro lado a função gaussiana 3.20 é amplamente utilizada; porém, é muito diferente do que seria a função de onda real, por isso utiliza-se uma combinação de gaussianas para gerar uma nova curva simulando as propriedades da

função de Slater se tornando mais adequada para descrever a função de onda do sistema com um custo computacional inferior comparado às funções de Slater.

A princípio, a equação 3.16 é composta de um somatório de infinito orbitais atômicos para melhor descrever os orbitais moleculares, porém, isso é computacionalmente impossível. Portanto, o tamanho das bases é um ponto crítico na precisão e no custo computacional do método. O conceito de base mínima é usado para descrever uma base que possui apenas uma função de base para cada tipo de orbital atômico no sistema estudado. É comum encontrar bases com duas, três ou mais funções de base para cada orbital comumente chamadas de "duplo- ζ ", "triplo- ζ " e assim por diante, respectivamente. A formação de ligação química geralmente envolve orbitais de valência e então foi criado um tipo diferente de base chamada de "split-valence" onde os orbitais mais internos são descritos por apenas uma função de base e os orbitais mais externos são descritos por um número maior de funções de base. Além destas bases apresentadas ainda existem vários outros tipos que incluem funções de polarização, funções difusas, inclusão de efeitos relativísticos, entre outros. Com estas bases bem estabelecidas resta apenas determinar os coeficientes da combinação linear de forma iterativa pelo HF-SCF.

Ao final do HF-SCF será obtido um conjunto spinors, para cada elétron do sistema será atribuído um spinor em ordem crescente de energia. Os spinors que sobraram, chamados de spinors virtuais, não serão incluídos no determinante de Slater quando o objetivo é obter a função de onda do sistema em seu estado fundamental (85), mas deve ser incluído para cálculos de estado excitado.

Embora seja um método repleto de falhas, a maioria dos métodos pós Hartree-Fock fazem referência ou utilizam a base do método Hartree-Fock original, pois representa um ótimo ponto de partida sendo apenas necessário reduzir o número de aproximações tornando o método mais preciso. Estas melhorias foram apresentadas nos métodos pós Hartree Fock como é o caso dos métodos de perturbação de Möller-Plesset (MP) e coupled-cluster (CC) (92), e embora eles sejam considerados mais precisos que o método Hartree Fock eles apresentam um custo computacional elevado.

3.3 Teoria do funcional da densidade (DFT)

A teoria do funcional da densidade é um modelo quântico computacional usado para investigar a estrutura eletrônica em sistemas de muitos corpos, em particular átomos e moléculas, geralmente em estado fundamental. A DFT parte do princípio que um sistema multieletrônico pode ter suas propriedades determinadas fazendo o uso de funcionais (funções de outras funções) dependentes da densidade eletrônica. Assim o nome teoria do funcional da densidade vem do uso de funcionais de densidade eletrônica. Muito popular para cálculos em física do estado sólido, só entrou efetivamente para o cenário da química quântica quando a teoria foi refinada para tratar melhor os efeitos de troca e correlação

eletrônica.

Dentro do contexto da DFT a estrutura eletrônica é avaliada a partir de um potencial aplicado ao sistema eletrônico. Esse potencial é construído a partir da soma de potenciais externos V_{ext} , dependente apenas da estrutura e composição do sistema, e o potencial efetivo V_{eff} que representa a interação inter-eletrônica. Com isso o sistema passa a ser representado como uma grande estrutura contendo n elétrons que podem ser estudados como um conjunto de equações de Schrödinger para sistemas mono-eletrônicos. Essas equações terão um formato bem similar ao método de Hartree-Fock porém, agora, denominadas equações de Kohn-Sham que também serão resolvidas por auto-consistência (94).

Assim como no método Hartree-Fock, a aproximação de Born-Oppenheimer é usada e gera um potencial estático externo V pelo qual os elétrons estão se movendo produzindo um hamiltoniano similar ao tratado em 3.10. A função de onda também será estacionária, ou seja, descrita apenas por um determinante de Slater satisfazendo a equação de Schrödinger independente do tempo 3.3. O termo problema associado a repulsão eletrônica (equação 3.9) ainda continua presente e para contornar isso a DFT utiliza a variável densidade eletrônica $n(\vec{r})$. A densidade eletrônica descreve a probabilidade de encontrar um elétron em uma posição específica no espaço. Esta variável está diretamente relacionada à função de onda (Ψ), que é uma descrição mais completa do sistema quântico, incluindo todas as informações sobre as posições e os estados dos elétrons. A densidade eletrônica é obtida a partir da função de onda ao integrar sobre todas as posições dos elétrons, exceto uma, como mostrado na equação:

$$n(\vec{r}) = \langle \Psi | \delta(\vec{r} - \vec{r}_i) | \Psi \rangle = N \int d^3r_2 \dots \int d^3r_N |\Psi(\vec{r}, \vec{r}_2, \dots, \vec{r}_N)|^2 \quad (3.21)$$

O operador delta ($\delta(\vec{r} - \vec{r}_i)$) na equação 3.21, conhecido como delta de Dirac, atua como uma função que “seleciona” a posição de um elétron específico (\vec{r}_i) na função de onda. Quando aplicado à função de onda, o operador delta integra a função de onda sobre todas as posições possíveis dos outros elétrons, resultando na densidade eletrônica em (\vec{r}). Este processo é essencial para reduzir a complexidade do problema quântico de muitos corpos para uma descrição em termos de uma única variável espacial. A DFT baseia-se na ideia de que a densidade eletrônica ($n_0(\vec{r})$) do estado fundamental determina unicamente a função de onda do estado fundamental ($\Psi_0(\vec{r}_1, \dots, \vec{r}_N)$). Isso significa que, conhecendo a densidade eletrônica, é possível calcular a função de onda correspondente. Hohenberg e Kohn fizeram uma generalização através dos teoremas (95):

1. Dois hamiltonianos distintos, em estados fundamentais não degenerados, não poderão fornecer densidade eletrônica igual, portanto é possível definir a energia do estado fundamental como sendo um funcional da densidade eletrônica $E_0 = E_0[n]$;
2. Partindo do princípio variacional de funções de onda para densidades eletrônicas, $E_0[n]$ será a energia mínima quando $n(\vec{r})$ for a densidade do estado fundamental real

do sistema, dentre todas as densidades eletrônicas acessíveis.

Portanto, a densidade eletrônica do estado fundamental contém todas as informações necessárias para determinar a função de onda do estado fundamental, e essa relação é única e reversível. Isso é fundamental para a DFT, pois permite que propriedades do sistema sejam calculadas diretamente a partir da densidade eletrônica, sem a necessidade de resolver a equação de Schrödinger para a função de onda de muitos corpos, o que é muito mais complexo computacionalmente. Com a DFT, o foco passa a ser a minimização do funcional de energia $E[n]$, que depende exclusivamente da densidade eletrônica. Surge, contudo, um novo desafio: a determinação exata do valor do funcional ($E[n]$) é inviável. Para contornar essa limitação, aproxima-se o funcional ($E[n]$) decompondo-o em componentes distintos: a energia cinética dos elétrons, o potencial eletrostático de atração entre elétrons e núcleos, o potencial eletrostático de repulsão entre elétrons, e a energia de troca e correlação (E_{xc}). A determinação do funcional de troca-correlação é um dos principais desafios na abordagem DFT (96). Os orbitais de Kohn-Sham (KS), que são orbitais virtuais derivados da hipótese de Kohn-Sham(97), são utilizados para construir a densidade eletrônica necessária para calcular o funcional de troca-correlação.

A hipótese de Kohn-Sham é baseada na proposição de um sistema contendo \mathbf{N} elétrons capazes de interagir entre si, sob ação de um potencial externo. Para esse sistema existirá um sistema teórico correspondente, com \mathbf{N} elétrons, desta vez, incapazes de interagir entre si. Mesmo sem interações elétron-elétron, esse sistema teórico deve produzir a mesma densidade eletrônica do sistema real, porém o potencial externo será diferente do potencial real. Para esse sistema teremos o potencial efetivo $\mathbf{V}_{\text{ef}}(\vec{\mathbf{r}})$ que deverá compensar o valor da interação elétron-elétron presente no sistema real. A energia cinética do sistema teórico descrita como um funcional da densidade eletrônica será dada por $\mathbf{T}_{\text{ef}}[\mathbf{n}]$ e a correção aplicada para a interação multieletrônica, $\mathbf{T}_{\text{c}}[\mathbf{n}]$, que na maior parte dos casos poderá ser desprezada com perdas mínimas de acurácia (98). O termo referente ao funcional de Coulomb será descrito, de forma similar, pelo termo de Hartree $\mathbf{E}_{\text{H}}[\mathbf{n}]$ e novamente para corrigir a interação inter-eletrônica será adicionado o termo $\Delta\mathbf{U}[\mathbf{n}]$ (94). Os termos referentes às correções desconhecidas, $\mathbf{T}_{\text{c}}[\mathbf{n}]$ e $\Delta\mathbf{U}[\mathbf{n}]$ serão agrupados em um único termo representando o funcional de troca-correlação eletrônica, $\mathbf{E}_{\text{XC}}[\mathbf{n}]$.

O funcional de troca-correlação, $\mathbf{E}_{\text{XC}}[\mathbf{n}]$, deverá abranger efeitos como o de repulsão entre elétrons de mesmo spin, ou repulsão de Pauli, assim como efeitos de correlação eletrônica, isto é, a propensão de elétrons com spin diferentes se evitarem (94). Esse funcional tem a seguinte expressão:

$$E_{XC}[n] = \int n(\vec{\mathbf{r}})\varepsilon_{XC}^h(n(\vec{\mathbf{r}}))d^3r \quad (3.22)$$

onde $\varepsilon_{XC}(n(\vec{\mathbf{r}}))$ é a energia de troca-correlação por elétron. O funcional de energia proposto por Kohn-Sham é apresentado na forma:

$$E[n] = T_{ef}[n] + E_H[n] + E_{XC}[n] + \int d^3r V_{ext}(\vec{r})n(\vec{r}) \quad (3.23)$$

O último termo apresentado na expressão do funcional de energia, eq. (3.23), $V_{ext}(\vec{r})$, representa o potencial externo aplicado ao sistema. Esse potencial não se limita apenas ao potencial dos núcleos atômicos; ele também descreve a interação entre os elétrons e qualquer campo externo presente. Portanto, enquanto $E_{ne}[n]$ descreve especificamente a interação entre elétrons e núcleos atômicos, e $V_{ext}(\vec{r})$ pode abranger outros potenciais externos que afetam o sistema eletrônico. Como o termo $T_{ef}[n]$ tem que ser análogo ao sistema real, ele deve conter orbitais ψ_i que fornecem a mesma densidade eletrônica $n(\vec{r})$ que o sistema com elétrons interagentes considerando a relação a seguir:

$$n(\vec{r}) = \sum_{i=1}^N |\psi_i(\vec{r})|^2 \quad (3.24)$$

A próxima etapa é minimizar o funcional de energia 3.23 respeitando a imposição de que o número total de partículas será sempre constante

$$N = \int n(\vec{r})d^3r \quad (3.25)$$

temos

$$\begin{aligned} \delta \left\{ E[n] - \mu \left| \int n(\vec{r})d^3r - N \right| \right\} &= 0 \\ \frac{\delta E}{\delta n} - \mu &= \frac{\delta T_{ef}}{\delta n} + \frac{\delta E_H}{\delta n} + \frac{\delta E_{ne}}{\delta n} + \frac{\delta E_{XC}}{\delta n} - \mu = 0 \\ \frac{\delta T_{ef}}{\delta n} + V_H(\vec{r}) + V_{ne}(\vec{r}) + V_{XC}(\vec{r}) - \mu &= 0 \end{aligned} \quad (3.26)$$

onde o potencial de Hartree é

$$V_H(\vec{r}) = \int \frac{n(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r' \quad (3.27)$$

e o potencial de troca-correlação dado por

$$V_{XC}(\vec{r}) = \frac{\delta E_{XC}}{\delta n} \quad (3.28)$$

O potencial efetivo $V_{ef}(\vec{r})$ ao qual os elétrons do sistema teórico estarão sujeitos é definido pela soma dos três potenciais em 3.26

$$V_{ef}(\vec{r}) = V_H(\vec{r}) + V_{ne}(\vec{r}) + V_{XC}(\vec{r}) \quad (3.29)$$

Com isso podemos descrever a equação de Schrödinger que fornecerá os orbitais ψ_i que

reproduzem a densidade 3.24, dada por:

$$\left(-\frac{1}{2}\nabla^2 + V_{ef}\right)\psi_i(\vec{r}) = \varepsilon_i\psi_i(\vec{r}) \quad (3.30)$$

De maneira similar à equação de Hartree-Fock, as equações de Kohn-Sham 3.24 e 3.30 também devem ser resolvidas de forma auto-consistente, pois $\mathbf{V}_{ef}(\vec{r})$ depende diretamente da densidade eletrônica $\mathbf{n}(\vec{r})$, que será obtida através dos orbitais de Kohn-Sham ψ_i . O processo de resolução se inicia com um potencial efetivo ou densidade inicial, usa-se 3.30 para calcular as novas funções de onda, que serão usadas para calcular uma nova densidade usando 3.24. O processo se repete de forma iterativa até que a diferença entre a densidade inicial e final seja igual ou inferior a um limite estipulado previamente. Em geral esse limite fica entre 10^{-4} a 10^{-6} unidades de Hartree, mas pode variar dependendo do software e das configurações específicas utilizadas. É importante notar que os autovalores de energia ε_i obtidos em 3.30 não são valores referentes ao sistema real, mas sim, ao sistema teórico introduzido para reproduzir a densidade eletrônica do sistema real, o que não impõe que suas funções de onda e autovalores sejam iguais. Para a implementação computacional do método DFT, o procedimento é similar ao método Hartree-Fock, sendo que na DFT os orbitais expandidos em função das bases são os orbitais de Kohn-Sham usando a equação 3.16, assim as equações abordadas nesta seção podem ser expressas na forma matricial e resolvidas, dentro do contexto de álgebra linear, por técnicas de diagonalização tradicionais (94).

Mesmo depois de todas as melhorias no DFT ainda há dificuldades em usar esta metodologia para descrever sistemas onde há forte caráter de interação intermolecular, transferências de carga e estados de transição. Para tentar resolver esse problema, vários métodos são propostos para melhorar a DFT, seja fazendo alterações nos funcionais (99), ou incluindo termos aditivos (100, 101). A princípio as equações de Kohn-Sham teriam resultado exato, porém o funcional $\mathbf{E}_{XC}[\mathbf{n}]$ é desconhecido, o que torna essas equações aproximadas. Em suma, a exatidão do método DFT depende diretamente da precisão em que o termo $\mathbf{E}_{XC}[\mathbf{n}]$ é descrito e apesar da forma exata ser desconhecida, muito se progrediu, ao longo dos anos, usando cada vez melhores aproximações para este funcional. Essas aproximações para o funcional de troca-correlação podem ser local ou não-local, sendo a primeira, a mais simples ditando que há uma relação linear entre o funcional e a densidade eletrônica seguindo o formato da expressão 3.22. Este tipo de aproximação é denominado *Local Density Approximation* (LDA) onde a energia de troca-correlação por elétron, ε_{XC}^h , de um gás homogêneo com densidade eletrônica $n(\vec{r})$ será igual à do sistema a ser estudado. Como esta é uma aproximação muito simplística, é de se esperar que ela seja melhor utilizada para descrever sistemas homogêneos, particularmente metálicos, pois estes apresentam uma densidade aproximadamente uniforme. A aproximação LDA falha muito quando o sistema é não homogêneo, por isso outra aproximação foi desenvolvida

intitulada *Generalized Gradient Approximation* (GGA) em que o funcional $E_{XC}^{GGA}[n]$ é expresso em termos da densidade eletrônica e de seu gradiente, dado por:

$$E_{XC}^{GGA}[n] = \int f(n(\vec{r}), \nabla n(\vec{r})) d^3r \quad (3.31)$$

Para os funcionais GGA existem diferentes parametrizações para o termo $f(n(\vec{r}), \nabla n(\vec{r}))$ e isso origina funcionais diferentes dentre os funcionais GGA, um dos mais conhecidos desenvolvido por Perdew, Burke e Ernzerhof, PBE (102). Além destas aproximações existem outras como: funcionais meta-GGA que utilizam a derivada de segunda ordem da densidade $\nabla^2 n(\vec{r})$; graus de liberdade adicionais e/ou densidade da energia cinética e funcionais híbridos que utilizam o termo de troca exato extraído do método Hartree-Fock com funcionais de troca-correlação GGA, tal como o caso do funcional B3LYP(103).

3.4 Potenciais interatômicos

Potenciais interatômicos químicos são modelos matemáticos utilizados para descrever as interações dos átomos em um determinado sistema. Utilizando estes modelos é possível calcular diversas propriedades de um material ou sistema como propriedades elásticas, térmicas ou a estrutura eletrônica. Podemos separar os potenciais atômicos em três grupos distintos, os potenciais interatômicos empíricos, os semi-empíricos e os *ab initio*.

Os potenciais *ab initio* são derivados da mecânica quântica como apresentado nas seções 3.2 e 3.3. Eles produzem os resultados mais precisos no estudo de estrutura eletrônica e propriedades para a maioria dos sistemas. Esta precisão vem com um alto custo computacional, portanto utiliza-se potenciais empíricos e semi-empíricos para estudar sistemas onde potenciais *ab initio* tem custo computacional proibitivos, principalmente devido ao tamanho. Potenciais empíricos são aproximações da superfície de energia potencial para um sistema físico. Estes potenciais são representados por equações em forma genérica cujo parâmetros são ajustados para corresponder a um determinado sistema e para ser facilmente calculável. Um exemplo desses potenciais empíricos facilmente calculável é o potencial de Lennard-Jones(104–106) dado pela seguinte equação:

$$V_{LJ}(r) = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (3.32)$$

onde $V_{LJ}(r)$ é a energia potencial entre um par de partículas separados pela distância r , ε é a profundidade do potencial e σ é a distância em que a energia potencial de interação partícula-partícula é 0. Este potencial modela as duas interações mais importantes: O termo repulsivo ($1/r^{12}$) descreve bem a repulsão do princípio de exclusão de Pauli para partículas interagindo em curta distância devido a sobreposição de orbitais eletrônicos. O termo atrativo ($1/r^6$) descreve a atração das partículas por interações de longa distância

(forças de dispersão de London) que tende a zero à medida que as partículas se afastam para o infinito. Devido a sua simplicidade matemática o potencial de Lennard-Jones tem sido extensamente utilizado desde os primórdios da computação(107–109) até a atualidade(110,111). Este potencial também é utilizado como base para construção de modelos moleculares (campos de força) para sistemas complexos(112,113).

O potencial de Lennard-Jones é baseado na interação de pares de partículas desconsiderando qualquer interação extra do meio, portanto este é um potencial de dois corpos. Um outro potencial muito utilizado é o potencial de Gupta(114–116), que foi modelado para reproduzir o comportamento da estrutura cristalina metálica no estado *bulk*(115). Este potencial é considerado um potencial semi empírico de muitos corpos baseado na teoria *tight binding*(115) considerando o momento de segunda ordem que, por sua vez, é baseado em métodos probabilísticos para a determinação dos momentos estocásticos de funções de várias variáveis(117,118). Neste modelo, as interações de longa distancia são desconsideradas, e a energia do sistema depende de parâmetros ajustados a partir de valores experimentais. O potencial de Gupta é escrito como:

$$V_{cluster} = \sum_i^N (V^r(i) - V^m(i)) \quad (3.33)$$

$$V^m(i) = \left[\sum_j^N \zeta^2 \exp \left(-2q \left(\frac{r_{ij}}{r_0} - 1 \right) \right) \right]^{1/2} \quad (3.34)$$

$$V^r(i) = \sum_j^N A \exp \left(-p \left(\frac{r_{ij}}{r_0} - 1 \right) \right) \quad (3.35)$$

A equação 3.33 representa a energia total do cluster $V_{cluster}$ e depende da energia de coesão do cluster V_i^m dada pela equação 3.34, e pela interação repulsiva de Born-Mayer, V_i^r , descrita pela equação 3.35. Os parâmetros A, r_0, p, q, ζ são determinados por ajuste de dados experimentais ao modelo, e r_{ij} é a distancia entre os átomos i e j .

Tabela 3.1: Parâmetros para o potencial de Gupta(119–121).

Parâmetros	A (eV)	ζ (eV)	p	q	r_0 (Å)
Ag-Ag	0.1031	1.1895	10.850	3.1800	2.8921
Au-Ag	0.1488	1.4874	10.494	3.6070	2.8885
Au-Au (em Au-Ag)	0.2061	1.790	10.229	4.0360	2.8840
Cu-Cu	0.0855	1.2240	10.960	2.2780	2.5560
Au-Cu	0.1539	1.5605	11.050	3.0475	2.5560
Au-Au (em Au-Cu)	0.2096	1.8153	10.139	4.033	2.8850
Ag-Pt	0.1750	1.7900	10.730	3.570	2.8330
Pd-Au	0.1900	1.7500	10.540	3.890	2.8160
Pd-Pt	0.2300	2.2000	10.740	3.870	2.7600
Pd-Pd	0.1746	1.7180	10.867	3.742	2.7485
Pt-Pt	0.2975	2.695	10.612	4.004	2.7747

A tabela 3.1 contém parâmetros para diversas interações metálicas e nota-se que para alguns casos, por exemplo, Au-Au, tem mais de um conjunto de parâmetros dependendo do tipo de metal que ele está interagindo. Esta é uma das desvantagens de utilizar potenciais empíricos e semi empíricos, pois eles estão sujeitos representar as características específicas do sistema para qual foi parametrizado, e, em específico para o Gupta, características do material *bulk*. Além do potencial de Gupta e Lennard-Jones existem diversos outros potenciais como Morse(122) e Finnis-Sinclair(15) que podem ser considerados opções adequadas dependendo do sistema a ser estudado. Mesmo que a estrutura e propriedades não sejam transferíveis para um tratamento *ab initio*, estes potenciais são capazes de produzir, muitas vezes, um ponto de partida adequado para um posterior tratamento *ab initio*(120,123,124).

Capítulo 4

Aprendizado de máquina

4.1 Introdução

Aprendizado de máquina ou *Machine learning*(ML) é uma área da inteligência artificial que se concentra em desenvolver algoritmos e técnicas para que computadores possam aprender sem serem explicitamente programados, utilizando-se apenas informações do problema ou dados. Isso permite que os computadores possam realizar tarefas como reconhecimento de padrões, tomada de decisão e predição de resultados, sem a necessidade de instruções explícitas.

Em 1943, McCulloch e Pitts(125) introduziram o conceito de neurônio artificial, considerado precursor das redes neurais modernas. A contribuição mais significativa deste trabalho foi a proposta de que as redes de neurônios poderiam ser representadas como circuitos lógicos digitais, estabelecendo assim uma conexão entre a computação e a atividade neural. Eles foram os primeiros a aplicar conceitos de lógica e computação para entender o funcionamento do cérebro, o que levou ao desenvolvimento da primeira teoria computacional da cognição(125). O conceito de algoritmos capazes de aprender surgiu em um livro publicado por Donald Hebb em 1949, onde ele apresentava algumas teorias de algoritmos com comportamentos similares ao funcionamento do neurônio humano(126).

Por volta de 1950, Arthur Lee Samuel(127) começou a usar esse conceito para criar um programa de computador capaz de jogar damas, que possuía mecanismos para melhorar seu desempenho ao longo das partidas, e em 1952 ele começou a usar o termo auto-aprendizagem em uma das primeiras utilizações do conceito de inteligência artificial (IA). Em 1957, Frank Rosenblatt propôs o perceptron, um modelo matemático deliberadamente feito para simular o funcionamento do neurônio humano, e durante a década de 1960, esse conhecimento foi expandido e surgiram novas propostas utilizando conjuntos de *perceptrons*, as vezes em várias camadas, em busca de atacar problemas cada vez mais complexos. Durante este período, *machine learning* poderia ser considerado sinônimo de redes neurais artificiais, pois era o algoritmo mais explorado(128–130).

No entanto, o final da década de 1970 e o início da década de 1980 marcaram um ponto de inflexão significativo. Os pesquisadores começaram a explorar novas abordagens que se concentravam na descoberta de padrões e na aplicação de técnicas de programação dinâmica, inaugurando assim o campo do aprendizado por reforço(131). Esta abordagem se destacou por sua capacidade de aprender comportamentos através da interação com um ambiente, refletindo um paradigma distinto dentro do *machine learning* que complementava as abordagens baseadas em redes neurais(132).

Na década de 1980, a inteligência artificial se expandiu para além das redes neurais artificiais, com um foco crescente em modelos lógicos e estatísticos em vez de algoritmos específicos(133). No entanto, o poder computacional limitado da época dificultou a implementação prática dessas técnicas, o que resultou em um declínio temporário no interesse por *machine learning* e redes neurais(134).

Na década de 1990 e 2000, vários fatores contribuíram para o ressurgimento do interesse em *machine learning*. Houve maior acessibilidade a computadores e um crescimento exponencial na capacidade de processamento. Além disso, surgiram novas arquiteturas de redes neurais e técnicas de aprendizado que não dependiam exclusivamente de redes neurais. Principalmente, houve um aumento na demanda por algoritmos capazes de processar e generalizar resultados a partir de grandes bancos de dados. Esses foram os precursores do reaquecimento e subsequente explosão no campo de pesquisa em *machine learning*(133,134). Hoje, *machine learning* faz parte do dia a dia de todos, estando presente nos mecanismos de busca moderno, reconhecimento de voz, tradução de idiomas, detecção de fraudes, diagnóstico de patologias entre vários outros(135).

Existem três tipos principais de aprendizado de máquina: aprendizado supervisionado, aprendizado não-supervisionado e aprendizado por reforço.

- Aprendizado supervisionado é quando os algoritmos de *machine learning* são treinados com dados rotulados, ou seja, com exemplos de entrada e saída esperada. Isso permite que o algoritmo aprenda a fazer previsões precisas a partir de novos dados. Exemplos comuns de tarefas de aprendizado supervisionado incluem classificação de imagens, detecção de spam e previsão de preços de ações;
- Aprendizado não-supervisionado é quando os algoritmos de *machine learning* são treinados com dados não rotulados, ou seja, sem exemplos de saída esperada. Isso permite que o algoritmo descubra estruturas e padrões nos dados. Exemplos comuns de tarefas de aprendizado não-supervisionado incluem agrupamento de dados, redução de dimensionalidade e detecção de anomalias;
- Aprendizado por reforço é quando os algoritmos de *machine learning* são treinados com uma recompensa ou punição, permitindo que o algoritmo aprenda a tomar decisões e realizar ações que maximizam essa recompensa. Exemplos comuns de

tarefas de aprendizado por reforço incluem jogos de computador, robótica e controle automático;

4.2 Aplicação na química

A química computacional é uma área interdisciplinar que combina conceitos da química e da computação para simular e prever propriedades físicas e químicas de clusters, moléculas e sistemas químicos. Nos últimos anos, o uso de técnicas de *machine learning* nesse campo tem crescido rapidamente, permitindo aos cientistas explorar novas áreas de pesquisa e resolver problemas complexos(136–141).

Uma aplicação específica do *machine learning* na química computacional é a otimização de processos químicos. Isso envolve o uso de algoritmos de aprendizado por reforço para controlar variáveis como, por exemplo, temperatura e pressão, visando maximizar a eficiência e o rendimento dos processos(61, 142).

Além disso, técnicas de aprendizado não supervisionado, como agrupamento e redução de dimensionalidade, são utilizadas para explorar e compreender grandes conjuntos de dados químicos, como dados de espectroscopia ou dinâmica molecular. Essas técnicas ajudam os pesquisadores a identificar novos padrões e relações, fundamentais para a descoberta de novos princípios e leis químicas(143, 144).

Uma das principais aplicações do *machine learning* na química computacional é a previsão de propriedades químicas, como energia, estrutura e estabilidade de moléculas. Isso é possível através do uso de algoritmos de aprendizado supervisionado, como redes neurais, que são treinados com dados de propriedades químicas conhecidas de moléculas específicas(136). Esses algoritmos podem então ser usados para estudar e modelar superfícies de energia potencial em diversos sistemas.

O uso de métodos de *machine learning* para modelar a Superfície de Energia Potencial (SEP) tem sido amplamente adotado para estimar propriedades de compostos ou estruturas desconhecidas devido à habilidade desses algoritmos em aprender padrões sem a necessidade de complementar o modelo com conhecimento químico específico. Por exemplo, não é necessário fornecer ou assumir informações como tipos de ligações ou interações químicas presentes no sistema.

Para construir a SEP, os dados adequados necessários são energia, força, ou ambos, geralmente obtidos a partir de cálculos *ab initio*. A maioria dos métodos de *machine learning* para construção da SEP baseia-se em redes neurais artificiais ou métodos baseados em *kernel*. Ambos os métodos exploram a linearização de problemas não-lineares, como prever energia a partir de posições nucleares, mapeando um espaço característico do sistema, geralmente multidimensional.

Métodos *kernel-based* utilizam o *kernel-trick*(145–147) para operar em um vetor característico implícito, sem a necessidade de utilizar explicitamente as coordenadas atômicas.

Já os métodos de *machine learning* baseados em redes neurais artificiais dependem de camadas de neurônios para mapear o espaço multidimensional característico, seguidas por transformações lineares usando parâmetros de aprendizado e uma função de ativação para quebrar a linearidade. Geralmente, múltiplas camadas sucessivas são utilizadas para reproduzir características de sistemas mais complexos.

4.3 Modelo *kernel-based*

Considere um conjunto de dados $(y_i; x_i)_{i=1}^N$ com N observações $y_i \in \mathbb{R}$ para uma dada entrada $x_i \in \mathbb{R}^D$, onde o conjunto \mathbb{R}^D representa um espaço euclidiano de D dimensões, ou seja, cada ponto neste espaço é descrito por um vetor de D números reais. O modelo *kernel*, visa estimar o valor desconhecido y_* para um determinado x_* . Para uma SEP, y geralmente é a energia total do sistema e \mathbf{x} é a representação da estrutura química (e.g. vetor de coordenadas internas, descritores moleculares como matriz de Coulomb(148), funções de simetria(149), entre outros(150–152)). Podemos assumir que para a relação $y = f(\mathbf{x})$, $f(\mathbf{x})$ pode sempre ser aproximada pela combinação linear(153):

$$f(\mathbf{x}) \approx \tilde{f}(\mathbf{x}) = \sum_{i=1}^N \alpha_i K(\mathbf{x}, \mathbf{x}_i) \quad (4.1)$$

onde α_i são coeficientes e $K(\mathbf{x}, \mathbf{x}')$ é a função *kernel*. A função $K(\mathbf{x}, \mathbf{x}')$ reproduz o *kernel* de um espaço de Hilbert \mathcal{H} , se o produto interno $\langle \phi(\mathbf{x}), \phi(\mathbf{x}') \rangle$ de \mathcal{H} pode ser expressado como $K(\mathbf{x}, \mathbf{x}')$ (154). Sendo que ϕ representa a conversão do conjunto de entrada \mathbb{R}^D para \mathcal{H} . Várias funções podem ser utilizadas como *kernel* sendo um exemplo a função polinomial

$$K(\mathbf{x}, \mathbf{x}') = \langle \mathbf{x}, \mathbf{x}' \rangle^d \quad (4.2)$$

onde $\langle \cdot, \cdot \rangle$ é o produto interno e d é o grau do polinômio. Outra função *kernel* que pode ser utilizada é a gaussiana dada por

$$K(\mathbf{x}, \mathbf{x}') = e^{-\gamma \|\mathbf{x} - \mathbf{x}'\|^2} \quad (4.3)$$

em que o hiper-parâmetro γ determina a amplitude da curva gaussiana e $\|\cdot\|$ representa a normalização L2, uma técnica utilizada para ajustar os valores de um vetor, de forma que a soma dos quadrados de seus elementos seja igual a 1. Também é possível incluir interações de longa distancia diretamente na função *kernel*(155, 156).

