

UNIVERSIDADE FEDERAL DE ALFENAS

GABRIEL DA SILVA MACEDO

**REDES NEURAS ARTIFICIAIS APLICADAS NA OTIMIZAÇÃO DE
PROPRIEDADES DE SISTEMAS SEMICONDUTORES CONTENDO POÇOS
QUÂNTICOS**

Alfenas/MG

2020

GABRIEL DA SILVA MACEDO

REDES NEURAIS ARTIFICIAIS APLICADAS NA OTIMIZAÇÃO DE PROPRIEDADES
DE SISTEMAS SEMICONDUTORES CONTENDO POÇOS QUÂNTICOS

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Física pela Universidade Federal de Alfenas. Área de concentração: Matéria Condensada.

Orientador: Prof. Dr. Anibal Thiago Bezerra

Alfenas/MG

2020

Dados Internacionais de Catalogação-na-Publicação (CIP)
Biblioteca da Unidade Educacional Santa Clara da Universidade Federal de
Alfenas

Macedo, Gabriel da Silva.

Redes neurais artificiais aplicadas na otimização de propriedade de sistemas semicondutores contendo poços quânticos / Gabriel da Silva Macedo. -- Alfenas, MG, 2020.

86 f.

Orientador: Anibal Thiago Bezerra.

Dissertação (Mestrado em Física) - Universidade Federal de Alfenas, 2020.

Bibliografia.

1. Redes Neurais (Computação). 2. Semicondutores. 3. Ciência dos Materiais. 4. Física do Estado Sólido. I. Bezerra, Anibal Thiago. II. Título.

CDD-530.41

REDES NEURAIS ARTIFICIAIS APLICADAS NA OTIMIZAÇÃO DE PROPRIEDADES DE SISTEMAS SEMICONDUTORES CONTENDO POÇOS QUÂNTICOS

A Banca examinadora abaixo-assinada aprova a Dissertação apresentada como parte dos requisitos para a obtenção do título de Mestre em Física pela Universidade Federal de Alfenas. Área de concentração: Física de Matéria Condensada

Aprovada em: 26 de novembro de 2020.

Prof. Dr. Anibal Thiago Bezerra
Instituição: Universidade Federal de Alfenas

Prof. Dr. Samuel Bueno Soltau
Instituição: Universidade Federal de Alfenas

Prof. Dr. Marcelo Gonçalves Vivas
Instituição: Universidade Federal de Alfenas



Documento assinado eletronicamente por **Anibal Thiago Bezerra, Professor do Magistério Superior**, em 27/11/2020, às 10:42, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Samuel Bueno Soltau, Professor do Magistério Superior**, em 27/11/2020, às 16:10, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Marcelo Gonçalves Vivas, Professor do Magistério Superior**, em 30/11/2020, às 13:57, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://sei.unifal-mg.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0426210** e o código CRC **95BD61AD**.

A todos que me ajudaram de alguma maneira a chegar ao fim desta jornada, em especial minha família.

AGRADECIMENTOS

Em primeiro agradeço a minha mãe Maria Aparecida da Silva Macedo e a meu pai Sebastião Valias Macedo por todo carinho, incentivo e puxões de orelha que tornaram quem eu sou e acreditarem em mim sempre. Muito obrigado. Ao meu irmão Evandro Custódio da Silva Macedo e a minha namorada Rayssa Aparecida de Souza por todo apoio.

Ao amigo Prof. Dr. Anibal Thiago Bezerra pela coragem e loucura de encarar esse desafio de me orientar, a Universidade Federal de Alfenas e a PIB-PÓS pelo incentivo financeiro durante os primeiros meses de minha caminhada. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. E a todos meu amigos, colegas e professores que me ajudaram a caminhar até aqui.

*It's a new dawn
It's a new day
It's a new life
For me
And I'm feeling good*

Feeling Good - Nina Simone

RESUMO

Com a descoberta dos transistores, em 1948, os dispositivos semicondutores começaram a ter uma importância incomensurável no avanço tecnológico, possibilitando a construção de diversos outros dispositivos. Inicialmente os estudos com semicondutores se baseavam apenas em um material, entretanto ao unir dois ou mais semicondutores geram-se estruturas que apresentam um comportamento eletrônico distinto de quando os materiais estão separados, estas estruturas são chamadas de heteroestruturas. Tais heteroestruturas são crescidas por meio de técnicas altamente precisas e possibilitam o confinamento dos portadores de carga em uma, duas e três dimensões. Contudo o estudo dessas estruturas, devido ao grande número de portadores, geralmente é realizado por meio de várias aproximações para que a equação de Schrödinger tenha solução com os recursos computacionais atuais. Paralelamente, outra área que está crescendo atualmente é a inteligência artificial que tem como um dos métodos as redes neurais artificiais. Nesse sentido, esta dissertação tem como objetivo utilizar as redes neurais artificiais no estudo de heteroestruturas semicondutoras que possuem poços quânticos. Utilizando a aproximação de massa efetiva, a relação de dispersão para elétron nessas heteroestruturas é parabólica no número de onda, igual a de um elétron livre, contudo essa não é a relação de dispersão mais completa. Levando em consideração fatores como o acoplamento entre bandas, a melhor relação de dispersão deve levar em consideração termos de ordem superior no número de onda, dando origem a bandas de energia não-parabólica. Podemos obter a relação de dispersão mais completa por meio de um modelo semi-empírico, onde a massa efetiva do elétron passa a depender da energia da banda. Nesse modelo, desacoplando as bandas de valência e condução, podemos usar um método iterativo para tratar as bandas não-parabólicas como desacopladas e obter as autoenergias da heteroestrutura. Utilizando este modelo foi construído o *dataset* utilizado para treinar as redes neurais artificiais e verificar que com um conjunto de dados limitado as redes neurais são capazes de encontrar padrões e retornar as energias não parabólicas para diversas heteroestruturas otimizando o processo de obtenção dessas energias.

Palavras-Chave: Heteroestruturas, Redes Neurais Artificiais, Efeitos de bandas não-parabólicas

ABSTRACT

With the discovery of transistors in 1948, semiconductor devices began to have immeasurable importance in technological advancement, enabling the construction of several other devices. Initially, studies with semiconductors were based only on one material, however, when joining two or more semiconductors, structures are generated that presents an electronic behavior different from when the materials are separated, these structures are called heterostructures. Such heterostructures are cultivated with high precision techniques and allow the confinement of carriers in one, two, and three dimensions. However, the study of these structures, due to a large number of carriers, is usually carried out using different approaches so that a Schrödinger equation has a solution with the current computational resources. At the same time, another area that is currently growing is artificial intelligence, which uses artificial neural networks as one of its methods. In this sense, this dissertation aims to use artificial neural networks in the study of semiconductor heterostructures that have quantum wells. Using the effective mass approximation, the dispersion to electron ratio in these heterostructures is parabolic in the wavenumber, equal to a free electron, however, this is not the most complete dispersion relationship. Taking into account factors such as coupling between bands, the best dispersion ratio should take into account the higher-order terms in the wavenumber, giving rise to bands of non-parabolic energy. We can obtain the most complete dispersion relation using a semi-empirical model, where an effective mass of the electron becomes dependent on the energy of the band. In this model, decoupling the valence and conduction bands, we can use an iterative method to treat non-parabolic bands as decoupled and obtain the energies of the heterostructured carriers. Using this model, a data set was used to train artificial neural networks and verify that with a limited data set as neural networks they are able to find patterns and return as non-parabolic energies to different heterostructures optimizing the process of obtaining these energies.

Keywords: Heterostructures, Artificial Neural Networks, Band Nonparabolicity Effects

LISTA DE FIGURAS

Figura 1 –	Modelo de Kronig–Penney para energia potencial periódica de um elétron em um cristal onde a periodicidade é representada por múltiplos poços quânticos.	23
Figura 2 –	Solução transcendental para o modelo de Kronig–Penney. Gráfico referente a Eq. (18) mostrando as regiões que são solução desta equação, que representa as energia permitidas e proibidas.	26
Figura 3 –	Classificação dos sólidos baseada no perfil de bandas. (a) Isolantes possuem a banda de valência completamente cheia e uma largo <i>gap</i> . (b) Semicondutores a banda de valência completamente cheia e um <i>gap</i> de poucos eV. (c) condutores a banda de valência é parcialmente cheia (d) condutores onde ocorre a superposição de bandas.	28
Figura 4 –	Influência da temperatura na distribuição de Fermi. Linha contínua representa $T = 0$ K onde a função possui uma queda abrupta para E igual a E_f . Linha tracejada $T > 0$ onde ocorreu uma suavização da função em torno da energia de E_f	28
Figura 5 –	Relação de dispersão da energia para elétrons livres (linha contínua) e para uma rede linear com parâmetro de rede igual à a (linha tracejada) para a primeira zona de brillouin.	33
Figura 6 –	Exemplo esquemático de como ocorre o processo de crescimento de heteroestruturas semicondutoras utilizando a técnica de <i>molecular beam epitaxy</i>	35
Figura 7 –	Classificação das heteroestruturas baseado na descontinuidade da interface dos materiais. Heteroestruturas do tipo I possui a banda de condução e de valência de um material inferior e superior do que do outro, do tipo II tanto a banda de condução quanto de valência são inferiores de um material em relação ao outro e do tipo III ocorre a descontinuidade dos <i>gaps</i>	36
Figura 8 –	Exemplo de função periódica (linha tracejada) modulado por uma função envelope (linha contínua).	38

Figura 9 –	Fluxograma do método iterativo para o cálculo na massa dependente da energia e da energia via Eq. de Schrödinger proposto por Machado (2019).	44
Figura 10 –	Representação da anatomia de um neurônio biológico destacando sua principais partes.	46
Figura 11 –	Representação de um neurônio artificial onde as entradas são representadas por x_i , w_i são os pesos sinápticos, Θ é chamado de <i>bias</i> , $g(u)$ a função de ativação e y a saída do neurônio. . .	47
Figura 12 –	Exemplos de funções de ativação utilizadas em neurônios artificiais e suas respectivas derivadas.	48
Figura 13 –	Diagrama de uma rede neural <i>perceptron</i> multicamadas profunda com 3 camadas ocultas.	49
Figura 14 –	Exemplos de dados de entrada para uma rede neural artificial.	50
Figura 15 –	Processo de treinamento <i>backpropagation</i> onde a linha continua com fluxo da direita para esquerda representando a fase de <i>forward</i> onde os pesos são aleatoriamente calculados. Linha tracejada com fluxo da esquerda para direita representa a fase de <i>backward</i> onde ocorre a alteração dos pesos realizando a descida do gradiente.	52
Figura 16 –	Exemplos de curvas de aprendizagem de uma rede neural artificial onde ocorreu (a) <i>Underfitting</i> onde a rede não encontra uma generalização para os dados. (b) Boa aprendizagem onde a rede não encontra um padrão para os dados. (c) <i>Overfitting</i> quando a rede "decora" certo conjunto de dados mas não encontra uma generalização para estes.	59
Figura 17 –	Curvas de erro quadrático médio de rede neurais com uma, duas e três camadas seguindo o método de Heaton (2008). (a) Método a): $n_{inp} \leq n_{oc} \leq n_{out}$ (b) Método b): $n_{oc} = 2/3 (n_{inp} + n_{out})$ (c) Método c): $n < 2n_{in}$	65
Figura 18 –	Curva de <i>MSE</i> para a rede com a melhor arquitetura utilizando o otimizador NADAM. (a) Sem aplicação da técnica <i>Early Stopping</i> . (b) Com <i>Early Stopping</i>	67

