

Caio Amaral Custódio

**Redes Neurais Artificiais e Teoria do Funcional da  
Densidade: otimização de funcionais para modelagem de  
nanomateriais**

Dissertação apresentada ao Instituto de Química, Universidade Estadual Paulista, como parte dos requisitos para obtenção do título de Mestre em Química

Universidade Estadual Paulista – UNESP  
Campus de Araraquara – Instituto de Química  
Programa de Pós-Graduação

Orientador: Prof<sup>a</sup> Dr<sup>a</sup> Vivian Vanessa França  
Coorientador: Prof<sup>a</sup> Dr<sup>a</sup> Érica Regina Filletti

Araraquara - SP

2019

FICHA CATALOGRÁFICA

C987r	<p>Custódio, Caio Amaral</p> <p>Redes neurais artificiais e teoria do funcional da densidade: otimização de funcionais para modelagem de nanomateriais / Caio Amaral Custódio. – Araraquara : [s.n.], 2019 48 f. : il.</p> <p>Dissertação (mestrado) – Universidade Estadual Paulista, Instituto de Química Orientador: Vivian Vanessa França Coorientador: Érica Regina Filletti</p> <p>1. Funcionais de densidade. 2. Redes neurais (Computação). 3. Nanoestruturas. 4. Hubbard, Modelo de. 5. Informação quântica. I. Título.</p>
-------	--

**CERTIFICADO DE APROVAÇÃO**

TÍTULO DA DISSERTAÇÃO: "Redes Neurais Artificiais e Teoria do Funcional da Densidade: otimização de funcionais para modelagem de nanomaterias"

**AUTOR: CAIO AMARAL CUSTÓDIO**

**ORIENTADORA: VIVIAN VANESSA FRANÇA HENN**

**COORIENTADORA: ERICA REGINA FILLETTI NASCIMENTO**

Aprovado como parte das exigências para obtenção do Título de Mestre em QUÍMICA, pela Comissão Examinadora:

Prof<sup>a</sup>. Dr<sup>a</sup>. VIVIAN VANESSA FRANÇA HENN  
Departamento de Físico Química / Instituto de Química - UNESP - Araraquara

Prof. Dr. RODRIGO FERNANDO COSTA MARQUES  
Departamento de Físico-Química / Instituto de Química - UNESP - Araraquara

Prof. Dr. SÉRGIO RICARDO MUNIZ  
Departamento de Física e Ciência dos Materiais / Instituto de Física de São Carlos - USP - São Carlos

Araraquara, 06 de setembro de 2019

## DADOS CURRICULARES

### Caio Amaral Custódio

---

#### Dados pessoais

**Nome** Caio Amaral Custódio  
**Nascimento** 05/12/1993 - Brasil

**Endereço profissional** Universidade Estadual Paulista Júlio de Mesquita Filho  
Rua Professor Francisco Degni  
Jardim Quitandinha - Araraquara  
14800060, SP - Brasil

**Endereço eletrônico** caiocustodio@gmail.com

---

#### Formação acadêmica/titulação

**2017 - 2019** Mestrado em Química.  
Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP, São Paulo, Brasil  
Orientador: Vivian Vanessa França  
Coorientador: Érica Regina Filletti

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.  
Processo: 33004030072P8

**2013 - 2017** Graduação em Química.  
Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP, São Paulo, Brasil  
Título: Otimização de Funcionais da Densidade via Redes Neurais Artificiais  
Orientador: Vivian Vanessa França

---

#### Formação complementar

**2017 - 2017** Curso de curta duração em School and Workshop on Density Functional Theory and Quantum Information. (Carga horária: 37h).  
ICTP South American Institute for Fundamental Research, ICTP-SAIFR, Brasil

---

#### Atuação profissional

##### 1. Universidade Estadual Paulista Júlio de Mesquita Filho - UNESP

---

###### Vínculo institucional

**2016 - 2016** Vínculo: Bolsista, Enquadramento funcional: Monitor de Física II, Carga horária: 12, Regime: Parcial

---

#### Produção bibliográfica

##### Artigos completos publicados em periódicos

1. CUSTÓDIO, CAIO A.; FILLETTI, ÉRICA R.; FRANÇA, VIVIAN V. Artificial neural networks for density-functional optimizations in fermionic systems. Scientific Reports. , v.9, p.1886 - , 2019.

### **Apresentação de trabalho e palestra**

1. CUSTÓDIO, C. A.; NASCIMENTO, E. R. F.; V. V. França  
**FVC DENSITY FUNCTIONAL OPTIMIZATION THROUGH ARTIFICIAL NEURAL NETWORKS**, 2017.  
(Congresso,Apresentação de Trabalho)
2. CUSTÓDIO, C. A.; NASCIMENTO, E. R. F.; V. V. França  
**Otimização de Funcionais da Densidade via Redes Neurais Artificiais**, 2016.  
(Congresso,Apresentação de Trabalho)
3. CUSTÓDIO, C. A.; NASCIMENTO, E. R. F.; V. V. França  
**Otimização de Funcionais da Densidade via Redes Neurais Artificiais**, 2016.  
(Outra,Apresentação de Trabalho)
4. CUSTÓDIO, C. A.; V. V. França  
**Estudo dos efeitos de campos magnéticos internos em nanoestruturas**, 2015.  
(Congresso,Apresentação de Trabalho)

### **Eventos**

#### **Participação em eventos**

1. Apresentação de Poster / Painel no(a) **46th World Chemistry Congress**, 2017. (Congresso)  
FVC Density Functional optimization through Artificial Neural Networks.
2. Apresentação de Poster / Painel no(a) **Mulheres na Ciência 2016: o Gênero na Pós-Graduação**, 2016. (Outra)  
Otimização de Funcionais da Densidade via Redes Neurais Artificiais.
3. Apresentação de Poster / Painel no(a) **XXVIII Congresso de Iniciação Científica da UNESP (CIC)**, 2016. (Congresso)  
Otimização de Funcionais da Densidade via Redes Neurais Artificiais.
4. Apresentação de Poster / Painel no(a) **XXVII Congresso de Iniciação Científica da UNESP (CIC)**, 2015. (Congresso)  
Estudo dos efeitos de campos magnéticos internos em nanoestruturas.
5. **XXVII Congresso de Iniciação Científica da UNESP (CIC)**, 2015. (Congresso)  
Estudo dos efeitos de campos magnéticos internos em nanoestruturas.

#### **Organização de evento**

1. V. V. França; NASCIMENTO, E. R. F.; **CUSTÓDIO, C. A.**  
**Mulheres na Ciência 2017: Educação sem Distinção por Gênero**, 2017. (Outro, Organização de evento)
2. **CUSTÓDIO, C. A.**  
**46ª Semana da Química e 1ª Semana da Engenharia Química**, 2016. (Outro, Organização de evento)

*Aos meus pais, familiares e amigos que me acompanharam nessa jornada*

# Agradecimentos

À Profa. Dra. Vivian Vanessa França Henn, pela sua orientação, amizade, ensinamentos e inspirações, especialmente por sua paciência e apoio durante meus momentos difíceis sob sua orientação.

À Profa. Dra. Erica Regina Filletti Nascimento, pela sua co-orientação, amizade e contribuições pessoais e profissionais em minha formação.

Aos meus pais Antônio e Janete, por todo seu apoio, criação e sacrifícios que me trouxeram até aqui, assim como todos meus familiares.

Aos meus companheiros do Laboratório de Simulações Quânticas, compartilhando conhecimentos e muito momentos felizes.

À todos os docentes, pesquisadores, funcionários e amigos do Instituto de Química da UNESP de Araraquara, que estiveram ao meu lado durante toda minha formação e de que algum modo contribuíram para meu crescimento pessoal.

Ao meu companheiro Murilo, por estar sempre ao meu lado, por todo seu carinho e especialmente por toda sua paciência durante a confecção desse trabalho.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. Processo: 33004030072P8

*"I believe there is no philosophical high-road in science, with epistemological signposts. No, we are in a jungle and find our way by trial and error, building our road behind us as we proceed."  
(Max Born)*



# Resumo

Nesse trabalho propomos o desenvolvimento de redes neurais artificiais capazes de fornecer a energia do estado fundamental do modelo de Hubbard para nanoestruturas fermiônicas interagentes e homogêneas. Uma vez otimizado o funcional via rede neural, este pode ser usado como *input* em cálculos de funcionais da densidade para sistemas heterogêneos. O modelo neural obtido mostrou um desempenho excelente, com desvios menores que  $\sim 0,2\%$ , recuperando todos os regimes de densidade, magnetização e uma vasta extensão de regimes de interação, quando comparado com resultados numéricos exatos. Comparado à funcionais analíticos, o modelo neural é mais preciso em todos os regimes de parâmetros, especialmente no regime de fraca interação, onde o funcional analítico mais recente apresenta um grande desvio:  $\sim 7\%$ , contra  $\sim 0,1\%$  para o nosso modelo neural. Aplicado em aproximações de densidade local para cálculos de DFT para cadeias finitas e com heterogeneidades, como impurezas localizadas e potenciais confinantes, nosso modelo neural se mostrou uma alternativa confiável e usando apenas uma fração dos recursos computacionais de outros tratamentos numéricos.

**Palavras-chave:** Teoria do Funcional da Densidade. Redes Neurais Artificiais. Nanoestruturas. Modelo de Hubbard. Informação Quântica.

# Abstract

In this work we propose an artificial neural network model to the ground-state energy of fermionic interacting particles in homogeneous chains described by the Hubbard model. Once the neural network functional is optimized, it can be used as input in density functional calculations for inhomogeneous systems. The neural network model obtained, showed excellent performance, deviating by less than  $\sim 0.2\%$ , recovering all regimes of density and magnetization and for a vast range of interactions when compared to exact numerical results. Compared to analytical functionals, the neural network is more accurate in all regimes of parameters, especially at the weakly interacting regime, where the most recent analytical parametrization fails the most:  $\sim 7\%$ , while only  $\sim 0.1\%$  for our neural network model. When applied in local density approximations for density functionals calculations for finite chains with inhomogeneities, such as localized impurities and confining potentials, our neural model has proven to be a reliable alternative, while using only a fraction of the computational resources from other numerical treatments.

**Keywords:** Density Functional Theory. Artificial Neural Networks. Nanostructures. Hubbard Model. Quantum Information.

# Lista de Figuras

Figura 1 – Cálculo auto-consistente das equações de Kohn-Sham . . . . .	21
Figura 2 – Representação de uma cadeia unidimensional com $L$ sítios do modelo de Hubbard. Cada sítio pode ser ocupado por dois elétrons de spin opostos que sofrem interações Coulombianas $U$ , onde o movimento desses elétrons é descrito pelo termo de <i>hopping</i> $t_{ij}$ , com $j = i \pm 1$ . . . . .	22
Figura 3 – O <i>Hopping</i> de um elétron do sítio $i$ para o sítio $j$ , através da operação $\hat{c}_j^\dagger \hat{c}_i$ , seguida de uma operação do movimento simétrico de buraco do sítio $j$ para o sítio $i$ , $\hat{b}_i^\dagger \hat{b}_j$ , onde $\hat{b}_i^\dagger = (-1)^i \hat{c}_i$ e $\hat{b}_i = (-1)^i \hat{c}_i^\dagger$ são os operadores de criação e aniquilação de buracos. . . . .	22
Figura 4 – Esquema de neurônio biológico . . . . .	25
Figura 5 – Neurônio não-linear baseado no <i>perceptron</i> . . . . .	25
Figura 6 – Funções de ativação comumente utilizadas. . . . .	26
Figura 7 – Rede neural <i>feed-forward</i> de uma única camada . . . . .	28
Figura 8 – Rede <i>feed-forward</i> de múltiplas camadas . . . . .	28
Figura 9 – Rede neural recorrente . . . . .	28
Figura 10 – Estrutura da rede neural de arquitetura 3-20-1 e seu desempenho durante treinamento . . . . .	35
Figura 11 – Energia do estado fundamental por sítio em função da densidade $n$ para os funcionais FVC e RNA em comparação com o resultado numericamente exato FN. Regime de interação $U = 6t$ , magnetização $m \sim 0$ . <i>Inset</i> : erro percentual $100 \left  (e_0^{FVC(RNA)} - e_0^{FN}) / e_0^{FN} \right $ . $D$ é o desvio médio percentual definido como $100 \sum \left  (e_0^{FVC(RNA)} - e_0^{FN}) / e_0^{FN} \right $ . . . . .	36
Figura 12 – Energia do estado fundamental por sítio em função da interação intrasítio $U$ (em unidades de $t$ ), para os funcionais FVC e RNA em comparação com o resultado numericamente exato FN. Regime de densidade $n \sim 0,5$ e magnetização $m \sim 0$ . <i>Inset</i> : erro percentual $100 \left  (e_0^{FVC(RNA)} - e_0^{FN}) / e_0^{FN} \right $ . $D$ é o desvio médio percentual definido como $100 \sum \left  (e_0^{FVC(RNA)} - e_0^{FN}) / e_0^{FN} \right $ . . . . .	36
Figura 13 – Energia do estado fundamental por sítio em função da magnetização $m$ , para os funcionais FVC e RNA em comparação com o resultado numericamente exato FN. Regime de densidade $n \sim 0,5$ . <i>Inset</i> : erro percentual $100 \left  (e_0^{FVC(RNA)} - e_0^{FN}) / e_0^{FN} \right $ . $D$ é o desvio médio percentual definido como $100 \sum \left  (e_0^{FVC(RNA)} - e_0^{FN}) / e_0^{FN} \right $ . . . . .	37

Figura 14 – (a) Perfil de densidade e (b) ilustração para uma cadeia finita de tamanho  $L = 40$ ,  $n = 0,8$ ,  $U = 1t$  e  $m \sim 0$  com condição de cadeia aberta. Os resultados exatos são via DMRG e os aproximados via DFT-LSDA com um dos funcionais: FN, FVC ou RNA.  $D$  refere-se ao desvio médio percentual definido como  $100 \sum \left| (e_0^{FN(FVC,RNA)} - e_0^{DMRG}) / e_0^{DMRG} \right|$ . . . . . 39

Figura 15 – (a) Perfil de densidade e (b) ilustração para um sistema com impureza localizada, de tamanho  $L = 15$ , intensidade da impureza  $V = 2t$ ,  $n = 0,4$ ,  $U = 1t$  e  $m \sim 0$  com condição de cadeia aberta. Os resultados exatos são via DMRG e os aproximados via DFT-LSDA com um dos funcionais: FN, FVC ou RNA. *Inset* traz os desvios percentuais definido como  $100 \left| (e_0^{FN(FVC,RNA)} - e_0^{DMRG}) / e_0^{DMRG} \right|$  e  $D$  refere-se ao desvio médio percentual definido como  $100 \sum \left| (e_0^{FN(FVC,RNA)} - e_0^{DMRG}) / e_0^{DMRG} \right|$ . . . . . 40