Os coeficientes α_i (equação 4.1) podem ser calculados para que $\tilde{f}(\mathbf{x}_i) = y_i$ para todo \mathbf{x}_i no conjunto de dados

$$\alpha = K^{-1}y \quad (4.4)$$

onde $\boldsymbol{\alpha} = [\alpha_1 \cdots \alpha_N]^T$ é o vetor de coeficientes, \mathbf{K} é uma matriz $N \times N$ contendo os dados de entrada (*input*) $K_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$, conhecida como matriz *kernel*(157, 158) e $\mathbf{y} = [y_0 \cdots y_N]^T$ é um vetor contendo as $N - 1$ observações, ou valores para o treinamento da rede referente aos valores de entrada. A matriz *kernel* é simétrica e positiva e portanto é possível utilizar a fatoração de Cholesky(159) para resolver a equação 4.4. Uma vez que os coeficientes $\boldsymbol{\alpha}$ foram determinados é possível estimar os valores desconhecidos y_* a partir dos valores de x_* através da equação 4.1 como sendo $y_* = \tilde{f}(\mathbf{x}_*)$.

Em essência, este método de aprendizado de máquina opera transformando os dados de entrada em um espaço de características de maior dimensão, onde se torna mais fácil encontrar relacionamentos ou padrões lineares. Uma vez que os dados são transformados neste novo espaço, métodos baseados em *kernel* podem realizar operações lineares, como classificação ou regressão, como se estivessem operando em um espaço de características de maior dimensão. Isso é conhecido como “*kernel trick*”, mencionado anteriormente, e a grande vantagem é que, enquanto os cálculos são realizados no espaço transformado, os resultados são implicitamente representados no espaço de entrada original, evitando a necessidade de calcular explicitamente a transformação(160).

O aprendizado de máquina baseado em *kernel* contém vários tipos diferentes de algoritmos, e um dos mais utilizados são as máquinas de Vetores de Suporte (*Support Vector Machines* ou SVMs)(161). As SVMs encontram um hiperplano ótimo que separa diferentes classes no espaço de características, maximizando a margem entre as classes. Embora as SVMs sejam amplamente reconhecidas por sua eficácia em tarefas de classificação, elas têm sido aplicadas com sucesso na química computacional em tarefas de classificação e regressão, como demonstrado em estudos recentes (162, 163).

O *kernel* gaussiano apresentado na equação 4.3 é exemplo popular de uma função de *kernel*, também conhecida como função de base radial (*Radial Basis Function* ou RBF). Ela mede a similaridade entre dois pontos de dados com base em sua distância no espaço de entrada, atribuindo maior similaridade a pontos próximos e menor similaridade a pontos distantes. Na próxima seção será discutido o aprendizado de máquina utilizando redes neurais artificiais que também contém um algoritmo que utiliza RBF, chamado redes neurais RBF, porém vale ressaltar que são dois conceitos diferentes. O *kernel* RBF é uma função *kernel* que será utilizada em um algoritmo como SVM, enquanto uma rede neural RBF utiliza RBF como função de ativação (conceito que será apresentado na seção 4.4).

Métodos baseados em *kernel* oferecem várias vantagens. Eles podem capturar efetivamente relacionamentos não-lineares nos dados sem definir explicitamente transformações complexas, fornecem uma medida de similaridade flexível e adaptável a diferentes tipos de dados, no entanto, a escolha da função de kernel é crucial, pois afeta diretamente o desempenho do modelo. Diferentes *kernel* têm propriedades diferentes, e selecionar o kernel apropriado para um problema específico é muitas vezes uma tarefa que requer conhecimento especializado e experimentação.

4.4 Modelo de redes neurais artificiais

A base para construção de modelos de redes neurais artificiais RNA são as “camadas densa de neurônios”(164–166), que transforma o vetor de entrada $\mathbf{x} \in \mathbb{R}^{n_{in}}$ linear no vetor de saída $\mathbf{y} \in \mathbb{R}^{n_{out}}$ a partir da equação

$$\mathbf{y} = \mathbf{W}\mathbf{x} + \mathbf{b} \quad (4.5)$$

onde os pesos $\mathbf{W} \in \mathbb{R}^{n_{out} \times n_{in}}$ e bias $\mathbf{b} \in \mathbb{R}^{n_{out}}$ são parâmetros do aprendizado, e n_{in} e n_{out} denotam a dimensão da entrada e da saída, respectivamente. Com uma camada de neurônios é possível representar apenas relações lineares entre entrada e saída. Para modelar relações não-lineares, são necessárias pelo menos duas camadas de neurônios (camadas intermediárias à entrada e saída são geralmente chamadas de camadas ocultas) combinado com uma função não linear σ (chamada função de ativação), como exemplificado

$$\mathbf{h} = \sigma(\mathbf{W}_1\mathbf{x} + \mathbf{b}_1) \quad (4.6)$$

$$\mathbf{y} = \sigma(\mathbf{W}_2\mathbf{h} + \mathbf{b}_2) \quad (4.7)$$

Este modelo de múltiplas camadas pode ser utilizado como função aproximadora generalizada, portanto, sendo capaz de correlacionar a entrada \mathbf{x} com a saída \mathbf{y} com precisão, contanto que a dimensionalidade das camadas seja grande o suficiente(167, 168). Com isso fica claro porque RNA são utilizadas para representar SEP, e.g., determinação da energia de uma determinada estrutura química (para este tipo de SEP a saída \mathbf{y} geralmente é unidimensional e contém apenas o valor da energia).

Embora RNA superficiais (apenas uma camada oculta) sejam, em princípio, suficiente para modelar qualquer sistema, na prática, RNA profundas com múltiplas camadas ocultas são exponencialmente mais eficientes(169). Em uma RNA profunda, são utilizadas l camadas ocultas em sequência

$$\begin{aligned} \mathbf{h}_1 &= \sigma(\mathbf{W}_1\mathbf{x} + \mathbf{b}_1) \\ \mathbf{h}_2 &= \sigma(\mathbf{W}_2\mathbf{h}_1 + \mathbf{b}_2) \\ &\vdots \\ \mathbf{h}_l &= \sigma(\mathbf{W}_l\mathbf{h}_{l-1} + \mathbf{b}_l) \\ \mathbf{y} &= \sigma(\mathbf{W}_{l+1}\mathbf{h}_l + \mathbf{b}_{l+1}) \end{aligned} \quad (4.8)$$

que manipulam o vetor de entrada \mathbf{x} , adicionando características cada vez mais complexas, até finalmente, correlacionar linearmente a camada oculta final \mathbf{h}_l com os dados de saída \mathbf{y} . A matriz \mathbf{W}_l e o vetor \mathbf{b}_l , que contém os parâmetros de aprendizado da RNA, são inicializados aleatoriamente e otimizados, por exemplo, via gradiente descendente, para

minimizar uma função custo que mede o desvio entre o resultado obtido pela RNA \mathbf{y}_* e um valor de referencia \mathbf{y} . Esta otimização de hiper-parâmetros é o processo de “aprendizagem” do algoritmo. O erro quadrático médio (RMSE) e erro absoluto médio (MAE) são funções de custo comum utilizadas em RNA. A Figura 4.1 exemplifica uma RNA genérica considerando um vetor de entrada \mathbf{x} contendo os elementos escalares $[x_1, \dots, x_n]$. Esta RNA contém apenas uma camada oculta \mathbf{h}_1 com um número arbitrário m de neurônios e a saída \mathbf{y} .

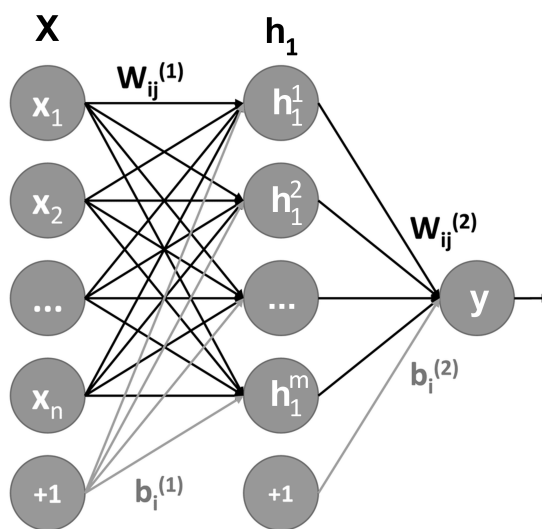


Figura 4.1: Exemplo de rede neural artificial com uma entrada, uma camada oculta e uma saída.

As primeiras SEP baseadas em RNA utilizavam um conjunto de coordenadas internas, e.g., distâncias e ângulos, como vetor de entrada para a RNA(56,170,171), porém, a partir apenas das coordenadas internas, estes parâmetros que definem a entrada do algoritmo podem ser diferentes para clusters iguais, e.g., rotações espaciais dos clusters e permuta de átomos iguais resulta em um conjunto de entrada diferente para representar o mesmo cluster. Outra grande desvantagem de utilizar as coordenadas internas como entrada é que uma RNA treinada para um determinado cluster ou molécula não pode ser utilizada para calcular estruturas de tamanho diferentes, porque requerem uma quantidade diferente de parâmetros para descrever sua geometria. Para contornar este problema em sistemas pequenos foram desenvolvidas RNA inspiradas na expansão de muitos corpos(172–174). Como este método requer uma RNA para cada termo da expansão, este sistema não reproduz com grande precisão sistemas maiores.

Para sistemas maiores, com dez ou mais átomos, uma arquitetura de RNA que tem sido bastante explorada é baseada na decomposição da energia total do sistema em contribuições de energias atômicas, portanto a energia total do sistema é calculada pela combinação linear do resultado de várias RNA (uma RNA por átomo). Esta arquitetura de rede proposta por Behler e Parrinello(149) é conhecida como rede neural atômística multidimensional

(*high dimensional neural network* (HDNN)). Este modelo é baseado na hipótese de que a contribuição energética de cada átomo depende do seu ambiente químico local.

Pode-se observar na Figura 4.2 que as coordenadas cartesianas dos átomos $[R_1, \dots, R_N]$ são utilizadas para gerar as funções de simetria, propostas por Behler e Parrinello(149), para cada átomo da molécula ou cluster $[G_1, \dots, G_N]$. As funções de simetria são utilizadas em RNA independentes do tipo exemplificado na Figura 4.1, com número de camadas e número de neurônios por camada arbitrário para cada sistema.

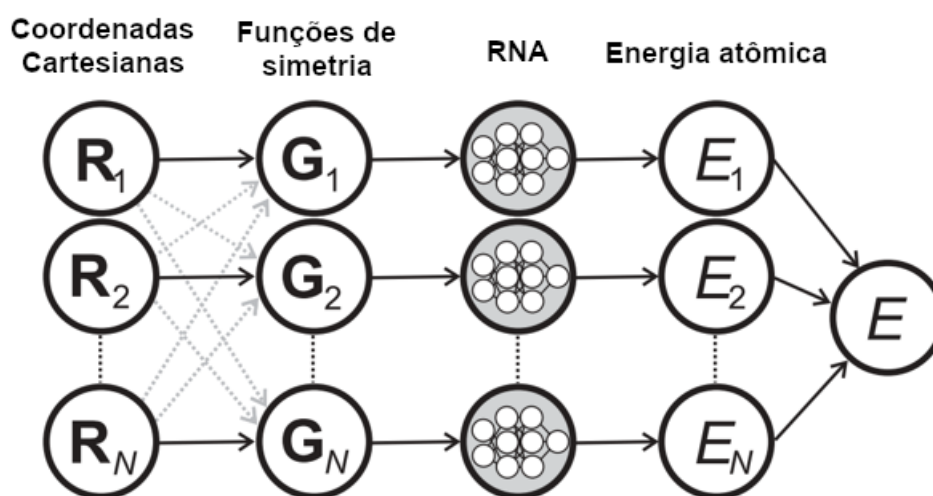


Figura 4.2: Exemplo de rede neural atômica multidimensional artificial(65).

O resultado de cada RNA individual é um valor escalar que representa a energia de cada átomo do sistema, denotado como $[E_1, \dots, E_N]$. Apesar de cada rede neural retornar o valor de energia para seus respectivos átomos, esses valores não possuem significado físico direto e servem apenas como intermediários para calcular a energia total do sistema. A energia total do sistema é determinada pelo somatório dessas energias, conforme mostrado na Equação 4.9, dada por:

$$E = \sum_i E_i \quad (4.9)$$

As funções de simetria mostradas na Figura 4.2 por $[G_1, \dots, G_N]$ representam um exemplo de descritor molecular que será detalhado logo à frente. Descritores moleculares desempenham um papel essencial na modelagem e análise de moléculas em química computacional, oferecendo uma maneira eficiente e eficaz de representar e explorar a estrutura e propriedades das moléculas de forma computacionalmente acessível(175). Esses descritores são utilizados para representar as estruturas químicas de compostos numericamente, permitindo que moléculas complexas sejam descritas matematicamente, o que facilita sua manipulação e comparação computacional. Além disso, são fundamentais

para avaliar a similaridade estrutural entre moléculas em bancos de dados de compostos, identificando análogos ou substitutos estruturais de interesse. Modelos estatísticos e de aprendizado de máquina, como HDNN, frequentemente usam descritores como variáveis independentes para estimar propriedades físico-químicas. Esses descritores são construídos convertendo as posições espaciais atômicas de um sistema em um conjunto de valores que são invariantes em relação à translação, rotação no espaço e permutação de átomos idênticos.

As funções de simetria centradas em átomos (*Atom-Centered Symmetry Functions - ACSF*) constituem um descritor molecular utilizado em conjunto com a arquitetura HDNN proposta. Essas funções são divididas em duas classes: funções radiais, que são funções de dois corpos, e funções angulares, que são funções de três corpos. Os parâmetros destas função são definidos arbitrariamente de forma a gerar um conjunto de valores que identificam unicamente cada sistema. A quantidade necessária de funções radiais e angulares para descrever um sistema varia com seu tamanho, sendo que sistemas maiores requerem um número maior de funções para uma descrição mais precisa (65). As funções de simetria são definidas pelas seguintes equações:

$$G_i^1 = \sum_j f_c(R_{ij}) \quad (4.10)$$

$$f_c(R_{ij}) = \begin{cases} 0.5 \left[\cos\left(\frac{\pi R_{ij}}{R_c} + 1\right) \right] & \text{para } R_{ij} \leq R_c \\ 0 & \text{para } R_{ij} > R_c \end{cases} \quad (4.11)$$

onde $f_c(R_{ij})$ é uma função de corte (Equação 4.11) que determina o grau de influencia, ou peso, que cada átomo vizinho j exerce sobre o átomo central i , e R_{ij} é a distancia entre os átomos i e j . Na Equação 4.11, R_c é o raio de corte, ou seja, todos átomos vizinhos fora desta esfera de raio R_c não terão influência sobre o átomo central. A esfera de influência descrita por f_c será utilizada tanto nas funções radiais quanto angulares. A Equação 4.12 que descreve a simetria radial, é dada por:

$$G_i^2 = \sum_j e^{-\eta(R_{ij}-R_s)^2} f_c(R_{ij}) \quad (4.12)$$

é correspondente a uma curva gaussiana, cujo parâmetro η determina a abertura da curva, que por sua vez, está diretamente correlacionado com crescimento e decaimento da influencia dos átomos vizinhos. R_s é um parâmetro para deslocar o centro da curva. A Equação 4.13 corresponde a uma curva de cosseno amortecido para descrever a influencia dos átomos vizinhos, descrita como:

$$G_i^3 = \sum_j \cos(\kappa R_{ij}) f_c(R_{ij}) \quad (4.13)$$

onde o parâmetro κ controla o período de oscilação. As funções angulares 4.14 e 4.15 são

dadas por:

$$G_i^4 = 2^{1-\zeta} \sum_{j,k \neq i} (1 + \lambda \cos \theta_{ijk})^\zeta e^{-\eta(R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} f_c(R_{ij}) f_c(R_{ik}) f_c(R_{jk}) \quad (4.14)$$

$$G_i^5 = 2^{1-\zeta} \sum_{j,k \neq i} (1 + \lambda \cos \theta_{ijk})^\zeta e^{-\eta(R_{ij}^2 + R_{ik}^2)} f_c(R_{ij}) f_c(R_{ik}) \quad (4.15)$$

$$\theta_{ijk} = \arccos \left(\frac{\mathbf{R}_{ij} \cdot \mathbf{R}_{ik}}{R_{ij} R_{ik}} \right) \quad (4.16)$$

onde o parâmetro ζ controla a resolução da função. Altos valores de ζ produzem menos funções angulares de valor 0. O parâmetro λ pode ter o valor $+1$ e -1 e é utilizado para deslocar o máximo da função em 180° . Similar à Equação 4.12, o parâmetro η controla a parte radial da função. θ_{ijk} (Equação 4.16) representa o ângulo entre os átomos i, j e k centrado no átomo i . A diferença entre as funções G^4 (4.14) e G^5 (4.15) é que G^5 não contém a contribuição R_{jk} , resultando em um número menor de funções angulares nulas em relação a G^4 .

Estas RNA multidimensionais baseadas em descritores moleculares para representar o sistema atômico e usar como entrada em RNA atômicas, foram usadas com bastante sucesso em diversos trabalhos(65, 176–178) e geraram algoritmos baseados neste mesmo princípio, como “*Accurate Neural Network Engine for Molecular Energies (ANAKIN-ME)*”(179) e “*TensorMol*”(180).

Uma outra variante de RNA utilizadas para modelar uma SEP sem utilizar descritor molecular é uma arquitetura proposta por Schütt et al.(181), que usa como entrada as coordenadas atômicas e cargas dos átomos. Nesta arquitetura, os átomos “troquem informações” entre eles através da rede neural para aprender sobre seus respectivos ambientes químicos(182). Esta metodologia foi posteriormente refinada, gerando algoritmos como, por exemplo, “*SchNet*”(183), “*HIP-NN*”(184), e “*PhysNet*”(185).

Capítulo 5

A metodologia *New Quantum Genetic Algorithm*(NQGA)

5.1 Introdução

A construção do algoritmo genético neste trabalho, chamado NQGA (*New Quantum Genetic Algorithm*), envolve desenvolver cada parte do algoritmo de forma modular. Isso significa que cada componente é projetado separadamente, facilitando tanto a escalabilidade quanto a manutenção do algoritmo. Cada módulo é configurável, permitindo ajustar ou substituir os operadores conforme necessário. Isso pode ser feito para o uso em sistemas específicos ou para melhorar o desempenho geral do algoritmo.

Desenvolver os algoritmos significa, neste contexto, construí-los de forma modular e, em princípio, refere-se a dividir o algoritmo genético em partes independentes (ou módulos) que desempenham funções específicas. Cada módulo pode ser desenvolvido, testado e ajustado separadamente. Isso não apenas simplifica a implementação inicial do algoritmo, mas também facilita alterações e melhorias futuras. Além disso, a modularidade permite que diferentes partes do algoritmo sejam reutilizadas em outros contextos ou módulos sem modificar o sistema inteiro, o que promove uma manutenção mais simples e eficiente.

O modelo matemático básico de um algoritmo genético é capaz de minimizar uma função utilizando principalmente operadores que utilizam o sistema binário para modificar e criar novos indivíduos da população. No caso de clusters químicos, ao invés de utilizar apenas números, utiliza-se a geometria espacial dos átomos e operações físicas no espaço tridimensional, e com isso, é possível propor e utilizar operadores regidos por uma intuição física mais apropriada, eliminando certos padrões que poderiam ser possíveis matematicamente, porém fisicamente improváveis.

Mas qual é a justificativa para a necessidade de criar e aperfeiçoar modelos matemáticos complexos para gerar a geometria das estruturas que serão objeto de estudo? À medida que um sistema químico aumenta em tamanho e complexidade, há um aumento não linear

correspondente no número de arranjos espaciais possíveis para este sistema. Em um trabalho realizado em colaboração para estudar as propriedades e estruturas eletrônicas de complexos Eu^{3+} ibuprofanato com ligantes doadores e receptores de elétrons(186), embora o foco do trabalho não seja a geometria dos complexos em si, todos os resultados dependem da acurácia dessas geometrias, pois haverá um impacto direto na estrutura eletrônica do complexo e, portanto, em suas propriedades.

A Figura 5.1 mostra uma representação genérica dos complexos estudados. É fácil notar que se trata de uma estrutura complexa com vários elementos diferentes e diversos graus de liberdade para movimentação de partes da estrutura, tais como o posicionamento dos ligantes e a rotação interna da cadeia radical **R**. Portanto, a escolha da geometria não é trivial e deve ser tratada com importância igual ou superior a outros aspectos do trabalho, como a escolha do método quântico, o escopo das propriedades a serem estudadas, entre outros.

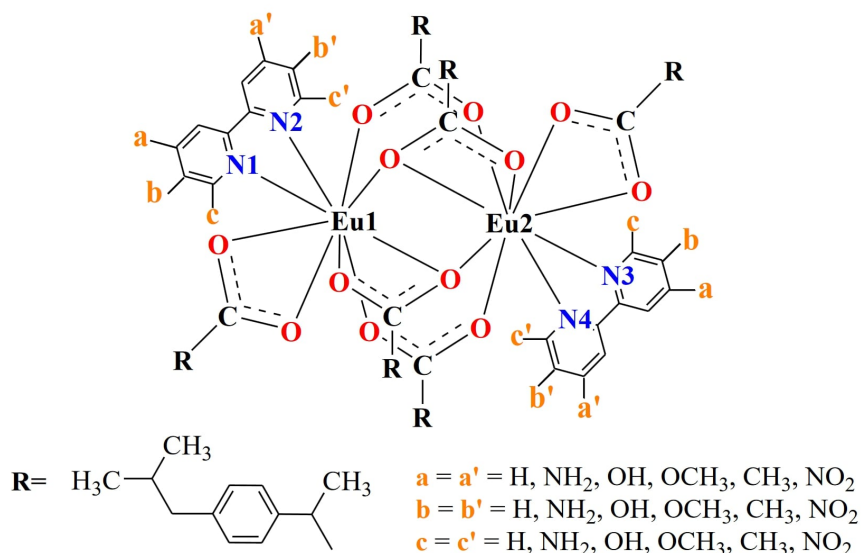


Figura 5.1: Estruturas dos complexos estudados(186).

Uma abordagem comum para propor as estruturas iniciais é realizar uma otimização por partes. Por exemplo, particularmente para o sistema em análise, cada parte da molécula foi separada, isto é, os ligantes, o radical e a parte metálica do complexo. Após a otimização individual, conectamos esses fragmentos usando intuição química. Quando há múltiplas possibilidades, elaboramos várias estruturas quimicamente válidas que são candidatas á estrutura de mínimo e analisamos qual delas possui a menor energia potencial. No entanto, mesmo com experiência e intuição química, há um limite para a capacidade humana de considerar todas as interações possíveis entre esses fragmentos em um espaço tridimensional. Em estudos teóricos sem dados experimentais para comparação, surgem muitas incertezas sobre a validade prática dos resultados obtidos. Eliminar a geometria do sistema dessas incertezas seria um avanço significativo para a química teórica computacional.

Por isso, há um interesse crescente em desenvolver métodos mais precisos para identificar ou gerar geometrias iniciais em estudos de sistemas químicos, similarmente à metodologia apresentada no presente trabalho. Utilizando o NQGA, foi possível simplificar o processo de otimização molecular e reduzir possíveis erros humanos associados à montagem do complexo em três dimensões. O processo de otimização por partes e a combinação das partes otimizadas foram integrados em uma única tarefa de otimização, resultando na obtenção das estruturas de menor energia para os complexos estudados.

A metodologia NQGA utiliza o GA genérico apresentado na Figura 2.1 como base, pode-se alterar o funcionamento dos operadores e adicionar novos operadores com objetivo de desenvolver um GA eficaz para otimização de clusters utilizando método quântico. Considerando apenas o GA, e não o cálculo quântico, há um número grande de variáveis que devem ser cautelosamente escolhidas e estudadas, tais como método de geração da população inicial, taxa de recombinação e taxa de mutação. Estas variáveis podem alterar consideravelmente o resultado e certamente irão alterar o tempo de cálculo total de execução do algoritmo uma vez que estas variáveis controlam indiretamente o número total de indivíduos gerados.

Um dos principais objetivos do algoritmo genético proposto é garantir sua escalabilidade e flexibilidade, permitindo sua aplicação em uma ampla gama de problemas químicos de otimização que envolvem cálculos quânticos, assegurando assim sua relevância contínua no futuro. O NQGA foi desenvolvido principalmente em *C++* devido ao seu desempenho computacional superior a linguagens como *Python* e recursos modernos de programação em comparação com linguagens como *C* e FORTRAN. Nosso algoritmo também utiliza o conceito de objetos na programação. Cada átomo é tratado como um objeto, contendo não apenas um conjunto de coordenadas, mas também número atômico, carga e funções relacionadas para modificar esses parâmetros dentro do algoritmo. Portanto, um átomo é representado por um conjunto de parâmetros e funções programáticas.

Para esclarecer, essas funções são operações de programação e não funções matemáticas, embora, em alguns casos, possam incluir operações matemáticas. Com isso, um cluster pode ser tratado como um objeto composto por vários subconjuntos de átomos. A Figura 5.2 ilustra esse modelo para o tratamento de átomos e clusters. Um átomo é representado por um objeto que contém informações como número atômico, carga e coordenadas, assim como funções para acessar essas informações. Da mesma forma, moléculas e clusters são formados por objetos que contêm vários subobjetos (átomos) e um conjunto de funções computacionais para acessar informações relacionadas.

Utilizando o mesmo conceito da Figura 5.2, então será criado um conjunto de clusters que irá conter subconjuntos de clusters. Este conjunto maior, que em essência é uma lista de clusters, será a população total do GA ao longo das iterações. O número de clusters dentro deste conjunto será o mesmo ao longo das iterações, porém, não será constante durante cada iteração, pois, os operadores genéticos vão gerar e adicionar novos clusters à

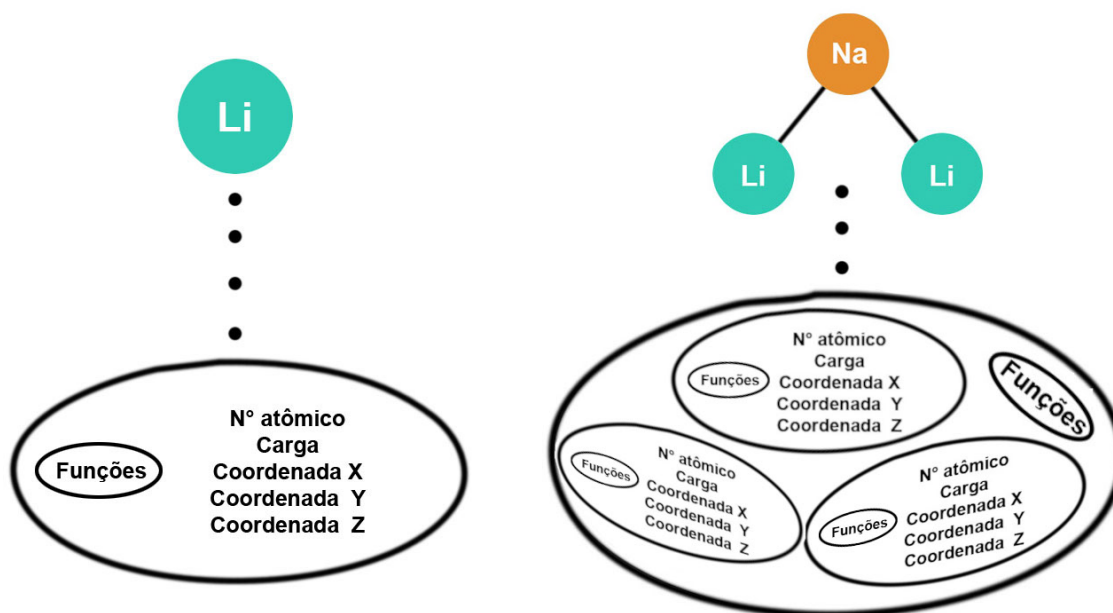


Figura 5.2: Conversão de átomos e clusters em objetos programáveis.

lista, assim como modificar clusters já presentes na população. Ao final de cada iteração, após ordenar a lista por ordem crescente de energia, os clusters de maior energia serão eliminados para manter o número de clusters total da população constante para a próxima iteração.

Partindo do exemplo de GA genérico (Figura 2.1), serão propostas modificações nos principais aspectos do GA, como geração da população inicial, operadores mutação, cruzamento e imigração. O principal aspecto deste projeto de doutorado reside no desenvolvimento de novos operadores, na modificação criteriosa de operadores existente na literatura e na definição da arquitetura de como estes operadores estarão interconectados aprimorar a precisão e eficiência dos cálculos em química quântica computacional. Essas inovações representam um avanço significativo no campo, trazendo uma perspectiva inventiva que desafia e expande os métodos convencionais. A aceitação do artigo científico(187), que expõe e detalha essas contribuições, não apenas valida o impacto prático e teórico deste trabalho, mas também sublinha a relevância do nosso trabalho na expansão das fronteiras do conhecimento científico. Os operadores e a forma como eles interagem são apresentados no seguinte fluxograma de funcionamento do NQGA.

Pode-se observar na Figura 5.3 o processo de funcionamento do algoritmo genético proposto. O primeiro passo é gerar uma população inicial para dar início ao processo iterativo. Após estabelecer uma população inicial, entram em ação operadores para criar a próxima geração de indivíduos. Os termos indivíduos, clusters e moléculas podem ser utilizados de maneira intercambiáveis neste contexto. As três etapas: Cruzamento, mutação e imigração podem ocorrer de forma independente, ou seja, em paralelo, e os novos indivíduos gerados são direcionados para o operador DGP que decidirá se os indivíduos

passarão por uma otimização local utilizando métodos quânticos ou serão excluídos do algoritmo. Após o operador Predador, que atua de forma similar à seleção natural ao final de cada iteração, dois novos operadores (*Machine Learning Predictor* e Simetrizador) poderão atuar para gerar novos indivíduos que serão avaliados na próxima iteração, se houver. Isto é apenas uma visão geral do NQGA e todas as etapas, operadores e condições de atuação serão discutidos com detalhes nas próximas seções.

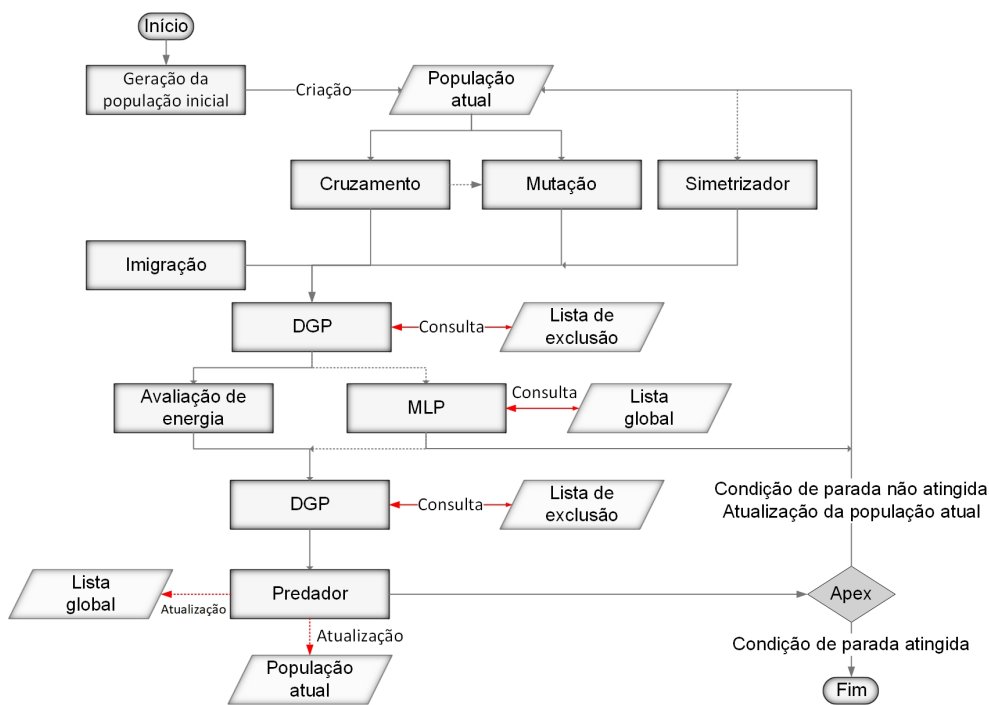


Figura 5.3: Fluxograma de funcionamento do NQGA. Linhas tracejadas representam caminhos opcionais. DGP e MLP são representativos dos operadores predador de sócias (*Doppelgänger Predator*) e previsão por *Machine learning*(*Machine Learning Predictor*), respectivamente.