Figura 19 –	Esquema da rede neural utilizando o mais amplo <i>dataset</i> . . .	68
Figura 20 –	Comparação dos 20% de dados de teste previstos pela rede neural com as respectivas saídas calculadas via matriz de diferenciação. (a) Todo o conjunto de teste. (b) Recorte dos dados centrais da Fig. (a).	69
Figura 21 –	Esquema da rede neural utilizado com a primeira variação do <i>dataset</i>	70
Figura 22 –	Comparação dos 20% de dados de teste previstos pela rede neural com as respectivas saídas para primeira redução do <i>dataset</i> (a) Todo o conjunto de teste. (b) Recorte dos dados centrais da Fig. (a).	71
Figura 23 –	Esquema de rede neural com <i>dataset</i> sem as energias parabólicas parâmetros obtidos via Schrödinger.	72
Figura 24 –	Comparação dos 20% de dados de teste previstos pela rede neural com as respectivas saídas para o <i>dataset</i> sem parâmetros obtidos via método iterativo. (a) Todo o conjunto de teste. (b) Recorte dos dados centrais da Fig. (a).	73
Figura 25 –	Esquema da rede neural utilizado o <i>dataset</i> com apenas parâmetros estruturais.	73
Figura 26 –	Comparação das energias não-parabólicas para a rede neural com o <i>dataset</i> com apenas parâmetros estruturais. (a) Todo conjunto de teste. (b) Recorte dos dados centrais da Fig (a). .	74
Figura 27 –	Comparação dos 20% de dados entre as previsões da rede neural com sua respectiva saída. (a) Energias não-parabólicas. (b) <i>Zoom</i> da Fig. (a). (c) Massa no poço. (d) Massa na barreira.	75
Figura 28 –	Comparação entre as previsões da rede neural com sua respectiva saída para o <i>dataset</i> sem parâmetros via Schrödinger. (a) Energias não-parabólicas. (b) <i>Zoom</i> da Fig. (a). (c) Massa no poço. (d) Massa na barreira.	76

Figura 29 –	Comparação entre as previsões da rede neural com sua respectiva saída para o <i>dataset</i> apenas com dados estruturais. (a) Energias não-parabólicas. (b) <i>Zoom</i> da Fig. (a). (c) Massa no poço. (d) Massa na barreira	77
Figura 30 –	Comparação entre as previsões da rede neural com sua respectiva saída para o <i>dataset</i> com dados estruturais e energias de poço infinito. (a) Energias não-parabólicas. (b) <i>Zoom</i> da Fig. (a). (c) Massa no poço. (d) Massa na barreira. (e) Energia parabólica. (f) <i>Zoom</i> da Fig. (e)	78
Figura 31 –	Comparação entre as previsões da rede neural com sua respectiva saída para o <i>dataset</i> com dados estruturais. (a) Energias não-parabólicas. (b) <i>Zoom</i> da Fig. (a). (c) Massa no poço. (d) Massa na barreira. (e) Energia parabólica. (f) <i>Zoom</i> da Fig. (e)	79

LISTA DE TABELAS

Tabela 1 – Comparação entre as energias parabólicas dadas por Nelson, Miller e Kleinman (1987) e usando as matrizes de Chebyshev para diversas larguras de poços quânticos.	62
Tabela 2 – Comparação entre as energias não-parabólicas dadas por Nelson, Miller e Kleinman (1987) e usando as matrizes de Chebyshev para diversas larguras de poços quânticos.	62
Tabela 3 – Heteroestruturas utilizadas para a construção <i>dataset</i> e seus respectivos <i>band offset</i>	63
Tabela 4 – Topologia das redes treinadas para os três modelos de Heaton (2008)	66
Tabela 5 – Comparação entre os valores de <i>MSE</i> e <i>MAE</i> de vários otimizadores.	66
Tabela 6 – Comparação entre as energias não-parabólicas calculadas via matriz de diferenciação e previstas pelas redes neurais artificiais para algumas larguras de poços quânticos.	68
Tabela 7 – Comparação entre as energias não-parabólicas calculadas via matriz de diferenciação e previstas pelas redes neurais artificiais com diminuição do <i>dataset</i>	70
Tabela 8 – Comparação entre as energias não-parabólicas calculadas via matriz de diferenciação e previstas pelas redes neurais artificiais com a segunda diminuição do <i>dataset</i>	72
Tabela 9 – Comparação entre as energias não-parabólicas usando conjunto de entrada com apenas dados estruturais.	74

LISTA DE ABREVIações E SIGLAS

IA	Inteligência Artificial
ML	<i>Machine Learning</i>
RNA	Redes Neurais Artificiais
HE	Heteroestrutura
DFT	<i>Density Functional Theory</i>
AFE	Aproximação de Função Envelope
DL	<i>Deep Learning</i>
MLP	<i>Multilayer Perceptron</i>
Mi	Massas Efetivas ¹
Eg _i	<i>Band gaps</i>
Ai	Parâmetros de Rede
Eis	Constante dielétrica estática
Eihf	Constante dielétrica para altas frequências
con	Concentração
Lqw	Largura do poço quântico
Eqw	<i>Band offset</i>
Ej_SNP	Níveis de energia parabólica ²
Ej_inf	Níveis de energia de um poço quântico infinito
E_CNP	Nível de energia não-parabólica
MSE	<i>Mean square error</i>
MAE	<i>Mean absolute error</i>
Mei	Massas Efetivas dependente da energia

¹O índice i pode ser w ou b representando poço e barreira.

²O índice j vai de 1 a 5 representando os níveis de energia.

SUMÁRIO

1	INTRODUÇÃO	17
2	ESTRUTURA DE BANDAS EM SÓLIDOS	21
2.1	MODELO DE KRONIG-PENNEY	23
3	DINÂMICA ELETRÔNICA EM UM CRISTAL MONOATÔMICO	30
3.1	APROXIMAÇÃO DE BORN-OPPENHEIMER	30
3.2	APROXIMAÇÕES DE HARTREE	31
3.3	APROXIMAÇÃO DE MASSA EFETIVA	32
4	DINÂMICA DE UM ELÉTRON EM UMA HETEROESTRUTURA	35
4.1	HETEROESTRUTURAS	35
4.2	APROXIMAÇÃO DE FUNÇÃO ENVELOPE	37
4.3	EQUAÇÃO DE SCHRÖDINGER COM MASSA EFETIVA DEPENDENTE DA POSIÇÃO	39
5	EFEITOS DE NÃO-PARABOLICIDADE	41
6	INTELIGÊNCIA ARTIFICIAL	45
6.1	REDES NEURAS ARTIFICIAIS	45
6.2	NEURÔNIO MATEMÁTICO	46
6.3	ARQUITETURA DAS REDES NEURAS ARTIFICIAIS	49
6.4	TREINAMENTO	50
6.5	<i>BACKPROPAGATION</i>	51
6.5.1	Ajuste dos pesos da camada de saída	52
6.5.2	Ajuste dos pesos das camadas intermediárias	54
7	METODOLOGIA	57
7.1	SOLUÇÃO MATRICIAL DA EQUAÇÃO DE SCHRÖDINGER	57
7.2	REDES NEURAS ARTIFICIAIS	58
8	RESULTADOS	61
8.1	CONSTRUÇÃO DO <i>DATASET</i>	62
8.2	APLICAÇÃO DA REDES NEURAS ARTIFICIAIS	64
8.3	PREVISÃO DAS ENERGIAS NÃO-PARABÓLICAS	67
8.3.1	<i>Dataset</i> com remoção das massas dependentes da energia	69
8.3.2	<i>Dataset</i> sem parâmetros obtidos via equação de Schrödinger	71
8.3.3	<i>Dataset</i> apenas com parâmetros estruturais	72

8.4	PREVISÃO DA ENERGIA NÃO-PARABÓLICA E DAS MASSAS DEPENDENTES DA ENERGIA	74
8.5	PREVISÃO DA ENERGIA NÃO-PARABÓLICA, ENERGIA PARABÓLICA DO ESTADO FUNDAMENTAL E DAS MASSAS DEPENDENTES DA ENERGIA	76
9	CONCLUSÃO	81
	REFERÊNCIAS	83

1 INTRODUÇÃO

A eletrônica da primeira metade do século XX era baseada em tubos de vácuo. Entretanto, em 1948 ocorreu uma grande mudança neste cenário devido aos transistores concebidos através de estruturas contendo semicondutores, inventados por William Shockley, John Bardeen e Walter Brattain. Deste ponto em diante a eletrônica passou a ser baseada nesses dispositivos que possuem características intermediárias aos metais e isolantes, ou seja, são dispositivos semicondutores (TIPLER; LLEWELLYN, 2010; SERWAY; JEWETT, 2013). A grande vantagem da utilização de semicondutores está em suas propriedades elétricas e óticas, como, por exemplo, condutividade, absorção ótica e dependência com temperatura que surgem devido a um arranjo bastante especial dos níveis de energia dos elétrons nesses materiais (EISBERG; RESNICK, 1979; CHIQUITO, 1999).

Os primeiros trabalhos com semicondutores foram realizados em sistemas de materiais únicos baseados em Silício, Germânio, Arseneto de Gálio etc. Entretanto, ao unir dois ou mais materiais a estrutura resultante apresenta propriedades eletrônicas e óticas distintas das apresentadas pelos materiais da liga separados em consequência da flexibilidade da estrutura de bandas. Tais estruturas são denominadas heteroestruturas e são usadas principalmente para confinar os portadores de carga, produzindo sistemas eletrônicos de baixa dimensionalidade³ como poços, fios e pontos quânticos (SINGH, 2003; CHIQUITO; LANCIOTTI, 1998).