Figura 16 – (a) Perfil de densidade e (b) ilustração para uma cadeia harmonicamente confinada de tamanho  $L = 40$ , curvatura harmônica  $k = 0,01$ ,  $n = 0,8$ ,  $U = 1t$  e  $m \sim 0$  com condição de cadeia aberta. Os resultados exatos são via DMRG e os aproximados via DFT-LSDA com um dos funcionais: FN, FVC ou RNA.  $D$  refere-se ao desvio médio percentual definido como  $100 \sum \left| (e_0^{FN(FVC,RNA)} - e_0^{DMRG}) / e_0^{DMRG} \right|$ . . . . . 41

Figura 17 – Energia por sítio do estado fundamental em função da curvatura harmônica confinada em cadeias de tamanho  $L = 40$ ,  $n = 0,8$ ,  $U = 1t$  e  $m \sim 0$  com condição de cadeia aberta. Os resultados exatos são via DMRG e os aproximados via DFT-LSDA com um dos funcionais: FN, FVC ou RNA. *Inset* traz os desvios percentuais definido como  $100 \left| (e_0^{FN(FVC,RNA)} - e_0^{DMRG}) / e_0^{DMRG} \right|$ .  $D$  refere-se ao desvio médio percentual definido como  $100 \sum \left| (e_0^{FN(FVC,RNA)} - e_0^{DMRG}) / e_0^{DMRG} \right|$ . . . . . 42

# Lista de tabelas

Tabela 1 – Arquitetura e desempenho dos modelos neurais treinados. O desempenho é medido pelo erro quadrado médio. . . . .	34
Tabela 2 – Desvios médios percentuais para energia do estado fundamental de sistemas homogêneos obtida via funcionais FVC e RNA. . . . .	43
Tabela 3 – Desvios médios percentuais para energia do estado fundamental de sistemas homogêneos obtida via funcionais FN, FVC e RNA. . . . .	44

# Lista de abreviaturas e siglas

DFT	<i>Density Funcional Theory</i>
DMRG	<i>Density Matrix Renormalization Group</i>
EQM	Erro Quadrado Médio
ES	Equação de Schrödinger
FN	<i>Fully Numerical</i>
HK	<i>Hohenberg-Kohn</i>
KS	<i>Kohn-Sham</i>
LDA	<i>Local Density Aproximation</i>
LM	<i>Levenberg-Marquardt</i>
LSDA	<i>Local-Spin-Density approximation</i>
OBC	<i>Open Boundary Conditions</i>
PBC	<i>Periodic Boundary Conditions</i>
RNA	Redes Neurais Artificiais

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>15</b>
<b>2</b>	<b>FUNDAMENTOS TEÓRICOS</b>	<b>17</b>
<b>2.1</b>	<b>Teoria do Funcional da Densidade</b>	<b>17</b>
2.1.1	Aproximações de Thomas-Fermi e LDA	18
2.1.2	Energia de troca e correlação	19
2.1.3	Esquema de Kohn-Sham	20
<b>2.2</b>	<b>Modelo de Hubbard</b>	<b>21</b>
<b>2.3</b>	<b>Redes Neurais Artificiais</b>	<b>24</b>
2.3.1	Modelos Matemáticos de Neurônios	25
2.3.1.1	Funções de Ativação	26
2.3.2	Arquiteturas de Redes Neurais Artificiais	27
2.3.3	Tipos de Aprendizagem	29
<b>3</b>	<b>MÉTODOS COMPUTACIONAIS</b>	<b>30</b>
<b>4</b>	<b>RESULTADOS E DISCUSSÕES</b>	<b>34</b>
<b>4.1</b>	<b>Sistemas homogêneos</b>	<b>34</b>
<b>4.2</b>	<b>Sistemas heterogêneos</b>	<b>37</b>
<b>5</b>	<b>CONCLUSÃO</b>	<b>43</b>
	<b>REFERÊNCIAS</b>	<b>45</b>

# 1 Introdução

Para descrevermos de forma completa a matéria, sua estrutura e suas propriedades, estudos teóricos à nível quântico são imprescindíveis. A abordagem usual envolve a solução da equação de Schrödinger (ES), processo que pode ser sumarizado do seguinte modo

$$v(r) \xrightarrow{ES} \Psi(r_1, r_2, \dots, r_N) \xrightarrow{\langle \Psi | \dots | \Psi \rangle} \text{observáveis}, \quad (1.1)$$

onde inicia-se especificando o sistema através do potencial  $v(r)$  e então resolve-se a ES para a função de onda  $\Psi(r_1, r_2, \dots, r_N)$  e em seguida calcula-se os observáveis de interesse.

Embora esse formalismo seja exato, ele torna-se computacionalmente inviável em sistemas complexos envolvendo muitas partículas, mesmo com poderosos recursos computacionais à disposição. Tratamentos exatos se tornam especialmente proibitivos quando adicionamos às simulações teóricas heterogeneidades espaciais, tais como: bordas, impurezas, desordem, entre outros. Muitas vezes inevitáveis, essas heterogeneidades são importantes para tornarmos nossas simulações teóricas mais realistas<sup>[1]</sup>.

Nesse contexto, a Teoria do Funcional da Densidade (DFT)<sup>[2-4]</sup>, apresenta-se como uma alternativa interessante. A DFT tem como pilar o teorema de Hohenberg-Kohn (HK)<sup>[2]</sup>, que promove a densidade de partículas  $n_0(r)$ , um dos observáveis do sistema, à variável principal, sendo possível obter todos os outros observáveis à partir de  $n_0(r)$ . A abordagem de DFT, resume-se em:

$$n_0(r) \implies \Psi_0(r_1, \dots, r_N) \implies v_0(r), \quad (1.2)$$

onde dada uma densidade  $n_0(r)$ , a função de onda  $\Psi_0$  e o potencial  $v(r)$  são univocamente determinados, logo, todos os observáveis podem em princípio ser determinados por  $n_0(r)$ , sem o conhecimento de  $\Psi_0$ . A simplificação computacional é imensa: substitui-se a função de onda  $\Psi_0$ , uma função  $3N$ -dimensional, pela densidade  $n_0(r)$  3-dimensional.

O esquema Kohn-Sham (KS)<sup>[5]</sup>, é um mapeamento comumente usado para obter a densidade de um sistemas interagentes. Ele mapeia um sistema interagente de muitos corpos em um sistema não-interagente fictício através do potencial de KS,  $v_{KS}$ . O potencial  $v_{KS}$  é construído para reproduzir a densidade  $n_0(r)$  do sistema interagente. Entretanto,  $v_{KS}$  depende da energia de troca e correlação, cujo funcional não é tipicamente conhecido de forma explícita. Deste modo, aproximações são necessárias e o desempenho dos resultados da DFT dependem das aproximações e funcionais utilizados.

Neste trabalho, o foco está no estudo de cadeias unidimensionais homogêneas e heterogêneas, que são descritas com sucesso pelo modelo de Hubbard<sup>[6]</sup>. O modelo de Hubbard, muito explorado no estudo de sistemas fortemente correlacionados, é o modelo



mais simples que descreve a itinerância (através do termo  $t$ ) e interação (através do termo  $U$ ) entre elétrons em uma cadeia (de densidade  $n$  e magnetização  $m$ ), como por exemplo, em nanoestruturas, sólidos e átomos frios<sup>[7–12]</sup>. Embora seja simples, é capaz de descrever fenômenos importantes<sup>[13]</sup>, como a transição de Mott de metal para isolante (para  $U > 0$  e  $n = 1$ )<sup>[14]</sup>, a transição de superfluido de Bardeen-Cooper-Schrieffer para condensado de Bose-Einstein (para  $U < 0$  e  $m = 0$ )<sup>[14]</sup> e superfluidez exótica (para  $U < 0$  e  $m \neq 0$ )<sup>[11, 15, 16]</sup>.

O modelo de Hubbard apresenta alguns funcionais analíticos da densidade de relativo sucesso para a obtenção da energia do estado fundamental por sítio:  $LSOC$ <sup>[17]</sup> e  $FVC$ <sup>[18]</sup>. O primeiro, apresenta desvios relativamente pequenos ( $\sim 6\%$ ), mas é limitado a sistemas não-magnetizados,  $m = 0$ . Já o funcional FVC, descreve adequadamente sistemas magnetizados e apresenta desvios menores que o  $LSOC$  ( $\sim 2\%$  para a maioria dos regimes); contudo, no regime de fraca interação, bastante comum em ligas e átomos frios, os desvios ainda são maiores do que o desejável ( $\sim 7\%$ ).

Assim, existe uma comunidade ativa de pesquisadores buscando aperfeiçoar e desenvolver novos funcionais da densidade. Ferramentas como as Redes Neurais Artificiais (RNA)<sup>[19, 20]</sup>, tem sido pouco exploradas para atacar esse tipo de problema. As RNAs são modelos computacionais que se referem ao funcionamento do cérebro humano e são capazes de calcular funções matemáticas não lineares. Assemelham-se ao cérebro biológico, pois são capazes de extrair informações de exemplos apresentados e generalizar esse conhecimento posteriormente<sup>[21]</sup>. Outras vantagens incluem: alta precisão, baixo custo computacional e fácil implementação. Tem sido usadas em muitas áreas: engenharia biomédica, neuro-computação, engenharia química, aplicações industriais, geologia, medicina, química e física<sup>[22–32]</sup>.

Neste trabalho, usamos conceitos de RNA para estimar a energia do estado fundamental de sistemas fermiônicos interagentes em cadeias homogêneas descritas pelo modelo de Hubbard. Obtivemos com sucesso funcionais baseados em RNAs, e os usamos como *input* em cálculos de DFT para o estudo de sistemas heterogêneos, como na presença de impurezas e potenciais confinantes. Nossos funcionais via RNAs mostraram desempenho superior aos funcionais analíticos em todos os regimes de parâmetros e se revelaram uma alternativa confiável e menos custosa computacionalmente para cálculos de DFT em cadeias finitas e heterogêneas.

## 2 Fundamentos Teóricos

### 2.1 Teoria do Funcional da Densidade

A Teoria do Funcional da Densidade (DFT)<sup>[2,4]</sup> é fundamentada no teorema de Hohenberg-Kohn (HK), que estabelece a relação unívoca entre a densidade  $n_0(\mathbf{r})$ , de um sistema de  $N$  partículas no estado fundamental sob a ação de um potencial externo  $v(\mathbf{r})$ , com a função de onda do sistema  $\Psi_0(r_1, r_2, \dots, r_N)$ . Desse modo, a função de onda é um *funcional da densidade*

$$\Psi_0(r_1, r_2, \dots, r_N) = \Psi_0[n_0(\mathbf{r})], \quad (2.1)$$

onde a densidade  $n_0(\mathbf{r})$  é definida como

$$n_0(\mathbf{r}) = N \int d^3r_2 \int d^3r_3 \dots \int d^3r_N \Psi_0^*(\mathbf{r}, r_2, \dots, r_N) \Psi_0(\mathbf{r}, r_2, \dots, r_N). \quad (2.2)$$

O teorema de HK tem como consequência que o valor esperado do estado fundamental de um dado observável  $\hat{O}$  é, também, um funcional da densidade

$$O_0 = O[n_0] = \langle \Psi_0[n_0] | \hat{O} | \Psi_0[n_0] \rangle. \quad (2.3)$$

Um dos observáveis mais importantes é a energia do estado fundamental

$$E_0 = E[n_0] = \langle \Psi[n_0] | \hat{H} | \Psi[n_0] \rangle, \quad (2.4)$$

onde  $\hat{H} = \hat{T} + \hat{U} + \hat{V}$  é a soma dos operadores de energia cinética  $\hat{T}$ , de interação Coulombiana  $\hat{U}$ , e do operador do potencial externo  $\hat{V}$ , que também são funcionais da densidade.

Os funcionais da densidade para a energia cinética,  $T[n]$ , e para interação Coulombiana,  $U[n]$ , são independentes do potencial externo  $v(\mathbf{r})$  e portanto, são determinados apenas pela estrutura dos operadores  $\hat{T}$  e  $\hat{U}$ . Desse modo, podem ser escritos como o funcional universal  $F[n]$

$$F[n] = T[n] + U[n]. \quad (2.5)$$

O funcional da densidade para o potencial externo  $V[n]$ , é dependente de  $v(\mathbf{r})$ , que define o sistema. É escrito como

$$V[n] = \int d^3r n(\mathbf{r}) v(\mathbf{r}), \quad (2.6)$$

onde, uma vez determinado o sistema através de  $v(\mathbf{r})$ ,  $V[n]$  é conhecido explicitamente.

A energia do estado fundamental pode ser obtida minimizando o funcional  $E[n]$

$$E[n] = F[n] + V[n]. \quad (2.7)$$

Em relação a densidade,

$$\begin{aligned} E_0 &= \min_n E[n] \\ &= \min_n [F[n] + V[n]]. \end{aligned} \quad (2.8)$$

A minimização do funcional da energia  $E[n]$  leva à

$$\begin{aligned} \frac{\delta E[n]}{\delta n(r)} &= \frac{\delta F[n]}{\delta n(r)} + \frac{\delta V[n]}{\delta n(r)} \\ 0 &= \frac{\delta F[n]}{\delta n(r)} + v(r), \end{aligned} \quad (2.9)$$

onde  $v(r) = \delta V[n]/\delta n(r)$  refere-se ao potencial externo.

Embora com o teorema de HK seja possível, em teoria, obter todos os observáveis do sistema, uma vez que são funcionais da densidade  $n_0(r)$ , na prática, o funcional  $F[n]$  não é conhecido explicitamente. Desse modo, são necessárias aproximações para a energia cinética e para a interação Coulombiana.

### 2.1.1 Aproximações de Thomas-Fermi e LDA

A aproximação mais simples para o funcional  $F[n]$  é a aproximação de Thomas-Fermi. A energia Coulombiana do sistema interagente  $U[n]$  é aproximada para a *energia de Hartree*,

$$U[n] \approx E_H[n] = \frac{e^2}{2} \int d^3r \int d^3r' \frac{n(r)n(r')}{|r-r'|}, \quad (2.10)$$

que é a energia eletrostática clássica.

A energia cinética  $T[n]$  é obtida através de duas aproximações. A primeira, a aproximação de densidade local (LDA – do inglês *local density approximation*)<sup>[17]</sup>

$$T[n(r)] \approx T^{LDA}[n] = \int d^3r t^{hom}(n)|_{n \rightarrow n(r)}, \quad (2.11)$$

onde aproxima-se o funcional da densidade  $T[n]$  de um sistema interagente heterogêneo, através da integração da densidade volumétrica de energia  $t^{hom}(n) = \lim_{V \rightarrow \infty} T^{hom}(n)/V$  de um sistema interagente homogêneo, substituindo localmente a densidade  $n$  por  $n(r)$ . A segunda aproximação consiste em assumir a energia cinética de um sistema homogêneo *não-interagente*, no lugar da densidade de energia do sistema homogêneo interagente,

$$t^{hom}(n) \approx t_s^{hom}(n) = \frac{3\hbar^2}{10m} (3\pi^2)^{2/3} n^{5/3}. \quad (2.12)$$

Logo,  $T[n]$  adquire a forma

$$T[n(r)] \approx T_s^{LDA}[n] = \int d^3r t_s^{hom}(n)|_{n \rightarrow n(r)}, \quad (2.13)$$

e desse modo, a energia total, Eq.(2.7), na aproximação de Thomas-Fermi é dada por

$$E[n] \approx E^{TF}[n] = T_s^{LDA}[n] + E_H[n] + V[n], \quad (2.14)$$

assim como o funcional universal  $F[n]$ , Eq.(2.5)

$$F[n] \approx F^{TF}[n] = T_s^{LDA}[n] + E_H[n]. \quad (2.15)$$

Embora a simplicidade da aproximação de Thomas-Fermi facilite sua implementação, as aproximações da interação, Eq.(2.10) e, da energia cinética, Eq.(2.13), ignoram os efeitos de correlação e portanto afetam a precisão.