5.2 Geração da população inicial

A primeira etapa do GA, gerar a população inicial, é tradicionalmente feita de maneira aleatória, e para os métodos de cálculo de energia de baixo custo computacional este método de geração não acrescenta nenhum problema e inclusive é desejável devido a sua natureza aleatória não tendenciosa. Apesar de aleatória, a criação das coordenadas dos átomos será regida por algumas regras para garantir a validade física do indivíduo criado. As coordenadas cartesianas dos átomos são geradas, porém o átomo será considerado válido se a menor distancia entre o novo átomo e os átomos já criado (d_{min}) é maior que um valor pré-estabelecido para evitar uma repulsão excessiva e conseqüentemente falha no cálculo do gradiente durante o processo de otimização local. De maneira similar, d_{min} também deve ser maior que outro valor pré-estabelecido para evitar a formação de clusters

ou átomos desligados. Caso o primeiro conjunto de coordenadas gerado não atenda a essas condições então um novo conjunto de coordenadas é gerado e um novo teste de validade é feito e assim sucessivamente até que todos os átomos tenham sido gerados. Este método resulta em vários agregados de átomos com pouco ou nenhum sentido químico que serão descartados ao longo das operações do GA. A grande desvantagem deste método é que durante a operação de otimização utilizando métodos quânticos, estes clusters sem sentido químico terão alto custo computacional agregado sem benefício direto para a otimização do cluster. Para combater este problema tem-se duas propostas:

- Tornar o algoritmo um processo de duas etapas;
- Propor uma forma alternativa de gerar um conjunto de átomos no espaço.

Transformar o GA em um processo de duas etapas se refere a fazer uma minimização utilizando uma geração aleatória, combinado com um método de baixo custo computacional, como por exemplo, potenciais empíricos e semi-empíricos. Uma nova otimização será feita, agora com o método quântico escolhido e população inicial refinada obtida na primeira otimização, assim como exemplificado na Figura 5.4.

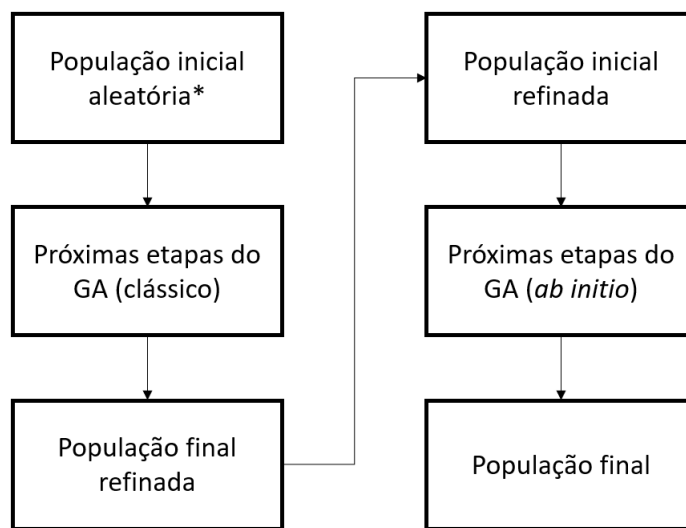


Figura 5.4: Fluxograma de funcionamento simplificado do NQGA utilizando duas etapas, incluindo métodos de avaliação de energia clássicos e quântico.

Embora seja possível desprezar o caminho proposto pela Figura 5.4 e iniciar um GA(quântico) a partir de uma população aleatória, esta alternativa aumenta consideravelmente o tempo necessário para obtenção do resultado devido aos recursos computacionais extras necessários para gerar a população inicial e subseqüentes adições via operador imigração. Em contrapartida, iniciar o GA(quântico) a partir de uma população otimizada por um método de cálculo de energia clássica cria uma tendência na população a formar

clusters extremamente simétricos, que podem ser um mínimo local profundo na superfície de energia potencial quântica, causando problemas na localização do mínimo global. Neste caso, espera-se que os operadores genéticos consigam escapar desta tendência e obter o mínimo global com um número menor de iterações e, conseqüentemente, de visitas à superfície de energia potencial quântica.

A segunda alternativa para criação dos clusters iniciais se refere a criar átomos no espaço de maneira aleatória, porém com maior restrição durante sua criação, com a expectativa de reduzir ou eliminar clusters com estrutura impossíveis, mantendo a natureza aleatória do algoritmo. Neste caso propõe-se o método de geração de átomos por preenchimento de volume esférico. Partindo de um primeiro átomo na origem do espaço cartesiano, o algoritmo de criação segue a seguinte lógica:

1. São criadas duas esferas concêntricas centradas neste átomo com raios r_{min} e r_{max} , respectivamente; onde r_{min} é a distancia mínima aceita para dois átomos e r_{max} uma distancia máxima de separação;
2. Um novo átomo é gerados contido no volume gerado pela subtração dos volumes das esferas, e certifica-se de que este átomo está respeitando as condições de distancia mínima entre os átomos já presentes no cluster. Caso contrário o átomo é recriado até que a condição seja atendida;
3. Repete-se o passo 2 até que a condição de distancia mínima não possa mais ser atendida;
4. Um novo átomo é aleatoriamente selecionado e repete-se a partir do passo 1 até que todos os átomos do cluster sejam criados.

A Figura 5.5 exemplifica este método de criação por volume esférico no espaço bidimensional para facilitar a visualização. Neste caso, devido ao espaço bidimensional, cada átomo é gerado na área entre as circunferências r_{min} e r_{max} até que não haja mais espaço para novos átomos. Um novo átomo é escolhido e o processo continua até que todos os átomos estejam distribuídos.

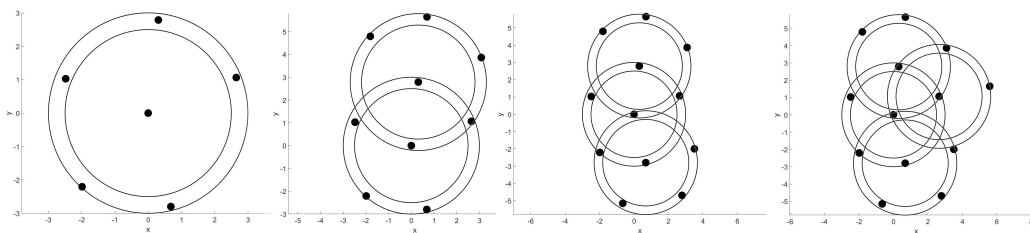


Figura 5.5: Método de geração aleatório por volume esférico representado no espaço 2D.

Os três métodos, geração aleatória, geração em duas etapas e geração aleatória por volume esférico, apresentam vantagens e desvantagens que podem mais pronunciadas,

dependendo do sistema a ser estudado. Como mencionado anteriormente, a geração aleatória será indicada para otimizações com baixo custo computacional. A geração em duas etapas é uma boa alternativa para otimizações utilizando metodologia quântica porém dependendo do sistema pode não haver nenhum potencial empírico ou semi-empírico que descreva o comportamento do sistema impossibilitando o uso desse método de geração em duas etapas. Embora a geração aleatória por volume esférico favoreça a criação de clusters iniciais mais favoráveis, ela pode contribuir negativamente na localização do mínimo em casos onde a geometria preferencial é planar(188). Este problema pode ser ligeiramente mitigado ajustando o método de geração para não saturar os átomos sequencialmente, ou seja, ao invés de criar um átomo e usá-lo como foco para tentar colocar todos os átomos subsequentes ao entorno do primeiro átomo, um novo sorteio será feito a cada átomo criado para decidir o átomo em foco facilitando assim a criação de clusters planares e lineares.

5.3 Operadores tradicionais

Pela Figura 5.3, é possível notar a primeira diferença do NQGA e outros GA's tradicionais: o operador mutação pode atuar tanto na população inicial quanto no resultado do operador cruzamento. Ao atuar em um determinado cluster, ele utiliza apenas a referência deste cluster para aplicar a mutação, caso ela ocorra. Isso implica que caso um indivíduo seja selecionado para mutação, o resultado desta mutação gera um novo cluster e mantém o indivíduo previamente selecionado intacto. Embora esse conceito implique em um número maior de indivíduos produzido pelo operador mutação, isso traz uma flexibilidade maior para o operador, uma vez que ele pode atuar em um número maior de indivíduos e evita que bons resultados de cruzamento sejam destruídos por uma mutação indesejada. Destaca-se que da maneira que o GA foi implementado é possível utilizar este operador no modo duplo (atua em ambas populações), simples (atua apenas na população inicial ou no cruzamento) e também poderá ser completamente omitido.

Dado que vários operadores seguem o mesmo princípio, como os diferentes operadores de mutação ou de cruzamento, podemos agrupar esses operadores em classes. Optamos por usar o termo “classe” em vez de “operador genético” para distinguir que uma classe pode conter dois ou mais operadores genéticos disponíveis para aplicação exclusiva (como na classe de cruzamento) ou inclusiva (como na classe de mutação). Cada uma das classes de operadores será discutida detalhadamente a seguir.

5.3.1 Imigração

A classe imigração é a mais simples e é responsável por empregar de maneira exclusiva um dos seguintes operadores:

- Gerar novos clusters aleatórios;

- Gerar novos clusters semialeatórios;

O objetivo desses operadores é aumentar a diversidade genética da população incluindo novas espécies. Como discutido anteriormente, utilizar clusters aleatórios não será viável quando o GA for utilizado para estudos quânticos, e para isso podemos utilizar a inclusão de clusters semialeatórios, que são clusters localmente otimizados por métodos computacionalmente mais baratos (clássico ou semiempírico) ou gerados pelo mesmo método de geração por volume esférico apresentado na seção de geração da população inicial.

5.3.2 Mutaç o

Um dos operadores mais tradicionais utilizados em algoritmos genéticos é o de muta o. Originou-se nas primeiras propostas de GA focado na otimiza o de fun es matemáticas e foi mantido desde os primeiros GA's focados em otimizar geometria e energia de clusters atômicos (41, 189); por m, apenas um tipo de muta o era utilizado nesta  poca. A classe muta o do algoritmo proposto neste trabalho cont m alguns operadores genéticos de muta o que podem ser utilizados de forma inclusiva:

- Muta o sigma ou deslocamento cartesiano(78);
- Modifica o de interior(79);
- Operador de superf cie angular(80);
- Muta o de permuta ou *flip*;
- Muta o rotacional;

Os operadores muta o sigma, modifica o de interior e superf cie angular foram apresentados e detalhados no Capitulo 2. Com a necessidade de estudar clusters poliatômicos o operador de muta o por permuta ou *flip*, ilustrado na figura 5.6 foi introduzido com objetivo de fazer a permuta de átomos do cluster com objetivo de reduzir o n mero de itera es totais, aumentar a diversidade genética e explorar a superf cie de energia potencial para os diferente hom topos.

Com exce o do operador muta o rotacional todos os operadores muta o mencionados est o dispon veis na literatura. O operador de muta o rotacional est  sendo proposto neste trabalho para contribuir com os outros operadores para melhorar a efici ncia do GA. Como operador muta o sua fun o, assim como os outros,   modificar a estrutura do cluster de maneira a ampliar a variabilidade, ou diversidade, genética da popula o e potencialmente gerar clusters melhores. Este operador   an logo a uma muta o sigma m ltipla e organizada. A partir do centro geométrico do cluster   criado um plano de normal aleat ria e os átomos sofrer o uma rota o arbitraria θ ao longo do eixo normal.

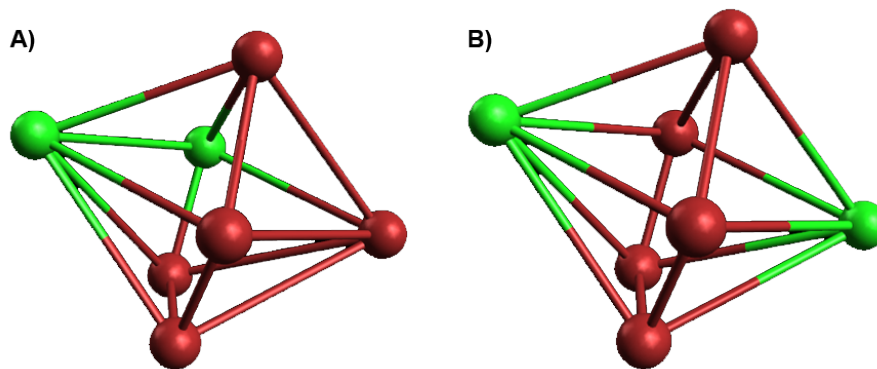


Figura 5.6: Aplicação do operador mutação *flip* no cluster A resultando no cluster B.

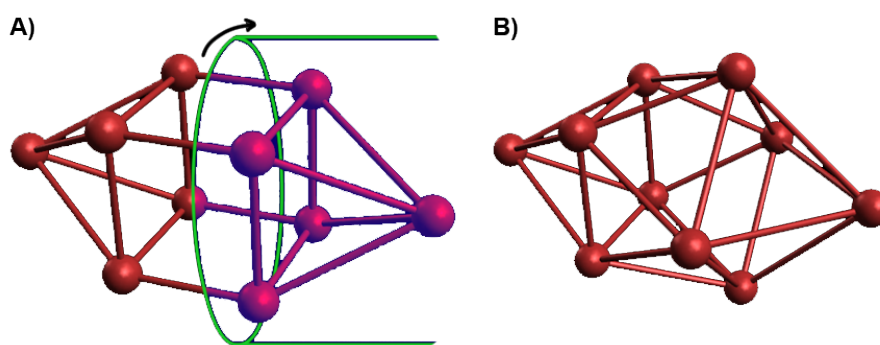


Figura 5.7: Aplicação do operador mutação rotacional no cluster A resultando no cluster B.

É mais fácil visualizar essa mutação como sendo a rotação dos átomos do cluster contidos em um cilindro criado ao longo da normal do plano gerado, como ilustrado na figura 5.7.

A classe de operadores mutação é baseada em material disponível e utilizado na literatura, porém, com modificações para que todos os operadores desta classe sejam passíveis de uso em qualquer tipo de sistema (clusters atômicos, poliatômicos, moleculares, etc.). Estas modificações incluem uma sondagem do cluster resultado após a aplicação do operador para certificar de que nenhuma distância atômica se tornou inferior ou superior aos limites pré estabelecidos, e isso é feito a partir do recálculo das distâncias internas do cluster e da interação com outro operador genético, DGP, que será discutido na seção 5.4.1. O operador *flip* apresenta algumas particularidades extras como, por exemplo, a restrição deste operador trocar apenas partículas de natureza diferente, ou seja, ele não pode trocar átomos iguais de posição. Além disso, há a possibilidade de fazê-lo optar por efetuar trocas que favoreçam com que átomos iguais mantenham proximidade, ao invés da seleção aleatória.

5.3.3 Cruzamento

A seguir estão os operadores disponíveis na classe cruzamento, sendo que ele deve ser escolhido de maneira exclusiva, ou seja, apenas um operador pode ser utilizado por vez durante o processo de otimização no NQGA:

- Cortar e emendar(C&S)(42);
- Cortar e emendar generalizado(genC&S)(74);
- Cortar e emendar modificado(modC&S);

O operador C&S apresentado pela primeira vez no trabalho de Deaven e Ho(42) mostrou uma maneira fisicamente intuitiva de usar partes de dois clusters, juntar e formar dois novos clusters. Inovador na época, porém, define que o plano de corte deve ser horizontal. Portanto, é necessário aplicar uma rotação aleatória nos clusters progenitores para evitar efeitos tendenciosos e reduzir a diversidade genética dos novos clusters gerados. Por este motivo no NQGA foi implementado também uma versão modificada deste operador, o modC&S, que é similar ao C&S, com a diferença de que o plano de corte deve passar por um ponto definido aleatoriamente em proximidade do centro geométrico do cluster e terá um vetor normal também gerado aleatoriamente.

Considerando trabalhos mais recentes, em 2007, o operador genC&S (74) foi proposto como uma opção superior ao C&S. Ele determina que a seleção e troca dos subclusters serão baseadas apenas na distância euclidiana, eliminando assim a necessidade de definir e utilizar um plano de corte para separar os clusters. Ambos os métodos, C&S e genC&S, não possuem diretrizes indicando como proceder no caso de cruzamento de clusters poliatômicos que resultam em um cluster com composição diferente do original. Portanto, nos três operadores, foi implementada uma abordagem inicialmente aleatória para garantir que os clusters formados tenham a mesma composição dos clusters originais. Isso é feito selecionando aleatoriamente átomos de elementos em excesso e convertendo-os em elementos que estão em falta até atingir a composição original.

Em sistemas moleculares, as moléculas são tratadas como pontos no espaço, e o processo ocorre de forma análoga ao sistema atômico. No entanto, como neste caso não é possível considerar o volume molecular utilizando esse modelo de pontos no espaço, após a criação do novo indivíduo, é necessário validar se as distâncias entre os átomos não estão menores do que o limite estabelecido. Caso estejam, rotações aleatórias e pequenos ajustes de distância são feitos para a criação do novo indivíduo.

5.4 Operadores adicionais

As classes de operadores exposta até aqui apresentam operadores comumente utilizados em GA's para otimização de clusters, porém, alguns operadores adicionais são necessários

para combater as fraquezas de um algoritmo genético quando a função a ser minimizada é muito computacionalmente custosa. Uma adição proposta neste trabalho tem o objetivo de fazer com que a população final seja composta de vários mínimos distintos e não uma população composta apenas por um cluster repetido várias vezes, ou seja, evitar a falha do algoritmo em manter a diversidade genética populacional. Durante um trabalho anterior (190) com outro algoritmo genético quântico QGA (191) foi notado que a população final, na maioria dos casos, era composta de n clusters muito semelhantes e de energia próxima. Contornar isso modificando apenas os parâmetros dos operadores genéticos tradicionais aumenta o custo computacional e não garante um bom resultado.

5.4.1 Predador de sócia (DGP)

O predador de sócia (*Doppelgänger* predador - DGP) é a principal ferramenta do NQGA para identificar clusters similares e energeticamente próximos. Esse operador compara cada novo cluster formado com todos os outros clusters gerados anteriormente. Se eles forem considerados iguais, dentro de uma margem de tolerância, o cluster será descartado sem a necessidade de gastar recursos computacionais aplicando outros operadores ou calculando sua energia. Essa abordagem reduz significativamente o número de visitas à superfície de energia potencial quântica, tornando o GA mais rápido e eficiente.

Comparar a estrutura de dois clusters pode parecer trivial, mas é desafiador quando se trata de clusters médios e grandes. Operações como translação, rotação e permutação de átomos idênticos podem fazer com que clusters aparentemente diferentes sejam considerados iguais. O número de isômeros e as diversas localizações no espaço tornam a classificação de clusters como iguais ou diferentes um tópico amplamente discutido na literatura, especialmente em aplicações como *machine learning* aplicado a sistemas químicos.

Embora seja um problema antigo, muitos algoritmos genéticos, como o QGA (191), tentam resolver este problema eliminando clusters cuja energia seja próxima à dos clusters da população atual. Algoritmos genéticos mais modernos adotam comparações geométricas em vez de energéticas(192). Embora o método indireto de comparação baseado na energia funcione, ele exclui a possibilidade de encontrar estruturas geométricas diferentes com energias próximas, o que poderia incluir o mínimo global real. Além disso, não existe uma lógica universal para definir o valor de energia de corte, e o sucesso do algoritmo depende desse valor, tornando essa abordagem menos robusta do que a comparação geométrica de clusters.

A nova classe DGP inclui três operadores de uso inclusivo:

- Distância interatômica simples;
- Funções de simetria centrada em átomo;
- USR - *Ultrafast Shape Recognition*.(193)

O DGP utilizando distância interatômica simples é computacionalmente mais eficiente, uma vez que os valores de distância interna já foram calculados previamente no algoritmo durante a etapa de criação dos clusters. No entanto, esse operador (DGP), é capaz de descrever com precisão apenas clusters pequenos com até 17 átomos(136,175).

Este operador funciona de maneira simples: primeiramente, constrói-se um vetor x^a contendo todas as distâncias internas do cluster que queremos comparar, organizadas em ordem crescente. O vetor x^a é subtraído dos outros vetores $x^{b..z}$ para cada cluster com o qual ele será comparado, gerando assim o vetor diferença y^a . O valor absoluto de cada elemento de y^a é comparado com o parâmetro predefinido γ , e se todos os elementos de $y^a \leq \gamma$, o cluster é considerado igual e descartado para a lista de exclusão. Essa lista será utilizada apenas para popular a lista global, e esse procedimento será detalhado na seção 5.4.2.

Como esperado, esse tipo de operador funciona apenas para clusters pequenos. Quanto maior o número de átomos, mais combinações possíveis de distâncias resultam em um mesmo valor, fazendo com que clusters diferentes sejam considerados iguais e eliminando vários clusters viáveis por efeito falso-positivo.

Para resolver esse problema, o outro operador DGP utiliza informações obtidas a partir da aplicação das funções de simetria centrada em átomos propostas por Behler e Parrinello (65,149). O objetivo é classificar clusters com base em interações de dois e três corpos, conforme detalhado na seção 4.4. A construção do vetor x^a é similar à descrita anteriormente, com a única diferença sendo que o vetor é populado pelo resultado das Equações 4.12 e 4.14. Espera-se que, desde que os parâmetros escolhidos ao aplicar as funções de simetria sejam coerentes, o resultado dessa comparação seja bom o suficiente para classificar qualquer cluster gerado.

Os valores calculados referentes às funções de simetria serão utilizados no operador MLP, que será discutido mais adiante. O fator γ determina o nível de aceitação com base na similaridade dos clusters. Quanto maior o valor de γ , mais clusters serão rejeitados, e quanto menor o valor de γ , mais clusters similares serão aceitos. Portanto, em alguns casos, é interessante variar o valor de γ ao longo das iterações do NQGA para restringir ou aumentar o espaço de busca do algoritmo.

Uma outra forma de se comparar clusters é utilizando o método USR (*Ultrafast Shape Recognition*) (193). Assim como as funções de simetria centradas em átomos o USR também é um descritor molecular; porém, USR parte do princípio de que as distâncias interatômicas são resultado direto da interação entre os átomos da molécula, e portanto, possuem informação o suficiente para descrever a molécula de forma única. O USR codifica a molécula utilizando os três momentos de uma distribuição estatística de distâncias atômicas para 4 locais moleculares resultando em 12 descritores válidos para qualquer número de átomos. São calculados os três momentos (μ_1, μ_2, μ_3) para o centro molecular(ctd), o átomo mais próximo do centro molecular(cst), o átomo mais longe do centro molecular(fct)

e para o átomo mais longe do $\text{fct}(\text{ftf})$. Ao calcular estes 12 descritores, temos um vetor para molécula, por exemplo, A \vec{M}^A . Podemos comparar duas moléculas A e B utilizando a seguinte equação:

$$S_q = \frac{1}{1 + \frac{1}{12} \sum_{i=1}^{12} |M_i^A - M_i^B|} \in (0, 1] \quad (5.1)$$

onde S_q é o fator de semelhança entre as moléculas A e B que pode assumir valores entre 0 e 1, onde 1 significa que as moléculas são idênticas. A princípio o método USR pode parecer superior às funções simetria. Porém, ao assumir que as interações atômicas definem as distâncias interatômicas, ele falha em casos que uma alteração foi feita de maneira arbitrária por algum operador antes de passar pela otimização local.

Na figura 5.8 temos o exemplo de um cluster que sofreu uma mutação *flip* e teve um de seus átomos trocados de posição; porém, ainda não passou pelo processo de otimização local. Como o processo de otimização local será feito utilizando um método quântico, é importante ter a certeza de que este cálculo será necessário e útil, portanto, é necessário saber se algo similar ou igual ao cluster B já foi criado previamente pelo NQGA. Supondo que ele nunca tenha sido criado, ao comparar o cluster B com o cluster A utilizando o método USR seria obtido um falso-positivo indicando que estes clusters são 100% iguais. Em termos de *performance*, o USR é um método mais rápido e menos computacionalmente custoso do que o de funções de simetria, portanto, é um candidato ideal para ser utilizado no operador MLP ignorando as comparações pós operadores para comparar milhares de clusters simultaneamente. Já as funções de simetria podem ser utilizadas em qualquer caso.

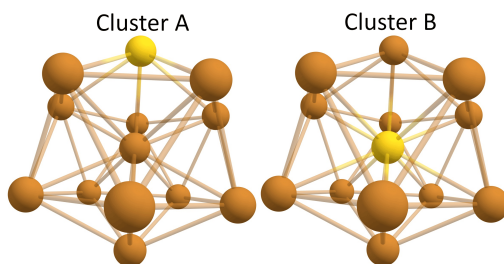


Figura 5.8: Cluster genérico A sofrendo uma mutação e se tornando o cluster B.

5.4.2 Predador

O predador é um operador simples, responsável por manter o tamanho da população constante ao longo das gerações. A população é organizada em ordem do melhor cluster (menor energia), para o pior cluster (maior energia) e os piores cluster são eliminados da população para manter seu tamanho constante. Mesmo que os piores clusters tenham sido removidos da população principal, eles serão mantidos na lista global. A lista global contém todos os clusters únicos que tiveram sua energia calculada. À medida que o NQGA

gera e processa mais clusters, todas estas informações convergem no operador predador. Ter uma lista que correlaciona geometria com energia quer dizer ter informação discreta sobre a SEP e, portanto, suas informações podem ser utilizadas posteriormente para criar uma SEP artificial para tentar extrair informações implícitas nestes dados. No NQGA isto será feito pelo operador MLP, explorado na seção 5.4.5.

5.4.3 Apex

Após aplicado todos os operadores, o algoritmo deve conter um ou mais critérios de parada, sinais que indicam que o algoritmo obteve algum resultado com forte tendência de se ter obtido, em princípio, o mínimo global. Existem três critérios de parada regendo o funcionamento do GA proposto:

1. Número máximo de iterações atingido;
2. Falta de diversidade genética populacional;
3. Permanência constante de um indivíduo no topo da lista de clusters;

Os diferentes critérios indicam os possíveis resultados que o algoritmo apresentou, no caso do primeiro tópico quer dizer que o número máximo de iterações não foi suficiente para que o algoritmo localizasse o mínimo global. A permanência da falta de diversidade genética também é um problema, uma vez que o algoritmo fica preso em uma região da superfície de energia potencial, consumindo ciclos e recursos computacionais sem avançar. Portanto, o algoritmo deverá ser reiniciado utilizando, talvez, outra configuração de operadores ou condições menos restritas, ou ainda melhorando a condição inicial da população. O terceiro item é o critério de parada mais desejável uma vez que o cluster que sobreviveu como melhor, por n gerações definidas, tem grandes chances de ser o mínimo real do sistema. Com base no terceiro critério que é possível variar o valor de γ (ou S_q caso o USR seja utilizado) para que no início das iterações o NQGA faça uma busca mais grosseira na SEP quântica e à medida que o algoritmo avança para o fim, a busca se torna cada vez mais refinada.

Caso nenhum dos critérios de parada tenham sido atingidos, o Apex é responsável por iniciar a nova geração e modificar parâmetros de comparação geométrica, se necessário. Este operador acompanha o melhor indivíduo da lista (cluster de menor energia), o chamado cluster Apex, ao longo das gerações para verificar se o algoritmo atende o último item dos critérios de parada. Esta permanência do cluster Apex ao longo de N gerações, onde N é um número arbitrário decidido no início da operação do algoritmo, pode ser entendida como um ciclo ao qual o NQGA busca finalizar. Se um novo cluster diferente assume este lugar, como novo cluster Apex então a sequência de gerações com sucesso é quebrada e o ciclo recomeça. O ciclo então pode ser definido como o número de gerações

que o cluster Apex deve sobreviver para que o algoritmo considere este um critério de parada válido e, por ser fundamentado na ideologia de sobrevivência do melhor indivíduo, não só é compatível com o algoritmo genético como também é o critério de parada mais desejável, contanto que N seja razoavelmente grande ao mesmo tempo que a diversidade genética seja mantida.

À medida que o NQGA avança no ciclo da vez, isso ocasionará alterações no parâmetro de comparação geométrica e, conseqüentemente, no comportamento do operador DGP. Essas mudanças podem ser estipuladas no início da execução do algoritmo ou, caso nada seja especificado, serão baseadas no comportamento da população. Em outras palavras, o comportamento do comparador geométrico pode ser rígido ou adaptativo.

O comportamento da população está principalmente relacionado à diversidade genética total da população. Se o NQGA não conseguir criar novos indivíduos capazes de se juntar à população (ou seja, novos indivíduos são gerados, mas não possuem energia menor do que os indivíduos da população atual), o operador DGP será ajustado para ampliar a área de busca através do parâmetro γ para explorar novas regiões da SEP.

5.4.4 Simetrizador

O Simetrizador é um operador opcional aplicado no Apex apenas uma vez por ciclo. Embora o momento de atuação do operador possa ser modificado, geralmente ele ocorre próximo ao final daquele ciclo. Devido à natureza aleatória do GA e seus operadores, às vezes é difícil encontrar uma geometria altamente simétrica utilizando metodologias quânticas. Nesse caso, o operador Simetrizador age como uma mutação no cluster Apex, forçando a estrutura para o grupo de simetria mais próximo. Um vetor de deslocamento é criado para cada átomo e cada grupo de simetria. Prevalece a simetria que resulta no menor deslocamento total dos átomos. O cluster Apex original é mantido, e este novo cluster simétrico criado é oferecido na próxima iteração. Pode ou não ser aceito na população, passando pelo processo de validação e cálculo de energia que todos os outros clusters passaram.

Adicionalmente, pode-se estender o efeito do simetrizador para os N melhores clusters da população. Como esses novos clusters serão sugestões para a próxima iteração e, portanto, serão obrigatoriamente filtrados pelo operador DGP, eles não deverão impactar negativamente na performance do NQGA.

5.4.5 Predição por *Machine learning* (MLP)

O operador de predição por *Machine learning* (MLP) também é opcional e aproveita todos os cálculos quânticos armazenados na lista global para treinar uma rede neural multidimensional atômica(65, 149). Essa rede utiliza as funções de simetria centradas em átomos, exemplificadas na Figura 4.2. A implementação desse operador foi baseada na

biblioteca *TensorFlow*, que é focada em desempenho e está disponível na linguagem de programação *C++*.

Se o NQGA estiver utilizando o operador DGP com as funções de simetria então a correlação entre funções de simetria e energia já existe. Portanto, pode-se fazer um treinamento supervisionado da rede neural com este conjunto de dados, caso contrário estas funções serão criadas *on-the-fly* com base nos cluster da lista global.

De posse do modelo treinado é possível iniciar uma nova instância isolada do NQGA que utilizará a função matemática representativa da rede neural atomística otimizada para gerar uma lista extra de cluster ótimos com base neste modelo. O MLP pode ser visto como uma recursão do NQGA, ou seja, um NQGA dentro do próprio NQGA com algumas diferenças. O MLP pode derivar suas configurações (Tamanho da população, quantidade de cruzamentos, mutações, critérios de parada, etc...) do NQGA base; porém, para aproveitar as vantagens na utilização de uma função parametrizada ao invés de um método quântico, é recomendado configurar o MLP de forma a varrer o maior espaço possível desta SEP virtual criada.

O resultado do MLP é uma lista de mínimos calculados nessa SEP virtual com base na SEP quântica. A partir dessa lista extra, é possível selecionar um ou mais clusters como sugestões de adição à lista principal do NQGA. Esse processo de treinamento e utilização do modelo resulta em um consumo adicional de recursos computacionais, que pode ser redirecionado para GPUs. No entanto, se comparado à avaliação das estruturas por métodos quânticos, esse consumo extra pode ser desconsiderado. Todo o processo de treinamento e utilização do MLP pode ser realizado em paralelo ao cálculo de energia quântica, visando um melhor aproveitamento do tempo de cálculo.

Embora a arquitetura da rede adquirir qualquer forma, baseado em outros trabalhos disponíveis na literatura, assume-se que a arquitetura com duas camadas ocultas de tamanho equivalente seja o suficiente para descrever a maioria dos sistemas atômicos(141, 192, 194–196). O treinamento da rede consiste na minimização do erro obtido pela rede comparado com o valor de referência. Esta diferença é monitorada utilizando as equações de erro médio absoluto e erro quadrático médio a seguir:

$$RMSE = \sqrt{\frac{1}{N} \sum^{clusters} [E^{ANN} - E^{ref}]^2} \quad (5.2)$$

$$MAE = \frac{1}{N} \sum^{clusters} |E^{ANN} - E^{ref}| \quad (5.3)$$

onde E^{ANN} é o valor de energia obtido para um determinado cluster pela rede neural, e E^{ref} é o valor de energia de referência obtida por método quântico.