A vantagem da utilização de heteroestruturas semicondutoras, além dos pontos mencionados acima, está no fato de que o crescimento destas estruturas pode ser precisamente controlado. Tal crescimento se faz por meio de técnicas que são altamente sofisticadas, como por deposição de camadas moleculares. Exemplos destas técnicas são Epitaxia de Feixe Molecular e Deposição de Vapor Químico Metal-Orgânico, as quais permitem que os cristais sejam depositados em um substrato, uma monocamada atômica por vez, com grande precisão. Essas técnicas tornaram possível a sintetização de estruturas cristalinas artificiais conhecidas como superredes e poços quânticos (LAMAS, 2004).

Assim, diversos dispositivos são constituídos de heteroestruturas semicondutoras como por exemplo, *lasers* de diodo, fotodetectores, células solares, entre outros. Entre-

³Com confinamento quântico dos portadores de carga na estrutura.

tanto para se estudar as propriedades físicas destas estruturas é necessário resolver a equação de Schrödinger que, devido ao grande número de constituintes do cristal, se torna extremamente complexa de modo que algumas aproximações são necessárias. A vasta quantidade de materiais, propriedades estruturais, composições das heteroestruturas leva a uma enorme variedade de possíveis dispositivos, com as mais diversas propriedades, tornando-se necessário o uso de ferramentas computacionais para prever tais propriedades.

Nessa perspectiva, uma área que vem ganhando enorme destaque nos dias atuais e pode ser utilizada como ferramenta computacional para a física de dispositivos semicondutores é a inteligência artificial. Uma área da ciência da computação voltada para o desenvolvimento de sistemas de computadores, que exibem as características que se assemelham à inteligência no comportamento humano, em aspectos como percepção, processamento de linguagem natural, resolução e planejamento de problemas, aprendizado e adaptação. (BARR; FEIGENBAUM, 1981).

A inteligência artificial consiste de diversos métodos para que os computadores consigam interpretar dados e obter conclusões, chamados de aprendizagem de máquina (ALPAYDIN, 2004). Alguns desses métodos são: redes neurais artificiais, algoritmos genéticos, lógica *fuzzy*, solução e planejamento de problemas, raciocínio não monotônico, programação lógica, processamento de linguagem natural e visão computacional (MELLIT; KALOGIROU, 2008).

Tais métodos já são utilizados na engenharia de materiais. Silva et al. (2013) faz uso das redes neurais artificiais para a síntese de semicondutores quaternários *InAlGaAs*, onde a inteligência artificial auxilia na escolha das concentrações dos elementos Índio, Gálio e Alumínio os quais são necessários para a construção do material. Tal processo é feito a partir de estudos anteriores ou através de tentativa e erro, assim as redes neurais, atuando na previsão destes parâmetros reduzem o tempo gasto em experimentação (SILVA et al., 2013).

Lima (2006), no estudo do processo de deposição de filmes finos poliméricos, usa as redes neurais para a obtenção das características finais dos filmes, uma vez que estes estão relacionados com os parâmetros de deposição, cuja relação não é simples, necessitando de um grande número de dados e experimentos. Com a utilização das redes foi possível criar sistemas que simulam e reproduzem os processos de deposição

de filmes finos (LIMA, 2006).

A análise de *lasers* de pontos quânticos, por exemplo, pode ser realizada por dois métodos, a descrição de microestados (GRUNDMANN, HEITZ, et al., 1997 apud SANTOS 2018) e o modelo baseado em equações de taxa (SUGAWARA, HATORI, et al., 2005 apud SANTOS 2018), os quais possuem um elevado custo computacional. Afim de contornar esse problema, Santos (2018) mostra que a simulação desses dispositivos pode ser realizada utilizando inteligência artificial que, comparado aos métodos usuais, possui um custo computacional muito menor. Nesse estudo, a potência e curvas de potência pela corrente, para diversos valores de corrente e temperatura, foi prevista (SANTOS, 2018).

Assim é possível notar que a inteligência artificial, em especial as redes neurais artificiais, está sendo aplicada na síntese de dispositivos e heteroestruturas semicondutoras como mostra Silva et al. (2013) e Lima (2006) e também estão começando a ser utilizadas no estudo dessas estruturas como propõe Santos (2018). Tal utilização tem se mostrado muito promissora e com resultados bastante consistentes.

Deste modo, o presente trabalho tem como objetivo a utilização das redes neurais artificiais (RNA) no estudo de heteroestruturas semicondutoras afim de obter a energia dos estados do poço quântico. Tais estados surgem na banda de condução devido ao confinamento quântico dos portadores de carga, dado o crescimento da heteroestrutura. Adicionalmente, buscou-se analisar a alteração na energia do estado fundamental ao se levar em conta os efeitos de não-parabolicidade nesta banda.

Todavia, obter a energia destes estados não é trivial. Como já mencionado, se faz necessário a solução da equação de Schrödinger para o Hamiltoniano que descreve a heteroestrutura (HE). Uma vez que a HE é composta por materiais diferentes o comportamento do elétrons em cada material é distinto. Em cada região de poço e dá barreira existe uma massa efetiva específica, sendo necessário uma renormalização do Hamiltoniano. Além deste fato devemos lembrar que em um cristal a quantidade de partículas é da ordem de 10^{23} átomos/cm³, tornando o problema extremamente complicado (YU; CARDONA, 2010), sendo necessário o uso de aproximações, tais quais as de Born-Oppenheimer, Hartree e de massa efetiva.

Com todo o aparato acima descrito, para as heteroestruturas semicondutoras de AlGaAs/GaAs, foi solucionada a equação de Schrödinger e determinados os níveis de

energia. Para encontrar a alteração dos níveis de energia, ao se levar em consideração os efeitos de não-parabolicidade, foi utilizado o modelo empírico proposto por Nelson, Miller e Kleinman (1987) e com este simular o comportamento dos elétrons em um HE semicondutora (NELSON; MILLER; KLEINMAN, 1987).

De posse destes resultados foi possível construir um conjunto de dados de tamanho suficiente para a construção de uma RNA de múltiplas camadas capaz de generalizar estes dados e prever quais seriam as energias para o sistema. Após o sucesso desta fase, onde a RNA conseguiu fazer excelentes previsões, foram inseridas ao conjunto de dados⁴ outras HE semicondutoras, afim de uma ampla generalização da rede.

A dissertação foi organizada da seguinte maneira. No capítulo 2, foi realizada a discussão a respeito da união de inúmeros átomos para o crescimento de um cristal, o que leva a formação de bandas de energia, tópico que define o que são semicondutores. No capítulo 3, foi discutido o comportamento eletrônico em um cristal singular e as simplificações aplicadas.

O comportamento dos elétrons em heteroestruturas foi descrito no capítulo 4, juntamente com a aproximação de função envelope e a descrição da massa dependente da posição. Efeitos de bandas não-parabólicas foram tratados no capítulo 5.

Nos capítulos 6 e 7 foi apresentado o que são redes neurais artificiais, o processo de treinamento e cuidados necessários na sua construção. Os resultados encontrados foram apresentados no capítulo 8 e, por fim, foi realizada uma discussão a respeito dos resultados obtidos.

⁴Também chamado de *dataset*

2 ESTRUTURA DE BANDAS EM SÓLIDOS

Um material sólido pode ser visto como a união de vários átomos em um arranjo cristalino periódico (KITTEL, 2006). A organização desses átomos gera uma energia potencial⁵ (V) que é periódica e que pode ser escrita em uma dimensão como,

$$V(z) = V(z + Na), \quad (1)$$

onde N é a quantidade de átomos e a a distância de separação entre eles⁶, de modo que o termo Na represente uma translação no sólido. A equação de Schrödinger independente do tempo para esse potencial é

$$\hat{H}\Psi = \frac{-\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial z^2} + V(z)\Psi = E\Psi, \quad (2)$$

onde \hat{H} é o Hamiltoniano que descreve o sistema, dado pela soma do operador energia cinética $\left(\frac{-\hbar^2}{2m} \frac{\partial^2}{\partial z^2}\right)$ e o operador energia potencial ($V(z)$). A constante de Planck reduzida, a massa do elétron e a energia total do sistema são representados respectivamente por \hbar , m e E .

A solução da Eq. (2) para o potencial descrito na Eq. (1) é dada pelo teorema de Bloch, de modo que os autoestados de são dados pelas funções de Bloch (GRIFFITHS, 2011; GROSSO; PARRAVICINI, 2013). Tais funções consistem de uma onda plana com vetor de onda k modulada por uma função que possui a periodicidade da rede cristalina,

$$\Psi_k(z) = e^{ikz} u_k(z), \quad (3)$$

onde

$$u_k(z + Na) = u_k(z). \quad (4)$$

A Eq. (3), mediante a uma translação, pode ser escrita como

$$\Psi_k(z + Na) = e^{ik(z+Na)} u_k(z + Na) \quad (5)$$

⁵Daqui em diante energia potencial será simplificado para potencial.

⁶Chamado de parâmetro de rede.

de modo análogo

$$\Psi_k(z + Na) = e^{ikz} e^{ikNa} u_k(z + Na), \quad (6)$$

e usando as Eq. (3) e (4), alcançamos uma relação que mostra que os valores da autofunção, $\Psi_k(z)$, avaliada em quaisquer dois pontos do espaço real separados por uma translação, Na , estão relacionados por uma fase e^{ikNa} . Assim

$$\Psi_k(z + Na) = e^{ikNa} \Psi_k(z), \quad (7)$$

a função de onda no cristal apresenta a mesma periodicidade da rede cristalina, a menos de uma fase.

Um sólido é composto por um grande número de átomos, $N \approx 10^{23}$ átomos/cm³. Por maior que seja N este ainda é finito. Assim, as extremidades do material interferem na periodicidade do sistema e precisam, em princípio, ser levadas em consideração. Entretanto, para contornar esse problema podem ser aplicadas as condições de Born-von Karman que são condições de contorno cíclica dada por (GROSSO; PARRAVICINI, 2013; ASHCROFT; MERMIN; OLIVEIRA, 2010).

$$\Psi_k(z + Na) = \Psi_k(z). \quad (8)$$

Usando a condições de Born-von Karman e a Eq. (7) obtemos

$$e^{ikNa} \Psi_k(z) = \Psi_k(z), \quad (9)$$

que implica em $e^{ikNa} = 1$. Usando a fórmula de Euler, alcançamos os valores permitidos de k os quais são dados por,

$$k_n = \frac{2\pi}{Na} n, \quad \forall n \in Z. \quad (10)$$

A Eq. (10) mostra que o vetor de onda é periódico e toda a informação sobre o sistema físico está contida em uma região específica no espaço recíproco, determinada pelo parâmetro de rede a . Essa região recebe o nome de primeira zona de Brillouin.