### 2.1.2 Energia de troca e correlação

O funcional da *energia de troca e correlação* é definido como

$$E_{XC}[n] = T[n] - T_s[n] + U[n] - E_H[n], \quad (2.16)$$

e contém as informações perdidas na aproximação de Thomas-Fermi,  $T[n] \approx T_s[n]$  e  $U[n] \approx E_H[n]$ . O funcional  $E_{XC}$ , pode ser decomposto em  $E_X + E_C$ , onde,

$$E_X = \frac{-q^2}{2} \sum_{jk} \int d^3r \int d^3r' \frac{\phi_j^*(r)\phi_k^*(r')\phi_j(r')\phi_k(r)}{|r-r'|} \quad (2.17)$$

refere-se a energia de troca, correspondente ao princípio de exclusão de Pauli, escrito em função dos orbitais  $\phi_i(r)$ , de um sistema não interagente. O funcional  $E_C$  refere-se às correlações e não é conhecido explicitamente.

Desse modo, é possível escrever o funcional da energia para sistemas interagentes

$$E[n] = T_s[n] + E_H[n] + E_{XC}[n] + V[n], \quad (2.18)$$

e sua minimização leva à

$$\begin{aligned} 0 &= \frac{\delta E[n]}{\delta n(r)} = \frac{\delta T_s[n]}{\delta n(r)} + \frac{\delta V[n]}{\delta n(r)} + \frac{\delta U_H[n]}{\delta n(r)} \\ &+ \frac{\delta E_{xc}[n]}{\delta n(r)} = \frac{\delta T_s[n]}{\delta n(r)} + v(r) + v_H(r) + v_{xc}(r). \end{aligned} \quad (2.19)$$

Como consequência da Eq.(2.6), o potencial externo  $v(r) = \delta V[n]/\delta n(r)$  é adicionado. O termo  $v_H(r) = \delta E_H[n]/\delta n(r)$  é o potencial de Hartree e  $v_{xc}(r) = \delta E_{XC}[n]/\delta n(r)$  é o potencial de troca e correlação.

Embora, as equações Eq.(2.18) e Eq.(2.19) sejam formalmente exatas, os funcionais  $T_s[n]$  e  $E_{XC}$  não são conhecidos de forma explícita. Desse modo, é possível abordar o problema através do esquema proposto por Kohn e Sham<sup>[5]</sup>, onde a minimização Eq.(2.19) é realizada indiretamente.

### 2.1.3 Esquema de Kohn-Sham

O esquema de Kohn-Sham (KS) propõe o mapeamento de um sistema interagente, com densidade  $n$  e energia  $E[n]$ , em um sistema auxiliar não-interagente, com a mesma densidade  $n$  e energia  $E_{KS}[n]$

$$E_{KS}[n] = T_{KS}[n] + V_{KS}[n], \quad (2.20)$$

uma vez que para um sistema não-interagente  $E_H = 0$  e  $E_{XC} = 0$ . A minimização da energia  $E_{KS}[n]$ , leva à

$$\begin{aligned} \frac{\delta E_{KS}[n]}{\delta n(\mathbf{r})} &= \frac{\delta T_{KS}[n]}{\delta n(\mathbf{r})} + \frac{\delta V_{KS}[n]}{\delta n(\mathbf{r})} \\ 0 &= \frac{\delta T_{KS}[n]}{\delta n(\mathbf{r})} + v_{KS}(\mathbf{r}) \end{aligned} \quad (2.21)$$

Igualando as equações Eq.(2.19) e Eq.(2.21), observamos que o potencial de Kohn-Sham é dado por:

$$v_{KS}(r) = v(r) + v_H(r) + v_{XC}(r). \quad (2.22)$$

Consequentemente, é possível calcular a densidade de um sistema interagente de muitos corpos sob efeito de um potencial  $v(r)$ , descrito por uma equação de Schrödinger, resolvendo-se a equação para um sistema não-interagente (ou de partícula única) sob efeito do potencial  $v_{KS}(r)$ . A equação de Schrödinger para o sistema auxiliar de KS, é dada por

$$\left[ \frac{-\hbar^2 \nabla^2}{2m} + v_{KS}(r) \right] \phi_i(r) = \epsilon_i \phi_i(r), \quad (2.23)$$

onde  $n(r) = n_{KS}(r)$  é dada por

$$n(r) = \sum_i^N |\phi_i(r)|^2, \quad (2.24)$$

e  $\phi_i(r)$  são os orbitais de Kohn-Sham.

O conjunto de equações Eqs.(2.22 – 2.24), constitui as equações de Kohn-Sham. Desse modo, substitui-se o problema da minimização de  $E[n]$  ao se resolver a equação de Schrödinger para um sistema não-interagente.

Uma vez que  $v_{XC}$  depende de  $\phi_i$ , que por sua vez depende de  $v_{KS}$ , a solução das equações de KS requer um processo auto-consistente, como ilustrado na Figura 1. Assim, inicia-se com uma suposição para a densidade  $n_0^k(r)$ , calcula-se o potencial de KS correspondente  $v_{KS}(r)$  e, então resolve-se a equação Eq.(2.23) para  $\phi_i$ . Em seguida, calcula-se uma nova densidade através da equação Eq.(2.24), que será usada em um novo ciclo do cálculo. Continua-se o ciclo até que haja convergência entre as densidades, isto é,  $|n_0^{k+1}(r) - n_0^k(r)| < \text{convergência}$ . Uma vez concluído o cálculo, é possível obter o funcional da energia do estado fundamental  $E_0[n_0(r)]$ .

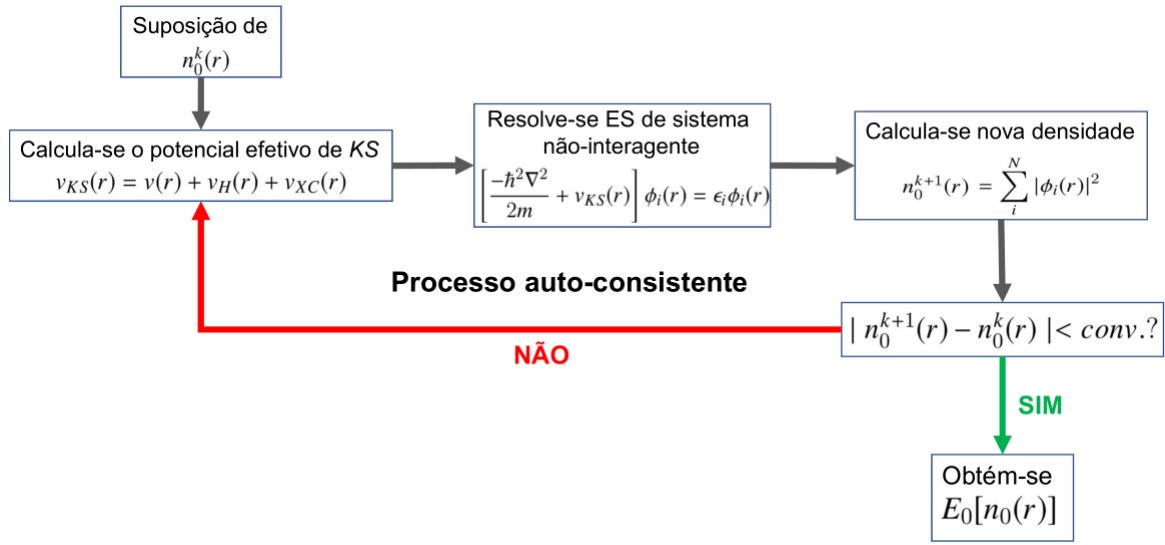


Figura 1 – Cálculo auto-consistente das equações de Kohn-Sham

## 2.2 Modelo de Hubbard

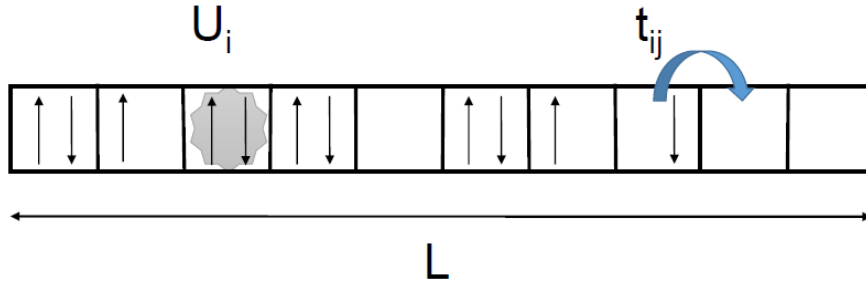
Dada uma cadeia unidimensional de tamanho  $L$ , o movimento e a interação de férmions, como os elétrons, são descritos pelo modelo de Hubbard como ilustrado na Figura 2. Cada sítio representa um orbital eletrônico não-degenerado e, seguindo o Princípio de Exclusão de Pauli, só pode ser ocupado por no máximo dois elétrons de spins opostos. Dois elétrons em um mesmo sítio  $i$  sofrem uma interação Coulombiana  $U_i$ . A energia cinética dessas partículas é dada pelo termo de *hopping*  $t_{ij}$ , que é a probabilidade de uma das partículas se mover do sítio  $i$  para o sítio vizinho  $j$ . O potencial externo que pode, em princípio ser distinto em cada sítio, é dado por  $v_{i\sigma}$ , onde  $i$  refere-se ao sítio em questão e  $\sigma$  é a orientação do spin (para cima ou para baixo). Desse modo, o Hamiltoniano para o modelo de Hubbard unidimensional é dado por

$$\hat{H} = - \sum_{ij\sigma} t_{ij} (\hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} + H.c.) + \sum_i U_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} + \sum_{i\sigma} v_{i\sigma} \hat{n}_{i\sigma}, \quad (2.25)$$

onde,  $\hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma}$ , são os operadores de criação e aniquilação de partículas de spin  $\sigma$  em um dado sítio  $i$  e  $H.c.$  é o operador Hermitiano Conjugado do termo  $\hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma}$ .

O operador densidade do sistema é  $\hat{n}_{i\sigma} = \hat{c}_{i\sigma}^\dagger \hat{c}_{i\sigma}$ , com  $\sum_{i\sigma} \langle \hat{n}_{i\sigma} \rangle = n = N/L$ ,  $N = \langle \hat{n}_\uparrow \rangle + \langle \hat{n}_\downarrow \rangle$  é o número total de partículas no sistema e  $m = \langle \hat{n}_\uparrow \rangle - \langle \hat{n}_\downarrow \rangle$  a magnetização.

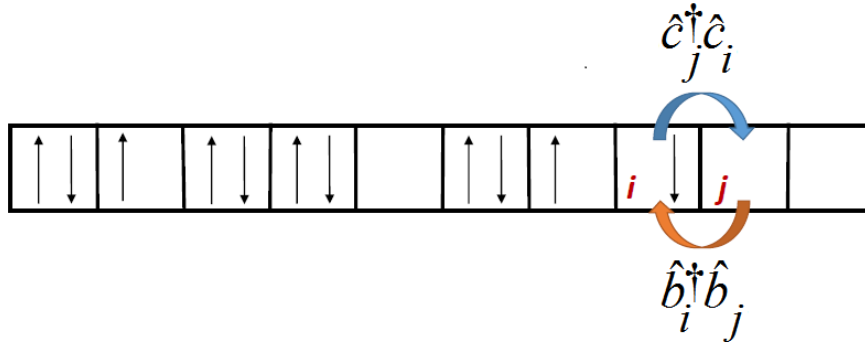
A itinerância de elétrons ao longo da cadeia depende das probabilidades de *hopping*  $t_{ij}$  entre os sítios  $i$  e  $j$ , assim como da interação Coulombiana  $U$ . Por exemplo, se em um sítio  $i$  há apenas um elétron com spin  $\sigma$ , a probabilidade de um elétron nos sítios vizinhos  $j = i \pm 1$  migrar para o sítio  $i$  será nula,  $t_{ij} = 0$ , caso tenha mesmo spin  $\sigma$ , e baixa, se estiver sozinho no sítio  $j$ , pois o custo energético de manter dois elétrons em um mesmo sítio é maior por causa da interação  $U$ . Entretanto, caso um sítio  $i$  esteja vazio e um sítio



**Figura 2** – Representação de uma cadeia unidimensional com  $L$  sítios do modelo de Hubbard. Cada sítio pode ser ocupado por dois elétrons de spin opostos que sofrem interações Coulombianas  $U$ , onde o movimento desses elétrons é descrito pelo termo de *hopping*  $t_{ij}$ , com  $j = i \pm 1$ .

vizinho  $j$  estiver duplamente ocupado, a probabilidade de *hopping* de um elétron do sítio  $j$  para o sítio  $i$  será alta, tendendo a atingir a configuração de menor energia.

Para o caso em que a interação é igual em todos os sítios,  $U_i = U$ , o movimento dos elétrons entre o sítio  $i$  e o sítio  $j$ , pode ser analisado de forma simétrica através do movimento de buracos ou ausência de partículas, como ilustrado na Figura 3.



**Figura 3** – O *Hopping* de um elétron do sítio  $i$  para o sítio  $j$ , através da operação  $\hat{c}_j^\dagger \hat{c}_i$ , seguida de uma operação do movimento simétrico de buraco do sítio  $j$  para o sítio  $i$ ,  $\hat{b}_i^\dagger \hat{b}_j$ , onde  $\hat{b}_i^\dagger = (-1)^i \hat{c}_i$  e  $\hat{b}_i = (-1)^i \hat{c}_i^\dagger$  são os operadores de criação e aniquilação de buracos.

O modelo de Hubbard também é capaz de descrever sistemas *homogêneos* e *heterogêneos*. Por exemplo, dado um sistema de tamanho infinito,  $L = \infty$ , sem a presença de potenciais externos  $v_{i\sigma} = 0$ , com a mesma interação Coulombiana  $U_i = U$  e energia cinética  $t_{ij} = t$  em todos os sítio, a distribuição de densidades  $n_i$ , será uniforme  $n_i = n = N/L$  e teremos um sistema homogêneo. Do mesmo modo, caso o sistema seja finito com condição de contorno periódica (PBC – do inglês *periodic boundary conditions*), isto é, sem a presença de bordas, também terá distribuição de densidades homogênea,  $n_i = n = N/L$ . Por outro lado, o sistema será heterogêneo caso apresente condição aberta de contorno (OBC – do inglês *open boundary conditions*), uma vez que a presença de bordas afeta o

perfil de densidade apresentando oscilações de Friedel<sup>[33]</sup> nos sítios próximos às bordas. O sistema também será heterogêneo caso apresente potencial externo  $v_{i\sigma} \neq 0$  a distribuição de densidades dependerá da forma de  $v_{i\sigma}$ .