O treinamento da rede neural é realizado utilizando o gradiente descendente. Embora o gradiente descendente não seja o melhor método de treinamento disponível atualmente, ele apresenta excelente custo-benefício em todos os casos. Portanto, optou-se por utilizá-

lo inicialmente e implementar outros métodos no futuro, como a eficiente metodologia BFGS(27–30).

O treinamento é feito independentemente da utilização da rede neural. Caso o treinamento resulte em um erro elevado, aquela superfície de energia potencial (SEP) virtual será descartada para evitar o desperdício de recursos computacionais e a introdução de erros no algoritmo.

O operador MLP pode fornecer sugestões de clusters que obtiveram menor energia potencial utilizando a rede neural, mas também pode sugerir clusters que tiveram o maior erro entre a rede treinada e a referência. É importante ressaltar que esse operador atua de forma independente. Todos os clusters gerados por ele devem passar obrigatoriamente pelo operador DGP e toda a cadeia de processamento do algoritmo, independentemente do quão pequeno seja o erro de treinamento.

A inspiração para utilizar os clusters com maior erro estatístico vem de técnicas de aprendizado ativo(141, 197), baseando-se na estratégia de consulta de amostragem de incerteza (*Uncertainty Sampling Query Strategy*). Essa abordagem sugere que os pontos com maior erro representam regiões da SEP que não ficaram claras durante o aprendizado do modelo, devido à falta de informações adicionais ou à elevada complexidade da região. Portanto, sugerir novas buscas centradas nessas regiões de incerteza pode esclarecer o comportamento da SEP, facilitando o aprendizado e levando a uma melhoria do modelo a cada nova iteração.

5.4.6 Avaliação de energia

Para o cálculo de energia, está implementado no NQGA potenciais clássicos como Lennard-Jones, Gupta e Morse; sendo necessário apenas fornecer os parâmetros utilizados para cada sistema. Além do cálculo de energia, é feita uma otimização local utilizando do método L-BFGS. Sem executar essa otimização local, é comum o GA ficar preso na região próxima ao mínimo da função, sem nunca de fato chegar nele. A necessidade da otimização local é algo bem conhecido, porém com a introdução de técnicas como *machine learning*, isso pode mudar para facilitar a exploração de superfícies de energia potencial para métodos de alto nível, cuja otimização local seja inviável.

Embora estejam disponíveis os vários potenciais parametrizados, o foco principal do NQGA é na otimização geométrica de clusters utilizando métodos quânticos. Para isso, os cálculos de energia são redirecionados para pacotes quânticos como GAMESS-US(33) e Orca(31, 32), pois são programas robustos capazes de utilizar diversas metodologias quânticas como DFT, MP2 e CCSD, das quais foram validadas em diversos trabalhos, além de serem pacotes disponíveis para o público de forma gratuita. Neste caso, é possível utilizar todas as funcionalidades disponíveis nos pacotes quânticos, incluindo otimização local utilizando o método BFGS e granularidade da convergência dos métodos disponíveis.

Capítulo 6

Resultados e discussão

6.1 Introdução

No capítulo anterior, discutimos a estrutura geral e o funcionamento da nova metodologia proposta, o NQGA. Embora os aspectos principais (operadores genéticos) sejam baseados em técnicas amplamente utilizadas em outros algoritmos genéticos com sucesso, a versão desses operadores implementada neste algoritmo introduz novas contribuições específicas para o NQGA, como a generalização dos operadores para lidar com qualquer tipo de sistema e parâmetros mutáveis durante as iterações(187). Mesmo que esses operadores sejam baseados em modelos desenvolvidos para outros algoritmos genéticos, é essencial garantir que ao estabelecer uma metodologia que utilize parcialmente esse conjunto de técnicas, irá produzir o resultado esperado.

Antes de aplicar esta nova metodologia proposta no capítulo anterior, e utilizar o NQGA em sistemas desconhecidos, é importante validar seu funcionamento em diversos sistemas cobrindo os vários possíveis usos para a metodologia proposta. Para isso, foram escolhidos três casos de validação:

- Nano-ligas de cobre-ouro e ouro-prata, onde foi possível estudar o comportamento do NQGA para sistemas com até 150 átomos;
- Clusters de lítio, para validação utilizando os métodos quânticos DFT e DLPNO-CCSD(T);
- Clusters moleculares, de água, para validação de pequenos clusters moleculares.

Esses casos foram escolhidos para testar as diversas condições aplicáveis da nova metodologia proposta com funções específicas do NQGA, como:

- Tamanho do sistema;
- Capacidade de otimização em uma SEP quântica;

- Capacidade de otimização de cluster moleculares.

Os resultados de validação do NQGA para todos os casos foram consistentes em termos de energia, comprimento de ligação e distribuição espacial dos átomos, além de apresentarem novos resultados que ainda não foram publicados na literatura e, portanto, os mínimos obtidos são, até outros novos estudos surgirem, os novos mínimos globais. Em particular, no nosso trabalho recentemente publicado, exploramos diversos aspectos do NQGA.(187)

Em resumo, os resultados de validação do NQGA demonstraram que a metodologia é capaz de produzir resultados consistentes com os resultados disponíveis na literatura. Com base no resultado obtido na etapa de validação, utilizou-se o NQGA para fazer um estudo da interação de clusters Mg_6 com moléculas de hidrogênio, para o qual obteve-se resultados que elucidam crescimento destes clusters perante a adição de H_2 . A absorção reversível de hidrogênio em nano-ligas metálicas é um campo de pesquisa interessante promete ser uma alternativa segura e com maior densidade de armazenamento comparado à adsorção de hidrogênio em zeólitas e nanoestruturas à base de carbono(198,199).

6.2 Clusters de ligas cobre-ouro e ouro-prata

Com base no algoritmo proposto, a metodologia do NQGA deve ser validada. Estudaremos sistemas previamente descritos na literatura para medir o desempenho, precisão e robustez do algoritmo em uma variedade de casos, os quais serão apresentados ao longo deste capítulo. O primeiro sistema estudado são clusters de nanoliga CuAu. Esses clusters foram extensivamente investigados utilizando várias metodologias, e há uma ampla gama de dados disponíveis na literatura sobre método, estrutura e energia do mínimo global para diferentes tamanhos de clusters CuAu(200–203).

Para explorar diferentes tamanhos de clusters, realizamos um teste preliminar com esses clusters utilizando o potencial Gupta como função de energia. Embora este potencial não seja adequado para estudar a estrutura eletrônica dos clusters, iniciar a busca pelo mínimo global com um potencial parametrizado e depois refinar a busca com um método quântico é uma estratégia amplamente utilizada na literatura para encontrar mínimos globais(1, 204, 205). Apesar de otimizações usando potenciais parametrizados seguidos de otimização por método quântico não serem o objetivo principal do NQGA, esses testes preliminares com o potencial de Gupta servem para avaliar a performance do algoritmo para clusters atômicos com até 150 átomos.

As Tabelas 6.1 e 6.2 contém os resultados utilizando o potencial Gupta (equações 3.33-3.35) como função de energia. A estrutura dos clusters e a energia de seus respectivos mínimos globais estão de acordo com estudo anteriores disponíveis na literatura(18–21), com valores de energia com erro inferior a 0.001%. O erro calculado em valores de

energia obtidos quando utiliza-se funções parametrizadas tende a ser bem baixo comparado a métodos quânticos por resultar em uma hiper-superfície de energia mais suave com mínimos melhor definidos. Este resultado positivo indica que o NQGA chegou no resultado adequado (geometria correta) para cada um dos clusters testados, com uma exceção no caso do $Cu_{12}Au_7$ onde o NQGA encontrou uma geometria de menor energia, ou seja, um candidato melhor ao mínimo global do que o previamente reportado na literatura.

Tabela 6.1: Testes preliminares feitos considerando Cu_nAu_m ($n + m \leq X$ para $X = 14, 19, 38, e 55$). O potencial de Gupta foi utilizado como função de energia [3.33], e a Tabela 2 compara alguns clusters puros e ligas com trabalhos da literatura.

Cluster	$V_{cluster}$ (eV)		Cluster	$V_{cluster}$ (eV)	
	NQGA	Ref(19)		NQGA	Ref(20)
Cu_{14}	-36,2642	-36,2642	Cu_{55}	-163,2132	-163,2131
Cu_7Au_7	-42,1905	-42,1905	Cu_{38}	-108,9676	-108,9676
Au_{14}	-45,9582	-45,9582	$Cu_{25}Au_{30}$	-184,1296	-184,1296
Cu_{19}	-51,2540	-51,2540	$Cu_{21}Au_{17}$	-121,6516	-121,6516
$Cu_{12}Au_7$	-56,9774	-56,8451	$Cu_{10}Au_{45}$	-191,3060	-191,3059
Cu_5Au_{14}	-62,0759	-62,0759	Au_{38}	-130,5400	-130,5400
Au_{19}	-63,4351	-63,4351	Au_{55}	-191,0309	-191,0308

Tabela 6.2: Testes preliminares feitos considerando Au_nAg_n ($n = 10, 20, 30, 40, 50, 60, 70, e 75$). O potencial de Gupta foi utilizado como função de energia [3.33], e a Tabela 2 compara alguns clusters puros e ligas com trabalhos da literatura.

Cluster	$V_{cluster}$ (eV)		Cluster	$V_{cluster}$ (eV)	
	NQGA	Ref(21)		NQGA	Ref(21)
$Au_{10}Ag_{10}$	-36,2642	-36,2642	$Au_{50}Ag_{50}$	-163,2132	-163,2131
$Au_{20}Ag_{20}$	-42,1905	-42,1905	$Au_{60}Ag_{60}$	-108,9676	-108,9676
$Au_{30}Ag_{30}$	-45,9582	-45,9582	$Au_{70}Ag_{70}$	-181,1118	-181,1118
$Au_{40}Ag_{40}$	-51,2540	-51,2540	$Au_{75}Ag_{75}$	-121,6516	-121,6516

Como o NQGA conseguiu encontrar corretamente a geometria de todos os clusters testados, pode-se dizer que ele obteve um alto índice de sucesso na busca global. Este sucesso pode ser atribuído principalmente a elevada taxa de mutação, especialmente a mutação *flip*, pois no caso de nanoligas, encontrar o homótopo ótimo (isômero com um arranjo diferente de dois tipos de átomos; porém, com mesma geometria e composição) é o fator importante com grande influência na capacidade do algoritmo em encontrar o mínimo global para uma determinada composição. Uma alta taxa de mutação é, geralmente, indesejada, pois degrada a qualidade dos clusters gerados por cruzamento e gera um alto número de novos indivíduos cuja energia deve ser calculada, inviabilizando a utilização de métodos quânticos em clusters maiores. Para mais detalhes sobre todos os parâmetros do NQGA utilizados nos sistemas apresentados nas tabelas acima, consulte os Apêndices A.1 e A.2.

Para o NQGA, o aumento da taxa de mutação *flip* não tem um impacto negativo grande graças ao operador DGP, filtrando qualquer homótopo repetido e certificando que apenas novos clusters tenham sua energia calculada. A degradação dos resultados do operador cruzamento também não é um problema para o NQGA, pois devido as mutações serem não-destrutivas (mutações são considerados novos clusters e não alteram nenhuma característica do cluster de origem), é possível reduzir o número de mutações geométricas mantendo um alto número de mutações *flip* para facilitar a determinação da localização do mínimo global.

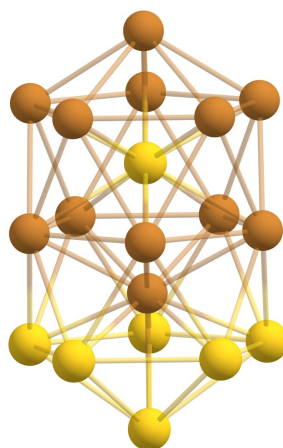


Figura 6.1: Novo mínimo global para o $Cu_{12}Au_7$.

Devido a estas características o NQGA encontrou um novo mínimo global para o $Cu_{12}Au_7$ ilustrado na Figura 6.1. A nova geometria encontrada tem conexões iguais a reportada na literatura com o icosaedro de cobre com átomo central de ouro, porém a versão do NQGA tem a distância de ligação média 0.003\AA menor do que a obtida por Lordeiro et al.(19). Esta aproximação dos átomos levou o cluster para uma região de menor energia da SEP ainda sem entrar na região completamente repulsiva. Embora não seja correto afirmar que este é o mínimo global definitivo para este cluster, pois devido a complexidade da superfície de energia potencial e ao fato que métodos heurísticos não dão certeza de que o mínimo encontrado é realmente o mínimo global, podemos dizer que até o presente momento, a estrutura da Figura 6.1 representa o mínimo global para o cluster $Cu_{12}Au_7$. Previamente o mínimo global reportado para o $Cu_{12}Au_7$ por Lordeiro et al.(19) apresenta a mesma distribuição de átomos com a energia total de $-56,8451\text{ eV}$. O novo mínimo global encontrado pelo NQGA tem energia total igual a $-56,9774\text{ eV}$. À medida que o número de átomos em um sistema aumenta, a complexidade do cálculo das superfícies de energia potencial também aumenta exponencialmente. Isso ocorre porque cada átomo adicionado ao sistema aumenta o número de graus de liberdade, resultando em uma hiper superfície de energia potencial de dimensão cada vez maior. Além disso, os métodos quânticos, que são baseados em princípios fundamentais sem a necessidade de parâmetros empíricos, são computacionalmente intensivos. Portanto, à medida que o número de átomos aumenta,

o tempo e os recursos computacionais necessários para realizar esses cálculos tornam-se proibitivos.

Ao longo da otimização do cluster $Cu_{12}Au_7$, o NQGA encontrou diversos mínimos locais previamente reportados na literatura, como observado na Figura 6.2, onde cada mudança de energia (eixo Y) representa um novo mínimo local.

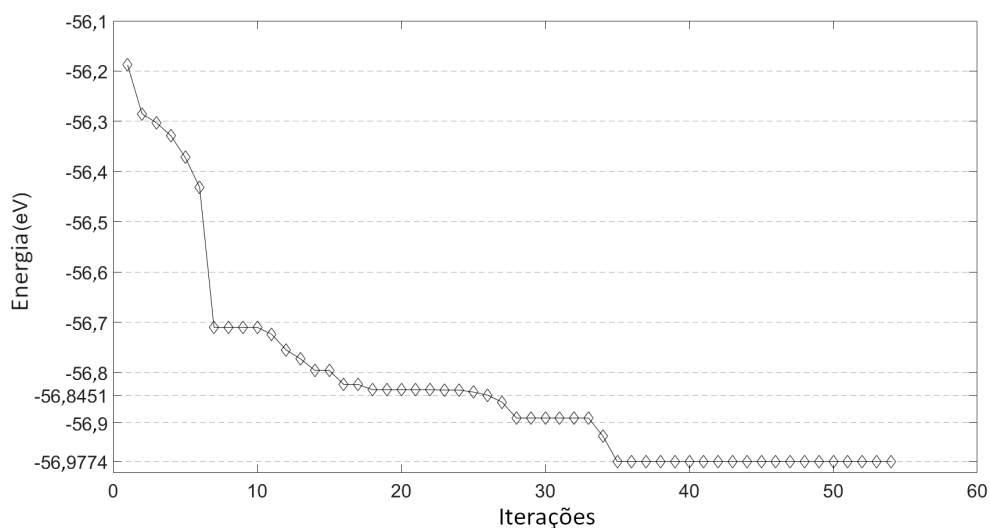


Figura 6.2: Progresso do NQGA ao otimizar o cluster $Cu_{12}Au_7$.

Todos os mínimos locais de menor energia encontrados são diferentes homótopos do mínimo global, como esperado de um algoritmo cujo resultado é uma população de soluções, como é o caso de um GA. Por exemplo, na Figura 6.3 observa-se que os clusters B e C têm um átomo de ouro no centro e um *cap* de ouro imperfeito comparado ao mínimo global.

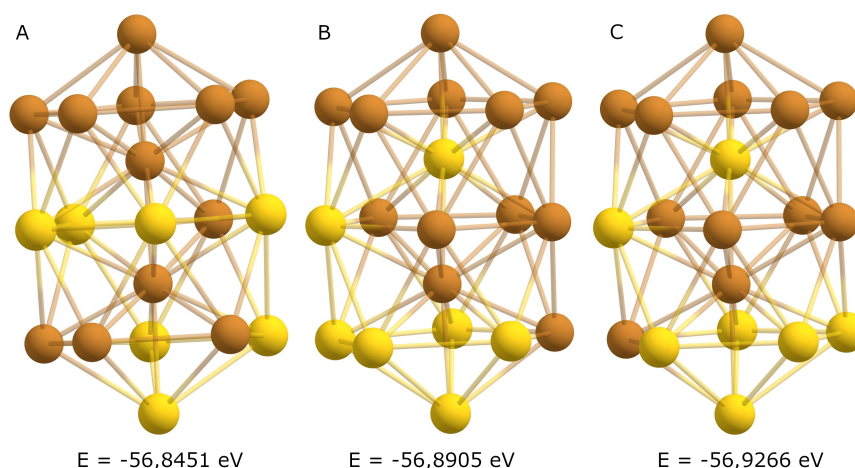


Figura 6.3: Os três mínimos locais de menor energia encontrados para o cluster $Cu_{12}Au_7$.

O cluster A não apresenta nenhuma dessas características, resultando no cluster de maior energia entre os apresentados. Isso ilustra bem a importância de um método eficiente para

localizar o melhor homótopo dentro de cada possível geometria, maximizando as chances do algoritmo localizar o mínimo global correto.

Como evidenciado anteriormente, aumentar o tamanho do cluster aumenta exponencialmente a dificuldade de se determinar mínimos na superfície de energia potencial, especialmente o mínimo global. Para clusters heteroatômicos, é particularmente mais difícil devido a adição das possíveis combinações entre os átomos diferentes, que também aumenta com o tamanho do cluster. Uma possível solução para varrer esse espaço de combinações atômicas e fazer uma busca local iterativa (*Iterated Local Search* - ILS)(206). Usando o ILS, é possível otimizar a ordem química dos átomos no cluster, porém requer a energia de um grande número de homótopos possíveis para uma determinada geometria e, portanto, teria um elevado custo computacional para uma otimização utilizando um método quântico, inviabilizando assim este método para otimizações quânticas em um GA. Por este motivo, o NQGA tem alta taxa de mutação *flip*, combinada com o operador DGP, para filtrar geometrias e homótopos repetidos.

Um dos problemas provindos da natureza aleatória de um algoritmo heurístico, como o GA, é que não há como garantir que o resultado de uma otimização é o melhor resultado possível. Não existe nenhum parâmetro convergente e, portanto, ao encontrar um mínimo global para um cluster uma vez não significa que o mínimo encontrado é global ou que ele será encontrado todas as vezes que o algoritmo for executado. Isso geralmente é devido à geração aleatória das condições iniciais. Uma alternativa que pode permitir a reprodutibilidade é usar um gerador aleatório que, se a semente for a mesma e ele iniciar da mesma forma, então garante-se a reprodução do caminho de busca do mínimo na SEP. Esta repetibilidade de resultado em um algoritmo de principio aleatório pode ser utilizado como parâmetro para estimar a eficiência do algoritmo e complexidade da SEP. Esta taxa de sucesso pode ser definida como o número de vezes que o NQGA encontrou o mínimo global correto dividido pelo número total de tentativas. Wu, Cai e Shao(20) reportaram uma taxa de sucesso média de 66,2%, em 20 execuções independentes, para os clusters Cu_nAu_{38-n} ($0 \leq n \leq 38$). O NQGA obteve uma taxa de sucesso média, considerando 30 execuções, para os clusters apresentados na Tabela 6.1, de 88,6%, e 72,8% para os clusters da Tabela 6.2. Estes resultados indicam que o NQGA, além de ser eficiente em localizar o mínimo global, ele é capaz de localizar estes mínimos com consistência, o que é extremamente desejável. Devido a natureza heurística do algoritmo genético, é comum fazer a busca pelo mínimo global utilizando diversas sementes para tentar garantir que o mínimo encontrado é global, porém o ideal seria obter o resultado com apenas uma semente aleatória, economizando tempo e recursos computacionais. Por isso, quanto maior a taxa de sucesso, melhor será o algoritmo. Para os clusters $AuAg$, é possível observar uma queda de desempenho, que é esperada, considerando que clusters maiores têm uma SEP mais complexa e há um grande aumento no número de mínimos locais profundos em tal superfície. É possível aumentar a taxa de sucesso modificando os parâmetros do GA

para operar com uma população maior e taxas mais elevadas de mutação e cruzamento. Essas modificações influenciam diretamente o número de avaliações de energia que são feitas por iteração e, quanto maior este número, maior a probabilidade de encontrar o mínimo global. Embora este aumento do número de cálculos de energia seja aceitável para otimizações com potenciais parametrizados, ao substituir estes potenciais por métodos quânticos, o aumento esperado na taxa de sucesso do NQGA vem associado a um elevado custo computacional. Portanto, deve haver um balanceamento na escolha dos parâmetros de operação do NQGA, levando em consideração o sistema estudado, o método de avaliação de energia escolhido e os recursos computacionais disponíveis.

6.3 Clusters de lítio

Para avaliar o desempenho do NQGA utilizando métodos quânticos escolheu-se os clusters Li_n ($3 \leq n \leq 8$). Ou seja, a busca pelo mínimo global será feita diretamente na superfície de energia potencial quântica, livre de qualquer tendência gerada por potenciais semi-empíricos. Clusters de lítio foram escolhidos por serem clusters amplamente estudados, com uma boa disponibilidade de resultados na literatura para diferentes tamanhos e metodologias quânticas. Este primeiro estudo foi feito utilizando a metodologia DFT/PBE, utilizando o conjunto de funções de base aug-cc-pVTZ. O Apêndice A.3 contém os parâmetros de entrada e detalha como eles são montados para a execução do NQGA na otimização dos clusters de lítio. A escolha do funcional e base foi feita de forma que os resultados de energia fossem passíveis de comparação com resultados obtidos da literatura(207).

A Figura 6.4 ilustra a geometria e energia dos clusters de lítio otimizados pelo NQGA. As estruturas encontradas pelo NQGA estão de acordo com as estruturas obtidas por Brito et al.(207), e a diferença de energia entre elas é menor que 1%. Brito et al.(207) nota que Li_5 e Li_8 tem isômeros com energia bem próximas do seus respectivos mínimos globais. O NQGA também identificou a mesma tendência para o Li_5 e Li_8 ; porém, para o Li_8 , a geometria do mínimo global foi divergente. Esta diferença está ligada a escolha do função de troca e correlação. A seleção do funcional exerce um papel crucial na determinação da geometria do mínimo global, pois a modelagem da superfície é dependente do funcional escolhido. Consequentemente, isso afeta a diferença de energia entre os mínimos locais e globais.

A Figura 6.5 mostra o mínimo global encontrado pelo NQGA para o cluster Li_8 como sendo Li_8^A , e o mínimo local de menor energia Li_8^B . Isso sugere que o Li_8^A deve ser o mínimo global utilizando o funcional PBE. Embora Brito et al.(207) conclua que o funcional de troca e correlação PBE descreva melhor as características do Li_8 , a partir de cálculos DMC (*Diffusion quantum Monte Carlo simulation*) e resultados experimentais, os autores concluem que Li_8^B é a estrutura do mínimo global para o cluster Li_8 . Isso apresenta

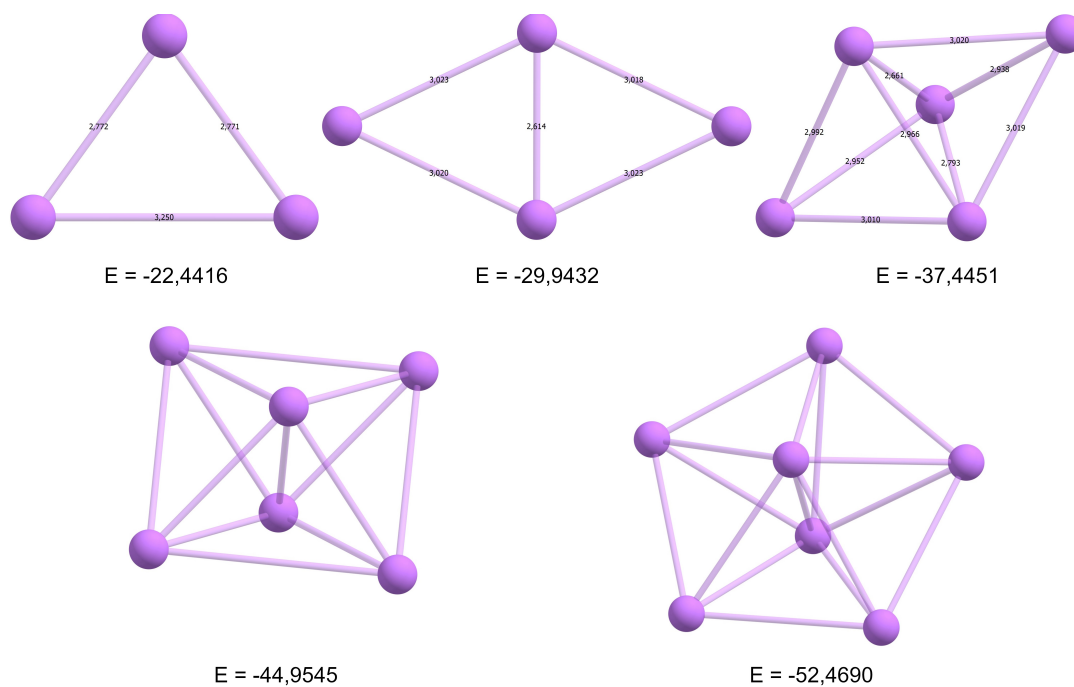


Figura 6.4: Estruturas otimizadas para os clusters Li_n (para $n=3-7$) e suas respectivas energias (a.u.), obtidas utilizando DFT/PBE com a base aug-cc-pVTZ.

uma boa oportunidade de testar as capacidades do NQGA de fazer uma otimização utilizando outros métodos quânticos. Foi feita uma otimização do cluster Li_8 utilizando a metodologia DLPNO-CCSD(T), confirmando que Li_8^B é o mínimo global com energia -60,0979 a.u. e Li_8^A um mínimo local profundo com energia -60,0967 a.u. A escolha do método utilizado para fazer a otimização é de extrema importância para obter resultados precisos, embora esta afirmação seja considerada trivial, ela reforça a necessidade do desenvolvimento de metodologias capazes de buscar a solução, neste caso a estrutura de mínima energia, diretamente na superfície de energia potencial quântica representada pelo método escolhido, e eliminar métodos intermediários que introduzem tendências e erros sistemáticos.

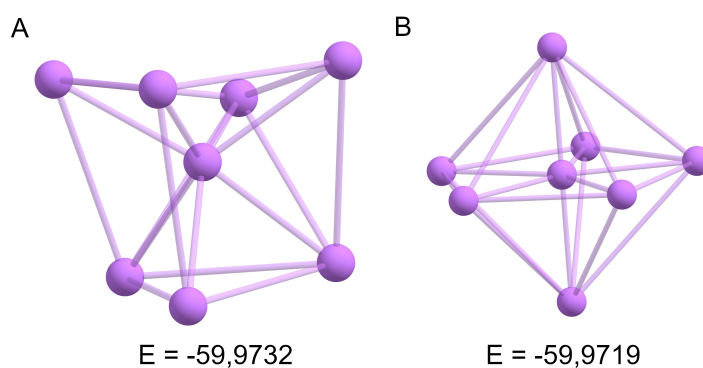


Figura 6.5: Estruturas de mínimo obtidas pelo NQGA para os clusters Li_8^A , Li_8^B e suas respectivas energias (a.u.), obtidas utilizando DFT/PBE com a base aug-cc-pVTZ.

O processo de otimização do Li_8 , considerando 10 execuções independentes, levou,

em média, 13 gerações. Durante estas execuções, uma média de 73 diferentes estruturas foram geradas e selecionadas para terem sua energia calculada via DFT. A capacidade do algoritmo em encontrar corretamente o mínimo global é crucial, especialmente para avaliar a energia mínima dos clusters. Para qualquer otimização que utilize métodos quânticos, o NQGA pode utilizar o operador MLP para acelerar o processo de minimização, reduzindo o número de visitas à superfície de energia potencial (SEP) e, conseqüentemente, o tempo total necessário para otimizar uma estrutura específica.

A Tabela 6.3 apresenta a diferença no número de avaliações de energia com e sem a influência do operador MLP. Os mesmos operadores e a mesma semente foi utilizada em ambos os casos para cada tamanho de cluster, para tentar isolar o efeito do operador MLP e avaliar sua influência. Pode-se inferir que o operador MLP reduz, em média, de 16 a 23% o número de avaliações de energia para os nanoclusters de lítio. Embora seja uma redução percentual significativa, espera-se uma redução ainda maior para clusters maiores, uma vez que, o número de visitas à SEP tende a aumentar à medida que o número de átomos aumenta, assim a precisão da rede neural contida no operador MLP também deve aumentar, resultando numa redução no total de cálculos quânticos, diretamente reduzindo os recursos computacionais necessários.

Tabela 6.3: O efeito do operador MLP no número de visitas a SEP quântica para os clusters de lítio Li_n ($n=3-8$).

Cluster	Número de avaliações de energia	
	com MLP	sem MLP
Li_3	53	63
Li_4	57	74
Li_5	64	79
Li_6	82	94
Li_7	106	122
Li_8	109	139

Do ponto de vista apenas de potenciais baseados em *machine learning*, o ganho de 23% poderia ser considerado muito pequeno; porém, estes potenciais são treinados utilizando grandes bancos de dados contendo milhares de estruturas selecionadas cuidadosamente para reproduzir uma SEP. Com dados o suficiente um potencial de rede neural artificial (RNA) é capaz de reproduzir uma SEP quântica com excelente precisão e erro quadrático médio de energia inferior a 1 kcal/mol(208). No caso do NQGA, o operador MLP é treinado em tempo real, paralelamente ao cálculo de energia quântico; geralmente, com um número bem reduzido de amostras contendo uma distribuição de estruturas que pode não ser ideal para reproduzir a SEP quântica. Este acumulado de condições adversas pode explicar o porque MLP não reduziu o número de avaliações de energia de maneira mais significativa. O uso do MLP gera um custo de recursos computacionais extra relevante; portanto, qualquer redução no número de avaliações por cálculos quânticos é desejável,

contanto que as estruturas geradas pelo MLP não prejudiquem os outros operadores. Neste caso o operador DGP novamente ficará responsável por filtrar todo resultado produzido pelo MLP e garantir que somente os resultados úteis sejam repassados para os próximos operadores.

Utilizar um potencial RNA criado *on-the-fly* confere algumas vantagens ao algoritmo, ao invés de utilizar uma metodologia convencional, como por exemplo, um potencial semi-empírico. Um potencial RNA irá reproduzir o comportamento e particularidades da SEP quântica no qual foi treinado, em nível de teoria equivalente, sem introduzir tendências de um potencial parametrizado ou até mesmo tendências ou imprecisões geradas por métodos DFT, em que o mesmo funcional pode apresentar comportamentos bem distintos para dois sistemas diferentes. Outra vantagem é que um potencial RNA pode ser treinado para qualquer sistema, independente da disponibilidade de potenciais ou parâmetros necessários para descrever o sistema escolhido. Em contrapartida, a maior desvantagem de um potencial RNA é ser extremamente dependente do seu treinamento e, portanto, dos dados coletados e seu tratamento, assim como da arquitetura escolhida.

Tendo como base poucos pontos para treinar um potencial com base na SEP construída com a metodologia quântica, poderá resultar em uma função de energia com erros de validação na ordem de 10^3 kcal/mol ou até mesmo superior; imediatamente invalidando qualquer dado obtido com esta RNA. Isto é inerente da falta de dados para reproduzir adequadamente a SEP; porém, esta limitação não causa nenhum problema para o NQGA, pois seu uso pode ser diretamente condicionado a sua precisão medida a cada iteração. Como mostrado na Figura 5.3, o operador MLP é opcional; portanto, só será utilizado quando há a possibilidade dele beneficiar o NQGA, produzindo indivíduos (clusters) qualitativamente diferentes e possivelmente melhores. Assim como os outros operadores, o MLP opera de forma independente e os resultados são sugestões de clusters para o NQGA; pois, mesmo com um pequeno erro de treinamento, considera-se a energia calculada pelo MLP distinta da energia obtida por método quântico. Portanto, esta divisão entre treinamento e utilização se mostra bem útil, principalmente nas primeiras iterações do NQGA, onde existem apenas uma pequena quantidade de pontos calculados na SEP quântica. Nestas primeiras iterações, o erro de validação é sempre elevado e, portanto, a etapa de utilização do operador MLP não será utilizada naquela iteração, apenas o treinamento, reduzindo assim a chance de gerar sugestões ruins, e consumir recursos computacionais desnecessariamente.