Para encontrar as funções de Bloch que são solução da equação de Schrödinger para um determinado vetor de onda, que obedecem a Eq. (10), tem-se que resolver

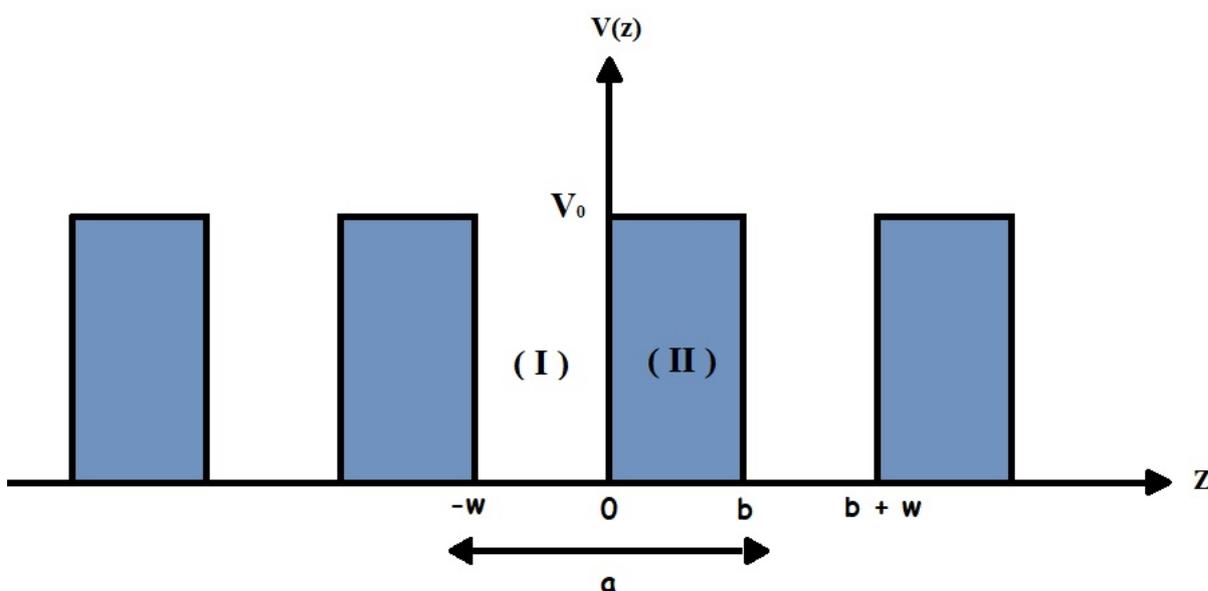
um problema de autovalor para determinar u_k , o que leva uma família de soluções (ASHCROFT; MERMIN; OLIVEIRA, 2010).

Todo o problema até aqui descrito foi feito usando uma energia potencial unidimensional, entretanto o problema real é tridimensional, mas as condições demonstradas acima são válidas para o caso tridimensional onde z se torne \mathbf{R} . O fato de a equação de Schrödinger ter soluções que dependem do vetor de onda k leva a crer que a energia também terá uma dependência, de modo que ela será dada por uma relação de dispersão $E_n(k)$, dando origem a bandas de energia. Tal propriedade é demonstrada de maneira qualitativa usando o modelo de Kronig-Penney.

2.1 MODELO DE KRONIG-PENNEY

A energia potencial sentida pelos elétrons em uma rede cristalina, devido aos íons presentes nela é a energia potencial coulombiana, entretanto a característica mais importante deste potencial não é o seu formato, mas sim o fato de ser periódico (TIPLER; LLEWELLYN, 2010). Assim é possível fazer a aproximação desta energia potencial por um perfil com vários poços quânticos (Fig. 1), essa aproximação é denominada modelo de Kronig-Penney.

Figura 1 – Modelo de Kronig-Penney para energia potencial periódica de um elétron em um cristal onde a periodicidade é representada por múltiplos poços quânticos.



Fonte: Do autor.

A solução da Eq. (2) para o potencial da Fig. 1 na região I (poço) é dada por

(GROSSO; PARRAVICINI, 2013),

$$\Psi_I(z) = Ae^{iqz} + Be^{-iqz}, \quad q(E) = \frac{\sqrt{2mE}}{\hbar^2} \quad (11)$$

e na região II (barreira) é,

$$\Psi_{II}(z) = Ce^{\beta z} + De^{-\beta z} \quad \beta(E) = \frac{\sqrt{2m(V_0 - E)}}{\hbar^2}, \quad (12)$$

sendo A, B, C e D coeficientes arbitrários. Estes coeficientes devem ser escolhidos de maneira a satisfazer as quatro condições de contorno a seguir,

$$\Psi_I(0) = \Psi_{II}(0); \quad (13)$$

$$\left(\frac{d\Psi_I}{dz} \right)_{z=0} = \left(\frac{d\Psi_{II}}{dz} \right)_{z=0}; \quad (14)$$

$$\Psi_{II}(b) = e^{ika} \Psi_I(-w); \quad (15)$$

$$\left(\frac{d\Psi_{II}}{dz} \right)_{z=b} = e^{ika} \left(\frac{d\Psi_I}{dz} \right)_{z=-w}. \quad (16)$$

As condições dadas nas Eq. (13) e (14) impõem a continuidade da função de onda e sua derivada em $x = 0$. Já as Eq. (15) e (16) conectam a função de onda e sua derivada em $x = -w$ e $x = b$ através do fator de fase, conforme exigido pela periodicidade no teorema de Bloch.

Aplicando as Eq. (11) e (12) e suas derivadas nas Eq. (13)-(16), temos o seguinte sistema de equações:

$$\begin{cases} A + B = C + D \\ Aiq - Biq = C\beta - D\beta \\ Ce^{\beta b} + De^{-\beta b} = e^{ika}[Ae^{-iqw} + Be^{+iqw}] \\ C\beta e^{\beta b} - D\beta e^{-\beta b} = e^{ika}[Aiqe^{-iqw} + Biqe^{+iqw}] \end{cases}$$

As quatro equações têm solução não trivial apenas se o determinante dos coeficientes

que multiplicam A, B, C, D desaparecer:

$$\begin{vmatrix} 1 & 1 & -1 & -1 \\ iq & -iq & -\beta & \beta \\ -e^{ika-iqw} & -e^{ika+iqw} & e^{\beta b} & e^{-\beta b} \\ -iqe^{ika-iqw} & iqe^{ika+iqw} & \beta e^{\beta b} & \beta e^{-\beta b} \end{vmatrix} = 0.$$

Calculando o determinante obtemos,

$$\frac{\beta^2 - q^2}{2q\beta} \sinh \beta b \sin qw + \cosh \beta b \cos qw = \cos ka. \quad (17)$$

A Eq. (17) é transcendental e pode ser solucionada graficamente. Entretanto, por critérios de simplificação, impõe-se que $b \rightarrow 0$ e $V_0 \rightarrow \infty$, de modo que o produto $b \times V_0$ se mantenha constante. O potencial, nesse caso, assume a forma periódica de funções δ . Aplicando esses limites para as relações de q e β nas Eqs. (11) e (12), temos que $\beta \gg q$ e $\beta b \ll 1$. Assim a Eq. (17) se reduz a

$$F(x) = \frac{P}{x} \sin x + \cos x = \cos ka, \quad (18)$$

onde

$$P = \frac{mV_0ba}{\hbar^2} \quad e \quad x = qa. \quad (19)$$

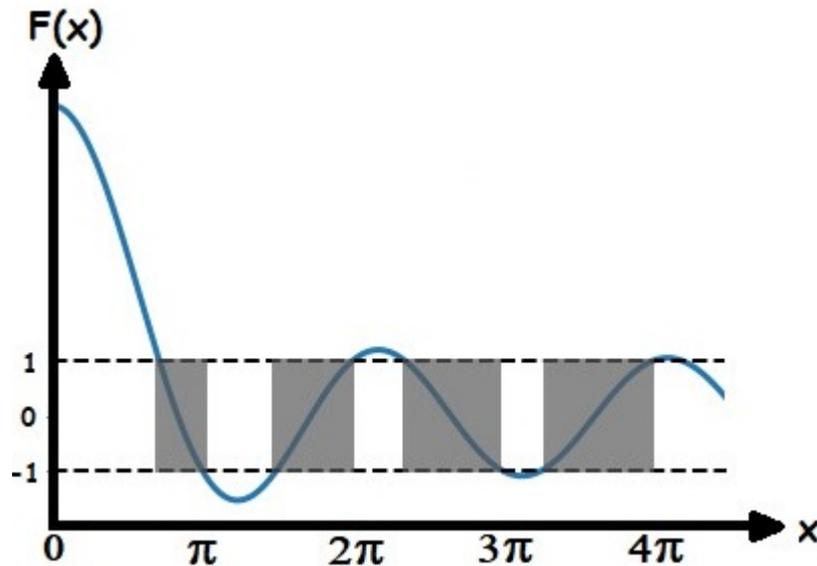
Dados os limites máximo e mínimo que as funções senoidais podem assumir, para valores fora do intervalos de 1 a -1 não é possível resolver a Eq. (18). Assim para valores de $F(x)$ que estão nesse intervalo encontra-se as bandas permitidas de energia, já as regiões fora deste intervalo representam energias proibidas, também chamadas de *gaps* de energia, como mostra a Fig. 2.

Do ponto de vista físico as bandas proibidas ocorrem devido a reflexões de Bragg das ondas progressivas que descrevem um elétron se propagando pela rede cristalina (EISBERG; RESNICK, 1979). A condição de Bragg de uma onda com vetor de onda k em uma dimensão se torna (KITTEL, 2006).

$$k = \pm \frac{n\pi}{a} \quad (20)$$

Uma onda progressiva que se desloca para o sentido positivo, ao incidir sobre as

Figura 2 – Solução transcendental para o modelo de Kroning-Penney. Gráfico referente a Eq. (18) mostrando as regiões que são solução desta equação, que representa as energia permitidas e proibidas.