O Hamiltoniano do modelo de Hubbard apresenta soluções exatas para alguns casos. Para sistemas pequenos,  $L \leq 20$ , a solução numérica exata é obtida via diagonalização de matrizes. No limite termodinâmico,  $L \rightarrow \infty$ , onde o sistema é homogêneo, a solução numérica exata é obtida através das integrais de Lieb-Wu<sup>[34,35]</sup>, porém as integrais não são funções diretas de  $n$  e  $m$ , assim os parâmetros não são diretamente controláveis. Para  $L < 200$  pode-se resolver via DMRG<sup>[36]</sup> (do inglês – *Density matrix renormalization group*), mas com elevado custo computacional. Por exemplo, para uma cadeia homogênea de tamanho  $L = 100$ , o cálculo via DMRG leva cerca de 10 horas utilizando-se computação paralela em 16 processadores do GridUNESP. Para cadeias de até cerca de  $L \sim 10^3$ , podemos utilizar a DFT e várias informações do Hamiltoniano podem ser obtidas, inclusive a energia por sítio, utilizando-se aproximações para o modelo. Cálculos via DFT podem ser realizados em computadores pessoais e levam poucos minutos para serem concluídos.

Para o modelo de Hubbard unidimensional, existem dois funcionais de relativo sucesso que retornam a energia do estado fundamental por sítio  $e_0$ : *LSOC*<sup>[17]</sup> e *FVC*<sup>[18]</sup>. O primeiro, apresenta desvios relativamente pequenos ( $\sim 6\%$ ), mas é limitado a sistemas não-magnetizados. A parametrização do funcional LSOC é dada por

$$e_0^{\text{LSOC}}(n, U) = -\frac{2\beta(U)}{\pi} \sin\left(\frac{\pi n}{\beta(U)}\right), \quad (2.26)$$

onde  $n = n_\uparrow + n_\downarrow$  é a densidade e a função  $\beta(U)$ , que é determinada pela equação transcendental

$$-\frac{2\beta(U)}{\pi} \sin\left(\frac{\pi}{\beta(U)}\right) = -4 \int_0^\infty dx \frac{J_0(x)J_1(x)}{x(1 + e^{Ux/2})}, \quad (2.27)$$

onde  $J_i(x)$  são funções de Bessel de ordem  $i$ .

O funcional FVC descreve adequadamente sistemas magnetizados com desvios menores que 2% para a maioria dos regimes, com exceção dos regimes de fraca interação,  $U \leq 7$ , com desvios maiores que o desejável ( $\sim 4\%$ )<sup>[16,37]</sup>. A parametrização do funcional FVC é dada por

$$e_0^{\text{FVC}}(n, m, U) = -\frac{2\beta(n, m, U)}{\pi} \sin\left(\frac{\pi n}{\beta(n, m, U)}\right) \cos\left(\frac{\pi m}{\gamma(n, m, U)}\right), \quad (2.28)$$



onde

$$\beta(n, m, U) = \beta(U)^{\alpha(n, m, U)}, \quad (2.29)$$

$$\gamma(n, m, U) = 2 \exp \left[ \frac{\sqrt{U}}{1 - (m/n)^{3/2}} \right], \quad (2.30)$$

$$\alpha(n, m, U) = \left[ \frac{n^2 - m^2}{n^{15/8}} \right] \sqrt[3]{U}. \quad (2.31)$$

Aqui,  $\beta(U)$  é o mesmo empregado na parametrização LSOC, isto é, obtido através da igualdade da Eq.(2.27).

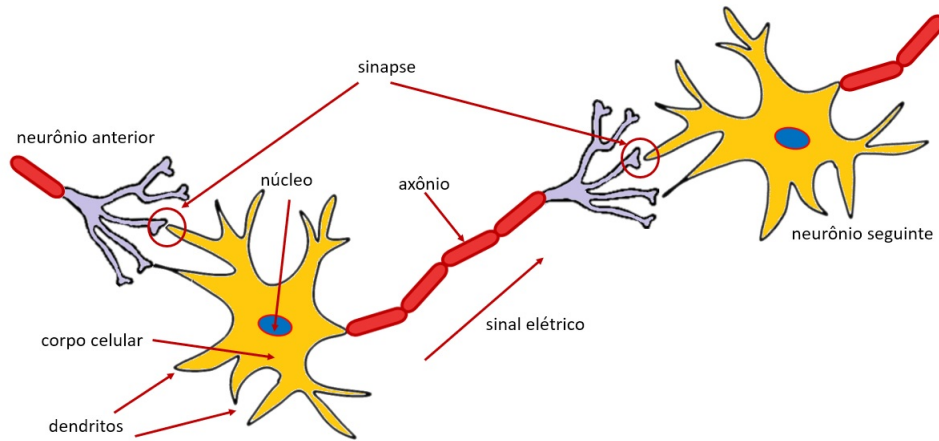
## 2.3 Redes Neurais Artificiais

Redes neurais artificiais (RNAs) são definidas como sistemas paralelos distribuídos compostos de unidades básicas de processamento, os neurônios, que calculam funções matemáticas, por exemplo, funções não-lineares. Esses neurônios são distribuídos em uma ou mais camadas.

As redes neurais artificiais têm como inspiração a estrutura biológica do cérebro. Assemelham-se principalmente na capacidade de adquirir conhecimento do ambiente através de um processo de aprendizagem e de armazenar o conhecimento em pesos sinápticos, que ponderam a informação recebida por cada neurônio de entrada.

A resolução de problemas utilizando RNAs se dá em duas etapas: uma etapa de *aprendizagem*, em que se alimenta a rede com exemplos e ela é treinada, extraíndo características pertinentes à descrição do problema, seguindo um algoritmo de aprendizado adequado; uma etapa de *generalização*, na qual a rede usa as características extraídas anteriormente para gerar respostas coerentes ao problema.

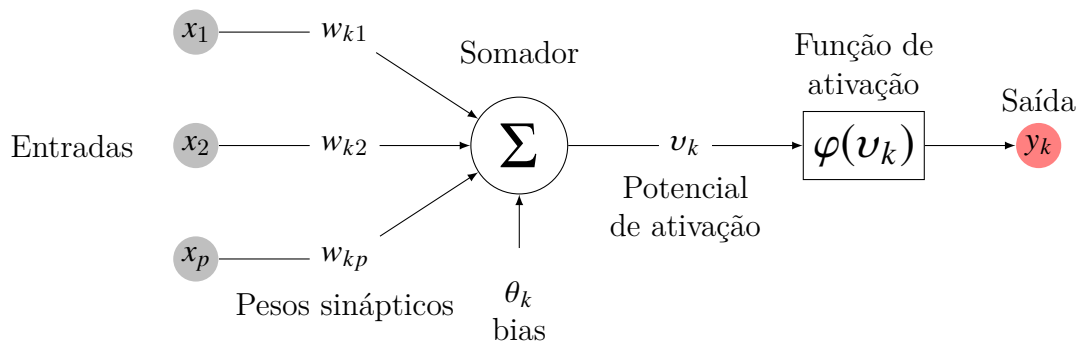
Nossos neurônios, constantemente utilizam padrões que aprendemos ao longo da vida para aplicá-los em situações novas. Nosso cérebro é formado por bilhões de neurônios, cada qual constituído pelo corpo central, pelos dendritos — que são os terminais de entrada — e pelo axônio, que contém os terminais de saída, como ilustra a Figura 4. Os dendritos recebem os estímulos (as informações) e os propagam ao corpo do neurônio, onde são processados até atingir um limite, quando são então transmitidos ao axônio, que os enviam ao neurônio vizinho através das sinapses. Assim, a informação é processada e transmitida através dos neurônios e nosso cérebro emite comandos para que executemos nossas tarefas.



**Figura 4** – Esquema de neurônio biológico

### 2.3.1 Modelos Matemáticos de Neurônios

Um *neurônio artificial* é uma unidade básica de processamento de informação, fundamental no funcionamento de redes neurais artificiais. A figura Figura 5 ilustra o modelo de um neurônio artificiais, que é definido por três elementos básicos:



**Figura 5** – Neurônio não-linear baseado no *perceptron*.

- (i) Conjunto de sinapses: cada qual com um *peso*  $w_{kj}$  característico. O sinal de entrada  $x_j$  da sinapse  $j$  conectada ao neurônio  $k$  é multiplicado pelo peso sináptico  $w_{kj}$ ;
- (ii) Somador: soma os sinais de entrada, cada qual ponderado pelas respectivas sinapses do neurônio;
- (iii) Função de ativação ou função de transferência: limita a amplitude da saída do neurônio.

O modelo do neurônio da Figura 5 também inclui um *bias* aplicado externamente, denotado por  $b_k$ . O *bias*  $b_k$  tem o efeito de aumentar ou diminuir a entrada da função de ativação, dependendo se a mesma é positiva ou negativa, respectivamente.

Em termos matemáticos, o neurônio  $k$  na Figura 5 é descrito pelas equações

$$u_k = \sum_{j=1}^p w_{kj} x_j, \quad (2.32)$$

$$v_k = u_k + b_k, \quad (2.33)$$

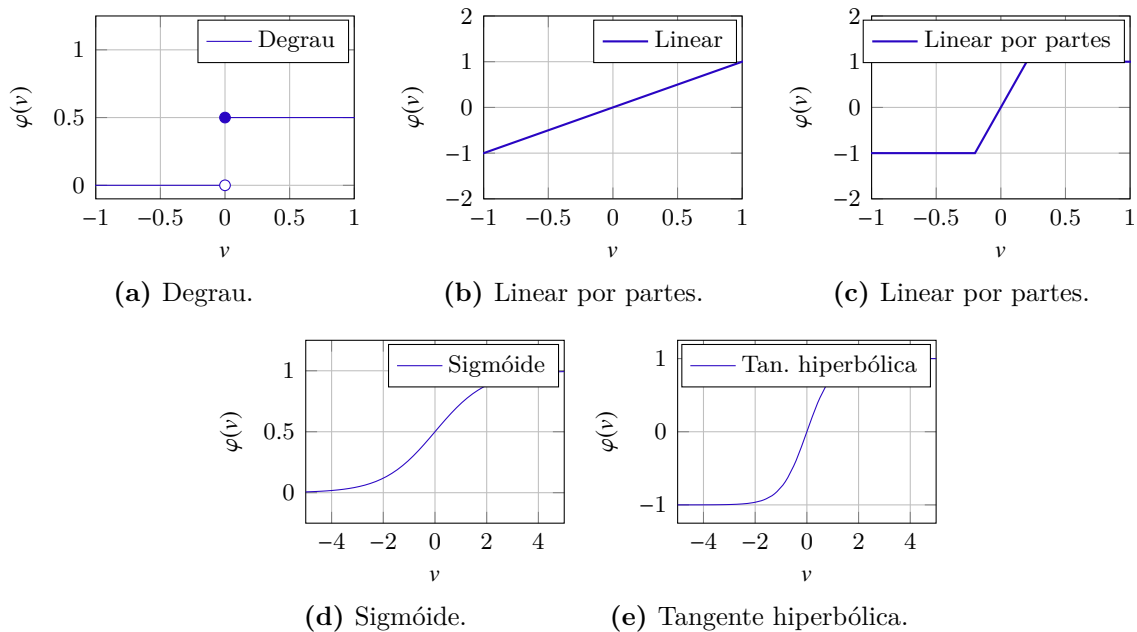
$$y_k = \varphi(u_k + b_k), \quad (2.34)$$

onde  $x_1, x_2, \dots, x_p$  são os sinais de entrada;  $w_{k1}, w_{k2}, \dots, w_{kp}$  são os respectivos pesos sinápticos do neurônio  $k$ ;  $u_k$ , não mostrado na Figura 5, é o somador devido os sinais de entrada;  $v_k$  é o *potencial de ativação*;  $b_k$  é o *bias*;  $\varphi(\cdot)$  é a *função de ativação*;  $y_k$  é o sinal de saída do neurônio.

Assim como os neurônios biológicos, Figura 4, no neurônio computacional, Figura 5, a informação chega nos canais de entrada (dentridos), é processada no somador (corpo celular) é transmita até a saída (sinapses) passando pela função de ativação (axônio).

### 2.3.1.1 Funções de Ativação

A função de ativação  $\varphi(v)$ , define a saída de um neurônio em termos do potencial de ativação  $v$ . Existem alguns tipos de funções de ativação. A Figura 6 apresenta algumas delas.



**Figura 6** – Funções de ativação comumente utilizadas.

Matematicamente, essas funções podem ser definidas como:

(a) Função de ativação degrau:

$$\varphi(v) = \begin{cases} 1, & \text{se } v \geq 0 \\ 0, & \text{se } v < 0 \end{cases} \quad (2.35)$$

(b) Função de ativação linear ou identidade:

$$\varphi(v) = v \quad (2.36)$$

(c) Função de ativação linear por partes:

$$\varphi(v) = \begin{cases} 1, & \text{se } v \geq a \\ v, & \text{se } -a < v < a \\ -1, & \text{se } v \leq -a \end{cases} \quad (2.37)$$

(d) Função de ativação sigmóide:

$$\varphi(v) = \frac{1}{1 + e^{-v}} \quad (2.38)$$

(e) Função de ativação tangente hiperbólica:

$$\varphi(v) = \frac{2}{1 + e^{-2v}} - 1 \quad (2.39)$$

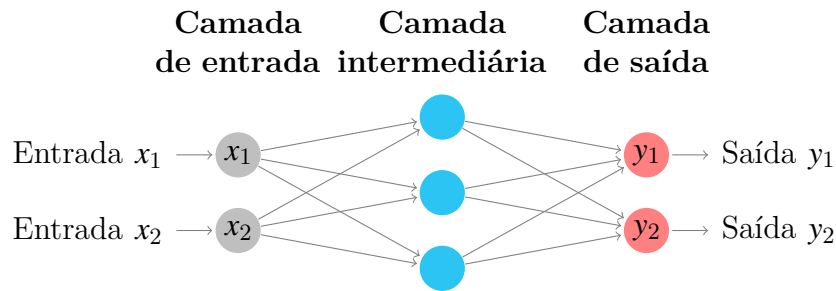
### 2.3.2 Arquiteturas de Redes Neurais Artificiais

As RNAs são definidas segundo três parâmetros que as constituem, sendo eles: a arquitetura, o tipo de neurônio e o tipo de aprendizagem.