A partir dos testes realizados com o operador MLP ativo, concluiu-se que o potencial RNA produz sugestões válidas quando o RMSE de validação é inferior a 10 kcal/mol. Considerou-se sugestões válidas aquelas em que o cluster sugerido passou pelo operador DGP e após ter sua energia calculada por método quântico, teve sua colocação na lista de energia entre o primeiro e o último cluster da iteração atual.

É possível utilizar a mesma metodologia com um nível de teoria superior para se estudar

clusters pequenos, como clusters de lítio. O custo computacional será exponencialmente superior, como ilustrado na figura 6.6, e acompanhado de alguns desafios que serão detalhados a seguir; porém, possibilitaria novos *insights* em métodos de otimizar a geometria de cluster neste nível de teoria e na SEP destes clusters.

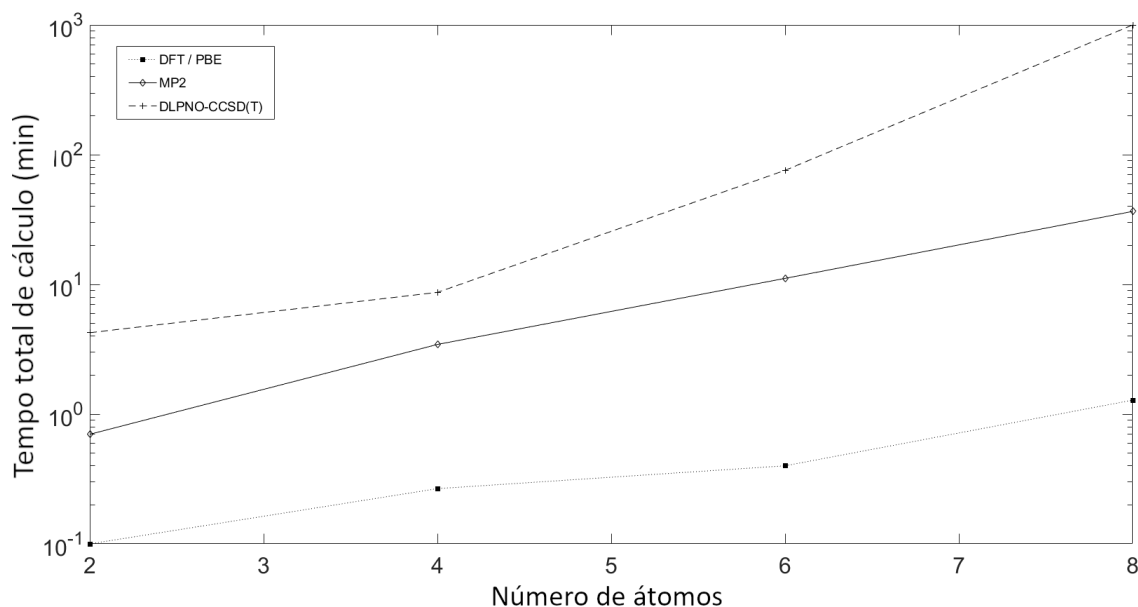


Figura 6.6: Tempo necessário para obter um cálculo de energia pontual para clusters de lítio. Dados obtidos utilizando o pacote quântico ORCA em um processador de 24 núcleos Intel Xeon E5-2650.

Dos pacotes quânticos testados, tanto o ORCA quanto o GAMESS-US não possuem implementação disponível nem capacidade para calcular o gradiente analítico em uma superfície de energia potencial no nível de teoria DLPNO-CCSD(T). A etapa de minimização local utilizando o método BFGS depende desse gradiente, sendo crucial para garantir a eficiência do algoritmo genético. Uma alternativa viável para contornar esse problema é utilizar o gradiente numérico para as minimizações locais. O principal desafio dessa abordagem é que o cálculo do gradiente numérico requer movimentos positivos e negativos para cada átomo em cada eixo cartesiano. Assim, o gradiente numérico demanda $6N$ cálculos de energia pontual individuais, onde N é o número de átomos no cluster. Esse requisito torna a otimização local proibitiva em sistemas médios e grandes, o que consequentemente inviabiliza a otimização global desses sistemas.

Uma outra opção é realizar as otimizações locais em um nível de teoria inferior que possua a capacidade de calcular o gradiente analítico, como DFT ou MP2, por exemplo. A escolha do nível de teoria para calcular esse gradiente deve ser cuidadosa, de modo a reproduzir com a maior precisão possível as características da SEP obtida com o método DLPNO-CCSD(T). Se o método DFT for escolhido, a seleção do funcional de troca e correlação também deve ser feita com atenção. A desvantagem dessa abordagem é que as minimizações locais são aproximações e podem falhar em alguns casos.

Devido aos novos operadores incorporados ao NQGA, há uma nova opção a ser explorada para esse tipo de minimização, que consiste em utilizar o gradiente numérico apenas na última iteração do operador Apex. Dessa forma, ao longo das iterações do NQGA, o algoritmo dependerá do MLP para realizar as minimizações locais. Uma vez que o MLP produz uma equação analítica representativa da superfície de energia potencial no nível de teoria desejado, o gradiente dessa superfície também deve ser representativo desse nível teórico e pode ser obtido com facilidade e baixo custo computacional.

A última iteração do Apex, como detalhado anteriormente (item 5.4.3), determina a finalização ou não do NQGA e pode ser executada várias vezes ao longo do algoritmo, tornando-a ideal para realizar uma verificação final da população e permitir uma minimização local utilizando o gradiente numérico. Isto significa em essência eliminar otimizações locais até que o algoritmo encontrar o melhor candidato a mínimo global. Somente neste ponto serão feitas otimizações locais para verificar se de fato foram encontrados mínimos na SEP. Isso possibilita um uso mais eficiente dos recursos computacionais disponíveis e torna viável otimizações globais em níveis de teoria superiores.

Embora essa técnica pareça promissora, ela ainda deve ser rigorosamente testada em outros sistemas para garantir que não impacte negativamente na eficiência do NQGA. No entanto, para obter resultados preliminares, foram escolhidos os clusters Li_3 , Li_4 , Li_2H e Li_3H como alvos para diferentes otimizações locais. Esses clusters são pequenos e relativamente simples, porém possuem dados teóricos e experimentais disponíveis na literatura, o que os torna um bom ponto de partida para elucidar a viabilidade desta técnica.

Os valores da distância de ligação experimentais obtidas por Huber e Herzberg(209) para o Li_2 e LiH são 2,673 e 1,596 Å, respectivamente, e os resultados obtidos utilizando DLPNO-CCSD(T), ilustrados na Figura 6.7, estão de acordo com resultados teóricos e experimentais(209, 210). Isso significa que os resultados obtidos são consistentes com as expectativas teóricas, com erro médio de $5E - 4$ Å para ambos casos, e distâncias quantitativamente iguais à observações experimentais. Na Figura 6.8, estão apresentados os resultados das geometrias otimizadas no NQGA utilizando DLPNO-CCSD(T) para os clusters Li_3 , Li_4 , Li_2H , e Li_3H .

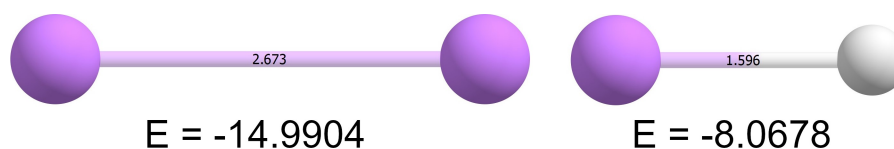


Figura 6.7: Distâncias de ligação e energia (em a.u.) para o Li_2 e LiH , obtidas no nível de teoria DLPNO-CCSD(T)/aug-CC-PVQZ.

Os valores de comprimento da ligação no cluster Li_3 estão quantitativamente iguais aos publicados na literatura(211). De forma similar, para os clusters Li_4 , Li_2H , Li_3H suas estruturas estão de acordo com os resultados obtidos por Brito et al.(212), com error

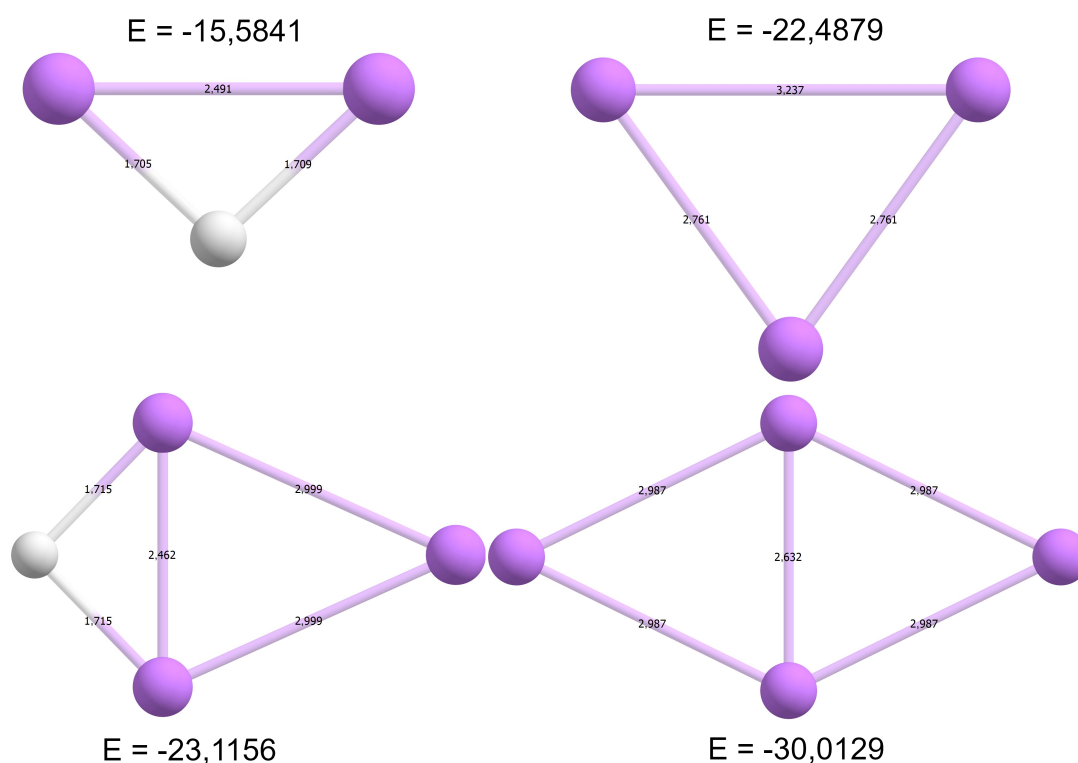


Figura 6.8: Estruturas otimizadas dos clusters Li_3 , Li_4 , Li_2H , Li_3H e suas energias (em a.u.), obtidas no nível de teoria DLPNO-CCSD(T)/aug-CC-PVQZ.

médio das distancias inferior a $1E - 4$ Å e $1,5E - 4$ a.u. de energia. Embora, neste caso, nenhum novo mínimo global tenha sido encontrado, isto já era esperado, considerando que clusters de 4 átomos não são particularmente difíceis de serem otimizados, e testes em clusters maiores devem ser feitos no futuro, pois isto pode ser uma opção viável para tratar clusters maiores. O principal objetivo deste experimento é testar a capacidade do NQGA de fazer uma minimização utilizando um nível de teoria superior, a princípio, nos casos mais simples, usando a hipótese de que não há necessidade de fazer minimizações locais computacionalmente dispendiosas em todas as iterações do algoritmo, e com base nos resultados positivos obtidos, será possível, no futuro investigar outros casos e clusters maiores.

6.4 Clusters de água

A metodologia NQGA não possui operadores genéticos exclusivos para o tratamento de clusters moleculares. Porém, todos os operadores são capazes de operar em átomos ou moléculas, conferindo ao algoritmo a capacidade de otimizar clusters atômicos, moleculares ou uma combinação de átomos e moléculas. Alguns fatores que diferenciam os operadores em clusters atômicos e moleculares são a capacidade de manter a integridade das moléculas e a composição do cluster intactas ao longo das iterações, e operadores de mutação também

podem aplicar rotação aos fragmentos moleculares que compõe o cluster. Os primeiros clusters escolhidos para validar a capacidade do NQGA em otimizar clusters moleculares são os cluster $(H_2O)_5$ e $(H_2O)_{11}$, por serem clusters bem estudados com boa variedade de dados disponíveis na literatura. O conjunto total de parâmetros de entrada para ambos os clusters podem ser encontrados no Apêndice A.5.

A Figura 6.9 mostra o mínimo global encontrado pelo NQGA para o cluster $(H_2O)_5$ usando o nível de teoria MP2 e conjunto de funções de base aug-cc-pVDZ. A energia obtida para a molécula de água foi -76,2641 a.u., e a energia de ligação (*binding energy*) encontrada para o cluster $(H_2O)_5$ foi -37,9639 kcal/mol.

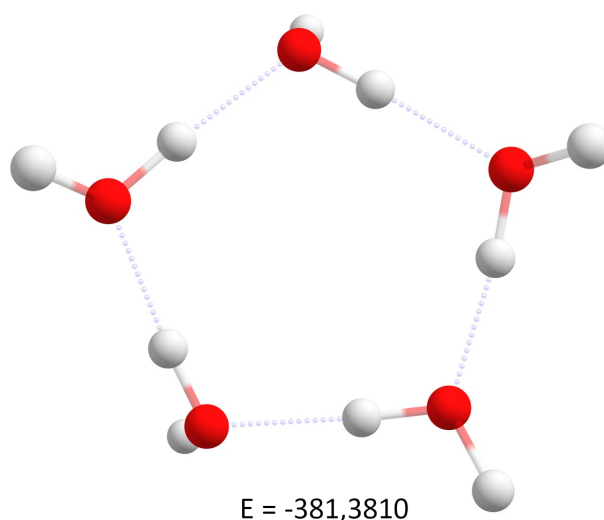


Figura 6.9: Estrutura otimizada do $(H_2O)_5$ e sua energia (em a.u.), obtidas no nível de teoria MP2/aug-cc-pVDZ.

A energia de ligação ilustra o ganho de estabilidade, ou redução de energia potencial do cluster, em relação as moléculas ou átomos isolados e pode ser calculada, para o cluster de água, através da seguinte equação:

$$E_b = E_{(H_2O)_n} - nE_{H_2O} \quad (6.1)$$

onde E_b é a energia de ligação, n é o número de moléculas de água que compõe o cluster, $E_{(H_2O)_n}$ é a energia total de um cluster composto por n moléculas de água e E_{H_2O} é a energia total de uma molécula de água. O resultado encontrado pelo NQGA para o cluster $(H_2O)_5$ é geometricamente similar e com mesma simetria ao descrito pela literatura, e também possui energia de ligação próxima. Temelso, Archer e Shields(213) reportaram energia de ligação para o cluster $(H_2O)_5$ como sendo -37,9012 kcal/mol, que um pouco mais alta do encontrado neste trabalho, com uma diferença média de $3,13E - 2$ kcal/mol. Silva, Duarte e Belchior(45) também apresenta geometria similar para este cluster, com energia de ligação ainda maior, -39,28 kcal/mol, com um erro médio em relação a este trabalho de $3,42E - 2$. Porém, não podemos comparar diretamente estes valores, porque neste

caso o nível de teoria utilizado foi diferente. Silva, Duarte e Belchior⁽⁴⁵⁾ tratou o sistema utilizando metodologia DFT, enquanto este trabalho e Temelso, Archer e Shields⁽²¹³⁾ foram feitos utilizando metodologia MP2, por isto a diferença quantitativa entre os valores calculados com MP2 são similares, porém a análise qualitativa da geometria é válida para os dois casos.

A Figura 6.10 mostra o mínimo global encontrado pelo NQGA usando o nível de teoria MP2 e conjunto de funções de base aug-cc-pVDZ. A energia obtida para a molécula de água foi -76,2641 a.u., e a energia de ligação (*binding energy*) encontrada para o cluster $(H_2O)_{11}$ foi -109,6968 kcal/mol. Bulusu et al.⁽²¹⁴⁾, utilizando o mesmo nível de teoria, encontrou um isômero similar ao ilustrado na Figura 6.10, denominado isômero 43'4, e reportaram a energia de ligação de -109,331 kcal/mol, sendo ligeiramente maior que a encontrada pelo NQGA. Esta diferença de energia pode estar associada a diferentes razões. As principais são diferenças nos parâmetros do cálculo quântico como grade de integração, parâmetros do SCF ou de convergência, assim como a estrutura encontrada pelo NQGA que é o novo mínimo global para o cluster $(H_2O)_{11}$. Este resultado demonstra principalmente a capacidade do NQGA de atacar problemas envolvendo clusters moleculares. É importante notar que o NQGA também conseguiu encontrar o segundo isômero de menor energia $(H_2O)_{11}$ 515, também reportado por Bulusu et al.⁽²¹⁴⁾. Takeuchi⁽²¹⁵⁾ também reportaram os isômeros 43'4 e 515 como mínimos globais, porém utilizando os potenciais TIP3P e TIP4P, respectivamente, corroborando com os resultados obtidos pelo NQGA. Este é um resultado extremamente importante para o NQGA pois indica a habilidade de distinguir mínimos de profundidade similar e identificar o mínimo global. O isômero ilustrado na Figura 6.10 correspondeu ao mínimo global três vezes consecutivas no NQGA, e não foi possível encontrar nenhum outro isômero de maior energia para o cluster $(H_2O)_{11}$, embora eles tenham sido reportados nos trabalhos referenciados^(214,215). Esta falha em encontrar os mínimos locais próximos pode estar associada a diversos fatores, como o pequeno número de indivíduos presente na população do NQGA, levando a estes indivíduos de maior energia serem eliminados da população, falha dos operadores genéticos em localizar tais estruturas ou escolha sub-ótima de parâmetros para operadores como DGP, permitindo que mínimos locais muito próximos ocupem vários lugares na população do NQGA.

Este resultado sugere um trabalho futuro focado na escolha dos parâmetros dos operadores genéticos, uma vez que a definição desses parâmetros é feita por meio de uma combinação de intuição química, reutilização de parâmetros de sistemas similares e tentativa e erro. No estado atual da metodologia NQGA, não existe uma fórmula específica para determinar os parâmetros ótimos que garantam uma população diversificada, composta por um mínimo global seguido de múltiplos mínimos locais, conforme relatado na literatura^(214,215).

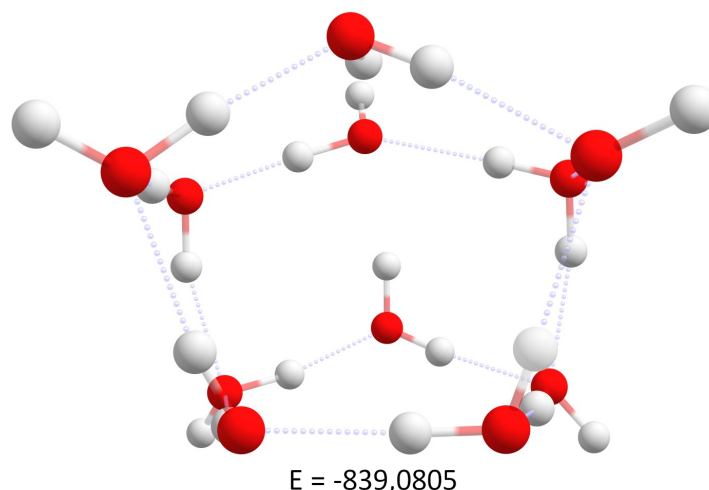


Figura 6.10: Estrutura otimizada do $(H_2O)_{11}$ e sua energia (em a.u.), obtidas no nível de teoria MP2/aug-cc-pVDZ.

6.5 Clusters atômicos e moleculares de magnésio-hidrogênio

Visto que o NQGA apresentou resultado satisfatório na identificação do mínimo global para um cluster molecular de 11 elementos, pode-se utilizar o mesmo procedimento para estudar casos de minimização onde moléculas, átomos e superfícies interajam de maneira controlada. Um desses casos seria clusters utilizados para adsorção de hidrogênio como maneira de entender e propor novos materiais como solução para armazenamento de hidrogênio. Partindo de dados disponíveis na literatura, fez-se uma validação prévia da metodologia para o caso do cluster Mg_6H_4 . Para esta tarefa, primeiramente foi feita uma otimização da estrutura do Mg_6 no NQGA a fim de determinar a estrutura que será usada para interagir com os átomos de hidrogênio ou moléculas H_2 . O conjunto total de parâmetros de entrada pode ser encontrado no Apêndice A.6.

A Figura 6.11 ilustra a estrutura de mínimo global, com energia -1200,3351 a.u., obtida pelo NQGA para o Mg_6 utilizando o nível de teoria MP2/aug-cc-pVDZ. Pode-se utilizar esta estrutura como base para interação com o hidrogênio; logo, o NQGA deve otimizar a interação da estrutura apresentada na Figura 6.11 com átomos de hidrogênio, mantendo a rigidez e integridade da cadeia de átomos durante as operações genéticas; porém, permitindo o relaxamento da estrutura durante a otimização local.

Mesmo com a complexidade química desta tarefa do ponto de vista algorítmico, esta é uma otimização de 5 partículas no espaço. O resultado obtido pelo NQGA está ilustrado na Figura 6.12 (A), e ao lado a estrutura (B) representa o melhor resultado encontrado na literatura para este cluster, reportado por Shen et al.(216). Este resultado corrobora com os resultados positivos obtidos nos experimentos anteriores e ainda apresenta uma nova estrutura de mínimo para o cluster Mg_6H_4 , que é mais um dos resultados inéditos obtidos

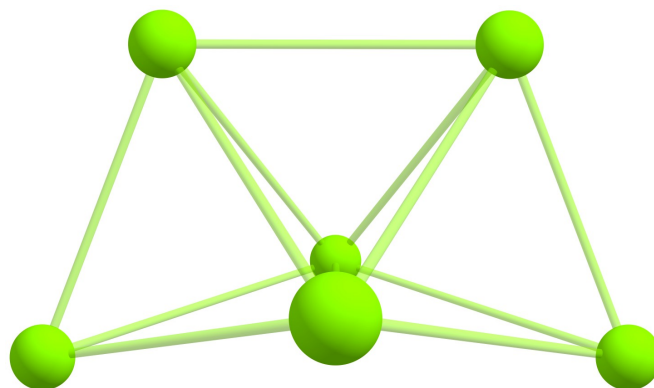


Figura 6.11: Estrutura otimizada do Mg_6 , obtida pelo NQGA no nível de teoria MP2/aug-cc-pVDZ.

pelo NQGA durante os testes de validação. É importante notar a técnica de construção do cluster partindo da estrutura do Mg_6 otimizado foi adotada para seguir a metodologia utilizada por Shen et al.(216), na expectativa de que os clusters de mínimo seriam iguais. Partir da premissa de que o cluster Mg_6H_4 tem a estrutura dos átomos de magnésio similar à estrutura do Mg_6 introduz uma tendência que induz a similaridade entre as estruturas do Mg_6H_4 e Mg_6 , que pode levar a uma estrutura de mínimo global errônea devido a mudanças que os átomos de hidrogênio podem causar na estrutura eletrônica do cluster, resultando em uma geometria de mínimo global completamente diferente. Para investigar isto, seria necessário fazer uma otimização sem restrição partindo de 6 átomos de magnésio e 4 átomos de hidrogênio, onde, neste caso, seria esperado uma estrutura igual ou melhor a estrutura (A) apresentada na Figura 6.12.

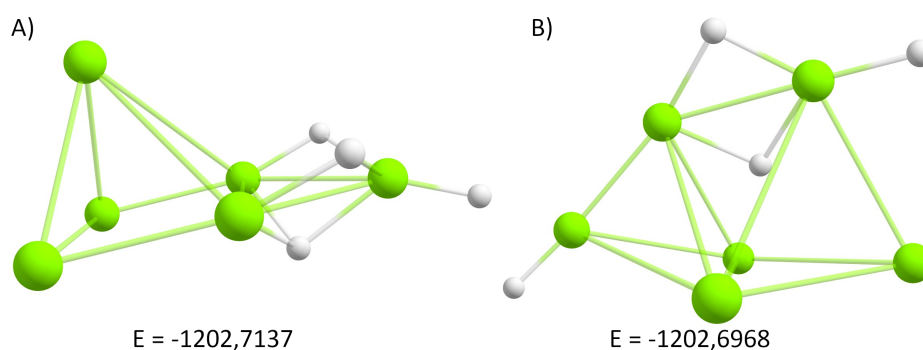


Figura 6.12: Estruturas de mínimo obtidas pelo NQGA (A) para o cluster Mg_6H_4 comparado com o mínimo reportado na literatura (B) e suas energias respectivas (em a.u.), obtidas no nível de teoria DFT/aug-cc-pVDZ com funcional de troca e correlação B3PW91 e correção de dispersão de Grimme's (DFT-D3).

O estudo do cluster Mg_6H_4 também pode ser utilizado para observar o efeito do operador MLP e sua influência no processo de otimização. Na Figura 6.13 está a relação entre o erro quadrático médio da energia do grupo de validação em relação aos clusters de referência durante o treinamento da rede neural gerada pelo operador MLP. Como esperado, o erro nas primeiras iterações é extremamente elevado, na faixa de 83,0 kcal/mol; portanto,

embora este potencial esteja pronto para ser utilizado, alocar recursos computacionais para otimizar um cluster com este erro elevado seria desperdício, como discutido anteriormente, e, portanto, neste ponto do algoritmo, esta RNA não é utilizada. A partir do momento em que o erro é considerado pequeno o suficiente para gerar uma SEP com precisão aceitável (na Figura 6.13, a partir da terceira iteração), o operador MLP começa a gerar clusters com base nesta SEP produzida pela RNA.

O critério de erro para a aceitação da rede neural de 10,0 kcal/mol foi selecionado com base no erro médio obtido pelo modelo de RNA adotado neste estudo, que foi de 4,8 kcal/mol. Esse valor é calculado como a média dos erros gerados em cada execução do NQGA, por sua vez sendo a média dos erros ao longo das iterações. Para referência, algumas redes neurais modernas conseguem atingir erros inferiores a 1,0 kcal/mol. É importante ressaltar que esse erro representa a média dos desvios associados à capacidade da rede neural em estimar corretamente, por exemplo, a energia de um determinado cluster. Em contraste, o erro associado à metodologia DFT pode ser de até 3,0 kcal/mol para algumas moléculas (217). Embora esses erros envolvam grandezas diferentes e não possam ser diretamente comparados, o erro da abordagem quântica influencia diretamente o desempenho da rede neural, sendo aditivo ao erro intrínseco da própria rede.

O valor de 4,8 kcal/mol pode ser considerado relativamente baixo, especialmente considerando que o conjunto de treinamento utilizado é uma pequena fração do utilizado em redes neurais de alto desempenho. Esse valor de erro é completamente aceitável, dado que esses clusters nunca seriam identificados como mínimos globais sem passar por uma otimização local no nível de teoria escolhido.

A partir da iteração em que o operador MLP começa a gerar novos indivíduos, é esperado que o erro da próxima iteração seja menor, representando uma redução assintótica a medida que novos clusters são adicionados a lista de treinamento. Porém, observa-se na Figura 6.13 que o erro varia e, em alguns casos, aumenta consideravelmente, a ponto de interromper uso do MLP novamente. Isto acontece porque este tipo de RNA interpola dados com precisão superior ao erro de treinamento da rede, porém tem grandes dificuldades em extrapolar resultados que são muito diferentes do grupo de dados em que o modelo foi treinado. Este efeito, associado com a variabilidade genética necessária em um algoritmo genético e a necessidade de expandir o espaço de busca na SEP quântica, possibilita o aumento do erro ao longo das iterações. Outro ponto a ser considerado é que, para evitar qualquer tendência a ser criada durante o processo de treinamento da RNA e propagar este erro ao longo das gerações, optou-se por treinar a RNA a partir de pesos aleatórios a cada iteração, ao invés de utilizar os pesos da última iteração. Embora existam outros métodos de inicialização de pesos em uma rede neural(218), a inicialização aleatória foi escolhida, a princípio, para manter a natureza aleatória do GA e testar a viabilidade do operador MLP no NQGA. Outros métodos de inicialização, reutilização de pesos, assim como outras arquiteturas, são possíveis pontos de investigação em um trabalho futuro com

objetivo de melhorar o desempenho do operador e, conseqüentemente, do NQGA.

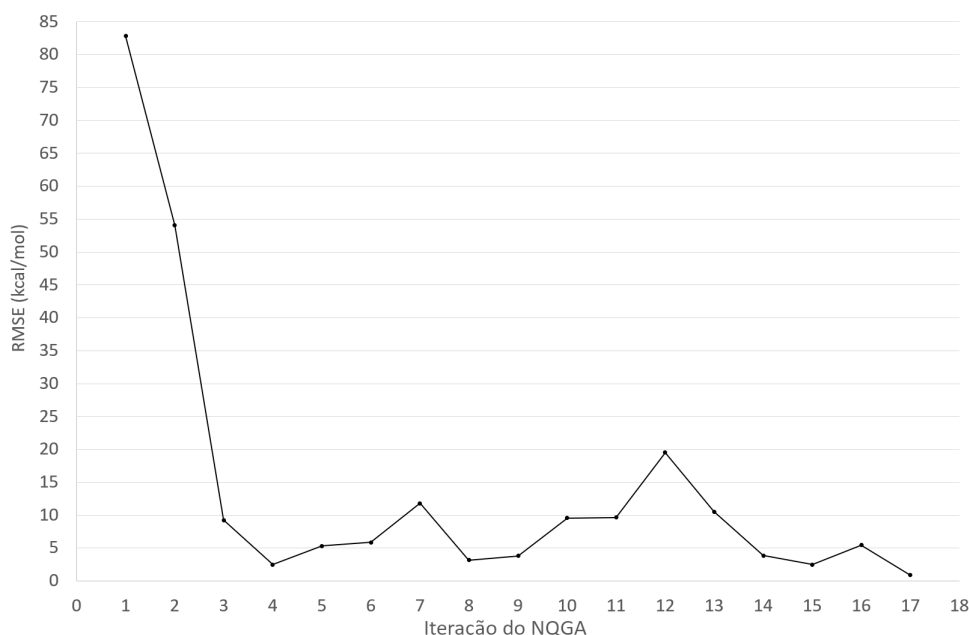


Figura 6.13: Erro quadrático médio (RMSE) do potencial gerado pelo operador MLP durante a otimização do cluster Mg_6H_4 .

De maneira simplificada, o operador MLP cumpre de forma satisfatória o papel de otimizar uma forma aproximada do método quântico escolhido. Observando o funcionamento do MLP e como são geradas as sugestões de estruturas para o NQGA, é natural pensar na possibilidade de fazer esta mesma operação utilizando um potencial parametrizado ou um método quântico que também demanda menos recursos computacionais. Para julgar a utilidade do operador MLP em relação a um outro método tradicional, utilizaremos o cluster Mg_6H_4 para fazer uma comparação de estruturas. Uma estrutura foi otimizada utilizando a RNA treinada pelo operador MLP, e o outro foi utilizado o método Hartree-Fock corrigido com a dispersão de Grimme's (HF-3c), que demanda menos recursos que o B3Pw91/aug-cc-pVDZ com DFT-D3 utilizado previamente e pode ser considerado mais preciso do que um método semi-empírico.

O NQGA, utilizando o potencial RNA treinado, achou uma estrutura com mesma ordem ligações e simetria similar a mostrada na Figura 6.12 (A), com energia igual a -1202,77 a.u.. A energia e os comprimentos de ligação não são muito relevantes neste caso, pois estas estruturas sempre sofrerão uma nova otimização local utilizando um método quântico. Após a re-otimização com o método quântico escolhido, B3Pw91/aug-cc-pVDZ (DFT-D3), esta estrutura converge exatamente para estrutura correta, como esperado de um método que simula o cálculo quântico. A otimização global do Mg_6H_4 utilizando as mesmas condições, sendo a única alteração o método de cálculo de energia, HF-3c, leva ao cluster apresentado na Figura 6.14 como mínimo global, e após re-otimização no método B3Pw91/aug-cc-pVDZ (DFT-D3), tem energia -1202,70, bem superior ao obtido

pela RNA.