Fonte: Do autor

barreiras, regiões onde se encontram os íons, serão parcialmente refletidas se movendo no sentido negativo. Tais ondas só estarão em fase se a condição de Bragg, Eq. (20) for satisfeita. Como a condição de Bragg fornece dois valores para um determinado vetor de onda, a função de onda total será obtida pela soma ou pela diferença destas. A função dada pela soma é,

$$\psi(+) = e^{i(\frac{\pi}{a})z} + e^{-i(\frac{\pi}{a})z} \quad (21)$$

e da diferença

$$\psi(-) = e^{i(\frac{\pi}{a})z} - e^{-i(\frac{\pi}{a})z} \quad (22)$$

A banda proibida se dá pois a densidade de probabilidade da Eq. (21) e (22) não são iguais. A densidade de probabilidade de $\psi(+)$ é proporcional ao $\cos^2(\pi z/a)$ e da função $\psi(-)$ à $\sin^2(\pi z/a)$ de modo que o maior valor de uma esteja no mesmo ponto em que o menor da outra, assim ao calcular o valor esperado para a energia utilizando essas funções, obtêm-se dois valores distintos de modo que o valor que está entre eles é definido como banda proibida ou *gap* de energia (EISBERG; RESNICK, 1979; KITTEL, 2006).

As regiões que são soluções para Eq. (18) são as bandas permitidas. Pela Eq. (10) temos que $ka = 2\pi n/N$, para cada banda existem N estados de energia, estes estão tão próximos (uma vez que N é grande) que podemos considerar a banda como sendo contínua (GRIFFITHS, 2011). A relação entre a energia e os valores possíveis para o número de onda determina a dispersão eletrônica para cada banda $E_n(k)$, ou seja, os estados que podem ser ocupados por elétrons no sistema.

A ocupação desses estados será governada pela distribuição de Fermi-Dirac (Eq. 23), uma vez que elétrons são férmions, partículas com *spin* semi-inteiro. Para $T = 0$ K os elétrons ocupam os estados de menor energia permitidos no cristal até um determinado nível, denominado nível de Fermi (E_f).

$$f(E) = \frac{1}{e^{\left(\frac{E-E_f}{kT}\right)} + 1}. \quad (23)$$

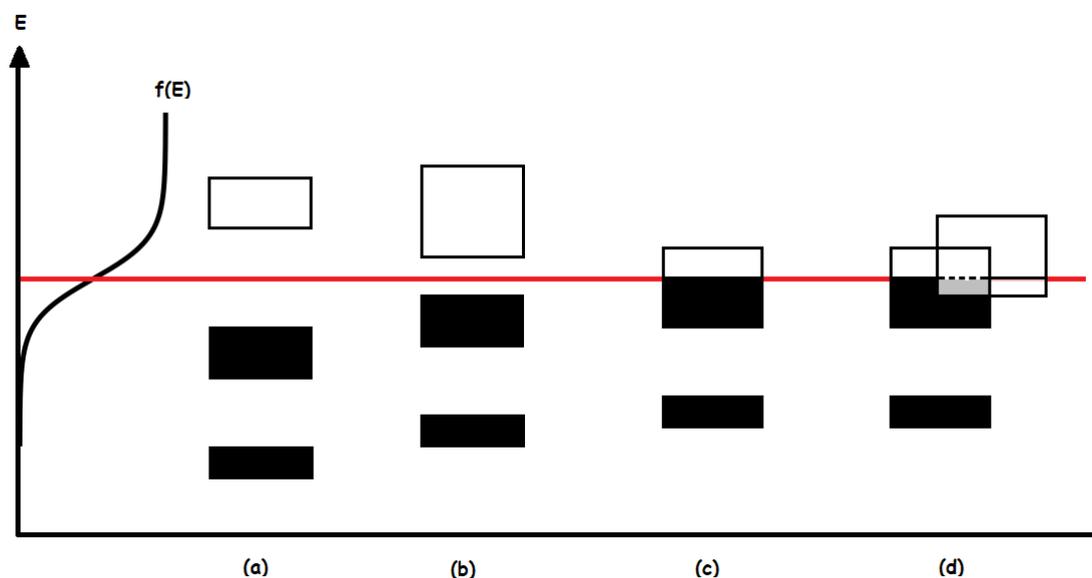
Assim, a ocupação das bandas de energia define a classificação do cristal. Analisamos, em geral, a última banda preenchida e a subsequente, chamadas de banda de valência e condução respectivamente. Cristais que possuem a $T = 0$ K a banda de valência completamente cheia são denominados isolantes (Fig. 3 (a)). Já os materiais que possuem a banda de valência parcialmente preenchida, ou que a banda de valência e de condução estejam superpostas⁷ (o mínimo da banda de condução possui uma energia menor do que o máximo da banda de valência), denominados condutores (Fig. 3 (c) e (d)).

Existe ainda uma classe de materiais que a $T = 0$ K são isolantes porém, devido ao pequeno *gap*, à temperatura ambiente, níveis superiores de energia podem ser ocupados. Tal fato acontece pois em $T = 0$ a distribuição de Fermi se assemelha à função degrau. Possui uma queda abrupta, mas ao aumentar a temperatura esta passa a ser suavizada em torno da E_f , fazendo com que estados antes vazios possam ser ocupados, como mostra a Fig. 4. Os materiais que apresentam esse comportamento são chamados de semicondutores (Fig. 3 (b)).

Estes materiais, semicondutores, são altamente interessantes devido às suas propriedades como por exemplo, sua condutividade. Uma vez que o *gap* de energia é

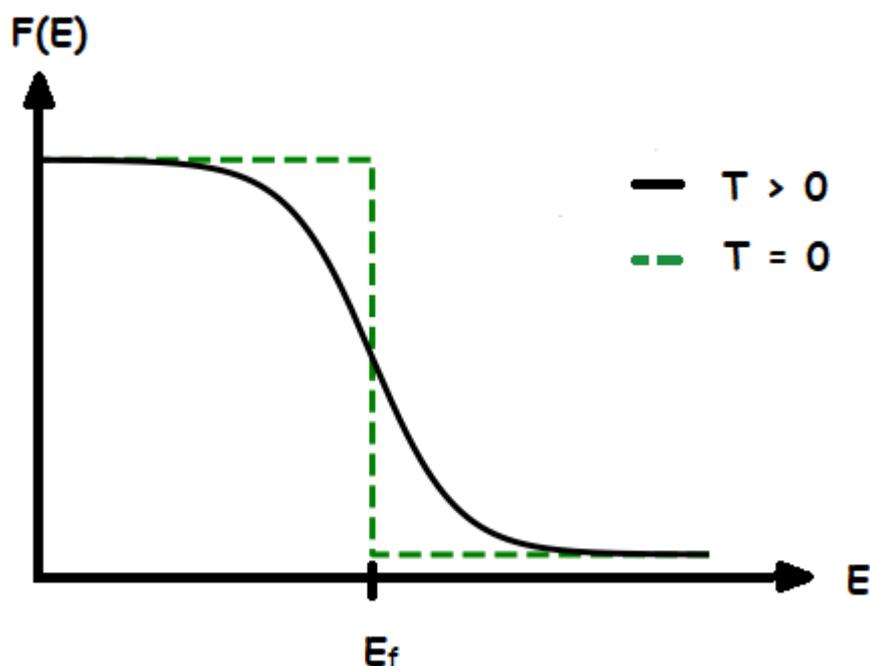
⁷Esse materiais com bandas superpostas também são chamados de semi-metais

Figura 3 – Classificação dos sólidos baseada no perfil de bandas. (a) Isolantes possuem a banda de valência completamente cheia e uma *large gap*. (b) Semicondutores a banda de valência completamente cheia e um *gap* de poucos eV. (c) condutores a banda de valência é parcialmente cheia (d) condutores onde ocorre a superposição de bandas.



Fonte: Do autor.

Figura 4 – Influência da temperatura na distribuição de Fermi. Linha contínua representa $T = 0$ K onde a função possui uma queda abrupta para E igual a E_f . Linha tracejada $T > 0$ onde ocorreu uma suavização da função em torno da energia de E_f .



Fonte: Do autor.

pequeno, inferior a 2 eV, à temperatura ambiente o número de elétrons na banda de condução é elevado de modo que acontecem excitações térmicas da banda de valência para a banda de condução e dobrando essa temperatura, indo de 300 K para 600 K cerca de um bilhão de elétrons são excitados(EISBERG; RESNICK, 1979).

Como a banda de valência está completamente cheia em 0 K, ao aumentar a temperatura os elétrons que forem excitados deixarão um estado vago denominado buraco, os quais funcionam como portadores de carga positiva na presença de campos elétricos e magnéticos, também reforçando a condutividade do material. Além dessa maneira existem outras formas de aumentar a condutividade, como a fotoexcitação, uma vez que a energia dos fótons na região do infravermelho do espectro eletromagnético é equivalente ao *gap* dos semicondutores. Por essas, e outras características, os semicondutores são a base de vários dispositivos eletrônicos (KITTEL, 2006; EISBERG; RESNICK, 1979).

3 DINÂMICA ELETRÔNICA EM UM CRISTAL MONOATÔMICO

O modelo descrito na Sec. 2 demonstra que ao unir íons para formar o cristal surgem bandas de energia, entretanto tal modelo é construído de maneira ilustrativa, não levando em conta todas as interações presentes em um cristal real. Sendo assim o Hamiltoniano descrito pela Eq. (2) deve ser reescrito levando em conta a maior parte possível das interações afim de obter informações mais precisas sobre a dinâmica eletrônica.

Em um cristal existem diversas interações que ocorrem entre os elétrons e os núcleos. Existem as interações elétron-elétron, elétron-núcleo e núcleo-núcleo. Além disso os elétrons e núcleos não estão em repouso, ou seja, possuem energia cinética. Com isso o Hamiltoniano de um cristal é dado por (YU; CARDONA, 2010; VIANNA; FAZZIO; CANUTO, 2004),

$$H = \sum_i \frac{p_i^2}{2m_i} + \sum_i \frac{P_i^2}{2M_i} + \frac{1}{2} \sum_i \frac{1}{4\pi\epsilon_0} \frac{Z_j Z_{j'} e^2}{|\mathbf{R}_j - \mathbf{R}_{j'}|} - \sum_i \frac{1}{4\pi\epsilon_0} \frac{Z_j e^2}{|\mathbf{r}_i - \mathbf{R}_j|} + \frac{1}{2} \sum_i \frac{1}{4\pi\epsilon_0} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_{i'}|}. \quad (24)$$

O primeiro e o segundo termos da Eq. (24) são as energias cinéticas dos elétrons e dos núcleos, respectivamente. Os termos restantes são a interação núcleo-núcleo, elétron-núcleo e elétron-elétron, nesta ordem. Em um cristal há cerca de 10^{23} átomos/cm³ (YU; CARDONA, 2010), assim as somas percorrem todos esse portadores.

Resolver a equação de Schrödinger para apenas um elétron já requer um esforço matemático considerável, resolvê-la para o Hamiltoniano dado pela Eq. (24) se torna praticamente impossível, de modo que algumas aproximações se fazem necessárias.