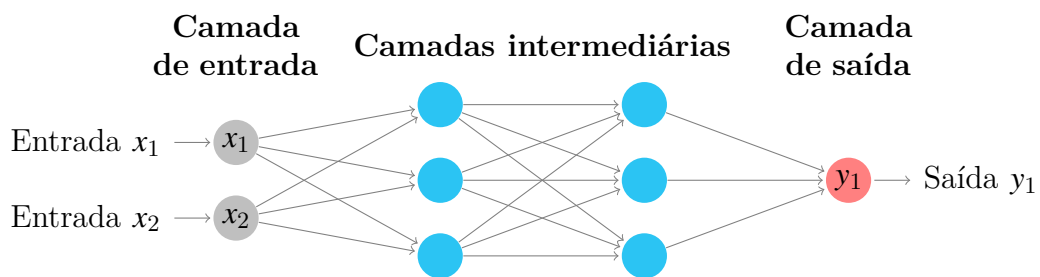
No que se diz respeito a arquitetura, a rede neural é formada por diferentes camadas. Usualmente, uma rede neural é constituída pelas seguintes camadas:

- (i) **Camada de Entrada:** onde informações são introduzidas à rede;
- (ii) **Camadas Intermediárias:** onde a maior parte do processamento é realizado;
- (iii) **Camada de Saída:** onde o resultado final é apresentado.

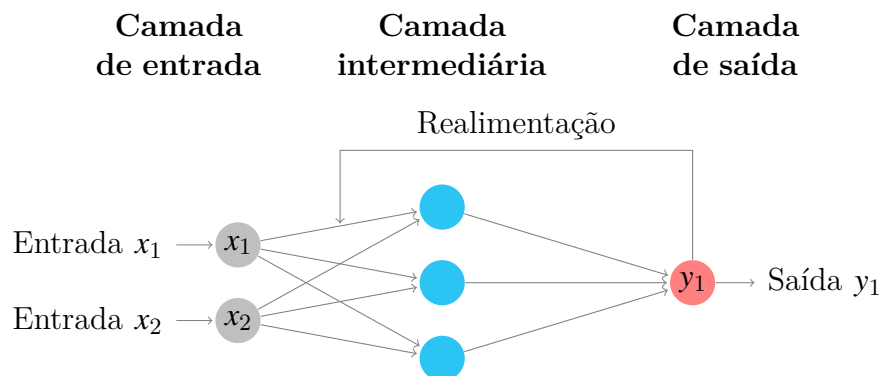
As arquiteturas das redes neurais são divididas em três classes fundamentais: redes *feed-forward* de uma única camada, redes *feed-forward* de múltiplas camadas e redes recorrentes, como mostram as Figuras 7 – 9. Nas redes neurais *feed-forward*, sejam elas de uma, Figura 7, ou mais camadas, Figura 8, a informação sempre se propaga para frente, passando por todos os neurônios e nunca no sentido oposto. A diferença entre as redes *feed-forward* de uma única ou múltiplas camadas, consiste no número de camadas intermediárias, que ficam entre a camada de entrada e a camada de saída.



**Figura 7** – Rede neural *feed-forward* de uma única camada



**Figura 8** – Rede *feed-forward* de múltiplas camadas



**Figura 9** – Rede neural recorrente

Nas redes neurais recorrentes, Figura 9, existem conexões de realimentação, onde as conexões de um neurônio podem estar ligadas em neurônios da mesma ou de camadas anteriores. As conexões de realimentação, agem na rede neural analogamente à *memória de curto prazo*.

As redes de múltiplas camadas, também podem ser totalmente conectadas, parcialmente conectadas ou localmente conectadas.

Nas redes totalmente conectadas, todos os neurônios de uma dada camada estão ligados a todos os neurônios da camada seguinte. Nas redes parcialmente conectadas, apenas parte dos neurônios de uma dada camada estão ligados à apenas uma parte dos neurônios da camada seguinte. Nas redes localmente conectadas, certas regiões de uma

camada se ligam a outras regiões da camada seguinte.

### 2.3.3 Tipos de Aprendizagem

Os processos de aprendizagem através dos quais as redes neurais funcionam podem ser classificados em três tipos:

- (i) **Aprendizagem supervisionada:** é o método de aprendizagem mais comum visto em redes neurais. A rede é treinada através do fornecimento de padrões de entrada e de saída desejados para a rede. O objetivo é ajustar os parâmetros da rede, de forma a encontrar uma ligação entre os pares de entrada e saída fornecidos. Assim, a cada saída gerada, compara-se com a saída desejada e efetua-se um ajuste nos pesos sinápticos de modo a minimizar este erro;
- (ii) **Aprendizagem não-supervisionada:** Nesse caso, não são fornecidos à rede pares de entrada e saída, apenas padrões de entrada. Desse modo, a rede procurar algum tipo de correlação ou redundância nos dados de entrada;
- (iii) **Aprendizagem por reforço:** Considerada como um caso particular de aprendizagem supervisionada. No aprendizado supervisionado, utiliza-se um critério de erro conhecido, enquanto que no caso por reforço, defini-se uma regra a priori para aquele sistema que possa indicar se a saída é boa ou ruim.

### 3 Métodos Computacionais

Consideramos cadeias unidimensionais homogêneas descritas pelo modelo de Hubbard, Eq. (2.25). Para o treinamento da rede neural, inicialmente foram obtidos 20.891 resultados de energia do estado fundamental através da solução das integrais de Lieb-Wu nos regimes  $0 \leq U \leq 10$ ,  $0 \leq n \leq 1$  e  $0 \leq m \leq n$ .

O treinamento do modelo neural se deu através do algoritmo de aprendizagem de Levenberg-Marquardt (LM)<sup>[38]</sup>, em uma rede do tipo *feedforward*, onde a informação se propaga para frente, camada à camada. O algoritmo utiliza uma modificação do método de Gauss-Newton<sup>[39]</sup>, bastante utilizado em aplicações que se reduzem a um problema de mínimos quadrados

$$f(W) = \frac{1}{2} \sum_{j \in C} E_j^2(k), \quad (3.1)$$

onde  $W$  são os pesos e  $C$  é conjunto de todos os neurônios da rede, onde  $E_j(k)$  é o erro relativo ao neurônio de saída  $j$  do exemplo de treinamento  $k$  dado por

$$E_j(k) = d_j(k) - y_j(k), \quad (3.2)$$

sendo  $d_j(k)$  se refere à resposta desejada para o neurônio  $j$ , aqui sendo a energia do estado fundamental exata  $\{e_0^{FN}(U, n, m)\}$  e  $y_j(k)$  à resposta da rede, aqui sendo a energia do estado fundamental via RNA,  $\{e_0^{RNA}(U, n, m)\}$ .

O desempenho da rede é calculado pelo *erro quadrado médio*

$$EQM = \frac{1}{K} \sum_{k=1}^K f(W), \quad (3.3)$$

onde  $K$  é o número total de dados para energia,  $\{e_0^{FN}(U, n, m)\}$ , fornecidos à rede.

A equação usada no método de Gauss-Newton para atualização dos pesos da RNA e consequente minimização do valor do EQM é

$$W_{i+1} = - [\nabla^2 f(W)]^{-1} \nabla f(W), \quad (3.4)$$

onde  $\nabla^2 f(W)$  é a matriz Hessiana

$$J \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_k} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_k \partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_k} & \cdots & \frac{\partial^2 f}{\partial x_k \partial x_k} \end{bmatrix}, \quad (3.5)$$

que é a matriz quadrada das derivadas parciais de segunda ordem, ou seja, a matriz jacobiana do gradiente.

A matriz Jacobiana  $J$  é

$$J(x) = \begin{bmatrix} \frac{\partial E_1(x)}{\partial x_1} & \frac{\partial E_1(x)}{\partial x_2} & \dots & \frac{\partial E_1(x)}{\partial x_k} \\ \frac{\partial E_2(x)}{\partial x_1} & \frac{\partial E_2(x)}{\partial x_2} & \dots & \frac{\partial E_2(x)}{\partial x_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial E_K(x)}{\partial x_1} & \frac{\partial E_K(x)}{\partial x_2} & \dots & \frac{\partial E_K(x)}{\partial x_k} \end{bmatrix}. \quad (3.6)$$

Se  $f(W) = \frac{1}{2} \sum_{j \in C} e_j^2(n)$  e  $\nabla f(W)$  é o gradiente de  $f(W)$ , então

$$\nabla f(W) = J^T \cdot e, \quad (3.7)$$

sendo  $J^T$  a matriz Jacobiana transposta.

Assim, podemos escrever a matriz Hessiana como

$$\nabla^2 f(W) = J^T J + S, \quad (3.8)$$

onde  $S$  é

$$S = \sum_{i=1}^I E_i \nabla^2 E_i. \quad (3.9)$$

Assumindo  $S \approx 0$  quando comparado ao produto da matriz Jacobiana, então a matriz Hessiana pode ser aproximada por

$$\nabla^2 f(W) \approx J^T(x)J(x). \quad (3.10)$$

O que nos dá o algoritmo de Gauss-Newton

$$W_{i+1} = W_i - \left( J^T J \right)^{-1} J^T \cdot E. \quad (3.11)$$

A modificação dada pelo algoritmo de Levenberg-Marquardt ao método de Gauss-Newton é dada pela adição do termo  $\mu$ , que torna possível a inversão da matriz Hessiana

$$W_{i+1} = W_i - \left( J^T J + \mu I \right)^{-1} J^T E, \quad (3.12)$$

onde  $I$  é a matriz identidade,  $E$  é o erro,  $J$  é a matriz Jacobiana e  $\mu$ , um parâmetro que torna a matriz Hessiana definida positiva.

O uso do algoritmo de Levenberg-Marquardt se dá da seguinte forma:

- 1) Apresenta-se à rede todos os padrões de entrada  $(U, n, m)$  com as saídas correspondentes  $\{e_0^{FN}(U, n, m)\}$ ;
- 2) A rede então calcula novas saídas,  $\{e_0^{RNA}(U, n, m)\}$ , referentes aos padrões de entrada;
- 3) Calcula-se o erro Eq.(3.2) e a soma dos erros quadráticos médios Eq.(3.3);



- 4) Calcula-se a matriz Jacobiana;
- 3) Resolve-se a equação Eq.(3.12);
- 4) Recalcula-se a soma dos erros quadráticos;
- 5) O treinamento converge quando o gradiente Eq.(3.7) atinge um valor menor que o pré-determinado ou quando o erro quadrático médio atinge o valor desejado.

O algoritmo é implementado na ferramenta *Neural Network Fitting Toolbox* do software *MATLAB*. O conjunto de dados utilizado para treinar a rede foi dividido aleatoriamente em três conjuntos. O primeiro conjunto, contendo 70% dos dados, é usado para computar o gradiente e atualizar os pesos e *bias*. O segundo conjunto, contendo 15%, é o de validação, onde o erro tende a diminuir junto com o erro do conjunto de treinamento durante o treino. Caso o erro da validação aumente consecutivamente para um número pré-estabelecido de épocas, o treinamento para. E o último conjunto, contendo 15% dos dados é usado para teste.

A densidade  $n$ , a interação intra-sítio  $U$  e a magnetização  $m$  são as entradas da rede neural  $(U, n, m)$ , seguida de camada(s) intermediária(s) com um certo número de neurônios e a saída da rede é a energia do estado fundamental  $\{e_0^{RNA}(U, n, m)\}$ . Considerou-se como critérios de parada do treinamento, o número de 1000 épocas e 6 falhas na validação. Esses critérios foram considerados com o intuito de evitar o fenômeno de *overfitting*, onde o erro do treinamento tem um valor muito pequeno, mas na etapa de generalização o erro é muito maior.

Uma vez obtido um modelo neural com um desempenho ótimo, o mesmo foi utilizado para gerar novos valores de energia do estado fundamental em diferentes regimes de  $U$ ,  $n$  e  $m$ , mas ainda dentro dos mesmos intervalos do conjunto de treinamento.

Para a descrição das nanoestruturas heterogêneas — como por exemplo na presença de impurezas, desordem e confinamento harmônico — usamos DFT implementada com a *local-spin-density approximation* (LSDA):

$$e_0^{hom} \approx e_0^{LSDA} = \frac{1}{L} \sum_{i=1}^L e_0^{hom}[n, m, U] \Big|_{\substack{n \rightarrow n_i \\ m \rightarrow m_i}} \quad (3.13)$$

onde  $e_0^{hom}$  é o funcional homogêneo para a energia total do estado fundamental e  $n_i, m_i$  são as densidades locais para o sistema heterogêneo. Assim, a aproximação da Eq. (3.13) permite aproximar a energia do estado fundamental do sistema heterogêneo através da energia do estado fundamental do sistema homogêneo, aproximando  $n, m$ , pelas densidades locais,  $n_i, m_i$ .

Utilizamos como *input* para a DFT-LSDA a energia do estado fundamental do sistema homogêneo obtida via os funcionais:  $e_0^{FN}$ ,  $e_0^{RNA}$  e  $e_0^{RNA}$ . Consideramos como

---

resultados exatos, aqueles obtidos via *DMRG*, uma vez que o funcional *FN* é exato apenas em sistemas infinitos homogêneos.

## 4 Resultados e Discussões

A Figura 10 apresenta a arquitetura 3–20–1 do modelo neural que obteve o melhor desempenho dentre àqueles treinados. Foram consideradas 1000 épocas de treinamento, com o intuito de evitar *over-fitting* do modelo neural, o que resultaria em uma pior generalização. O desempenho, medido através do erro quadrado médio, foi de  $2,70E - 07$ . Os neurônios da camada intermediária apresentam função de ativação sigmoide e o neurônio de saída apresenta função de ativação linear. A Tabela 1 apresenta o desempenho do treino de todas as arquiteturas consideradas. Para a rede 3–5–1, o treinamento parou em 172 épocas, após 6 falhas consecutivas na validação, com um erro quadrado médio de  $5,67E - 06$ . Para a rede 3–10–1, o treino parou após 1000 épocas de treinamento com melhora em uma ordem de grandeza no erro quadrado médio e em mais uma ordem de grandeza para a rede 3–20–1. Para as arquiteturas com duas camadas intermediárias, 3–10–10–1 e 3–12–12–1 não foram observadas melhoras apreciáveis em relação à rede 3–20–1. Para todos os casos, o treinamento foi completado em poucos minutos.