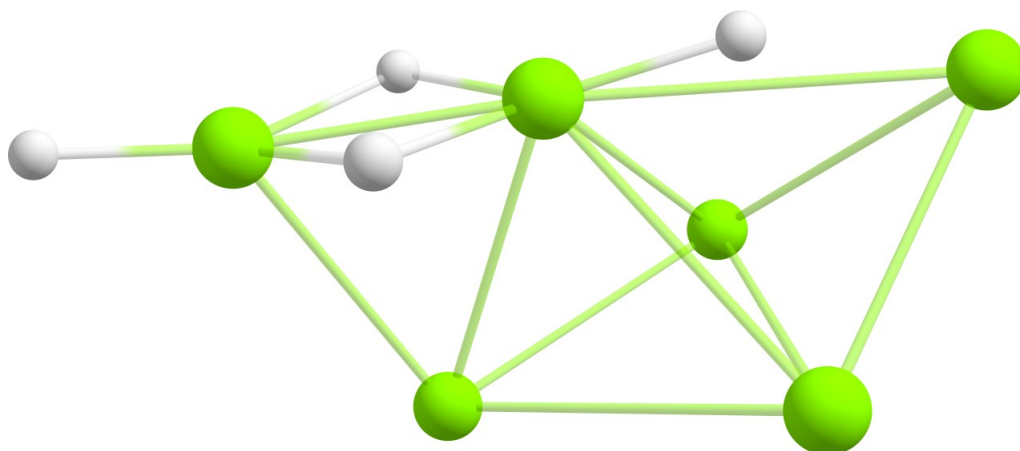


Figura 6.14: Mínimo global do cluster Mg_6H_4 utilizando o método HF-3c para otimização global.

Outro ponto negativo desta abordagem é que, embora o método HF-3c seja menos custoso do que o método DFT escolhido, ele ainda consumirá recursos que poderiam ser gastos com a própria otimização por método quântico, enquanto uma RNA pode utilizar, em maior parte, recursos de GPU, ao invés de CPU. Uma alternativa seria utilizar potenciais parametrizados que demandam menos recursos, até mesmo do que uma RNA; porém, eles falham em representar a SEP quântica, resultando no mesmo problema mencionado, além de que para muitos casos de clusters e moléculas menos estudados há dificuldade em achar parâmetros. Para este problema, o MLP representa uma boa solução, com desvantagens mínimas.

Como mencionado anteriormente, a otimização do cluster Mg_6H_4 poderia ser realizada por meio da combinação atômica, ou seja, partindo de átomos isolados de Mg e H , ao invés de usar o cluster Mg_6 com moléculas de H_2 . No entanto, para um estudo de adsorção de moléculas de hidrogênio em clusters de magnésio e para permitir a comparação com resultados da literatura, foi necessário otimizar a interação entre duas estruturas rígidas: o cluster Mg_6 com as moléculas de H_2 .

A Figura 6.15 mostra a estrutura de mínima energia para o cluster $Mg_6(H_2)$ obtida utilizando o NQGA, ao lado de um resultado obtido na literatura(216). Shen et al.(216) estudaram o cluster Mg_6H_2 . Portanto, permite que a molécula de hidrogênio seja adsorvida pelo cluster Mg_6 através de átomos individuais, ao invés da molécula H_2 . Pode-se notar que a energia da estrutura A, -1201,51 a.u., é consideravelmente inferior à energia da estrutura obtida por Shen et al.(216), sugerindo que o NQGA foi capaz de encontrar um novo mínimo de menor energia para este cluster.

A partir deste resultado, surgem novos questionamentos, tais como: Qual seria o cluster de mínima energia encontrado pelo NQGA, caso não houvesse a imposição de manter a molécula de hidrogênio intacta? O processo de otimização foi feito, desta vez, para o

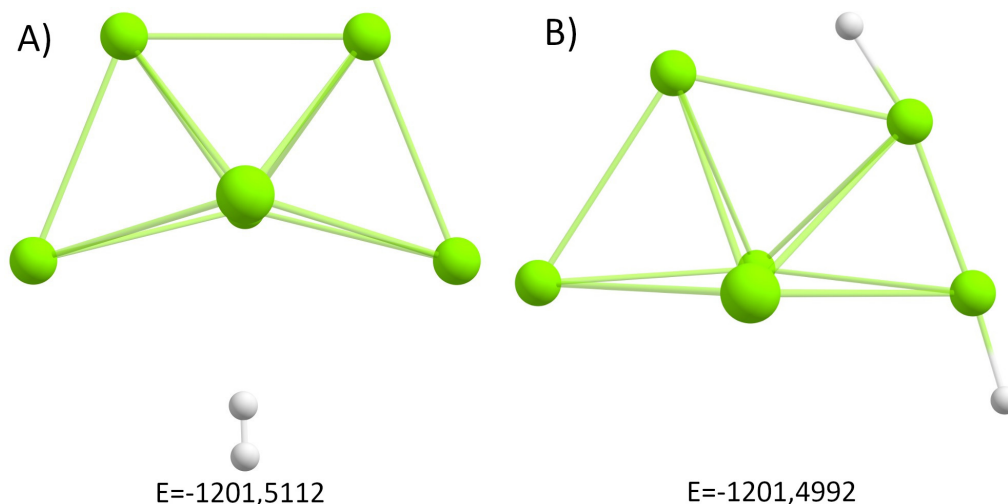


Figura 6.15: Mínimo global do cluster $Mg_6(H_2)$, em A obtido pelo NQGA e B obtido na literatura e suas respectivas energias, em a.u. utilizando o nível de teoria DFT/aug-cc-pVDZ com funcional de troca e correlação B3PW91 e correção de dispersão de Grimme's (DFT-D3).

cluster Mg_6H_2 , e o resultado é o mesmo do cluster $Mg_6(H_2)$ mostrado na Figura 6.15, apontando mais uma vez para o cluster A como sendo o mínimo global. Outro aspecto relevante é investigar se, neste caso, há interação entre o cluster de magnésio e a molécula de hidrogênio. Para isso, foi conduzida uma varredura variando a distância da molécula de hidrogênio em relação ao centro geométrico do cluster Mg_6 ao longo do eixo da ligação H_2 .

A Figura 6.16 mostra o resultado da varredura feita no cluster $Mg_6(H_2)$, variando a distância da molécula de hidrogênio e o centro do cluster Mg_6 entre 2.35 e 5.35 Å. É possível notar claramente que há um ganho de estabilidade a medida que a molécula H_2 aproxima do cluster Mg_6 , até um ponto que a energia começa a aumentar exponencialmente devido às forças de repulsão. Entre a região de repulsão e atração, existe um ponto de menor energia, onde se encontra a estrutura de maior estabilidade.

A partir dos dados apresentados, pode-se afirmar com um bom grau de certeza que o NQGA encontrou um novo mínimo global para a estrutura do Mg_6H_2 e $Mg_6(H_2)$ que são iguais. É possível fazer uma análise similar a da Equação 6.1 para esta adsorção, seguindo a seguinte equação:

$$E_{ads} = E_{(Mg_6H_2)} - (E_{Mg_6} + E_{H_2}) \quad (6.2)$$

onde E_{ads} é a energia de estabilização do clusters ganha ao adicionar a molécula H_2 ao cluster Mg_6 , $E_{(Mg_6H_2)}$ é a energia do cluster Mg_6H_2 , $E_{(Mg_6)}$ e $E_{(H_2)}$ são referentes a energia do cluster Mg_6 e H_2 isolados. A partir dos resultados encontrados, o valor de E_{ads} é de -1,045 kcal/mol.

Para o cluster Mg_6H_2 , a adsorção física apresenta forma mais estável; porém, isto não representa uma tendência à medida que novas moléculas de H_2 são adicionadas. Para comparar a adsorção física da molécula de H_2 com a adsorção de átomos de hidrogênio,

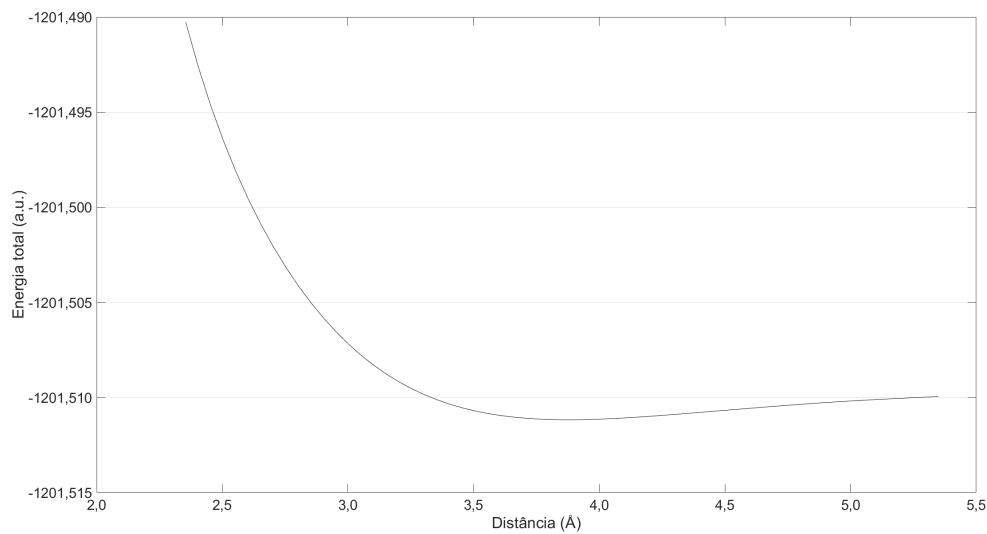


Figura 6.16: Varredura da distância da molécula de hidrogênio ao centro do cluster Mg_6 em função da energia total do cluster Mg_6H_2 .

podemos usar a seguinte equação:

$$E_{ads} = \frac{E_{(Mg_6H_x)} - (E_{Mg_6} + xE_H)}{1 + x} \quad (6.3)$$

Pode-se notar, através da Figura 6.17, que apenas Mg_6H_2 tem adsorção física favorável. Adições posteriores são minimamente favorecidas e, portanto, uma adsorção química, onde a quebra das ligações de hidrogênio será mais favorável. A configuração mais estável será a Mg_6H_{12} , corroborando com os resultados obtidos por Shen et al.(216), e é um resultado esperado, pois a adsorção física é mais fraca, composta majoritariamente de interações eletrostáticas, resultando em um cluster mais volumoso e menos estável, como pode-se observar na figura 6.18 onde é comparado com os clusters Mg_6H_{12} e $Mg_6(H_2)_6$.

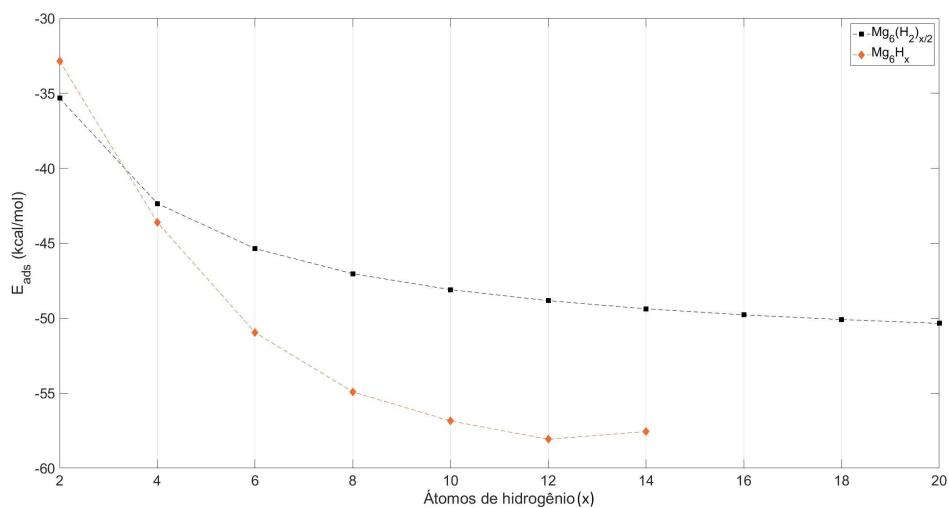


Figura 6.17: Comparativo da energia de adsorção do $Mg_6(H_2)_{x/2}$ com Mg_6H_x (216) utilizando de 1 a 10 pares de hidrogênio.

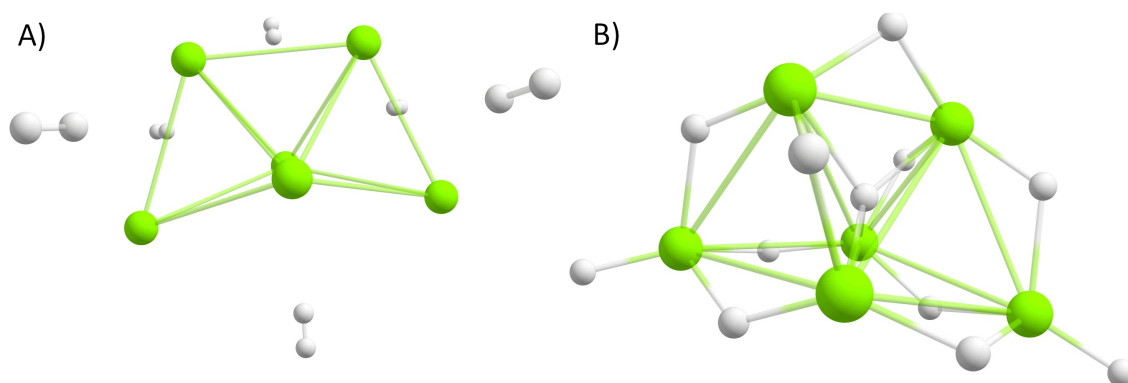


Figura 6.18: Estruturas dos clusters $Mg_6(H_2)_6$ (A) e Mg_6H_{12} (B).

Capítulo 7

Conclusão

A proposta do NQGA, como algoritmo e metodologia, é ser robusto e capaz de localizar estruturas de energia mínima para uma ampla gama de sistemas atômicos e moleculares, independentemente da condição inicial. Embora o método de avaliação de clusters e moléculas utilizando energia possa ser feito com qualquer tipo de função ou método matemático, o NQGA foca na utilização direta da superfície de energia potencial proveniente de métodos quânticos. Para que isso fosse possível, destacamos vários pontos fundamentais:

- Desenvolvemos um método eficiente para a geração inicial da população baseado na geração de esferas, evitando inicializações totalmente aleatórias;
- Implementamos operadores genéticos generalizados que podem operar em átomos, moléculas, fragmentos ou combinações destes;
- Criamos um método eficiente de comparar as geometrias dos clusters gerados e mensurar sua similaridade, o que resultou no desenvolvimento do operador DGP, que ainda serve como base para criação de uma RNA pelo operador MLP;
- Minimizar o número de visitas à SEP quântica utilizando os operadores DGP e MLP para reduzir drasticamente o número de clusters a terem sua energia calculada;
- Aumentar a chance de sucesso do algoritmo variando os parâmetros dos operadores ao longo das iterações em resposta a diversidade genética da população.

A metodologia NQGA obteve sucesso em todos os testes de validação realizados, encontrando a estrutura de mínimo global reportada na literatura ou uma estrutura mais adequada para o mínimo global. Para validar a capacidade do NQGA de minimizar clusters de vários tamanhos, foram realizados testes com clusters de cobre, ouro, cobre-ouro e ouro-prata contendo até 150 átomos utilizando o potencial de Gupta. Para testar a capacidade do algoritmo de operar com níveis de teoria superiores, como CCSD(T), foram otimizadas as estruturas de pequenos clusters de lítio. Embora sejam estruturas simples, esses testes

serviram para verificar a capacidade dos operadores em trabalhar com este nível de teoria e seus desafios. Além disso, foram conduzidos testes em clusters moleculares, incluindo clusters de água e clusters de Magnésio-Hidrogênio, nos quais resultados excelentes foram obtidos. Durante a validação para o sistema Magnésio-Hidrogênio, além dos resultados positivos, obtivemos resultados preliminares interessantes sobre o comportamento da interação do cluster Mg_6 com hidrogênio no contexto de armazenamento de hidrogênio.

Mesmo com todas as propostas especificamente desenvolvidas para reduzir o número de visitas à SEP quântica, este ainda é o gargalo do algoritmo. Isso reforça a importância deste trabalho e a necessidade de continuar desenvolvendo métodos mais inteligentes e menos restritivos para selecionar quando realizar essas avaliações na SEP quântica. Um reflexo disso foi observado durante a otimização do sistema $(H_2O)_{11}$, onde foi possível encontrar o mínimo global reportado na literatura, porém não foi possível localizar todos os mínimos locais reportados.

A proposta do operador DGP em filtrar clusters similares não é inovadora no contexto de algoritmos genéticos na química computacional, mas seu método de operação e natureza adaptativa, combinados com os resultados obtidos, indicam que ele é a peça principal do NQGA e o fator decisivo para o sucesso ou fracasso do algoritmo. O principal ponto negativo deste operador está justamente na necessidade de selecionar os parâmetros principais das equações 4.10 a 4.16, que têm influência direta na capacidade de comparação do algoritmo e, portanto, na precisão do resultado.

O operador MLP também merece destaque pela sua capacidade de reproduzir a SEP quântica de forma computacionalmente viável, reduzindo o número de avaliações de energia quântica em otimizações de geometria com métodos quânticos de alto nível. Também foi possível observar que em certas condições o MLP reproduz a SEP quântica com maior precisão do que outros métodos quânticos de menor nível. Este operador, ainda em estágios iniciais de desenvolvimento, apresenta um potencial significativo para exploração futura.

7.1 Perspectivas futuras

O operador DGP se revelou crucial, mas assim como os outros operadores, requer uma seleção cuidadosa dos parâmetros para otimizar a precisão do algoritmo. Para minimizar a necessidade de escolher uma grande variedade de parâmetros e evitar erros devido à má escolha desses parâmetros, seria importante desenvolver ferramentas que permitam ao NQGA monitorar o desempenho dos operadores e ajustar seus parâmetros para operar sempre com máxima eficiência. O operador MLP se destaca na reprodução da SEP quântica, reduzindo significativamente o número de visitas à SEP quântica real. Este operador abre diversas possibilidades para investigações futuras, como:

- Comparar o desempenho do operador MLP em otimizações com e sem cálculo de

gradiente para minimizações locais para estruturas maiores e mais complexas;

- Estudar como diferentes arquiteturas de RNA influenciam os resultados do algoritmo, dada a diversidade de capacidades e propriedades;
- Testar diferentes algoritmos de *machine learning* para classificação dos clusters.

Com base nos casos estudados, o NQGA demonstrou ser uma metodologia promissora para acelerar a determinação de estruturas de cluster, permitindo um estudo mais detalhado das propriedades eletrônicas. Portanto, planeja-se aplicar o NQGA em estudos adicionais de clusters com confiança na capacidade do algoritmo de produzir resultados robustos.

REFERÊNCIAS

- 1 RODRIGUES, D. D. et al. Global optimization analysis of $\text{Cu}(n)\text{Au}(m)$ ($n+m=38$) clusters: Complementary ab initio calculations. *Chemical Physics*, v. 349, n. 1, p. 91 – 97, 2008. ISSN 0301-0104. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0301010408001730>.
- 2 SALEH, T. A. Overview of surface and interface science. In: _____. *Surface Science of Adsorbents and Nanoadsorbents*. Elsevier, 2022. p. 1–38. ISBN 9780128498767. Disponível em: <http://dx.doi.org/10.1016/B978-0-12-849876-7.00001-4>.
- 3 LARK-HOROVITZ, K.; MEISSNER, K. W. The optical properties of semiconductors. i. the reflectivity of germanium semiconductors. *Physical Review*, American Physical Society (APS), v. 76, n. 10, p. 1530–1530, nov. 1949. ISSN 0031-899X. Disponível em: <http://dx.doi.org/10.1103/PhysRev.76.1530>.
- 4 AGUADO, A.; LÓPEZ, J. M. Structure determination in 55-atom Li-Na and Na-K nanoalloys. *The Journal of Chemical Physics*, v. 133, n. 9, p. 094302, 2010. Disponível em: <https://doi.org/10.1063/1.3479396>.
- 5 MOSELER, M. et al. Structure and magnetism of neutral and anionic palladium clusters. *Phys. Rev. Lett.*, American Physical Society, v. 86, p. 2545–2548, Mar 2001. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.86.2545>.
- 6 BALETTO, F.; FERRANDO, R. Structural properties of nanoclusters: Energetic, thermodynamic, and kinetic effects. *Rev. Mod. Phys.*, American Physical Society, v. 77, p. 371–423, May 2005. Disponível em: <https://link.aps.org/doi/10.1103/RevModPhys.77.371>.
- 7 LEE, B.; LEE, G. W. Comparative study of Ti and Ni clusters from first principles. *The Journal of Chemical Physics*, v. 127, n. 16, p. 164316, 2007. Disponível em: <https://doi.org/10.1063/1.2800026>.
- 8 SUN, H.; ZHANG, W.; XU, N. Density functional calculation of structural and electronic properties of Sn-XAl_x ($n=2-8, 13, x=0-n$) clusters. *Journal of Physics and Chemistry of Solids*, v. 118, p. 126 – 136, 2018. ISSN 0022-3697. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0022369717310478>.
- 9 ANDREEV, A. S.; KUZNETSOV, V. N.; CHIZHOV, Y. V. Atomic hydrogen activated TiO_2 nanocluster: DFT calculations. *The Journal of Physical Chemistry C*, v. 116, n. 34, p. 18139–18145, 2012. Disponível em: <https://doi.org/10.1021/jp3009805>.
- 10 ZHICHENG, Y. et al. Stability and electronic properties of WnX ($n=1-7, X=\text{Cu, Mo}$) clusters. *The European Physical Journal Plus*, v. 132, n. 11, p. 451, Nov 2017. Disponível em: <https://doi.org/10.1140/epjp/i2017-11719-x>.

- 11 LAZAUSKAS, T. et al. Thermodynamically accessible titanium clusters tin, $n = 2-32$. *Phys. Chem. Chem. Phys.*, The Royal Society of Chemistry, v. 20, p. 13962–13973, 2018. Disponível em: <http://dx.doi.org/10.1039/C8CP00406D>.
- 12 DOYE, J. P. K.; WALES, D. J. Structural consequences of the range of the interatomic potential a menagerie of clusters. *J. Chem. Soc., Faraday Trans.*, The Royal Society of Chemistry, v. 93, p. 4233–4243, 1997. Disponível em: <http://dx.doi.org/10.1039/A706221D>.
- 13 SAMULTSEV, D. O.; SEMENOV, V. A.; KRIVDIN, L. B. On the accuracy of the gao-dft calculation of $15n$ nmr chemical shifts of the nitrogen-containing heterocycles – a gateway to better agreement with experiment at lower computational cost. *Magnetic Resonance in Chemistry*, v. 52, n. 5, p. 222–230, 2014. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/mrc.4055>.
- 14 CLEMENTI, E. Ab initio computational chemistry. *The Journal of Physical Chemistry*, v. 89, n. 21, p. 4426–4436, 1985. Disponível em: <https://doi.org/10.1021/j100267a007>.
- 15 FINNIS, M. W.; SINCLAIR, J. E. A simple empirical n-body potential for transition metals. *Philosophical Magazine A*, Taylor, Francis, v. 50, n. 1, p. 45–55, 1984. Disponível em: <https://doi.org/10.1080/01418618408244210>.
- 16 ERRATUM. *Philosophical Magazine A*, Taylor, Francis, v. 53, n. 1, p. 161–161, 1986. Disponível em: <https://doi.org/10.1080/01418618608242815>.
- 17 ELLIOTT, J.; SHIBUTA, Y.; WALES, D. Global minima of transition metal clusters described by finnis–sinclair potentials: A comparison with semi-empirical molecular orbital theory. *Philosophical Magazine*, Taylor, Francis, v. 89, n. 34-36, p. 3311–3332, 2009. Disponível em: <https://doi.org/10.1080/14786430903270668>.
- 18 HSU, P. J.; LAI, S. K. Structures of bimetallic clusters. *The Journal of Chemical Physics*, v. 124, n. 4, p. 044711, 2006. Disponível em: <https://doi.org/10.1063/1.2147159>.
- 19 LORDEIRO, R. A. et al. Determination of main structural compositions of nanoalloy clusters of cuxauy ($x + y \leq 30$) using a genetic algorithm approach. *International Journal of Quantum Chemistry*, v. 95, n. 2, p. 112–125, 2003. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/qua.10660>.
- 20 WU, X.; CAI, W.; SHAO, X. Optimization of bimetallic cu–au and ag–au clusters by using a modified adaptive immune optimization algorithm. *Journal of Computational Chemistry*, v. 30, n. 13, p. 1992–2000, 2009. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.21197>.
- 21 CHEN, F.; JOHNSTON, R. L. Energetic, electronic, and thermal effects on structural properties of ag–au nanoalloys. *ACS Nano*, v. 2, n. 1, p. 165–175, 2008. Disponível em: <https://doi.org/10.1021/nn700226y>.
- 22 MICHAELIAN, K.; RENDÓN, N.; GARZÓN, I. L. Structure and energetics of ni, ag, and au nanoclusters. *Phys. Rev. B*, American Physical Society, v. 60, p. 2000–2010, Jul 1999. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevB.60.2000>.

- 23 PACHECO-CONTRERAS, R. et al. Empirical-potential global minima and dft local minima of trimetallic aglaumptn (l+m+n=13, 19, 33, 38) clusters. *Computational Materials Science*, v. 141, p. 30–40, 2018. ISSN 0927-0256. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0927025617305025>.
- 24 PAZ-BORBÓN, L. O. et al. A mixed structural motif in 34-atom pd-pt clusters. *The Journal of Physical Chemistry C*, v. 111, n. 7, p. 2936–2941, 2007. Disponível em: <https://doi.org/10.1021/jp0674165>.
- 25 BRUMA, A. et al. Dft study of the structures and energetics of 98-atom aupd clusters. *Nanoscale*, The Royal Society of Chemistry, v. 5, p. 646–652, 2013. Disponível em: <http://dx.doi.org/10.1039/C2NR32517A>.
- 26 ALCANTAR-MEDINA, K. et al. Evolution of the structural and electronic properties of small alkali metal-doped aluminum clusters. *Computational and Theoretical Chemistry*, v. 1099, p. 55–63, 2017. ISSN 2210-271X. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2210271X16304479>.
- 27 BROYDEN, C. G. The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations. *IMA Journal of Applied Mathematics*, v. 6, n. 1, p. 76–90, 03 1970. ISSN 0272-4960. Disponível em: <https://doi.org/10.1093/imamat/6.1.76>.
- 28 FLETCHER, R. A new approach to variable metric algorithms. *The Computer Journal*, v. 13, n. 3, p. 317–322, 01 1970. ISSN 0010-4620. Disponível em: <https://doi.org/10.1093/comjnl/13.3.317>.
- 29 GOLDFARB, D. A family of variable-metric methods derived by variational means. *Mathematics of Computation*, American Mathematical Society, v. 24, n. 109, p. 23–26, 1970. ISSN 00255718, 10886842. Disponível em: <http://www.jstor.org/stable/2004873>.
- 30 SHANNO, D. F. Conditioning of quasi-newton methods for function minimization. *Mathematics of Computation*, American Mathematical Society, v. 24, n. 111, p. 647–656, 1970. ISSN 00255718, 10886842. Disponível em: <http://www.jstor.org/stable/2004840>.
- 31 NEESE, F. The orca program system. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, v. 2, p. 73–78, 2012.
- 32 NEESE, F. Software update: the orca program system, version 4.0. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, v. 8, p. 1327, 2017.
- 33 SCHMIDT, M. W. et al. General atomic and molecular electronic structure system. *J. Comput. Chem.*, v. 14, p. 1347, 1993.
- 34 WALES, D. J.; DOYE, J. P. K. Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms. *The Journal of Physical Chemistry A*, v. 101, n. 28, p. 5111–5116, 1997. Disponível em: <https://doi.org/10.1021/jp970984n>.
- 35 LEARY, R. Global optimization on funneling landscapes. *Journal of Global Optimization*, v. 18, p. 367–383, 01 2000.

- 36 ZHANG, J.; DOLG, M. Global optimization of clusters of rigid molecules using the artificial bee colony algorithm. *Phys. Chem. Chem. Phys.*, The Royal Society of Chemistry, v. 18, p. 3003–3010, 2016. Disponível em: <http://dx.doi.org/10.1039/C5CP06313B>.
- 37 ZHANG, J.; DOLG, M. Abcluster: the artificial bee colony algorithm for cluster global optimization. *Phys. Chem. Chem. Phys.*, The Royal Society of Chemistry, v. 17, p. 24173–24181, 2015. Disponível em: <http://dx.doi.org/10.1039/C5CP04060D>.
- 38 CALL, S. T.; ZUBAREV, D. Y.; BOLDYREV, A. I. Global minimum structure searches via particle swarm optimization. *Journal of Computational Chemistry*, v. 28, n. 7, p. 1177–1186, 2007. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.20621>.
- 39 CHEN, H.-Z. et al. Predicting new tio₂ phases with low band gaps by a multiobjective global optimization approach. *The Journal of Physical Chemistry C*, v. 118, n. 5, p. 2333–2337, 2014. Disponível em: <https://doi.org/10.1021/jp411437f>.
- 40 LI, Z.; SCHERAGA, H. A. Structure and free energy of complex thermodynamic systems. *Journal of Molecular Structure: THEOCHEM*, v. 179, n. 1, p. 333–352, 1988. ISSN 0166-1280. Disponível em: <https://www.sciencedirect.com/science/article/pii/0166128088801337>.
- 41 HARTKE, B. Global geometry optimization of clusters using genetic algorithms. *The Journal of Physical Chemistry*, v. 97, n. 39, p. 9973–9976, 1993. Disponível em: <https://doi.org/10.1021/j100141a013>.
- 42 DEAVEN, D. M.; HO, K. M. Molecular geometry optimization with a genetic algorithm. *Phys. Rev. Lett.*, American Physical Society, v. 75, p. 288–291, Jul 1995. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.75.288>.
- 43 CRAENEN, B.; EIBEN, A.; MARCHIORI, E. How to handle constraints with evolutionary algorithms. In: *The Practical Handbook of Genetic Algorithms*. [S.l.]: Chapman and Hall/CRC, 2000.
- 44 GUIMARÃES, F. F. et al. Global optimization analysis of water clusters (h₂o)_n (11<n<13) through a genetic evolutionary approach. *The Journal of Chemical Physics*, v. 116, n. 19, p. 8327–8333, 2002. Disponível em: <https://aip.scitation.org/doi/abs/10.1063/1.1471240>.
- 45 SILVA, E. S. de Abreu e; DUARTE, H. A.; BELCHIOR, J. C. An approach based on genetic algorithms and dft for studying clusters: (h₂o)_n (2<n<13) cluster analysis. *Chemical Physics*, v. 323, n. 2, p. 553 – 562, 2006. ISSN 0301-0104. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0301010405005379>.
- 46 GUIMARÃES, F. F. et al. Global optimization analysis of water clusters (h₂o)_n (11≤n≤13) through a genetic evolutionary approach. *The Journal of Chemical Physics*, v. 116, n. 19, p. 8327–8333, 2002. Disponível em: <https://aip.scitation.org/doi/abs/10.1063/1.1471240>.
- 47 RODRIGUES, D. D. et al. Global optimization analysis of cunaum (n+m=38) clusters: Complementary ab initio calculations. *Chemical Physics*, v. 349, n. 1, p. 91–97, 2008. ISSN 0301-0104. Electron Correlation and

Molecular Dynamics for Excited States and Photochemistry. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0301010408001730>.

48 SILVA, M.; GALVÃO, B.; BELCHIOR, J. Growth analysis of sodium-potassium alloy clusters from 7 to 55 atoms through a genetic algorithm approach. *Journal of molecular modeling*, v. 20, p. 2421, 09 2014.

49 DAVIS, J. B. A.; HORSWELL, S. L.; JOHNSTON, R. L. Global optimization of 8–10 atom palladium–iridium nanoalloys at the dft level. *The Journal of Physical Chemistry A*, v. 118, n. 1, p. 208–214, 2014. PMID: 24328270. Disponível em: <https://doi.org/10.1021/jp408519z>.

50 SILVA, F. T. et al. Application of a quantum genetic algorithm and qtaim analysis in the study of structural and electronic properties of neutral bimetallic clusters $naxliy$ ($4 \leq x + y \leq 10$). *Journal of Molecular Modeling*, v. 26, n. 11, p. 317, Oct 2020. ISSN 0948-5023. Disponível em: <https://doi.org/10.1007/s00894-020-04576-1>.