3.1 APROXIMAÇÃO DE BORN-OPPENHEIMER

O Hamiltoniano de um cristal não é separável, pois a parte nuclear e eletrônica estão acopladas devido a interação coulombiana existente no cristal. Com o intuito de desacoplar o movimento eletrônico do nuclear utiliza-se a aproximação denominada de separação de Born-Oppenheimer. Proposta em 1927 por J. Oppenheimer e por seu orientador M. Born, a aproximação apresentada a seguir, torna possível obter soluções relevantes para a equação de Schrödinger de modo mais simplificado (ALESSIO, 2009).

A massa de um núcleo é muito maior que massa de um elétron. Tal fato tem como consequência que os movimentos dos núcleos (vibracional e rotacional) são lentos em comparação ao movimento eletrônico. Assim, pode-se tratar os núcleos como sendo estáticos, as energias cinéticas dos núcleos são desprezadas e a interação núcleo-núcleo se torna constante. A única variável nesse tratamento é a posição dos núcleos. Desse modo, pode-se reescrever o Hamiltoniano do cristal separado, nas contribuições eletrônicas (H_e) e nucleares (H_N), como (LEVINE, 2013; GEORG, 2006)

$$H = H_e + H_N, \quad (25)$$

onde

$$H_e = \sum_i \frac{p_i^2}{2m_i} - \sum_i \frac{1}{4\pi\epsilon_0} \frac{Z_j e^2}{|\mathbf{r}_i - \mathbf{R}_{j'}|} + \frac{1}{2} \sum_i \frac{1}{4\pi\epsilon_0} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_{i'}|} \quad (26)$$

e

$$H_N = \frac{1}{2} \sum_i \frac{1}{4\pi\epsilon_0} \frac{Z_j Z_{j'} e^2}{|\mathbf{R}_j - \mathbf{R}_{j'}|}. \quad (27)$$

Agora como não há mais o acoplamento, pode-se resolver a parte eletrônica do cristal, a qual nos é de interesse obtendo a função de onda eletrônica, que reproduza as características periódicas do cristal, conforme indicado nas seções anteriores.

3.2 APROXIMAÇÕES DE HARTREE

Ainda que da Eq. (24) para a Eq. (26) tenha-se uma grande simplificação no Hamiltoniano, ainda assim é necessário levar em conta a interação entre todos os elétrons do cristal se fazendo necessárias mais aproximações.

Entretanto, se considerarmos que não há interação entre os elétrons e que, ao invés disso, cada elétron está sujeito a um potencial médio devido a todo o cristal ($V_{\text{médio}}^{(i)}$), incluindo os demais elétrons, é possível reescrever o Hamiltoniano total do sistema como sendo a soma dos Hamiltonianos de cada elétron h_i (MEDEIROS, 2011; CUSTODIO, 2015),

$$H = \sum_i h_i, \quad (28)$$

onde o Hamiltoniano individual será $h_i = T_i + V_{\text{médio}}^{(i)}$, ou seja, será dado pela soma da energia cinética do i -ésimo elétron com o potencial médio devido ao cristal.

Assim, a função de onda eletrônica se torna separável e pode ser escrita como um produto de funções de onda monoelétrônicas ortogonais, de tal modo que cada elétron interage com o campo médio dos outros elétrons do sistema. Nesta aproximação, a função de onda eletrônica total do sistema (Ψ) é escrita como o produto das funções de onda dos elétrons individuais (ϕ) (OLIVEIRA, 2008),

$$\Psi_{\mathbf{R}}(\mathbf{r}) = \prod_i \phi_i(\mathbf{r}_i). \quad (29)$$

Assim, o problema de inúmeros corpos é reescrito como inúmeros problemas de um único corpo, sendo a equação de Schrödinger independente do tempo, para um elétron dada por

$$h_i \phi_n^{(i)} = E_n^{(i)} \phi_n^{(i)}, \quad (30)$$

ou, explicitando o Hamiltoniano,

$$\left[-\frac{\hbar^2}{2m} \nabla_i^2 + V_{medio}^{(i)} \right] \phi_n^{(i)}(\mathbf{r}) = E_n^{(i)} \phi_n^{(i)}(\mathbf{r}). \quad (31)$$

3.3 APROXIMAÇÃO DE MASSA EFETIVA

Novamente, por mais que o problema tenha sido extremamente simplificado, não se sabe o potencial médio ao qual o elétron está sujeito, $V_{medio}^{(i)}$. Esse potencial pode ser obtido, por exemplo, iterativamente nos modelos de primeiros princípios como o Teoria do funcional da densidade (*DFT*, sigla do inglês *density functional theory*) (YU; CARDONA, 2010). Porém para contornar esse problema se pode utilizar a aproximação de massa efetiva.

A aproximação de massa efetiva consiste em transformar o problema do elétron se movimentando no cristal na presença de um potencial médio, no movimento de um elétron livre onde todo o efeito do potencial está contido no parâmetro da aproximação, denominado massa efetiva (m^*) (PACOBAYHA, 2008).

A aproximação, supõe-se que o elétron se move como se estivesse livre ($V = 0$) porém com uma massa efetiva que é diferente da massa inercial do elétron no vácuo (m_0). Assim, o Hamiltoniano é reescrito com um operador de energia cinética efetivo

$$(T_{eff}) \quad \left[-\frac{\hbar^2}{2m^*} \nabla^2 \right] \phi_n^{(i)}(\mathbf{r}) = E_n^{(i)} \phi_n^{(i)}(\mathbf{r}), \quad (32)$$

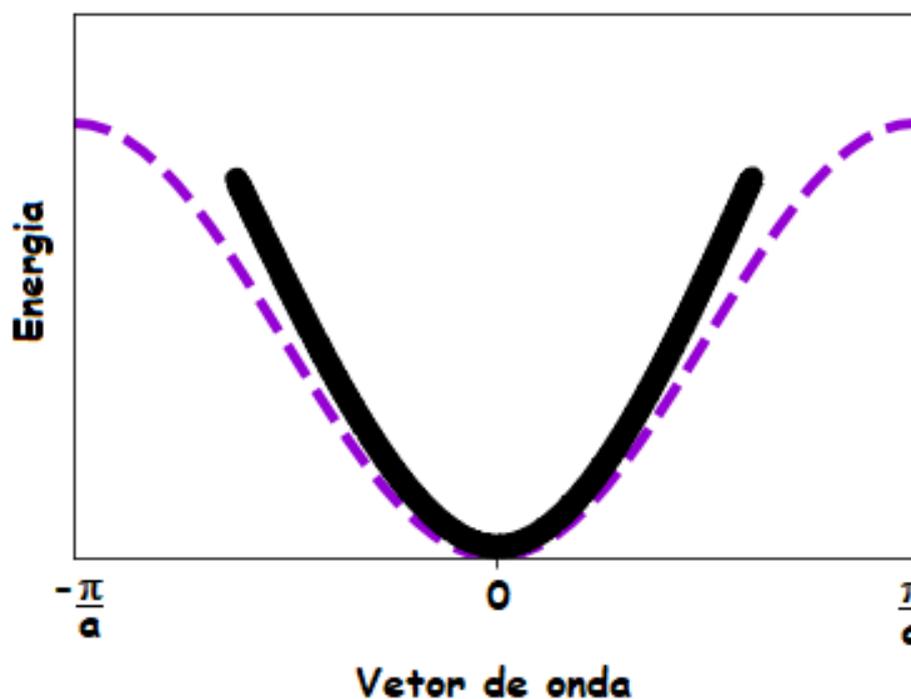
onde a massa efetiva contém toda informação do potencial ao qual o portador está submetido. Dada a necessidade de observar a periodicidade imposta pela rede cristalina, o Hamiltoniano descrito pela Eq. (32) possui autovalores que no espaço recíproco são funções não triviais do vetor de onda k , conforme veremos nas próximas seções.

Na aproximação de massa efetiva, temos que a relação de dispersão, solução da Eq. (32), é dada por

$$E_n^{(i)}(k) = \frac{\hbar^2 k^2}{2m^*}, \quad (33)$$

que é uma relação de dispersão parabólica como mostra a linha contínua da Fig. 5

Figura 5 – Relação de dispersão da energia para elétrons livres (linha contínua) e para uma rede linear com parâmetro de rede igual à a (linha tracejada) para a primeira zona de Brillouin.



Fonte: Do autor.

Ao se analisar a Eq. (33) vê-se que a curvatura da função $E_n^{(i)}(k)$ é determinada pelo coeficiente de k^2 , invertendo o raciocínio pode-se afirmar que a massa é determinada pela curvatura de $E_n^{(i)}(k)$.

Nesse sentido, a massa efetiva passa a ser um parâmetro na obtenção das pro-

priedades eletrônicas, de modo que seu valor pode ser obtido por meio de experimentos ou até mesmo através de métodos de primeiros princípio (como *DFT*). Essa aproximação tem sido amplamente utilizadas na descrição de materiais sólidos reais (BASTARD, 1988; KITTEL, 2006; SINGH, 2003), com resultados muito satisfatórios.

Porém esta aproximação é válida apenas para valores de k próximos de zero, pois para valores maiores a relação de dispersão apresenta regiões de alta curvatura (Fig. 5 linha tracejada), principalmente perto de $k = \pm\pi/a$, ou seja, próximos do *gap* de energia (KITTEL, 2006). Ainda assim, para pontos próximos de $k = 0$, ponto de simetria Γ no centro da primeira zona de Brillouin, pode se dizer que a banda é aproximadamente parabólica, tornando válida a relação acima. O que impõe a condição de que os autovalores de h_i sejam iguais aos de T_{eff} perto do ponto Γ (PACOBAYHA, 2008).