**Tabela 1** – Arquitetura e desempenho dos modelos neurais treinados. O desempenho é medido pelo erro quadrado médio.

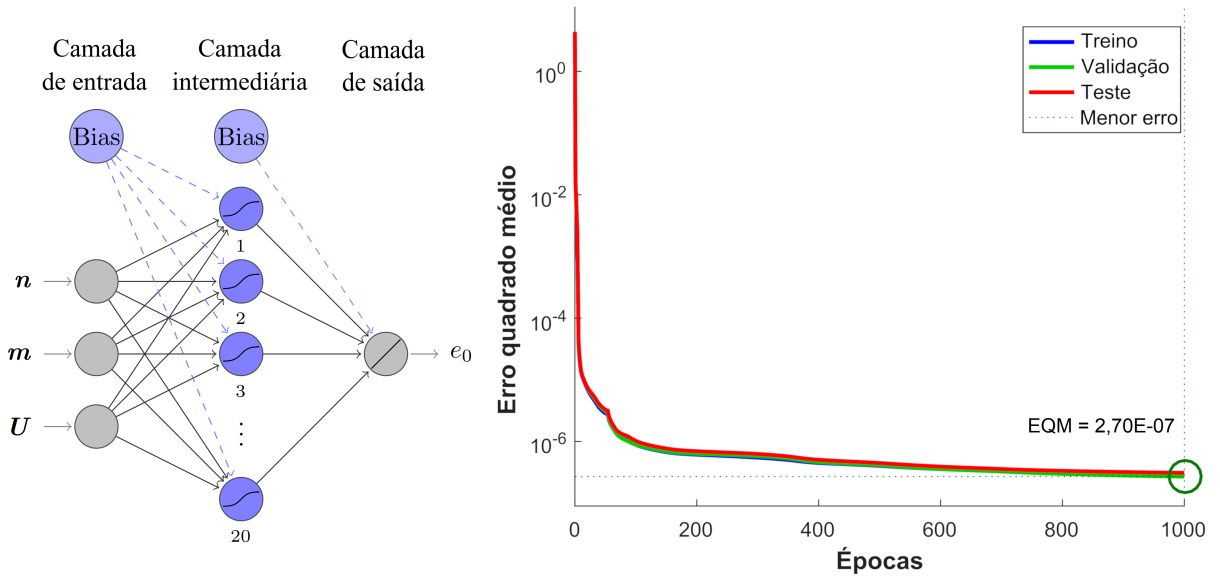
Arquitetura	Épocas de treinamento	Erro Quadrado Médio
3–5–1	172	$5,67E - 05$
3–10–1	1000	$1,25E - 06$
3–20–1	1000	$2,70E - 07$
3–10–10–1	1000	$9,84E - 08$
3–12–12–1	1000	$1,18E - 07$

**Fonte** – Produzido pelos autores

### 4.1 Sistemas homogêneos

Nessa seção comparamos o desempenho do nosso modelo neural com a parametrização FVC e o quanto ambas divergem da energia numericamente exata obtida via integrais de Lieb-Wu, aqui chamada de "FN" (*Fully numerical*). Consideramos cadeias homogêneas unidimensionais de tamanho  $L = \infty$  em diferentes regimes de interação, densidade e magnetização.

A Figura 11 apresenta o comportamento da energia por sítio do estado fundamental para o modelo de Hubbard em relação a densidade  $n$ . O regime estudado é de um sistema metálico fortemente correlacionado,  $U = 6t$  e não magnetizado  $m \sim 0$ . Para pequenas densidades, até  $n \sim 0,6$ , é possível observar diminuição na energia por sítio do estado

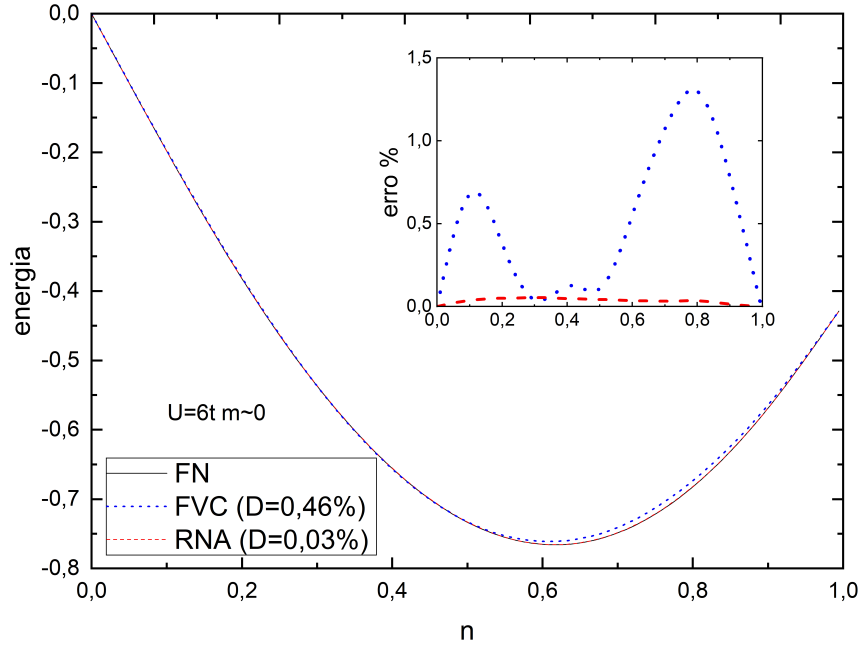


**Figura 10** – Estrutura da rede neural de arquitetura 3-20-1 e seu desempenho durante treinamento

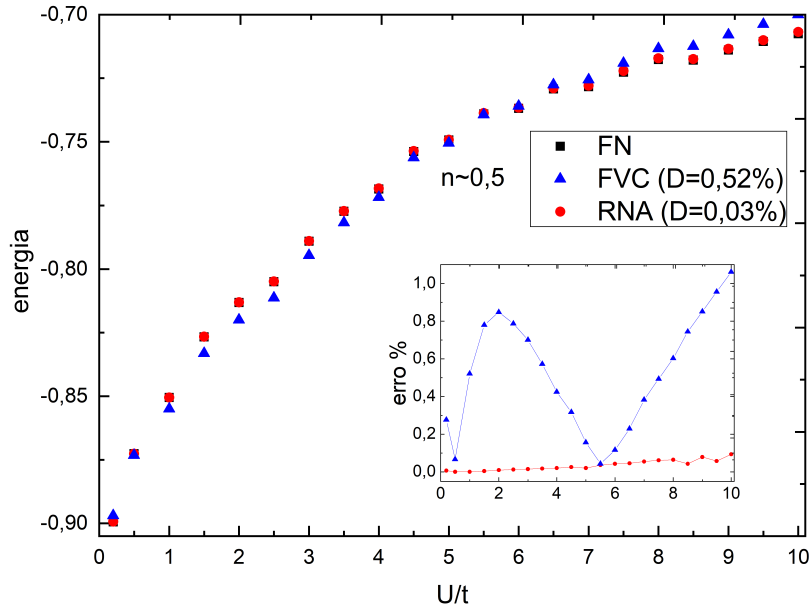
fundamental à medida em que a cadeia é preenchida. Isso ocorre porque para baixas densidades há grande mobilidade dos elétrons no metal e, portanto, o termo negativo da energia cinética é mais relevante que o positivo de interação. Em  $n \sim 0,6$  a energia por sítio atinge um mínimo e rapidamente aumenta a medida que o sistema se aproxima de  $n = 1$ , onde o metal se torna um isolante de Mott, que é a localização das partículas induzida pela interação Coulombiana elétron-elétron. Comparamos os resultados da energia obtidos via RNA e via FVC com o resultado numericamente exato FN. É possível observar que ambas as abordagens, RNA e FVC, conseguem descrever com sucesso o perfil energético desse sistema. A RNA tem o melhor desempenho, com desvios menores que 0,1%, enquanto a FVC, considerada suficientemente precisa, tem desvios menores que  $\sim 1,5\%$ .

Na Figura 12 apresentamos a energia no estado fundamental por sítio para o modelo de Hubbard em função da interação  $U$  em um sistema não-magnético,  $m = 0$ . Observa-se que, com o aumento da interação  $U$ , a energia do estado fundamental por sítio aumenta. Esse comportamento se deve ao fato de que com uma interação repulsiva mais intensa, a itinerância de elétrons de um sítio  $i$  duplamente ocupado, para um sítio  $j$ , requer mais energia, refletindo no perfil energético para esse sistema. Observa-se que a RNA apresenta resultados excelentes não só em regimes fortemente correlacionados, como também em todo o intervalo  $0 \leq U \leq 10$ . No *inset* observa-se desvios percentuais menores que  $\sim 0,08\%$ , enquanto o desvio percentual médio da RNA é de 0,03%. Para o funcional FVC temos desvio percentual médio de 0,5% e no *inset* vemos desvios de até  $\sim 1\%$ .

A Figura 13 mostra a energia do estado fundamental por sítio em função da magnetização  $m$ , dada pela população desbalanceada de *spins*, para diversos regimes de interação  $U$ . O comportamento observado é o do aumento na energia por sítio do estado

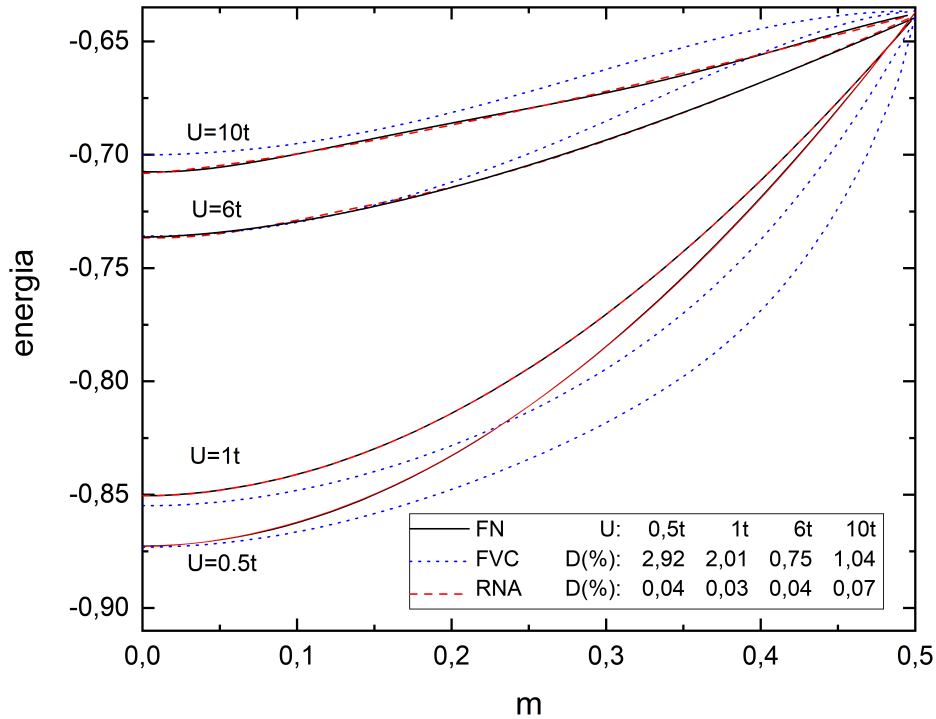


**Figura 11** – Energia do estado fundamental por sítio em função da densidade  $n$  para os funcionais FVC e RNA em comparação com o resultado numericamente exato FN. Regime de interação  $U = 6t$ , magnetização  $m \sim 0$ . *Inset*: erro percentual  $100 \left| (e_0^{FVC(RNA)} - e_0^{FN}) / e_0^{FN} \right|$ .  $D$  é o desvio médio percentual definido como  $100 \sum \left| (e_0^{FVC(RNA)} - e_0^{FN}) / e_0^{FN} \right|$ .



**Figura 12** – Energia do estado fundamental por sítio em função da interação intra-sítio  $U$  (em unidades de  $t$ ), para os funcionais FVC e RNA em comparação com o resultado numericamente exato FN. Regime de densidade  $n \sim 0,5$  e magnetização  $m \sim 0$ . *Inset*: erro percentual  $100 \left| (e_0^{FVC(RNA)} - e_0^{FN}) / e_0^{FN} \right|$ .  $D$  é o desvio médio percentual definido como  $100 \sum \left| (e_0^{FVC(RNA)} - e_0^{FN}) / e_0^{FN} \right|$ .

fundamental à media em que aumenta-se a magnetização do sistema. Isso acontece pois com o aumento da magnetização, um maior número de elétrons é forçado em uma mesma orientação de *spin*. Para sistemas com menor interação elétron-elétron, em  $U = 0,5t$  e  $U = 1t$ , a variação da energia é maior enquanto para forte interação, em  $U = 6t$  e  $U = 10t$ , apresenta uma variação menor. Isso é devido a maior probabilidade de itinerância em sistemas com fraca interação em relação àqueles com forte interação. Comparando-se os funcionais RNA e FVC, o funcional FVC apresenta desvios grandes e falha até mesmo na descrição qualitativa do perfil de energia. Os maiores desvios encontram-se no regime de fraca interação e maior magnetização. A RNA, por outro lado, apresenta um ótimo desempenho em todos os regimes de interação e magnetização, com desvios percentuais médios menores que 0,1%.



**Figura 13** – Energia do estado fundamental por sítio em função da magnetização  $m$ , para os funcionais FVC e RNA em comparação com o resultado numericamente exato FN. Regime de densidade  $n \sim 0,5$ . *Inset*: erro percentual  $100 \left| (e_0^{FVC(RNA)} - e_0^{FN}) / e_0^{FN} \right|$ .  $D$  é o desvio médio percentual definido como  $100 \sum \left| (e_0^{FVC(RNA)} - e_0^{FN}) / e_0^{FN} \right|$ .

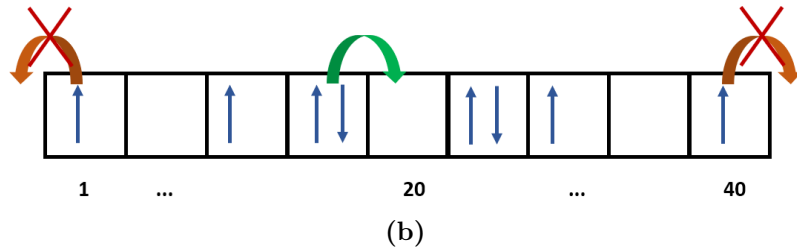
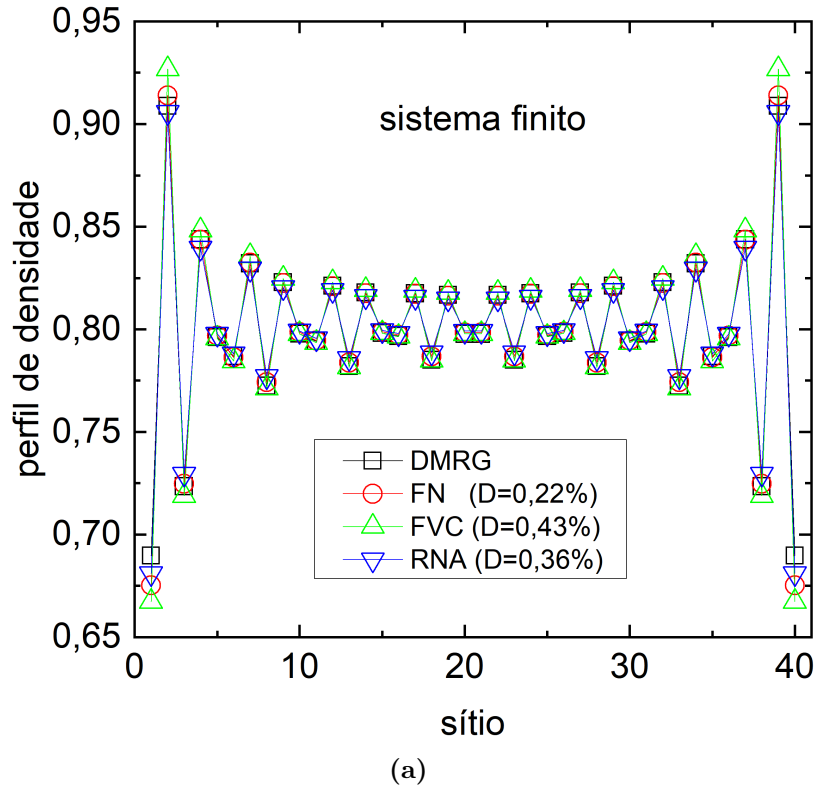
## 4.2 Sistemas heterogêneos

Nessa seção comparamos o desempenho dos funcionais FN, FVC e RNA com DMRG em cálculos DFT-LSDA para sistemas heterogêneos. Simulamos sistemas finitos, com impureza localizada e sistemas harmonicamente confinados, alterando o termo do potencial externo  $v_{i\sigma}$  do Hamiltoniano do modelo de Hubbard Eq. 2.25.