51 JONES, R. O. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.*, American Physical Society, v. 87, p. 897–923, Aug 2015. Disponível em: <https://link.aps.org/doi/10.1103/RevModPhys.87.897>.

52 SILVA, F. T.; SILVA, M. X.; BELCHIOR, J. C. A new genetic algorithm approach applied to atomic and molecular cluster studies. *Frontiers in chemistry*, v. 7, p. 707, 2019. ISSN 2296-2646. Disponível em: <https://europemc.org/articles/PMC6848380>.

53 VILHELMOSEN, L. B.; HAMMER, B. Systematic study of au_6 to au_{12} gold clusters on $mgo(100)$ f centers using density-functional theory. *Phys. Rev. Lett.*, American Physical Society, v. 108, p. 126101, Mar 2012. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.108.126101>.

54 LAZAUSKAS, T.; SOKOL, A. A.; WOODLEY, S. M. An efficient genetic algorithm for structure prediction at the nanoscale. *Nanoscale*, The Royal Society of Chemistry, v. 9, p. 3850–3864, 2017. Disponível em: <http://dx.doi.org/10.1039/C6NR09072A>.

55 GASTEIGER, J.; ZUPAN, J. Neural networks in chemistry. *Angewandte Chemie International Edition in English*, Wiley, v. 32, n. 4, p. 503–527, abr. 1993. ISSN 0570-0833. Disponível em: <http://dx.doi.org/10.1002/anie.199305031>.

56 BLANK, T. B. et al. Neural network models of potential energy surfaces. *The Journal of Chemical Physics*, AIP Publishing, v. 103, n. 10, p. 4129–4137, set. 1995. Disponível em: <https://doi.org/10.1063/1.469597>.

57 ZOU, J.; HAN, Y.; SO, S.-S. Overview of artificial neural networks. In: _____. *Artificial Neural Networks*. Humana Press, 2008. p. 14–22. ISBN 9781603271011. Disponível em: http://dx.doi.org/10.1007/978-1-60327-101-1_2.

58 AGGARWAL, C. C. *Neural Networks and Deep Learning: A Textbook*. Springer International Publishing, 2018. ISBN 9783319944630. Disponível em: <http://dx.doi.org/10.1007/978-3-319-94463-0>.

59 GALUSHKA, M. et al. Prediction of chemical compounds properties using a deep learning model. *Neural Computing and Applications*, Springer Science and Business

- Media LLC, v. 33, n. 20, p. 13345–13366, jun. 2021. ISSN 1433-3058. Disponível em: <http://dx.doi.org/10.1007/s00521-021-05961-4>.
- 60 MILJKOVIĆ, F.; RODRÍGUEZ-PÉREZ, R.; BAJORATH, J. Impact of artificial intelligence on compound discovery, design, and synthesis. *ACS Omega*, v. 6, n. 49, p. 33293–33299, 2021. Disponível em: <https://doi.org/10.1021/acsomega.1c05512>.
- 61 HEIN, J. E. Machine learning made easy for optimizing chemical reactions. *Nature*, Springer Science and Business Media LLC, v. 590, n. 7844, p. 40–41, fev. 2021. Disponível em: <https://doi.org/10.1038/d41586-021-00209-6>.
- 62 COVA, T. F. G. G.; PAIS, A. A. C. C. Deep learning for deep chemistry: Optimizing the prediction of chemical patterns. *Frontiers in Chemistry*, Frontiers Media SA, v. 7, nov. 2019. Disponível em: <https://doi.org/10.3389/fchem.2019.00809>.
- 63 UNKE, O. T. et al. High-dimensional potential energy surfaces for molecular simulations: from empiricism to machine learning. *Machine Learning: Science and Technology*, IOP Publishing, v. 1, n. 1, p. 013001, fev. 2020. Disponível em: <https://doi.org/10.1088/2632-2153/ab5922>.
- 64 AMABILINO, S. et al. Training neural nets to learn reactive potential energy surfaces using interactive quantum chemistry in virtual reality. *The Journal of Physical Chemistry A*, American Chemical Society (ACS), v. 123, n. 20, p. 4486–4499, mar. 2019. Disponível em: <https://doi.org/10.1021/acs.jpca.9b01006>.
- 65 BEHLER, J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. *The Journal of Chemical Physics*, v. 134, n. 7, p. 074106, 2011. Disponível em: <https://doi.org/10.1063/1.3553717>.
- 66 LEMM, D.; RUDORFF, G. F. von; LILIENFELD, O. A. von. Machine learning based energy-free structure predictions of molecules, transition states, and solids. *Nature Communications*, Springer Science and Business Media LLC, v. 12, n. 1, jul. 2021. Disponível em: <https://doi.org/10.1038/s41467-021-24525-7>.
- 67 DIRISIO, R. et al. DIFFUSION MONTE CARLO USING MACHINE LEARNING POTENTIAL ENERGY SURFACES. University of Illinois at Urbana-Champaign, jun. 2021. Disponível em: <https://doi.org/10.15278/isms.2021.wk02>.
- 68 SILVA, F. et al. Exploring the mp2 energy surface of nanoalloy clusters with a genetic algorithm: Application to sodium–potassium. *Chemical Physics Letters*, v. 639, p. 135–141, 2015. ISSN 0009-2614. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0009261415006946>.
- 69 SILVA, M. X. et al. A genetic algorithm survey on closed-shell atomic nitrogen clusters employing a quantum chemical approach. *Journal of Molecular Modeling*, Springer Science and Business Media LLC, v. 24, n. 8, jul. 2018. ISSN 0948-5023. Disponível em: <http://dx.doi.org/10.1007/s00894-018-3724-6>.
- 70 MICHALEWICZ, Z. *Genetic Algorithms + Data Structures = Evolution Programs*. Springer Berlin Heidelberg, 1996. ISBN 9783662033159. Disponível em: <http://dx.doi.org/10.1007/978-3-662-03315-9>.

- 71 MEYER-BAESE, A.; SCHMID, V. Chapter 5 - genetic algorithms. In: MEYER-BAESE, A.; SCHMID, V. (Ed.). *Pattern Recognition and Signal Analysis in Medical Imaging (Second Edition)*. Second edition. Oxford: Academic Press, 2014. p. 135 – 149. ISBN 978-0-12-409545-8. Disponível em: <http://www.sciencedirect.com/science/article/pii/B9780124095458000054>.
- 72 WHITLEY, D. A genetic algorithm tutorial. *Statistics and Computing*, v. 4, n. 2, p. 65–85, Jun 1994. ISSN 1573-1375. Disponível em: <https://doi.org/10.1007/BF00175354>.
- 73 LORDEIRO, R. A. et al. Determination of main structural compositions of nanoalloy clusters of Cu_xAg_y (x + y < 30) using a genetic algorithm approach. *International Journal of Quantum Chemistry*, v. 95, n. 2, p. 112–125, 2003. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/qua.10660>.
- 74 PEREIRA, F. B. et al. Designing efficient evolutionary algorithms for cluster optimization: A study on locality. In: _____. *Advances in Metaheuristics for Hard Optimization*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. p. 223–250. ISBN 978-3-540-72960-0. Disponível em: https://doi.org/10.1007/978-3-540-72960-0_11.
- 75 CHEN, Z. et al. A sphere-cut-splice crossover for the evolution of cluster structures. *The Journal of Chemical Physics*, v. 138, n. 21, p. 214303, 2013. Disponível em: <https://doi.org/10.1063/1.4807091>.
- 76 JOHNSTON, R. L. Evolving better nanoparticles: Genetic algorithms for optimising cluster geometries. *Dalton Trans.*, The Royal Society of Chemistry, p. 4193–4207, 2003. Disponível em: <http://dx.doi.org/10.1039/B305686D>.
- 77 WALES, D. J.; DOYE, J. P. K. Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *The Journal of Physical Chemistry A*, v. 101, n. 28, p. 5111–5116, 1997. Disponível em: <https://doi.org/10.1021/jp970984n>.
- 78 RONDINA, G. G.; SILVA, J. L. F. D. Revised basin-hopping Monte Carlo algorithm for structure optimization of clusters and nanoparticles. *Journal of Chemical Information and Modeling*, v. 53, n. 9, p. 2282–2298, 2013. PMID: 23957311. Disponível em: <https://doi.org/10.1021/ci400224z>.
- 79 TAKEUCHI, H. Novel method for geometry optimization of molecular clusters: Application to benzene clusters. *Journal of Chemical Information and Modeling*, v. 47, n. 1, p. 104–109, 2007. PMID: 17238254. Disponível em: <https://doi.org/10.1021/ci600336p>.
- 80 YE, T.; XU, R.; HUANG, W. Global optimization of binary Lennard-Jones clusters using three perturbation operators. *Journal of Chemical Information and Modeling*, v. 51, n. 3, p. 572–577, 2011. PMID: 21332209. Disponível em: <https://doi.org/10.1021/ci1004256>.
- 81 JOHNSTON, R. L.; ROBERTS, C. Genetic algorithms for the geometry optimization of clusters and nanoparticles. In: *Soft Computing Approaches in Chemistry*. Springer Berlin Heidelberg, 2003. p. 161–204. Disponível em: https://doi.org/10.1007/978-3-540-36213-5_7.

- 82 KATOCH, S.; CHAUHAN, S. S.; KUMAR, V. A review on genetic algorithm: past, present, and future. *Multimedia Tools and Applications*, Springer Science and Business Media LLC, v. 80, n. 5, p. 8091–8126, out. 2020. Disponível em: <https://doi.org/10.1007/s11042-020-10139-6>.
- 83 ZETTILI, N. *Quantum Mechanics: Concepts and Applications*. Paperback. Wiley, 2009. 688 p. ISBN 978-0470026793. Disponível em: <https://lead.to/amazon/com/?op=bt&la=en&cu=usd&key=0470026790>.
- 84 PRATT, C. J. *Quantum physics for beginners: From Wave Theory to Quantum Computing. Understanding How Everything Works by a Simplified Explanation of Quantum Physics and Mechanics Principles*. Hardcover. Carl J. Pratt, 2021. 104 p. ISBN 978-1802356588. Disponível em: <https://lead.to/amazon/com/?op=bt&la=en&cu=usd&key=1802356584>.
- 85 SZABO, N. S. O. A. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. [S.l.]: Dover Publications, 1996.
- 86 GRIFFITHS, D. J. *Introduction to Quantum Mechanics*. [S.l.]: Pearson Prentice Hall, 2004.
- 87 HARTREE, D. R. The wave mechanics of an atom with a non-coulomb central field. part i. theory and methods. *Mathematical Proceedings of the Cambridge Philosophical Society*, Cambridge University Press, v. 24, n. 1, p. 89–110, 1928. Disponível em: <https://doi.org/10.1017/S0305004100011919>.
- 88 HARTREE, D. R. The wave mechanics of an atom with a non-coulomb central field. part ii. some results and discussion. *Mathematical Proceedings of the Cambridge Philosophical Society*, Cambridge University Press, v. 24, n. 1, p. 111–132, 1928. Disponível em: <https://doi.org/10.1017/S0305004100011920>.
- 89 FADDEEV L.A. KHALFIN, I. K. L. *V.A. Fock - Selected Works: Quantum Mechanics and Quantum Field Theory*. [S.l.]: CRC Press, 2004.
- 90 Dirac, P. A. M. A new notation for quantum mechanics. *Proceedings of the Cambridge Philosophical Society*, v. 35, p. 416, 1939. Disponível em: <https://ui.adsabs.harvard.edu/abs/1939PCPS...35..416D>.
- 91 ROTHHAAN, C. C. J. New developments in molecular orbital theory. *Rev. Mod. Phys.*, American Physical Society, v. 23, p. 69–89, Apr 1951. Disponível em: <https://link.aps.org/doi/10.1103/RevModPhys.23.69>.
- 92 AN Introduction to Hartree-Fock Molecular Orbital Theory. 2000. <http://vergil.chemistry.gatech.edu/notes/hf-intro/hf-intro.pdf>.
- 93 CRAMER, C. J. *Essentials of Computational Chemistry: Theories and Models*. [S.l.]: Wiley; 2 edition, 2004.
- 94 FINOCCHI, F. Density functional theory for beginners: Basic principles and practical approaches. *Institut des NanoSciences de Paris (INSP) CNRS and University Pierre et Marie Curie*, Oct 2011. Disponível em: http://www.attacalite.com/wp-content/uploads/2017/04/pdf_DFT4beginners.pdf.

- 95 HOHENBERG, P.; KOHN, W. Inhomogeneous electron gas. *Phys. Rev.*, American Physical Society, v. 136, p. B864–B871, Nov 1964. Disponível em: <https://link.aps.org/doi/10.1103/PhysRev.136.B864>.
- 96 CHERMETTE, H. Density functional theory: A powerful tool for theoretical studies in coordination chemistry. *Coordination Chemistry Reviews*, v. 178-180, p. 699 – 721, 1998. ISSN 0010-8545. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0010854598001799>.
- 97 KOHN, W.; SHAM, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, American Physical Society, v. 140, p. A1133–A1138, Nov 1965. Disponível em: <https://link.aps.org/doi/10.1103/PhysRev.140.A1133>.
- 98 KATO, T. On the eigenfunctions of many-particle systems in quantum mechanics. *Communications on Pure and Applied Mathematics*, v. 10, n. 2, p. 151–177, 1957. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.3160100201>.
- 99 GRIMME, S. Semiempirical hybrid density functional with perturbative second-order correlation. *The Journal of Chemical Physics*, v. 124, n. 3, p. 034108, 2006. Disponível em: <https://doi.org/10.1063/1.2148954>.
- 100 ZIMMERLI, U.; PARRINELLO, M.; KOUMOUTSAKOS, P. Dispersion corrections to density functionals for water aromatic interactions. *The Journal of Chemical Physics*, v. 120, n. 6, p. 2693–2699, 2004. Disponível em: <https://doi.org/10.1063/1.1637034>.
- 101 GRIMME, S. Accurate description of van der waals complexes by density functional theory including empirical corrections. *Journal of Computational Chemistry*, v. 25, n. 12, p. 1463–1473, 2004. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.20078>.
- 102 PERDEW, J. P.; BURKE, K.; ERNZERHOF, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, American Physical Society, v. 77, p. 3865–3868, Oct 1996. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.77.3865>.
- 103 ŚMIGA, S.; CONSTANTIN, L. A. Unveiling the physics behind hybrid functionals. *The Journal of Physical Chemistry A*, American Chemical Society (ACS), v. 124, n. 27, p. 5606–5614, jun. 2020. ISSN 1520-5215. Disponível em: <http://dx.doi.org/10.1021/acs.jpca.0c04156>.
- 104 ON the determination of molecular fields.—I. From the variation of the viscosity of a gas with temperature. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, The Royal Society, v. 106, n. 738, p. 441–462, out. 1924. Disponível em: <https://doi.org/10.1098/rspa.1924.0081>.
- 105 ON the determination of molecular fields. —II. From the equation of state of a gas. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, The Royal Society, v. 106, n. 738, p. 463–477, out. 1924. Disponível em: <https://doi.org/10.1098/rspa.1924.0082>.
- 106 LENNARD-JONES, J. E. Cohesion. *Proceedings of the Physical Society*, IOP Publishing, v. 43, n. 5, p. 461–482, set. 1931. Disponível em: <https://doi.org/10.1088/0959-5309/43/5/301>.

- 107 WOOD, W. W.; PARKER, F. R. Monte carlo equation of state of molecules interacting with the lennard-jones potential. i. a supercritical isotherm at about twice the critical temperature. *The Journal of Chemical Physics*, AIP Publishing, v. 27, n. 3, p. 720–733, set. 1957. Disponível em: <https://doi.org/10.1063/1.1743822>.
- 108 ROSENBLUTH, M. N.; ROSENBLUTH, A. W. Further results on monte carlo equations of state. *The Journal of Chemical Physics*, AIP Publishing, v. 22, n. 5, p. 881–884, maio 1954. Disponível em: <https://doi.org/10.1063/1.1740207>.
- 109 ALDER, B. J.; WAINWRIGHT, T. E. Studies in molecular dynamics. i. general method. *The Journal of Chemical Physics*, AIP Publishing, v. 31, n. 2, p. 459–466, ago. 1959. Disponível em: <https://doi.org/10.1063/1.1730376>.
- 110 STEPHAN, S. et al. Thermophysical properties of the lennard-jones fluid: Database and data assessment. *Journal of Chemical Information and Modeling*, American Chemical Society (ACS), v. 59, n. 10, p. 4248–4265, out. 2019. Disponível em: <https://doi.org/10.1021/acs.jcim.9b00620>.
- 111 STEPHAN, S.; STAUBACH, J.; HASSE, H. Review and comparison of equations of state for the lennard-jones fluid. *Fluid Phase Equilibria*, Elsevier BV, v. 523, p. 112772, nov. 2020. Disponível em: <https://doi.org/10.1016/j.fluid.2020.112772>.
- 112 JORGENSEN, W. L.; MAXWELL, D. S.; TIRADO-RIVES, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, American Chemical Society (ACS), v. 118, n. 45, p. 11225–11236, nov. 1996. Disponível em: <https://doi.org/10.1021/ja9621760>.
- 113 STEPHAN, S. et al. MolMod – an open access database of force fields for molecular simulations of fluids. *Molecular Simulation*, Informa UK Limited, v. 45, n. 10, p. 806–814, abr. 2019. Disponível em: <https://doi.org/10.1080/08927022.2019.1601191>.
- 114 CLERI, F.; ROSATO, V. Tight-binding potentials for transition metals and alloys. *Physical Review B*, American Physical Society (APS), v. 48, n. 1, p. 22–33, jul. 1993. Disponível em: <https://doi.org/10.1103/physrevb.48.22>.
- 115 ROSATO, V.; GUILLOPE, M.; LEGRAND, B. Thermodynamical and structural properties of f.c.c. transition metals using a simple tight-binding model. *Philosophical Magazine A*, Informa UK Limited, v. 59, n. 2, p. 321–336, fev. 1989. Disponível em: <https://doi.org/10.1080/01418618908205062>.
- 116 GUPTA, R. P. Lattice relaxation at a metal surface. *Physical Review B*, American Physical Society (APS), v. 23, n. 12, p. 6265–6270, jun. 1981. Disponível em: <https://doi.org/10.1103/physrevb.23.6265>.
- 117 SUTTON, A. P. et al. The tight-binding bond model. *Journal of Physics C: Solid State Physics*, IOP Publishing, v. 21, n. 1, p. 35–66, jan. 1988. ISSN 0022-3719. Disponível em: <http://dx.doi.org/10.1088/0022-3719/21/1/007>.
- 118 HALDAR, A.; MAHADEVAN, S. *Probability, reliability, and statistical methods in engineering design*. Nashville, TN: John Wiley & Sons, 1999.

- 119 HSU, P. J.; LAI, S. K. Structures of bimetallic clusters. *The Journal of Chemical Physics*, AIP Publishing, v. 124, n. 4, p. 044711, jan. 2006. Disponível em: <https://doi.org/10.1063/1.2147159>.
- 120 CHEN et al. Structure, melting, and thermal stability of 55 atom ag-au nanoalloys. *The Journal of Physical Chemistry C*, American Chemical Society (ACS), v. 111, n. 26, p. 9157–9165, jun. 2007. Disponível em: <https://doi.org/10.1021/jp0717746>.
- 121 PAZ-BORBÓN, L. O. et al. Structural motifs, mixing, and segregation effects in 38-atom binary clusters. *The Journal of Chemical Physics*, AIP Publishing, v. 128, n. 13, p. 134517, abr. 2008. Disponível em: <https://doi.org/10.1063/1.2897435>.
- 122 MORSE, P. M. Diatomic molecules according to the wave mechanics. II. vibrational levels. *Physical Review*, American Physical Society (APS), v. 34, n. 1, p. 57–64, jul. 1929. Disponível em: <https://doi.org/10.1103/physrev.34.57>.
- 123 WU, X. et al. Structural optimization of ag–pd clusters based on different potential parameterizations. *Chemical Physics*, Elsevier BV, v. 390, n. 1, p. 36–41, nov. 2011. Disponível em: <https://doi.org/10.1016/j.chemphys.2011.09.023>.
- 124 BRUMA, A. et al. DFT study of the structures and energetics of 98-atom AuPd clusters. *Nanoscale*, Royal Society of Chemistry (RSC), v. 5, n. 2, p. 646–652, 2013. Disponível em: <https://doi.org/10.1039/c2nr32517a>.
- 125 MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, Springer Science and Business Media LLC, v. 5, n. 4, p. 115–133, dez. 1943. ISSN 1522-9602. Disponível em: <http://dx.doi.org/10.1007/BF02478259>.
- 126 HEBB, D. O. *The organization of behavior*. Philadelphia, PA: Psychology Press, 2002.
- 127 SAMUEL, A. Computing bit by bit or digital computers made easy. *Proceedings of the IRE*, Institute of Electrical and Electronics Engineers (IEEE), v. 41, n. 10, p. 1223–1230, out. 1953. Disponível em: <https://doi.org/10.1109/jrproc.1953.274271>.
- 128 ROSENBLATT, F. *The perceptron - A perceiving and recognizing automaton*. Ithaca, New York, 1957.
- 129 ROSENBLATT, F. *Two Theorems of Statistical Separability in the Perceptron: (Project Para) Contract Nonr-2381(00)*. Cornell Aeronautical Laboratory, 1958. (Cornell Aeronautical Laboratory, Inc). Disponível em: <https://books.google.com.br/books?id=5Q0-AQAAIAAJ>.
- 130 ROSENBLATT, F. Perceptron simulation experiments. *Proceedings of the IRE*, v. 48, p. 301–309, 1960.
- 131 ANDREAE, J.; CASHIN, P. A learning machine with monologue. *International Journal of Man-Machine Studies*, Elsevier BV, v. 1, n. 1, p. 1–20, jan. 1969. ISSN 0020-7373. Disponível em: [http://dx.doi.org/10.1016/S0020-7373\(69\)80008-8](http://dx.doi.org/10.1016/S0020-7373(69)80008-8).
- 132 SUTTON, R. S.; BARTO, A. G. *Reinforcement Learning*. Cambridge, MA: Bradford Books, 1998. (Adaptive Computation and Machine Learning series).

- 133 FRADKOV, A. L. Early history of machine learning. *IFAC-PapersOnLine*, Elsevier BV, v. 53, n. 2, p. 1385–1390, 2020. Disponível em: <https://doi.org/10.1016/j.ifacol.2020.12.1888>.
- 134 WANG, H.; RAJ, B. *On the Origin of Deep Learning*. 2017. Disponível em: <https://arxiv.org/abs/1702.07800>.
- 135 SHINDE, P. P.; SHAH, S. A review of machine learning and deep learning applications. In: *2018 Fourth International Conference on Computing Communication Control and Automation (ICCCUBEA)*. IEEE, 2018. Disponível em: <https://doi.org/10.1109/iccubea.2018.8697857>.
- 136 KEITH, J. A. et al. Combining machine learning and computational chemistry for predictive insights into chemical systems. *Chemical Reviews*, American Chemical Society (ACS), v. 121, n. 16, p. 9816–9872, jul. 2021. Disponível em: <https://doi.org/10.1021/acs.chemrev.1c00107>.
- 137 PINHEIRO, G. A. et al. Machine learning prediction of nine molecular properties based on the smiles representation of the qm9 quantum-chemistry dataset. *The Journal of Physical Chemistry A*, American Chemical Society (ACS), v. 124, n. 47, p. 9854–9866, nov. 2020. ISSN 1520-5215. Disponível em: <http://dx.doi.org/10.1021/acs.jpca.0c05969>.
- 138 SCHÜTT, K. T. et al. Unifying machine learning and quantum chemistry with a deep neural network for molecular wavefunctions. *Nature Communications*, Springer Science and Business Media LLC, v. 10, n. 1, nov. 2019. ISSN 2041-1723. Disponível em: <http://dx.doi.org/10.1038/s41467-019-12875-2>.
- 139 WESTERMAYR, J. et al. Perspective on integrating machine learning into computational chemistry and materials science. *The Journal of Chemical Physics*, AIP Publishing, v. 154, n. 23, jun. 2021. ISSN 1089-7690. Disponível em: <http://dx.doi.org/10.1063/5.0047760>.
- 140 OLIVEIRA, E. C. L. de et al. Predicting cell-penetrating peptides using machine learning algorithms and navigating in their chemical space. *Scientific Reports*, Springer Science and Business Media LLC, v. 11, n. 1, abr. 2021. ISSN 2045-2322. Disponível em: <http://dx.doi.org/10.1038/s41598-021-87134-w>.
- 141 LOURENCO, M. P. et al. Qmlmaterial-a quantum machine learning software for material design and discovery. *Journal of Chemical Theory and Computation*, American Chemical Society (ACS), v. 19, n. 17, p. 5999–6010, ago. 2023. ISSN 1549-9626. Disponível em: <http://dx.doi.org/10.1021/acs.jctc.3c00566>.
- 142 SHOKRY, A. et al. A machine learning-based methodology for multi-parametric solution of chemical processes operation optimization under uncertainty. *Chemical Engineering Journal*, Elsevier BV, v. 425, p. 131632, dez. 2021. Disponível em: <https://doi.org/10.1016/j.cej.2021.131632>.
- 143 GLIELMO, A. et al. Unsupervised learning methods for molecular simulation data. *Chemical Reviews*, American Chemical Society (ACS), v. 121, n. 16, p. 9722–9758, maio 2021. Disponível em: <https://doi.org/10.1021/acs.chemrev.0c01195>.

- 144 TETEF, S.; GOVIND, N.; SEIDLER, G. T. Unsupervised machine learning for unbiased chemical classification in x-ray absorption spectroscopy and x-ray emission spectroscopy. *Phys. Chem. Chem. Phys.*, The Royal Society of Chemistry, v. 23, p. 23586–23601, 2021. Disponível em: <http://dx.doi.org/10.1039/D1CP02903G>.
- 145 BOSER, B. E.; GUYON, I. M.; VAPNIK, V. N. A training algorithm for optimal margin classifiers. In: *Proceedings of the fifth annual workshop on Computational learning theory*. ACM, 1992. Disponível em: <https://doi.org/10.1145/130385.130401>.
- 146 SCHÖLKOPF, B.; SMOLA, A.; MÜLLER, K.-R. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, MIT Press - Journals, v. 10, n. 5, p. 1299–1319, jul. 1998. Disponível em: <https://doi.org/10.1162/089976698300017467>.
- 147 PREFACE. In: *PATTERN Recognition*. [S.l.]: Elsevier, 2009. p. xv–xvii.
- 148 RUPP, M. et al. Fast and accurate modeling of molecular atomization energies with machine learning. arXiv, 2011. Disponível em: <https://arxiv.org/abs/1109.2618>.
- 149 BEHLER, J.; PARRINELLO, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.*, American Physical Society, v. 98, p. 146401, Apr 2007. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.98.146401>.
- 150 BARTÓK, A. P.; KONDOR, R.; CSÁNYI, G. On representing chemical environments. *Physical Review B*, American Physical Society (APS), v. 87, n. 18, maio 2013. Disponível em: <https://doi.org/10.1103/physrevb.87.184115>.
- 151 FABER, F. A. et al. Alchemical and structural distribution based representation for universal quantum machine learning. *The Journal of Chemical Physics*, AIP Publishing, v. 148, n. 24, p. 241717, jun. 2018. Disponível em: <https://doi.org/10.1063/1.5020710>.
- 152 CHRISTENSEN, A. S. et al. FCHL revisited: Faster and more accurate quantum machine learning. *The Journal of Chemical Physics*, AIP Publishing, v. 152, n. 4, p. 044107, jan. 2020. Disponível em: <https://doi.org/10.1063/1.5126701>.
- 153 SCHÖLKOPF, B.; HERBRICH, R.; SMOLA, A. J. A generalized representer theorem. In: *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2001. p. 416–426. Disponível em: https://doi.org/10.1007/3-540-44581-1_27.
- 154 BERLINET, A.; THOMAS-AGNAN, C. *Reproducing kernel Hilbert spaces in probability and statistics*. 2004. ed. New York, NY: Springer, 2011.
- 155 HOLLEBEEK, T.; HO, T.-S.; RABITZ, H. Constructing multidimensional molecular potential energy surfaces from ab initio data. *Annual Review of Physical Chemistry*, Annual Reviews, v. 50, n. 1, p. 537–570, out. 1999. Disponível em: <https://doi.org/10.1146/annurev.physchem.50.1.537>.
- 156 SOLDÁN, P.; HUTSON, J. M. On the long-range and short-range behavior of potentials from reproducing kernel hilbert space interpolation. *The Journal of Chemical Physics*, AIP Publishing, v. 112, n. 9, p. 4415–4416, mar. 2000. Disponível em: <https://doi.org/10.1063/1.480988>.

- 157 MULLER, K.-R. et al. An introduction to kernel-based learning algorithms. *IEEE Transactions on Neural Networks*, Institute of Electrical and Electronics Engineers (IEEE), v. 12, n. 2, p. 181–201, mar. 2001. Disponível em: <https://doi.org/10.1109/72.914517>.
- 158 HOFMANN, T.; SCHÖLKOPF, B.; SMOLA, A. J. Kernel methods in machine learning. *The Annals of Statistics*, Institute of Mathematical Statistics, v. 36, n. 3, jun. 2008. Disponível em: <https://doi.org/10.1214/009053607000000677>.
- 159 GOLUB, G. H.; LOAN, C. F. V. *Matrix Computations (Johns Hopkins Studies in the Mathematical Sciences, 3)*. Hardcover. Johns Hopkins University Press, 2013. 784 p. ISBN 978-1421407944. Disponível em: <https://lead.to/amazon/com/?op=bt&la=en&cu=usd&key=1421407949>.
- 160 CHEN, N. et al. *Support Vector Machine in Chemistry*. Hardcover. World Scientific Publishing Company, 2004. 344 p. ISBN 978-9812389220. Disponível em: <https://lead.to/amazon/com/?op=bt&la=en&cu=usd&key=9812389229>.
- 161 CORTES, C.; VAPNIK, V. Support-vector networks. *Machine Learning*, Springer Science and Business Media LLC, v. 20, n. 3, p. 273–297, set. 1995. Disponível em: <https://doi.org/10.1007/bf00994018>.
- 162 LI, W. et al. Machine learning corrections for DFT noncovalent interactions. In: *Computational Materials, Chemistry, and Biochemistry: From Bold Initiatives to the Last Mile*. Springer International Publishing, 2021. p. 183–212. Disponível em: https://doi.org/10.1007/978-3-030-18778-1_10.
- 163 BALABIN, R. M.; LOMAKINA, E. I. Support vector machine regression (LS-SVM)—an alternative to artificial neural networks (ANNs) for the analysis of quantum chemistry data? *Physical Chemistry Chemical Physics*, Royal Society of Chemistry (RSC), v. 13, n. 24, p. 11710, 2011. Disponível em: <https://doi.org/10.1039/c1cp00051a>.
- 164 KOHONEN, T. An introduction to neural computing. *Neural Networks*, Elsevier BV, v. 1, n. 1, p. 3–16, jan. 1988. Disponível em: [https://doi.org/10.1016/0893-6080\(88\)90020-2](https://doi.org/10.1016/0893-6080(88)90020-2).
- 165 M., C. *Neural Networks for Pattern Recognition*. Oxford, England: Clarendon Press, 1995.
- 166 HAYKIN, S. O. *Neural Networks and Learning Machines*. 3. ed. Upper Saddle River, NJ: Pearson, 2008.
- 167 CYBENKO, G. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*, Springer Science and Business Media LLC, v. 2, n. 4, p. 303–314, dez. 1989. Disponível em: <https://doi.org/10.1007/bf02551274>.
- 168 HORNIK, K. Approximation capabilities of multilayer feedforward networks. *Neural Networks*, Elsevier BV, v. 4, n. 2, p. 251–257, 1991. Disponível em: [https://doi.org/10.1016/0893-6080\(91\)90009-t](https://doi.org/10.1016/0893-6080(91)90009-t).
- 169 ELDAN, R.; SHAMIR, O. *The Power of Depth for Feedforward Neural Networks*. arXiv, 2015. Disponível em: <https://arxiv.org/abs/1512.03965>.