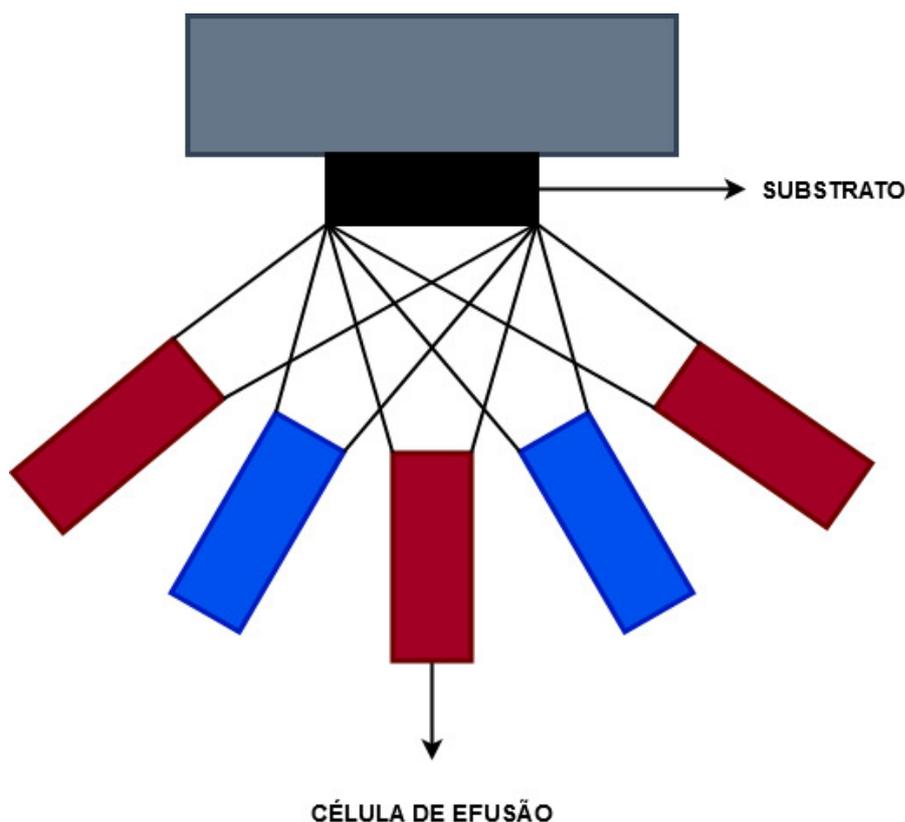
4 DINÂMICA DE UM ELÉTRON EM UMA HETEROESTRUTURA

Todas as aproximações mostradas anteriormente são aplicadas a cristais monoatômicos. Ao se trabalhar com heteroestruturas são necessárias algumas aproximações a mais. A seguir discutimos o que são heteroestruturas, como é realizado seu crescimento e as aproximações utilizadas.

4.1 HETEROESTRUTURAS

Ao se colocar dois semicondutores distintos, mas que possuem parâmetros de rede compatíveis, em contato obtemos uma heteroestrutura semicondutora, uma estrutura cristalina formada por dois diferentes materiais (BEZERRA, 2010).

Figura 6 – Exemplo esquemático de como ocorre o processo de crescimento de heteroestruturas semicondutoras utilizando a técnica de *molecular beam epitaxy*.



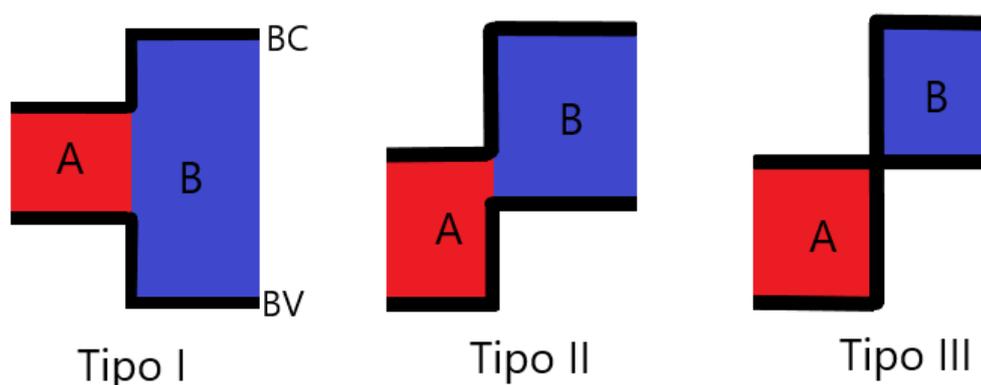
Fonte: Do autor.

O crescimento de heteroestruturas é, em geral, realizado a partir de métodos como Epitaxia de Feixe Molecular (*MBE*, do inglês *Molecular Beam Epitaxy*) e Deposição de Vapor Químico Metal-Orgânico (*MOCVD*, do inglês *Metal-Organic Chemical Vapour*

Deposition) (BEZERRA, 2010; SILVA, 2010). A MBE, por exemplo, consiste de uma câmara de ultra-alto vácuo onde os materiais são aquecidos e evaporados. Os materiais são colocados em células de efusão onde são evaporados, atingindo o substrato onde ocorre o crescimento da estrutura. Na frente de cada célula existem obturadores que bloqueiam determinado feixe de átomos de modo que possa se construir camadas de diferentes materiais, com precisão de monocamadas como exemplificado na Fig. 6 (LAMAS, 2004).

No crescimento de determinadas estruturas pode ocorrer o confinamento eletrônico em uma, duas ou três dimensões, dando origem a poços, fios e pontos quânticos (SILVA, 2010). Dependendo do alinhamento entre as bandas de valência (BV) e condução (BC) nas diferentes camadas (*band offset*), as estruturas são classificadas em três tipos, conforme mostra a Fig. 7. Para a heteroestrutura do tipo I, a BC e a BV de um dos semicondutores são respectivamente superior e inferior às bandas correspondentes do outro semicondutor. Na heteroestrutura do tipo II tanto a BV quanto a BC de um material são inferiores a do outro, e a última que se assemelha muito a do tipo dois porém ocorre uma descontinuidade dos *gaps* (BEZERRA, 2010; ROUL et al., 2015).

Figura 7 – Classificação das heteroestruturas baseado na descontinuidade da interface dos materiais. Heteroestruturas do tipo I possui a banda de condução e de valência de um material inferior e superior do que do outro, do tipo II tanto a banda de condução quanto de valência são inferiores de um material em relação ao outro e do tipo III ocorre a descontinuidade dos *gaps*.



Fonte: Do autor.

Como cada material possui seu perfil de bandas único, o crescimento de estruturas do tipo BAB onde o material A possui o *gap* de energia menor que do material B, possibilita realizar uma engenharia no perfil de potencial e obter estruturas com poços quânticos, os quais são a base de dispositivos como *lasers* de diodo, CCDs, células

solares, entre outros, de modo que o estudo dessas estruturas são de grande interesse. Entretanto, além das discussões realizadas anteriormente é necessário levar em conta mais algumas aproximações no caso de heteroestruturas as quais serão discutidas a seguir.

4.2 APROXIMAÇÃO DE FUNÇÃO ENVELOPE

As aproximações acima são aplicadas a semicondutores periódicos perfeitos, monoatômicos, de tamanho infinito, ou também denominados de *bulk*. Em heteroestruturas, dada a existência de interfaces, usaremos a aproximação de função envelope (AFE) para descrever os estados eletrônicos.

Assumindo que A e B são os materiais constituintes da heteroestrutura, a AFE consiste de duas hipóteses que são (ANGHINONI, 2012; BASTARD, 1988):

- Dentro de cada material as funções de onda são expandidas em termos das partes periódicas das funções de Bloch em torno de um determinado ponto k (k_0 nesse caso). Temos que

$$\phi_A(\mathbf{r}) = \sum_l f_l^{(A)}(\mathbf{r}) u_{l,k_0}^A(\mathbf{r}), \quad (34)$$

é a função de onda no material A e

$$\phi_B(\mathbf{r}) = \sum_l f_l^{(B)}(\mathbf{r}) u_{l,k_0}^B(\mathbf{r}) \quad (35)$$

corresponde à função de onda no material B.

- A parte periódica das funções de Bloch são assumidas idênticas nos dois materiais

$$u_{l,k_0}^A = u_{l,k_0}^B, \quad (36)$$

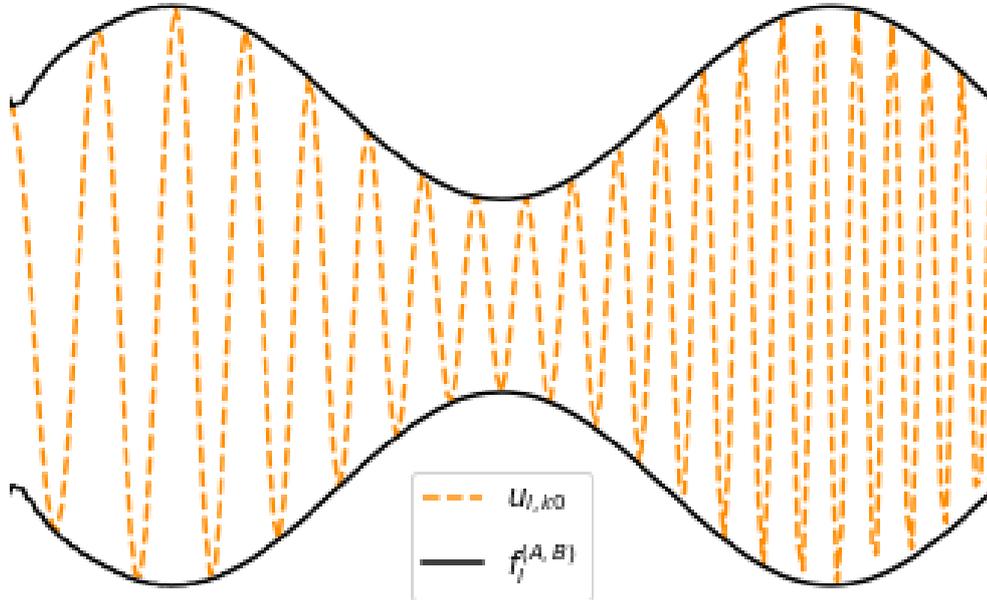
sendo a função de onda da heteroestrutura dada por

$$\phi(\mathbf{r}) = \sum_l f_l^{(A,B)}(\mathbf{r}) u_{l,k_0}(\mathbf{r}). \quad (37)$$

Assim, a função de onda é dada pelo produto de duas funções, uma que varia rapidamente u_{l,k_0} com a periodicidade da rede cristalina e a função envelope $f_l^{(A,B)}$, que varia lentamente (Fig. 8), determinada pelo perfil de potencial gerado pelas interfaces

entre os materiais.

Figura 8 – Exemplo de função periódica (linha tracejada) modulado por uma função envelope (linha contínua).



Fonte: Do autor.

Em estruturas tipo I, as quais definem poços quânticos, os portadores estão confinados em uma direção. Assumindo o crescimento das camadas atômicas dado na direção z , no plano $z = z_0$ ⁸ pelo fato das funções de Bloch serem independentes e a função envelope $f_l^{(A,B)}$ ser contínua para \hat{z} , temos que na interface entre os materiais

$$f_l^{(A)}(\mathbf{r}_\perp, z_0) = f_l^{(B)}(\mathbf{r}_\perp, z_0). \quad (38)$$

Dada a invariância translacional da heteroestrutura, devido a proximidade dos parâmetros de rede dos materiais, a função de onda envelope $f_l^{(A,B)}(\mathbf{r}_\perp, z_0)$ pode ser fatorada em

$$f_l^{(A,B)}(\mathbf{r}_\perp, z_0) = \frac{1}{\sqrt{S}} e^{i\mathbf{k}_\perp \cdot \mathbf{r}_\perp} \psi_l^{A,B}(z) \quad (39)$$

onde S é a área da amostra no plano x - y , $\mathbf{r}_\perp = (r_x, r_y)$ é o vetor posição perpendicular à direção de crescimento e $\mathbf{k}_\perp = (k_x, k_y)$ é o vetor de onda nessa mesma direção. Dessa forma, no plano x - y a função de onda é representada por ondas planas e o problema torna-se separável, nos levando a uma equação de Schrödinger

⁸O ponto $z = z_0$ define a interface entre as diferentes camadas de materiais semicondutores.

unidimensional.