Na Figura 14 (a), temos o perfil de densidade para uma cadeia finita de tamanho  $L = 40$ , em regime de densidade  $n = 0,8$ , interação  $U = 1t$  e sem magnetização em condição de cadeia aberta (OBC). A Figura 14 (b) ilustra essa cadeia. O perfil de densidade para essa cadeia finita, não-confinada e, sem a presença de impurezas, apresenta oscilações de Friedel<sup>[33]</sup> nas densidades locais. As oscilações são maiores próximo as bordas, que atuam como um tipo de impureza localizada de densidade infinita. Nas bordas o número de sítios disponíveis para a movimentação de elétrons também é menor, o que contribui com oscilações maiores nessas regiões. Os funcionais FN, FVC e RNA recuperam com sucesso o comportamento qualitativo do perfil de densidade. O desvio percentual médio para ambos os funcionais são pequenos: o funcional FVC ( $D = 0,43$ ) apresenta o maior desvio, seguido do funcional RNA ( $D = 0,36\%$ ) e por fim do FN ( $D = 0,22\%$ ). Mesmo com desvio médio percentual pequeno para o funcional FVC, ao analisar o perfil de densidade ponto à ponto, podemos observar que este funcional ora subestima, ora sobrestima os valores do perfil de densidade, especialmente nas regiões onde as oscilações são maiores. Por outro lado, o funcional FN e RNA, se mantem consistentes ao longo de toda a cadeia.

A Figura 15 (a) mostra o perfil de densidade para um sistema de tamanho  $L = 15$  com impureza localizada ( $V_i = V\delta_{i,i_V}$ ) de intensidade  $V = 2t$  em regimes de interação  $U = 1t$  e não-magnetizado, com condição de cadeia aberta. A Figura 15 (b) ilustra essa cadeia. Observa-se que, para uma impureza repulsiva,  $V = 2t$ , localizada no meio da cadeia, a densidade é praticamente nula no ponto da impureza, dividindo a cadeia em dois subsistemas. Desse modo, para elétrons se moverem de um subsistema para outro, obrigatoriamente devem passar pela impureza. Também é possível observar oscilações de Friedel<sup>[33]</sup> no perfil de densidade surgindo ao redor da impureza e nas bordas. Em sistemas metálicos qualquer perturbação localizada gera oscilações, assim como nas bordas de cadeias finitas abertas, que também agem como impurezas. Comparamos os resultados DFT-LSDA obtidos através dos funcionais FN, FVC e RNA com os resultados exatos via DMRG. Observa-se que todos os funcionais mostraram-se capazes de recuperar qualitativamente o perfil de densidade para este sistema. O *inset* apresenta o erro percentual de cada funcional em relação à DMRG. É possível observar que os funcionais apresentam desvios maiores para a densidade no sítio da impureza. Entretanto, o funcional FVC apresenta o maior desvio ( $\sim 10\%$ ). O desvio médio percentual é também ligeiramente maior para o funcional FVC em relação aos outros.

Na Figura 16 (a), comparamos o perfil de densidade obtido à partir dos funcionais em cálculos DFT-LSDA com resultados via DMRG para um sistema cujo potencial externo é parabólico ou harmônico,  $V_i = k(i - i_0)^2$ . A Figura 16 (b) ilustra essa cadeia. Este potencial simula, por exemplo, o confinamento de átomos ultra-frios via lasers e tem como característica o acúmulo das partículas nos sítios centrais. Assim a densidade é máxima em torno do sítio central e mínima nas bordas. Em torno de densidade  $n \sim 0,6$ , observamos a formação de um pequeno platô, que reflete a maior probabilidade de dupla ocupação.

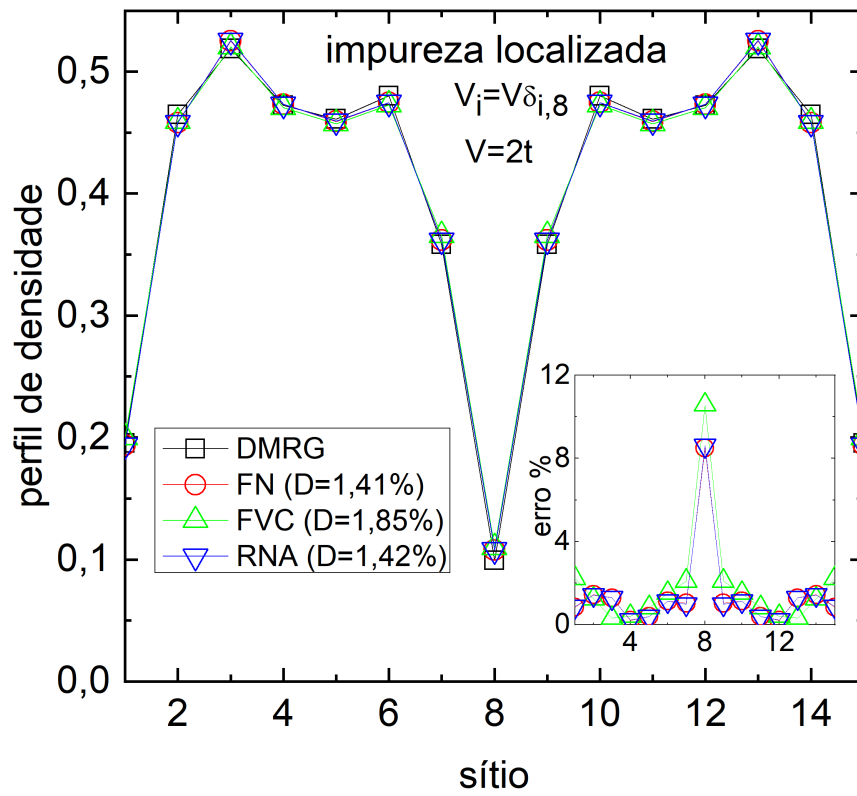


**Figura 14** – (a) Perfil de densidade e (b) ilustração para uma cadeia finita de tamanho  $L = 40$ ,  $n = 0,8$ ,  $U = 1t$  e  $m \sim 0$  com condição de cadeia aberta. Os resultados exatos são via DMRG e os aproximados via DFT-LSDA com um dos funcionais: FN, FVC ou RNA.  $D$  refere-se ao desvio médio percentual definido como  $100 \sum \left| (e_0^{FN(FVC, RNA)} - e_0^{DMRG}) / e_0^{DMRG} \right|$ .

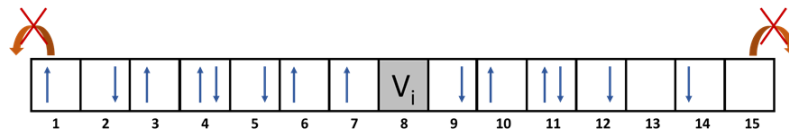
Observamos que para uma dada curvatura  $k$  do potencial, todos os funcionais recuperam qualitativamente o perfil de densidade, mas enquanto o desvio médio da RNA é comparável ao do FN, da ordem de 0,5%, o desvio médio da FVC é consideravelmente maior,  $\sim 1,7\%$ .

Na Figura 17 analisamos o perfil energia em função da intensidade  $k$  do potencial do sistema descrito na Figura 16. Apesar dos desvios médios dos três funcionais serem baixos em comparação à DMRG, verifica-se que o funcional FVC apresenta sistematicamente desvios maiores que os demais, como pode ser observado no *inset*.



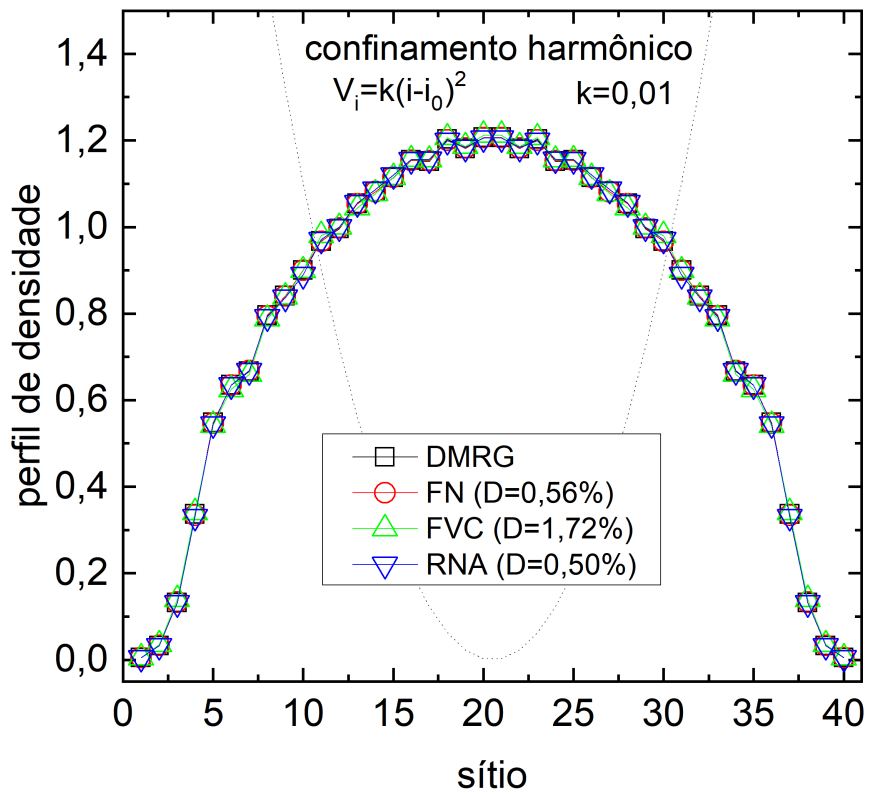


(a)

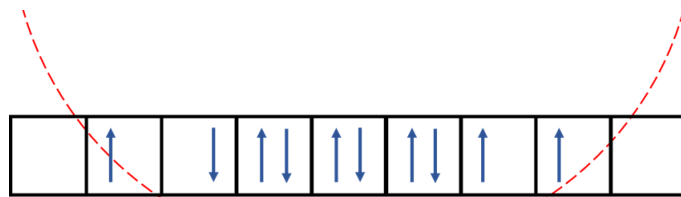


(b)

**Figura 15** – (a) Perfil de densidade e (b) ilustração para um sistema com impureza localizada, de tamanho  $L = 15$ , intensidade da impureza  $V = 2t$ ,  $n = 0,4$ ,  $U = 1t$  e  $m \sim 0$  com condição de cadeia aberta. Os resultados exatos são via DMRG e os aproximados via DFT-LSDA com um dos funcionais: FN, FVC ou RNA. *Inset* traz os desvios percentuais definido como  $100 \left| (e_0^{FN(FVC, RNA)} - e_0^{DMRG}) / e_0^{DMRG} \right|$  e  $D$  refere-se ao desvio médio percentual definido como  $100 \sum \left| (e_0^{FN(FVC, RNA)} - e_0^{DMRG}) / e_0^{DMRG} \right|$ .

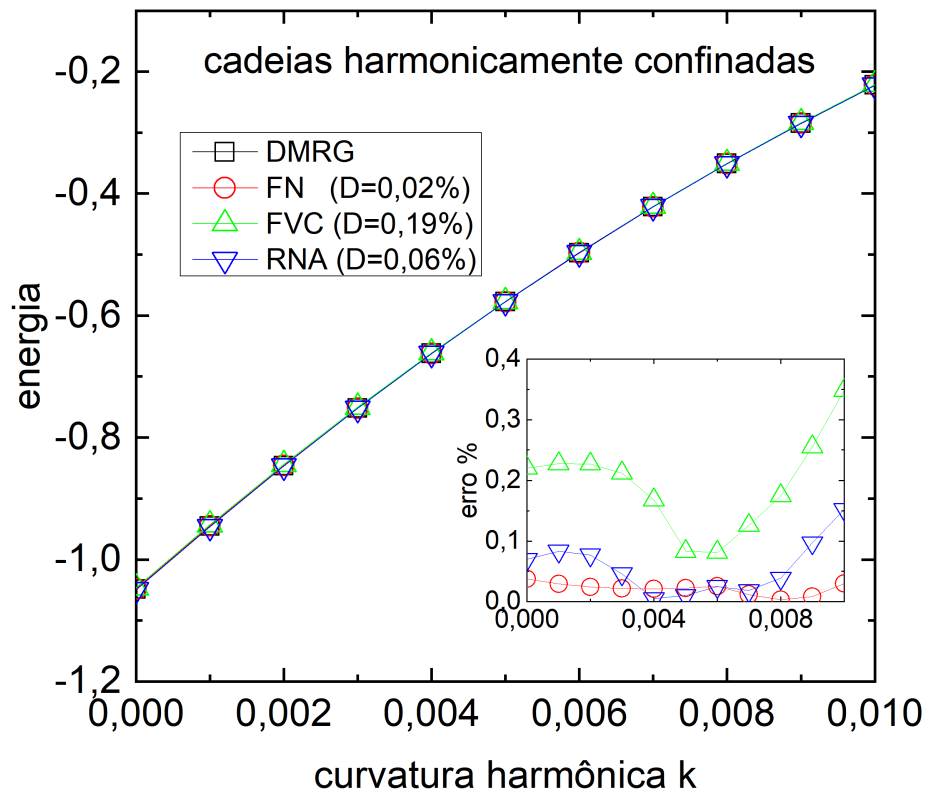


(a)



(b)

**Figura 16** – (a) Perfil de densidade e (b) ilustração para uma cadeia harmonicamente confinada de tamanho  $L = 40$ , curvatura harmônica  $k = 0,01$ ,  $n = 0,8$ ,  $U = 1t$  e  $m \sim 0$  com condição de cadeia aberta. Os resultados exatos são via DMRG e os aproximados via DFT-LSDA com um dos funcionais: FN, FVC ou RNA.  $D$  refere-se ao desvio médio percentual definido como  $100 \sum \left| (e_0^{FN(FVC, RNA)} - e_0^{DMRG}) / e_0^{DMRG} \right|$ .



**Figura 17** – Energia por sítio do estado fundamental em função da curvatura harmônica confinada em cadeias de tamanho  $L = 40$ ,  $n = 0,8$ ,  $U = 1t$  e  $m \sim 0$  com condição de cadeia aberta. Os resultados exatos são via DMRG e os aproximados via DFT-LSDA com um dos funcionais: FN, FVC ou RNA. *Inset* traz os desvios percentuais definido como  $100 \left| (e_0^{FN(FVC, RNA)} - e_0^{DMRG}) / e_0^{DMRG} \right|$ .  $D$  refere-se ao desvio médio percentual definido como  $100 \sum \left| (e_0^{FN(FVC, RNA)} - e_0^{DMRG}) / e_0^{DMRG} \right|$ .