- 170 BROWN, D. F. R.; GIBBS, M. N.; CLARY, D. C. Combining ab initio computations, neural networks, and diffusion monte carlo: An efficient method to treat weakly bound molecules. *The Journal of Chemical Physics*, AIP Publishing, v. 105, n. 17, p. 7597–7604, nov. 1996. Disponível em: <https://doi.org/10.1063/1.472596>.
- 171 TAFEIT, E. et al. Neural networks as a tool for compact representation of ab initio molecular potential energy surfaces. *Journal of Molecular Graphics*, Elsevier BV, v. 14, n. 1, p. 12–18, fev. 1996. Disponível em: [https://doi.org/10.1016/0263-7855\(95\)00087-9](https://doi.org/10.1016/0263-7855(95)00087-9).
- 172 MANZHOS, S.; CARRINGTON, T. A random-sampling high dimensional model representation neural network for building potential energy surfaces. *The Journal of Chemical Physics*, AIP Publishing, v. 125, n. 8, p. 084109, ago. 2006. Disponível em: <https://doi.org/10.1063/1.2336223>.
- 173 MANZHOS, S.; CARRINGTON, T. Using redundant coordinates to represent potential energy surfaces with lower-dimensional functions. *The Journal of Chemical Physics*, AIP Publishing, v. 127, n. 1, p. 014103, jul. 2007. Disponível em: <https://doi.org/10.1063/1.2746846>.
- 174 MALSHE, M. et al. Development of generalized potential-energy surfaces using many-body expansions, neural networks, and moiety energy approximations. *The Journal of Chemical Physics*, AIP Publishing, v. 130, n. 18, p. 184102, 2009. Disponível em: <https://doi.org/10.1063/1.3124802>.
- 175 GRISONI, F. et al. Molecular descriptors for structure–activity applications: A hands-on approach. In: _____. *Computational Toxicology: Methods and Protocols*. New York, NY: Springer New York, 2018. p. 3–53. ISBN 978-1-4939-7899-1. Disponível em: https://doi.org/10.1007/978-1-4939-7899-1_1.
- 176 KHORSHIDI, A.; PETERSON, A. A. Amp: A modular approach to machine learning in atomistic simulations. *Computer Physics Communications*, Elsevier BV, v. 207, p. 310–324, out. 2016. Disponível em: <https://doi.org/10.1016/j.cpc.2016.05.010>.
- 177 ARTRITH, N.; URBAN, A.; CEDER, G. Efficient and accurate machine-learning interpolation of atomic energies in compositions with many species. *Physical Review B*, American Physical Society (APS), v. 96, n. 1, jul. 2017. Disponível em: <https://doi.org/10.1103/physrevb.96.014112>.
- 178 UNKE, O. T.; MEUWLY, M. A reactive, scalable, and transferable model for molecular energies from a neural network approach based on local information. *The Journal of Chemical Physics*, AIP Publishing, v. 148, n. 24, p. 241708, jun. 2018. Disponível em: <https://doi.org/10.1063/1.5017898>.
- 179 SMITH, J. S.; ISAYEV, O.; ROITBERG, A. E. ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost. *Chemical Science*, Royal Society of Chemistry (RSC), v. 8, n. 4, p. 3192–3203, 2017. Disponível em: <https://doi.org/10.1039/c6sc05720a>.
- 180 YAO, K. et al. The TensorMol-0.1 model chemistry: a neural network augmented with long-range physics. *Chemical Science*, Royal Society of Chemistry (RSC), v. 9, n. 8, p. 2261–2269, 2018. Disponível em: <https://doi.org/10.1039/c7sc04934j>.

- 181 SCHÜTT, K. T. et al. Quantum-chemical insights from deep tensor neural networks. *Nature Communications*, Springer Science and Business Media LLC, v. 8, n. 1, jan. 2017. Disponível em: <https://doi.org/10.1038/ncomms13890>.
- 182 GILMER, J. et al. *Neural Message Passing for Quantum Chemistry*. arXiv, 2017. Disponível em: <https://arxiv.org/abs/1704.01212>.
- 183 SCHÜTT, K. T. et al. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. arXiv, 2017. Disponível em: <https://arxiv.org/abs/1706.08566>.
- 184 LUBBERS, N.; SMITH, J. S.; BARROS, K. Hierarchical modeling of molecular energies using a deep neural network. *The Journal of Chemical Physics*, AIP Publishing, v. 148, n. 24, p. 241715, jun. 2018. Disponível em: <https://doi.org/10.1063/1.5011181>.
- 185 UNKE, O. T.; MEUWLY, M. PhysNet: A neural network for predicting energies, forces, dipole moments, and partial charges. *Journal of Chemical Theory and Computation*, American Chemical Society (ACS), v. 15, n. 6, p. 3678–3693, maio 2019. Disponível em: <https://doi.org/10.1021/acs.jctc.9b00181>.
- 186 SOUZA, L. A. D. et al. Theoretical study of the influence of electron push or pull bipyridine ligands on the electronic structure of eu3+ ibuprofenate complexes. *Computational and Theoretical Chemistry*, Elsevier BV, v. 1232, p. 114459, fev. 2024. ISSN 2210-271X. Disponível em: <http://dx.doi.org/10.1016/j.comptc.2023.114459>.
- 187 REZENDE, U. L.; SOUZA, L. A. D.; BELCHIOR, J. C. An approach based on genetic algorithms and machine learning coupled for studying alloy and molecular clusters by optimizing quantum energy surfaces. *Journal of Computational Chemistry*, Wiley, v. 44, n. 24, p. 1956–1969, jun. 2023. ISSN 1096-987X. Disponível em: <http://dx.doi.org/10.1002/jcc.27174>.
- 188 KOSIMOV, D. P.; DZHURAKHALOV, A. A.; PEETERS, F. M. Carbon clusters: From ring structures to nanographene. *Physical Review B*, American Physical Society (APS), v. 81, n. 19, maio 2010. Disponível em: <https://doi.org/10.1103/physrevb.81.195414>.
- 189 XIAO, Y.; WILLIAMS, D. E. Genetic algorithm: a new approach to the prediction of the structure of molecular clusters. *Chemical Physics Letters*, v. 215, n. 1, p. 17–24, 1993. ISSN 0009-2614. Disponível em: <https://www.sciencedirect.com/science/article/pii/000926149389256H>.
- 190 REZENDE, U. L. E. M. *Explorando a estrutura de nanoclusters de titânio (até 10 átomos) e os encapsulamentos em nanotubos de carbono*. Tese (Doutorado) — Universidade Federal de Minas Gerais, 2019.
- 191 EXPLORING the MP2 energy surface of nanoalloy clusters with a genetic algorithm: Application to sodium–potassium. *Chemical Physics Letters*, v. 639, p. 135 – 141, 2015. ISSN 0009-2614. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0009261415006946>.
- 192 LOURENÇO, M. P. et al. Gamaterial - a genetic-algorithm software for material design and discovery. *Journal of Computational Chemistry*, Wiley, v. 44, n. 7, p. 814–823, nov. 2022. ISSN 1096-987X. Disponível em: <http://dx.doi.org/10.1002/jcc.27043>.

- 193 BALLESTER, P. J.; RICHARDS, W. G. Ultrafast shape recognition to search compound databases for similar molecular shapes. *Journal of Computational Chemistry*, Wiley, v. 28, n. 10, p. 1711–1723, mar. 2007. Disponível em: <https://doi.org/10.1002/jcc.20681>.
- 194 KOCER, E.; KO, T. W.; BEHLER, J. *Neural Network Potentials: A Concise Overview of Methods*. arXiv, 2021. Disponível em: <https://arxiv.org/abs/2107.03727>.
- 195 CHIRIKI, S. et al. Correlation of structure with UV-visible spectra by varying SH composition in au-SH nanoclusters. *The Journal of Chemical Physics*, AIP Publishing, v. 149, n. 7, p. 074307, ago. 2018. Disponível em: <https://doi.org/10.1063/1.5031478>.
- 196 CAO, L.-Z. et al. Artificial neural network potential for gold clusters. *Chinese Physics B*, IOP Publishing, v. 29, n. 11, p. 117304, out. 2020. Disponível em: <https://doi.org/10.1088/1674-1056/abc15d>.
- 197 LOURENÇO, M. P. et al. A new active learning approach for global optimization of atomic clusters. *Theoretical Chemistry Accounts*, v. 140, n. 6, p. 62, May 2021. ISSN 1432-2234. Disponível em: <https://doi.org/10.1007/s00214-021-02766-5>.
- 198 ZALUSKA, A.; ZALUSKI, L.; STRÖM-OLSEN, J. Structure, catalysis and atomic reactions on the nano-scale: a systematic approach to metal hydrides for hydrogen storage. *Applied Physics A Materials Science amp; Processing*, Springer Science and Business Media LLC, v. 72, n. 2, p. 157–165, fev. 2001. ISSN 0947-8396. Disponível em: <http://dx.doi.org/10.1007/s003390100783>.
- 199 MA, W.; JING, C. First-principles study on hydrogen storage in al-, ca-, mn-doped mgni clusters. *International Journal of Modern Physics B*, World Scientific Pub Co Pte Lt, v. 31, n. 26, p. 1730002, out. 2017. ISSN 1793-6578. Disponível em: <http://dx.doi.org/10.1142/S021797921730002X>.
- 200 ZHANG, M.; FOURNIER, R. Structure of 55-atom bimetallic clusters. *Journal of Molecular Structure: THEOCHEM*, v. 762, n. 1, p. 49–56, 2006. ISSN 0166-1280. A Collection of Papers Dedicated to Professor Annick Goursot on the Occasion of her 60th Birthday. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0166128005007724>.
- 201 DARBY, S. et al. Theoretical study of cu–au nanoalloy clusters using a genetic algorithm. *The Journal of Chemical Physics*, v. 116, n. 4, p. 1536–1550, 2002. Disponível em: <https://doi.org/10.1063/1.1429658>.
- 202 YIN, X. et al. Designing stable bimetallic nanoclusters via an iterative two-step optimization approach. *Mol. Syst. Des. Eng.*, The Royal Society of Chemistry, v. 6, p. 545–557, 2021. Disponível em: <http://dx.doi.org/10.1039/D1ME00027F>.
- 203 FERNÁNDEZ, E. M. et al. Structural properties of bimetallic clusters from density functional calculations. *International Journal of Modern Physics B*, v. 19, n. 15n17, p. 2339–2344, 2005. Disponível em: <https://doi.org/10.1142/S0217979205030943>.
- 204 ARSLAN, H.; GARIP, A. K.; JOHNSTON, R. L. Theoretical study of the structures and chemical ordering of cobalt–palladium nanoclusters. *Physical Chemistry Chemical Physics*, Royal Society of Chemistry (RSC), v. 17, n. 42, p. 28311–28321, 2015. Disponível em: <https://doi.org/10.1039/c5cp01029b>.

- 205 SHAYEGHI, A. et al. Pool-BCGA: a parallelised generation-free genetic algorithm for the ab initio global optimisation of nanoalloy clusters. *Physical Chemistry Chemical Physics*, Royal Society of Chemistry (RSC), v. 17, n. 3, p. 2104–2112, 2015. Disponível em: <https://doi.org/10.1039/c4cp04323e>.
- 206 GENDREAU, M.; POTVIN, J.-Y. *Handbook of metaheuristics*. [S.l.]: Springer US, 2010.
- 207 BRITO, B. et al. Binding energies of small lithium clusters: A comparison of different theoretical calculations. *Chemical Physics Letters*, v. 616-617, p. 212–216, 2014. ISSN 0009-2614. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0009261414008975>.
- 208 UNKE, O. T.; MEUWLY, M. *Machine Learning Potential Energy Surfaces*. arXiv, 2019. Disponível em: <https://arxiv.org/abs/1909.08027>.
- 209 HUBER, K. P.; HERZBERG, G. Constants of diatomic molecules. In: _____. *Molecular Spectra and Molecular Structure: IV. Constants of Diatomic Molecules*. Boston, MA: Springer US, 1979. p. 8–689. ISBN 978-1-4757-0961-2. Disponível em: https://doi.org/10.1007/978-1-4757-0961-2_2.
- 210 WHEELER, S. E. et al. Binding energies of small lithium clusters (lin) and hydrogenated lithium clusters (linh). *The Journal of Chemical Physics*, v. 120, n. 10, p. 4683–4689, 2004. Disponível em: <https://doi.org/10.1063/1.1645242>.
- 211 DUGOURD, P. et al. Vibronic structure of the Li_3 ground state. *Chemical Physics Letters*, v. 175, n. 6, p. 555–560, 1990. ISSN 0009-2614. Disponível em: <https://www.sciencedirect.com/science/article/pii/0009261490855806>.
- 212 BRITO, B. et al. Binding energies of small lithium clusters: A comparison of different theoretical calculations. *Chemical Physics Letters*, v. 616-617, p. 212–216, 2014. ISSN 0009-2614. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0009261414008975>.
- 213 TEMELSO, B.; ARCHER, K. A.; SHIELDS, G. C. Benchmark structures and binding energies of small water clusters with anharmonicity corrections. *The Journal of Physical Chemistry A*, American Chemical Society (ACS), v. 115, n. 43, p. 12034–12046, out. 2011. Disponível em: <https://doi.org/10.1021/jp2069489>.
- 214 BULUSU, S. et al. Lowest-energy structures of water clusters $(h_2o)_{11}$ and $(h_2o)_{13}$. *The Journal of Physical Chemistry A*, American Chemical Society (ACS), v. 110, n. 42, p. 11781–11784, out. 2006. Disponível em: <https://doi.org/10.1021/jp0655726>.
- 215 TAKEUCHI, H. Development of an efficient geometry optimization method for water clusters. *Journal of Chemical Information and Modeling*, American Chemical Society (ACS), v. 48, n. 11, p. 2226–2233, nov. 2008. Disponível em: <https://doi.org/10.1021/ci800238w>.
- 216 SHEN, D. et al. Investigation of properties of mg_n clusters and their hydrogen storage mechanism: A study based on DFT and a global minimum optimization method. *The Journal of Physical Chemistry A*, American Chemical Society (ACS), v. 119, n. 15, p. 3636–3643, abr. 2015. Disponível em: <https://doi.org/10.1021/acs.jpca.5b01474>.

217 BOGOJESKI, M. et al. Quantum chemical accuracy from density functional approximations via machine learning. *Nature Communications*, Springer Science and Business Media LLC, v. 11, n. 1, out. 2020. ISSN 2041-1723. Disponível em: <http://dx.doi.org/10.1038/s41467-020-19093-1>.

218 GLOROT, X.; BENGIO, Y. Understanding the difficulty of training deep feedforward neural networks. In: TEH, Y. W.; TITTERINGTON, M. (Ed.). *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. Chia Laguna Resort, Sardinia, Italy: PMLR, 2010. (Proceedings of Machine Learning Research, v. 9), p. 249–256. Disponível em: <https://proceedings.mlr.press/v9/glorot10a.html>.

Apêndice A

Parâmetros de entrada do NQGA

Neste apêndice, são apresentados todos os parâmetros de entrada utilizados para obter os resultados detalhados no Capítulo 6. Esses parâmetros podem ser utilizados junto com a metodologia NQGA para reproduzir os resultados obtidos ou como referência em estudos futuros.

A.1 Cobre, Ouro, liga Cobre-Ouro

Sistemas: Cu_nAu_m ($n + m \leq X$ para $X = 14, 19, 38, \text{ e } 55$)

Tamanho da população: 20 clusters

Método de geração da população inicial: Geração aleatória por volume esférico

Fragmentos: Cu, Au

Operador imigração:

Adição por iteração: 3 clusters

Método para gerar novos clusters: Geração aleatória por volume esférico

Operador cruzamento:

Método de seleção dos candidatos: Aleatório

Tipo de cruzamento: genC&S

Taxa de cruzamento: 25%

Operador mutação:

Alvo de atuação: Clusters da população e novos clusters gerados por cruzamento

Mutação sigma:

Taxa de ocorrência: 25%

Número máximo de átomos afetados: 3

Modificação de interior:

Taxa de ocorrência: 10%

Operador de superfície angular:

Taxa de ocorrência: 10%

Mutação de permuta:

Taxa de ocorrência: 20%

Mutação rotacional:

Taxa de ocorrência: 5%

Operador Predador:

Tamanho máximo da população: 20 clusters

Operador MLP: Não utilizado

Operador DGP:

Método de comparação: ACSF

Parâmetro do fator de corte R_c Eq. 4.11: 7.2

Parâmetros de simetria radial Eq. 4.12:

R_s : 0.2

η : [0.1, 0.2, 0.3]

Parâmetros de simetria angular Eq. 4.14:

λ : [-1, 1]

η : [0.01, 0.05, 0.1]

ζ : [3, 1]

γ : 0.01 à 10

Operador Apex:

Número máximo de iterações: 400

Iterações para o melhor indivíduo ser considerado mínimo global: 20

Operador Simetrizador:

Método de atuação: Grupo de simetria mais próximo

Tipo de seleção: Melhores clusters da população

Número de indivíduos afetados: 3

Condição de ativação: Última iteração do ciclo

Avaliação de energia:

Método para o cálculo de energia: Equação de Gupta (3.33)

Parâmetros utilizados: Tabela 3.1

Algoritmo de otimização local: L-BFGS

Alvo de otimização local: Todos os clusters

A.2 Liga Ouro-Prata

Sistemas: Au_nAg_n ($n = 10, 20, 30, 40, 50, 60, 70, e 75$)

Tamanho da população: 20 clusters

Método de geração da população inicial: Geração aleatória por volume esférico

Fragmentos: Au, Ag

Operador imigração:

Adição por iteração: 3 clusters

Método para gerar novos clusters: Geração aleatória por volume esférico

Operador cruzamento:

Método de seleção dos candidatos: Aleatório

Tipo de cruzamento: genC&S

Taxa de cruzamento: 30%

Operador mutação:

Alvo de atuação: Clusters da população e novos clusters gerados por cruzamento

Mutação sigma:

Taxa de ocorrência: 20%

Número máximo de átomos afetados: 5

Modificação de interior:

Taxa de ocorrência: 15%

Operador de superfície angular:

Taxa de ocorrência: 15%

Mutação de permuta:

Taxa de ocorrência: 30%

Mutação rotacional:

Taxa de ocorrência: 10%

Operador Predador:

Tamanho máximo da população: 20 clusters

Operador MLP: Não utilizado

Operador DGP:

Método de comparação: ACSF

Parâmetro do fator de corte R_c Eq. 4.11: 12.5

Parâmetros de simetria radial Eq. 4.12:

R_s : 0.2

η : [0.1, 0.2, 0.3, 0.4, 0.5, 1, 1.5]

Parâmetros de simetria angular Eq. 4.14:

λ : [-1, 1]

η : [0.01, 0.05, 0.1, 0.15]

ζ : [10, 5, 3, 1, 0.5]

γ : 0.01 à 10

Operador Apex:

Número máximo de iterações: 400

Iterações para o melhor individuo ser considerado mínimo global: 20

Operador Simetrizador:

Método de atuação: Grupo de simetria mais próximo

Tipo de seleção: Melhores clusters da população

Número de indivíduos afetados: 3

Condição de ativação: Última iteração do ciclo

Avaliação de energia:

Método para o cálculo de energia: Equação de Gupta (3.33)

Parâmetros utilizados: Tabela 3.1

Algoritmo de otimização local: L-BFGS

Alvo de otimização local: Todos os clusters

A.3 Lítio (DFT)

Sistema: $Li_n (3 \leq n \leq 8)$

Tamanho da população: 5 clusters

Método de geração da população inicial: Geração aleatória por volume esférico

Fragmentos: Li

Operador imigração:

Adição por iteração: 1 cluster

Método para gerar novos clusters: Geração aleatória por volume esférico

Operador cruzamento:

Método de seleção dos candidatos: Aleatório

Tipo de cruzamento: genC&S

Taxa de cruzamento: 5%

Operador mutação:

Alvo de atuação: Clusters da população e novos clusters gerados por cruzamento

Mutação sigma:

Taxa de ocorrência: 3%

Número máximo de átomos afetados: 1

Modificação de interior:

Taxa de ocorrência: 3%

Operador de superfície angular:

Taxa de ocorrência: 3%

Mutação de permuta:

Taxa de ocorrência: 0%

Mutação rotacional:

Taxa de ocorrência: 0%

Operador Predador:

Tamanho máximo da população: 5 clusters

Operador MLP:

Arquitetura da rede neural: 14-10-10-1

Parâmetros do NQGA usado no operador MLP: Apêndice A.7

Operador DGP:

Método de comparação: ACSF

Parâmetro do fator de corte R_c Eq. 4.11: 5.7

Parâmetros de simetria radial Eq. 4.12:

R_s : 0.0

η : [0.1, 0.2, 0.3, 0.4, 0.5, 1]

Parâmetros de simetria angular Eq. 4.14:

λ : [-1, 1]

η : [0.05, 0.1]

ζ : [1, 0.5]

γ : 0.01 à 2

Operador Apex:

Número máximo de iterações: 400

Iterações para o melhor indivíduo ser considerado mínimo global: 10

Operador Simetrizador:

Método de atuação: Grupo de simetria mais próximo

Tipo de seleção: Melhores clusters da população

Número de indivíduos afetados: 2

Condição de ativação: Última iteração do ciclo

Avaliação de energia:

Método para o cálculo de energia: DFT através do pacote ORCA

Funcional de troca-correlação: PBE

Conjunto de funções de base: Aug-CC-pVTZ

Conjunto de parâmetros de convergência do SCF: *Tight*

Conjunto de parâmetros de convergência de otimização: *Tight*

Algoritmo de otimização local: BFGS através do pacote ORCA

Alvo de otimização local: Todos os clusters

A.4 Lítio (DLPNO-CCSD(T))

Sistemas: Li_3 , Li_4 , Li_2H e Li_3H

Tamanho da população: 5 clusters

Método de geração da população inicial: Geração aleatória por volume esférico

Fragmentos: Li , H

Operador imigração:

Adição por iteração: 1 cluster

Método para gerar novos clusters: Geração aleatória por volume esférico

Operador cruzamento:

Método de seleção dos candidatos: Aleatório

Tipo de cruzamento: genC&S

Taxa de cruzamento: 15%

Operador mutação:

Alvo de atuação: Clusters da população e novos clusters gerados por cruzamento

Mutação sigma:

Taxa de ocorrência: 8%

Número máximo de átomos afetados: 1

Modificação de interior:

Taxa de ocorrência: 0%

Operador de superfície angular:

Taxa de ocorrência: 0%

Mutação de permuta:

Taxa de ocorrência: 0%

Mutação rotacional:

Taxa de ocorrência: 0%

Operador Predador:

Tamanho máximo da população: 5 clusters

Operador MLP:

Arquitetura da rede neural: 14-10-10-1

Parâmetros do NQGA usado no operador MLP: Apêndice A.7

Operador DGP:

Método de comparação: ACSF

Parâmetro do fator de corte R_c Eq. 4.11: 5.7

Parâmetros de simetria radial Eq. 4.12:

R_s : 0.0

η : [0.1, 0.2, 0.3, 0.4, 0.5, 1]

Parâmetros de simetria angular Eq. 4.14:

λ : [-1, 1]

η : [0.05, 0.1]

ζ : [1, 0.5]

γ : 0.01 à 2

Operador Apex:

Número máximo de iterações: 400

Iterações para o melhor individuo ser considerado mínimo global: 10

Operador Simetrizador:

Método de atuação: Grupo de simetria mais próximo

Tipo de seleção: Melhores clusters da população

Número de indivíduos afetados: 5

Condição de ativação: Penúltima iteração do ciclo

Avaliação de energia:

Método para o cálculo de energia: DLPNO-CCSD(T) através do pacote ORCA

Conjunto de funções de base: Aug-CC-pVQZ

Conjunto de parâmetros de convergência do SCF: *VeryTight*

Conjunto de parâmetros de convergência de otimização: *VeryTight*

Algoritmo de otimização local: BFGS através do pacote ORCA

Alvo de otimização local: Todos os clusters da população apenas na última iteração do APEX

A.5 Água

Sistemas: $(H_2O)_5$ e $(H_2O)_{11}$

Tamanho da população: 5 clusters

Método de geração da população inicial: Geração aleatória por volume esférico

Fragmentos: H_2O

Operador imigração:

Adição por iteração: 1 cluster

Método para gerar novos clusters: Geração aleatória por volume esférico

Operador cruzamento:

Método de seleção dos candidatos: Aleatório

Tipo de cruzamento: genC&S

Taxa de cruzamento: 5%

Operador mutação:

Alvo de atuação: Clusters da população e novos clusters gerados por cruzamento

Mutação sigma:

Taxa de ocorrência de translação molecular: 7%

Taxa de ocorrência de rotação molecular: 5%

Número máximo de moléculas afetados: 1

Modificação de interior:

Taxa de ocorrência de translação molecular: 3%

Taxa de ocorrência de rotação molecular: 3%

Operador de superfície angular:

Taxa de ocorrência de translação molecular: 3%

Taxa de ocorrência de rotação molecular: 3%

Mutação de permuta:

Taxa de ocorrência: 0%

Mutação rotacional:

Taxa de ocorrência de translação molecular: 3%

Taxa de ocorrência de rotação molecular: 8%

Operador Predador:

Tamanho máximo da população: 5 clusters

Operador MLP:

Arquitetura da rede neural: 46-10-10-1

Parâmetros do NQGA usado no operador MLP: Apêndice A.7

Operador DGP:

Método de comparação: ACSF

Parâmetro do fator de corte R_c Eq. 4.11: 8.5

Parâmetros de simetria radial Eq. 4.12:

R_s : 0.3

η : [0.1, 0.2, 0.3, 0.4, 0.5, 1]

Parâmetros de simetria angular Eq. 4.14:

λ : [-1, 1]

η : [0.01, 0.05, 0.1, 0.15]

ζ : [10, 5, 3, 1, 0.5]

γ : 0.01 à 2

Operador Apex:

Número máximo de iterações: 400

Iterações para o melhor indivíduo ser considerado mínimo global: 15

Operador Simetrizador:

Método de atuação: Grupo de simetria mais próximo

Tipo de seleção: Melhores clusters da população

Número de indivíduos afetados: 2

Condição de ativação: Última iteração do ciclo

Avaliação de energia:

Método para o cálculo de energia: MP2 através do pacote ORCA

Conjunto de funções de base: Aug-CC-pVDZ

Conjunto de parâmetros de convergência do SCF: *Tight*

Conjunto de parâmetros de convergência de otimização: *Tight*

Algoritmo de otimização local: BFGS através do pacote ORCA

Alvo de otimização local: Todos os clusters

A.6 Magnésio-Hidrogênio

Sistema: Mg_6 , Mg_6H_{2x} ($x = 1$ à 7) e $Mg_6(H_2)_x$ ($x = 1$ à 10)

Tamanho da população: 5 clusters

Método de geração da população inicial: Geração aleatória por volume esférico

Fragmentos:

Mg_6 : Mg

Mg_6H_{2x} ($x = 1$ à 7): Mg_6, H

$Mg_6(H_2)_x$ ($x = 1$ à 10): Mg_6, H_2

Operador imigração:

Adição por iteração: 1 cluster

Método para gerar novos clusters: Geração aleatória por volume esférico

Operador cruzamento:

Método de seleção dos candidatos: Aleatório

Tipo de cruzamento: genC&S

Taxa de cruzamento: 10%

Operador mutação:

Alvo de atuação: Clusters da população e novos clusters gerados por cruzamento

Mutação sigma:

Taxa de ocorrência: 5%

Número máximo de átomos afetados: 2

Modificação de interior:

Taxa de ocorrência: 5%

Operador de superfície angular:

Taxa de ocorrência: 5%

Mutação de permuta:

Taxa de ocorrência: 0%

Mutação rotacional:

Taxa de ocorrência: 3%

Operador Predador:

Tamanho máximo da população: 5 clusters

Operador MLP:

Arquitetura da rede neural: 24-10-10-1

Parâmetros do NQGA usado no operador MLP: Apêndice A.7

Operador DGP:

Método de comparação: ACSF

Parâmetro do fator de corte R_c Eq. 4.11: 7.0

Parâmetros de simetria radial Eq. 4.12:

R_s : 0.0

η : [0.1, 0.2, 0.3, 0.4, 0.5, 1]

Parâmetros de simetria angular Eq. 4.14:

λ : [-1, 1]

η : [0.05, 0.1, 0.5]

ζ : [1, 0.5, 0.1]

γ : 0.01 à 2

Operador Apex:

Número máximo de iterações: 400

Iterações para o melhor indivíduo ser considerado mínimo global: 15

Operador Simetrizador:

Método de atuação: Grupo de simetria mais próximo

Tipo de seleção: Melhores clusters da população

Número de indivíduos afetados: 3

Condição de ativação: Última iteração do ciclo

Avaliação de energia:

Para o Mg_6 :

Método para o cálculo de energia: MP2 através do pacote ORCA

Conjunto de funções de base: Aug-CC-pVTZ

Para o Mg_6H_{2x} ($x = 1$ à 7) e $Mg_6(H_2)_x$ ($x = 1$ à 10):

Método para o cálculo de energia: DFT através do pacote ORCA

Funcional de troca-correlação: B3PW91

Correção de dispersão: Grimme's DFT-D3

Conjunto de funções de base: Aug-CC-pVDZ

Comum para todos os sistemas:

Conjunto de parâmetros de convergência do SCF: *Tight*

Conjunto de parâmetros de convergência de otimização: *Tight*

Algoritmo de otimização local: BFGS através do pacote ORCA

Alvo de otimização local: Todos os clusters

A.7 MLP

Sistemas: Li_n ($3 \leq n \leq 8$), Li_3 , Li_4 , Li_2H , Li_3H , $(H_2O)_5$, $(H_2O)_{11}$, Mg_6 , Mg_6H_{2x} ($x = 1$ à 7) e $Mg_6(H_2)_x$ ($x = 1$ à 10)

Tamanho da população: 20 clusters

Método de geração da população inicial: Geração aleatória por volume esférico

Operador imigração:

Adição por iteração: 3 clusters

Método para gerar novos clusters: Geração aleatória por volume esférico

Operador cruzamento:

Método de seleção dos candidatos: Aleatório

Tipo de cruzamento: genC&S

Taxa de cruzamento: 25%

Operador mutação:

Alvo de atuação: Clusters da população e novos clusters gerados por cruzamento

Mutação sigma:

Taxa de ocorrência: 25%

Número máximo de átomos afetados: 2

Modificação de interior:

Taxa de ocorrência: 8%

Operador de superfície angular:

Taxa de ocorrência: 8%

Mutação de permuta:

Taxa de ocorrência: 15%

Taxa de ocorrência para mutação rotacional:

Taxa de ocorrência: 5%

Operador Predador:

Tamanho máximo da população: 20 clusters

Operador MLP: Não utilizado

Operador DGP:

Método de comparação: ACSF

Parâmetro do fator de corte R_c Eq. 4.11: 10.0

Parâmetros de simetria radial Eq. 4.12:

R_s : 0.25

η : [0.1, 0.2, 0.3, 0.4, 0.5, 1, 1.5]

Parâmetros de simetria angular Eq. 4.14:

λ : [-1, 1]

η : [0.01, 0.05, 0.1, 0.15]

ζ : [10, 5, 3, 1, 0.5]

γ : 0.01 à 10

Operador Apex:

Número máximo de iterações: 400

Iterações para o melhor indivíduo ser considerado mínimo global: 20

Operador Simetrizador:

Método de atuação: Grupo de simetria mais próximo

Tipo de seleção: Melhores clusters da população

Número de indivíduos afetados: 2

Condição de ativação: Última iteração do ciclo

Avaliação de energia:

Método para o cálculo de energia: Rede neural artificial (4.8)

Algoritmo de otimização local: L-BFGS

Alvo de otimização local: Todos os clusters

Funções de ativação: Sigmoide - Sigmoide - Linear

Função erro: Erro quadrático médio

Inicialização de pesos: Aleatória

Algoritmo para otimização dos pesos: ADAM

Erro mínimo: 10 kcal/mol

Proporção treinamento/validação: 80/20

Clusters selecionados para melhoria do treinamento: 2 clusters

Apêndice B

Publicações

TRABALHOS PUBLICADOS:

An approach based on genetic algorithms and machine learning coupled for studying alloy and molecular clusters by optimizing quantum energy surfaces

Disponível em: <https://doi.org/10.1002/jcc.27174>

Theoretical study of the influence of electron push or pull bipyridine ligands on the electronic structure of Eu^{3+} ibuprofenate complexes

Disponível em: <https://doi.org/10.1016/j.comptc.2023.114459>

CONGRESSOS:

“Carbon dioxide capture through functionalized graphenes with alkaline earth metal oxides” de autoria de Gabriel Salomon, Umar Lucio Rezende, Leonardo de Souza, Jadson Claudio Belchior apresentado no XXII Simpósio Brasileiro de Química Teórica (SBQT) no período de 23 a 26 de outubro de 2023 em Niterói-RJ.