## 5 Conclusão

Nesse trabalho propusemos o desenvolvimento de um funcional da densidade para a energia do estado fundamental de cadeias fermiônicas unidimensionais descritas pelo modelo de Hubbard, utilizando conceitos de redes neurais artificiais. Posteriormente, aplicamos o funcional obtido em cálculos DFT-LSDA para explorar sistemas heterogêneos.

O modelo neural obtido de arquitetura 3 – 20 – 1 descreve com sucesso sistemas homogêneos em todos os regimes de interação, densidade e magnetização estudados nesse trabalho. Na Tabela 2 temos o comparativo da RNA e FVC em relação ao resultado numericamente exato FN. É possível observar que a RNA mantém desvios médios percentuais menores que 0,07% em todos os regimes do sistema homogêneo. Por outro lado o funcional FVC apresenta desvios maiores principalmente no regime de fraca interação, com desvios que variam entre  $\sim 2\%$  à  $\sim 3\%$ .

**Tabela 2** – Desvios médios percentuais para energia do estado fundamental de sistemas homogêneos obtida via funcionais FVC e RNA.

	Desvio médio %		
	FVC	RNA	
Energia em relação a densidade	0,46	0,03	
Energia em relação a interação	0,52	0,03	
	U = 0,5t	2,92	0,04
	U = 1t	2,01	0,03
Energia em relação a magnetização	U = 6t	0,75	0,04
	U = 10t	1,04	0,07

**Fonte** – Produzido pelos autores

Para sistemas heterogêneos, simulamos uma cadeia finita, uma cadeia com impureza localizada e cadeia com confinamento harmônico. Utilizamos como *input* na aproximação LSDA a energia do estado fundamental obtida via FN, FVC e RNA e comparamos com a energia do estado fundamental exata obtida via DMRG. Na Tabela 3 temos o comparativo do desvio médio percentual para os três funcionais. Observamos que os desvios médios percentuais são pequenos para ambos os funcionais. Entretanto, para o funcional FVC o desvio médio percentual não mede o desempenho real em todos os regimes. Analisando ponto à ponto, é possível observar que o este funcional ora subestima, ora sobrestima os valores da energia do estado fundamental. Por outro lado, tanto FN, quanto RNA, se mantêm consistentes. Entretanto, o funcional FN tem desvantagens em relação à RNA, pois não temos total liberdade da escolha dos parâmetros.

Comparando os métodos explorados nesse trabalho para a obtenção da energia do

**Tabela 3** – Desvios médios percentuais para energia do estado fundamental de sistemas homogêneos obtida via funcionais FN, FVC e RNA.

	Desvio médio %		
	FN	FVC	RNA
Sistema finito	0,22	0,43	0,36
Impureza localizada	1,41	1,85	1,42
Confinamento harmônico	0,56	1,72	0,5
Energia x curvatura harmônica	0,02	0,19	0,06

**Fonte** – Produzido pelos autores

estado fundamental de sistemas homogêneos e heterogêneos, concluímos que o funcional obtido via redes neurais artificiais se sobressai em relação aos outros métodos explorados nesse estudo. A RNA apresenta desvios pequenos e consistentes, com precisão na ordem de  $\sim 0,1\%$  da DMRG, com custo computacional de  $\sim 10^{-3}$  da DMRG. Também apresenta total liberdade na escolha dos parâmetros de densidade, interação e magnetização, diferentemente do funcional FN.

Por fim, os resultados obtidos nesse trabalho foram publicados em artigo na revista *Scientific Reports*<sup>[40]</sup>.

# Referências

- 1 BRÜNNER, T. et al. Entanglement enhancement in spatially inhomogeneous many-body systems. *Physical Review A*, American Physical Society, v. 87, n. 3, p. 32311, mar 2013. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.87.032311>>. 15
- 2 HOHENBERG, P.; KOHN, W. Inhomogeneous Electron Gas. *Physical Review*, American Physical Society, v. 136, n. 3B, p. B864–B871, nov 1964. ISSN 0031-899X. Disponível em: <<http://link.aps.org/doi/10.1103/PhysRev.136.B864>>. 15, 17
- 3 KOHN, W. Nobel Lecture: Electronic structure of matter - wave functions and density functionals. *Rev. Mod. Phys.*, American Physical Society, v. 71, n. 5, p. 1253–1266, 1999. Disponível em: <<http://link.aps.org/doi/10.1103/RevModPhys.71.1253>>. 15
- 4 CAPELLE, K. A Bird's-Eye View of Density-Functional Theory. *Brazilian Journal of Physics*, v. 36, n. 4A, p. 1318–1343, 2006. ISSN 0103-9733. 15, 17
- 5 KOHN, W.; SHAM, L. J. Self-consistent equations including exchange and correlation effects. *Physical Review*, v. 140, n. 4A, 1965. ISSN 0031899X. 15, 19
- 6 HUBBARD, J. Electron Correlations in Narrow Energy Bands. *Source: Proceedings of the Royal Society of London. Series A Mathematical and Physical Sciences*, v. 276, n. 1365, p. 238–257, 1963. Disponível em: <<http://www.jstor.org/stable/2414761>>. 15
- 7 COE, J. P.; FRANÇA, V. V.; D'AMICO, I. Hubbard model as an approximation to the entanglement in nanostructures. *Physical Review A*, American Physical Society, v. 81, n. 5, p. 052321, may 2010. ISSN 1050-2947. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.81.052321>>. 16
- 8 COE, J. P.; FRANÇA, V. V.; D'AMICO, I. Feasibility of approximating spatial and local entanglement in long-range interacting systems using the extended Hubbard model. *EPL (Europhysics Letters)*, v. 93, n. 1, p. 10001, 2011. Disponível em: <<http://stacks.iop.org/0295-5075/93/i=1/a=10001>>. 16
- 9 COE, J. P.; FRANÇA, V. V.; D'AMICO, I. Approximation of the entanglement in quantum dot chains using Hubbard models. *Journal of Physics: Conference Series*, v. 286, n. 1, p. 12048, 2011. Disponível em: <<http://stacks.iop.org/1742-6596/286/i=1/a=012048>>. 16
- 10 FRANÇA, V. V.; CAPELLE, K. Entanglement in spatially inhomogeneous many-fermion systems. *Physical Review Letters*, v. 100, n. 7, p. 1–4, 2008. ISSN 00319007. 16
- 11 FRANÇA, V. V.; HÖRNDLEIN, D.; BUCHLEITNER, A. Fulde-Ferrell-Larkin-Ovchinnikov critical polarization in one-dimensional fermionic optical lattices. *Physical Review A*, American Physical Society, v. 86, n. 3, p. 033622, sep 2012. ISSN 1050-2947. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.86.033622>>. 16
- 12 JOHNSON, T. et al. Hubbard Model for Atomic Impurities Bound by the Vortex Lattice of a Rotating Bose-Einstein Condensate. *Physical Review Letters*, American

- Physical Society, v. 116, n. 24, p. 240402, jun 2016. ISSN 0031-9007. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.116.240402>>. 16
- 13 CAPELLE, K.; CAMPO, V. L. Density functionals and model Hamiltonians: Pillars of many-particle physics. *Physics Reports*, v. 528, n. 3, p. 91–159, 2013. ISSN 03701573. 16
- 14 FRANÇA, V. V.; CAPELLE, K. Entanglement of strongly interacting low-dimensional fermions in metallic, superfluid, and antiferromagnetic insulating systems. *Physical Review A*, American Physical Society, v. 74, n. 4, p. 42325, oct 2006. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.74.042325>>. 16
- 15 FRANÇA, V. V. Entanglement and exotic superfluidity in spin-imbalanced lattices. *Physica A: Statistical Mechanics and its Applications*, v. 475, p. 82–87, 2017. ISSN 0378-4371. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378437117301309>>. 16
- 16 De Picoli, T.; D'Amico, I.; FRANÇA, V. V. Metric-space approach for distinguishing quantum phase transitions in spin-imbalanced systems. p. 1–5. Disponível em: <<https://arxiv.org/pdf/1803.05216.pdf>>. 16, 23
- 17 LIMA, N. A. et al. Density Functionals Not Based on the Electron Gas: Local-Density Approximation for a Luttinger Liquid. *Physical Review Letters*, v. 90, n. 14, p. 146402+, 2003. ISSN 0031-9007. Disponível em: <<http://dx.doi.org/10.1103/physrevlett.90.146402>>. 16, 18, 23
- 18 FRANÇA, V. V.; VIEIRA, D.; CAPELLE, K. Simple parameterization for the ground-state energy of the infinite Hubbard chain incorporating Mott physics, spin-dependent phenomena and spatial inhomogeneity. *New Journal of Physics*, v. 14, 2012. ISSN 13672630. 16, 23
- 19 BISHOP, C. M. *Neural networks for pattern recognition*. [S.l.]: Clarendon Press, 1995. 482 p. ISBN 0198538642. 16
- 20 HAYKIN, S. *Neural Networks and Learning Machines*. 2008. 906 p. 16
- 21 NAFEY, A. S. Neural network based correlation for critical heat flux in steam-water flows in pipes. *International Journal of Thermal Sciences*, Elsevier Masson SAS, v. 48, n. 12, p. 2264–2270, 2009. ISSN 12900729. Disponível em: <<http://dx.doi.org/10.1016/j.ijthermalsci.2009.04.010>>. 16
- 22 MEIRELES, M. R. G.; ALMEIDA, P. E. M.; SIMÕES, M. G. A comprehensive review for industrial applicability of artificial neural networks. *IEEE Transactions on Industrial Electronics*, v. 50, n. 3, p. 585–601, 2003. ISSN 02780046. 16
- 23 BALABIN, R. M.; LOMAKINA, E. I. Neural network approach to quantum-chemistry data: Accurate prediction of density functional theory energies. *Journal of Chemical Physics*, v. 131, n. 7, 2009. ISSN 00219606. 16
- 24 POZUN, Z. D. et al. Optimizing transition states via kernel-based machine learning. *The Journal of Chemical Physics*, American Institute of Physics, v. 136, n. 17, p. 174101, may 2012. ISSN 0021-9606. Disponível em: <<http://aip.scitation.org/doi/10.1063/1.4707167>>. 16

- 25 MCGIBBON, R. T.; PANDE, V. S. Learning Kinetic Distance Metrics for Markov State Models of Protein Conformational Dynamics. *Journal of Chemical Theory and Computation*, American Chemical Society, v. 9, n. 7, p. 2900–2906, jul 2013. ISSN 1549-9618. Disponível em: <<https://doi.org/10.1021/ct400132h>>. 16
- 26 FILLETTI, É. R.; MARIA, J.; FERREIRA, V. G. Predicting of the Fibrous Filters Efficiency for the Removal Particles from Gas Stream by Artificial Neural Network. n. July, p. 317–327, 2015. 16
- 27 FILLETTI, É. R.; ROQUE, W. L. Estimating the mechanical competence parameter of the trabecular bone : a neural network approach. v. 32, n. 2, p. 137–143, 2016. 16
- 28 ORTEGA-ZAMORANO, F. et al. FPGA Implementation of Neurocomputational Models: Comparison Between Standard Back-Propagation and C-Mantec Constructive Algorithm. *Neural Processing Letters*, Springer US, v. 46, n. 3, p. 1–16, 2017. ISSN 1573773X. 16
- 29 JANET, J. P.; KULIK, H. J. Predicting electronic structure properties of transition metal complexes with neural networks. *Chem. Sci.*, Royal Society of Chemistry, v. 8, n. 7, p. 5137–5152, 2017. ISSN 2041-6520. Disponível em: <<http://xlink.rsc.org/?DOI=C7SC01247K>>. 16
- 30 BROCKHERDE, F. et al. Bypassing the Kohn-Sham equations with machine learning. *Nature Communications*, v. 8, n. 1, p. 872, 2017. ISSN 2041-1723. Disponível em: <<https://doi.org/10.1038/s41467-017-00839-3>>. 16
- 31 RAMIL, A. et al. A computer vision system for identification of granite-forming minerals based on RGB data and artificial neural networks. *Measurement*, v. 117, p. 90–95, 2018. ISSN 0263-2241. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0263224117307777>>. 16
- 32 WALCZAK, S.; VELANOVICH, V. Improving prognosis and reducing decision regret for pancreatic cancer treatment using artificial neural networks. *Decision Support Systems*, v. 106, p. 110–118, 2018. ISSN 0167-9236. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167923617302361>>. 16
- 33 VIEIRA, D. et al. Friedel oscillations in one-dimensional metals: From Luttinger’s theorem to the Luttinger liquid. *Journal of Magnetism and Magnetic Materials*, v. 320, n. 14, p. e418–e420, 2008. ISSN 0304-8853. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0304885308001996>>. 23, 38
- 34 LIEB, E. H.; WU, F. Y. Absence of Mott Transition in an Exact Solution of the Short-Range, One-Band Model in One Dimension. *Physical Review Letters*, American Physical Society, v. 20, n. 25, p. 1445–1448, jun 1968. ISSN 0031-9007. 23
- 35 OGATA, M.; SHIBA, H. Bethe-ansatz wave function, momentum distribution, and spin correlation in the one-dimensional strongly correlated Hubbard model. *Physical Review B*, American Physical Society, v. 41, n. 4, p. 2326–2338, feb 1990. ISSN 0163-1829. Disponível em: <<http://link.aps.org/doi/10.1103/PhysRevB.41.2326>>. 23
- 36 MALVEZZI, A. L. *An introduction to numerical methods in low-dimensional quantum systems* . [S.l.]: scielo, 2003. 55–72 p. 23



- 37 COE, J. P.; D'AMICO, I.; FRANÇA, V. V. Uniqueness of density-to-potential mapping for fermionic lattice systems. *EPL (Europhysics Letters)*, v. 110, n. 6, p. 63001, 2015. Disponível em: <<http://stacks.iop.org/0295-5075/110/i=6/a=63001>>. 23
- 38 HAGAN, M. T.; MENHAJ, M. B. Training Feedforward Networks with the Marquardt Algorithm. *IEEE Transactions on Neural Networks*, v. 5, n. 6, p. 989–993, 1994. ISSN 19410093. 30
- 39 BURKE, J. V.; FERRIS, M. C. A Gauss—Newton method for convex composite optimization. *Mathematical Programming*, v. 71, n. 2, p. 179–194, 1995. ISSN 1436-4646. Disponível em: <<https://doi.org/10.1007/BF01585997>>. 30
- 40 CUSTÓDIO, C. A.; FILLETTI, É. R.; FRANÇA, V. V. Artificial neural networks for density-functional optimizations in fermionic systems. *Scientific Reports*, Nature Publishing Group, v. 9, n. 1, p. 1886, dec 2019. ISSN 20452322. Disponível em: <<http://www.nature.com/articles/s41598-018-37999-1>>. 44