O Hamiltoniano da heteroestrutura passa a ser escrito como

$$H(z) = \frac{p^2}{2m^*} + V_{hetero}(z), \quad (40)$$

onde $V_{hetero}(z)$ define o perfil de potencial gerado pela descontinuidade dos *band gaps* nas diferentes camadas

Uma vez que o potencial varia lentamente em relação à periodicidade cristalina, a função de onda, Eq. (37), aplicada ao Hamiltoniano da Eq. (40), retorna uma equação de autovalores de autovetores unidimensional

$$\left[\frac{p^2}{2m^*} + V_{hetero} \right] \psi_l(\mathbf{z}) = E_l \psi_l(\mathbf{z}). \quad (41)$$

onde V_{hetero} é descrito através de um potencial de poço quântico. Dessa forma, a descrição das propriedades de heteroestruturas semicondutoras com confinamento em uma dimensão (poço quântico), pode ser tratada pela solução da equação de Schrödinger independente do tempo unidimensional, para um elétron sujeito a ação de um potencial definido pela descontinuidade dos *band gaps*. Como podemos notar, a contribuição da rede cristalina é determinada tanto pelos *band offsets* quanto pela massa efetiva, parâmetros esses que devem ser conhecidos de antemão, obtidos à partir de resultados experimentais ou simulações de primeiros princípios.

4.3 EQUAÇÃO DE SCHRÖDINGER COM MASSA EFETIVA DEPENDENTE DA POSIÇÃO

Todo estudo realizado nas seções anteriores foi aplicado em cristais formados por apenas um material sendo assim, sua massa é constante e para as heteroestruturas não se fez distinção entre as massas de modo que a Eq. (41) é descrita para uma massa efetiva constante, entretanto ao se tratar de heteroestruturas como cada material possui uma estrutura de bandas particular, implica que os portadores possuam massas efetivas diferentes em cada material. Assim dizemos que o material possui uma massa variável na direção de crescimento da heteroestrutura, ou seja, $m(z)$ ⁹.

Como o operador de energia cinética tem dependência direta com a massa, o

⁹No restante desta seção será omitido o índice * na massa efetiva por simplicidade de notação

Hamiltoniano com esse operador retorna autovalores não-hermitianos sendo necessário reescrevê-lo. Roos (1983) propôs um operador de energia cinética o qual retorna autovalores hermitianos para problemas com massa variável, tal operador é dado por (ROOS, 1983)

$$\hat{T} = -\frac{\hbar^2}{4} \left(m^\alpha \frac{\partial}{\partial z} m^\beta \frac{\partial}{\partial z} m^\gamma + m^\gamma \frac{\partial}{\partial z} m^\beta \frac{\partial}{\partial z} m^\alpha \right), \quad (42)$$

onde m é massa variável e os parâmetros α , β e γ são constantes tal que $\alpha + \beta + \gamma = -1$.

Escolhendo $\alpha + \gamma = 0$ e $\beta = -1$ e utilizando álgebra elementar a Eq. (42) é reescrita de modo simetrizado como

$$\hat{T} = -\frac{\hbar^2}{2} \frac{\partial}{\partial z} \frac{1}{m} \frac{\partial}{\partial z}. \quad (43)$$

ou ainda

$$\hat{T} = -\frac{\hbar^2}{2} \left[\left(\frac{d}{dz} \frac{1}{m} \right) \frac{d}{dz} + \frac{1}{m} \frac{d^2}{dz^2} \right]. \quad (44)$$

Substituindo a Eq. (44) na Eq. (41) temos que a equação de Schrödinger para uma heteroestrutura retorna

$$-\frac{\hbar^2}{2} \left[\left(\frac{d}{dz} \frac{1}{m} \right) \frac{d}{dz} + \frac{1}{m} \frac{d^2}{dz^2} \right] \psi_l + V_{hetero}(z) \psi_l = E_l \psi_l. \quad (45)$$

Para verificar a legitimidade da Eq. (45), basta supor uma massa constante e verificar que ela retorna a Eq. (41) acrescida do potencial que define o perfil da heteroestrutura.

5 EFEITOS DE NÃO-PARABOLICIDADE

Ao calcular os níveis de energia permitidos em um poço quântico semiconductor, devemos levar em consideração efeitos de bandas não-parabólicas (PERSSON; COHEN, 1988). Nas heteroestruturas efeitos como mistura de bandas, confinamento quântico de sub-bandas, e as densidades dos portadores e seus graus de influência variam dependendo dos materiais e da estrutura, fazendo com que a dispersão não seja apenas parabólica (ALHARBI, 2008). A relação completa é dada pela expansão da série em potências de k

$$E = \sum_{i=0}^{\infty} a_i k^i \quad (46)$$

Para estruturas cristalinas com simetria de inversão espacial, a relação de dispersão é uniformemente simétrica em qualquer direção, independentemente da complexidade da estrutura. Então, apenas os componentes pares de k contribuem para a expansão (ALHARBI, 2008), portanto, a Eq. (46) é reduzida para

$$E = \sum_{i=0}^{\infty} a_{2i} k^{2i}. \quad (47)$$

Geralmente, os cristais semicondutores dotados de poços quânticos são intrinsecamente inversos assimétricos. No entanto, por decomposição, a porção assimétrica é muito pequena em comparação com a parte simétrica. Desse modo, os efeitos assimétricos de inversão podem ser ignorados. Os componentes de baixa ordem em k são os mais significativos (EISENSTEIN et al., 1984), portanto, a equação acima pode ser reescrita como (ALHARBI, 2008; NELSON; MILLER; KLEINMAN, 1987)

$$E = \frac{\hbar^2 k^2}{2m^*} (1 - \gamma k^2), \quad (48)$$

onde γ é um parâmetro definido pela curvatura da banda, chamado de parâmetro de não-parabolicidade.

Nelson, Miller e Kleinman (1987) demonstra que é possível resolver a não-parabolicidade das bandas de maneira desacoplada com uma equação de Schrödinger para banda de condução e outra para a valência, com uma massa dependente da energia. Para tanto ele introduziu o conceito de banda fictícia, que é devida ao acoplamento e renormaliza o *gap* de energia do material (NELSON; MILLER; KLEINMAN, 1987; LEAVITT, 1991).

Levando em consideração que as propriedades eletrônicas são mais relevantes para os elétrons nas bandas de valência e condução, devemos nos ater ao acoplamento entre tais bandas. Para isso, podemos escrever o Hamiltoniano do sistema na base das autoenergias das bandas de condução e de valência, o que nos leva a uma representação matricial na forma de um Hamiltoniano 2 x 2, dado por

$$\mathbf{H} = \begin{pmatrix} E_c & \frac{i\hbar^2 k}{2m^* \sqrt{\gamma}} \\ -\frac{i\hbar^2 k}{2m^* \sqrt{\gamma}} & E_c - \frac{\hbar^2}{2m^* \gamma} \end{pmatrix}, \quad (49)$$

onde os elementos de matriz H_{11} e H_{22} representam a banda de condução e valência, respectivamente. Os elementos H_{12} e H_{21} representam o acoplamento entre estas bandas.

Adicionando o potencial devido á heteroestrutura, $V_{hetero}(z)$ e usando as relações $k = p/\hbar$ e $p = -i\hbar \frac{d}{dz}$ o Hamiltoniano da Eq. (49) pode ser reescrito como

$$\mathbf{H} = \begin{pmatrix} E_c + V(z) & \frac{\hbar^2}{2m^* \sqrt{\gamma}} \frac{d}{dz} \\ -\frac{\hbar^2}{2m^* \sqrt{\gamma}} \frac{d}{dz} & E_c + V(z) - \frac{\hbar^2}{2m^* \gamma} \end{pmatrix}. \quad (50)$$

As bandas de valência e condução com autofunções, dadas respectivamente por ψ_c e ψ_v , são determinadas pelos autovetores da matriz que descreve o Hamiltoniano

$$\begin{pmatrix} E_c + V(z) - E & \frac{\hbar^2}{2m^* \sqrt{\gamma}} \frac{d}{dz} \\ -\frac{\hbar^2}{2m^* \sqrt{\gamma}} \frac{d}{dz} & E_c + V(z) - \frac{\hbar^2}{2m^* \gamma} - E \end{pmatrix} \begin{pmatrix} \psi_c \\ \psi_v \end{pmatrix} = 0. \quad (51)$$

A partir da Eq. (51) obtemos um sistema de equações

$$[E_c + V(z) - E]\psi_c + \frac{\hbar^2}{2m^* \sqrt{\gamma}} \frac{d}{dz} \psi_v = 0, \quad (52)$$

$$-\frac{\hbar^2}{2m^* \sqrt{\gamma}} \frac{d}{dz} \psi_c + \left(E_c + V(z) - \frac{\hbar^2}{2m^* \gamma} - E \right) \psi_v = 0. \quad (53)$$

Manipulando a Eq. (53) encontra-se ψ_v

$$\psi_v = -\frac{m^* \sqrt{\gamma}}{m^*(E)} \frac{d\psi_c}{dz} \quad (54)$$

onde

$$m^*(E) = m^* \left(1 + \frac{2m^*\gamma[E - E_c - V]}{\hbar^2} \right). \quad (55)$$

Substituindo a relação (54) em (52) e com alguns manipulações matemáticas, alcançamos a equação de Schrödinger para a banda de condução, onde o parâmetro $m(E)$, contém as informações da não-parabolicidade da banda em questão.

$$-\frac{\hbar^2}{2} \frac{d}{dz} \left[\frac{1}{m^*(E)} \frac{d\psi_c}{dz} \right] + V(z)\psi_c = (E - E_c)\psi_c. \quad (56)$$

Uma vez que estamos tratando materiais semicondutores com perfil de poço quântico, temos que na região do poço, $V = 0$, e na região da barreira, $V = V_0$, assim temos a massa dependente da energia no poço e na barreira dadas por

$$m_w^*(E) = m_w^* \left(1 + \frac{E - E_c}{E_w} \right) \quad (57)$$

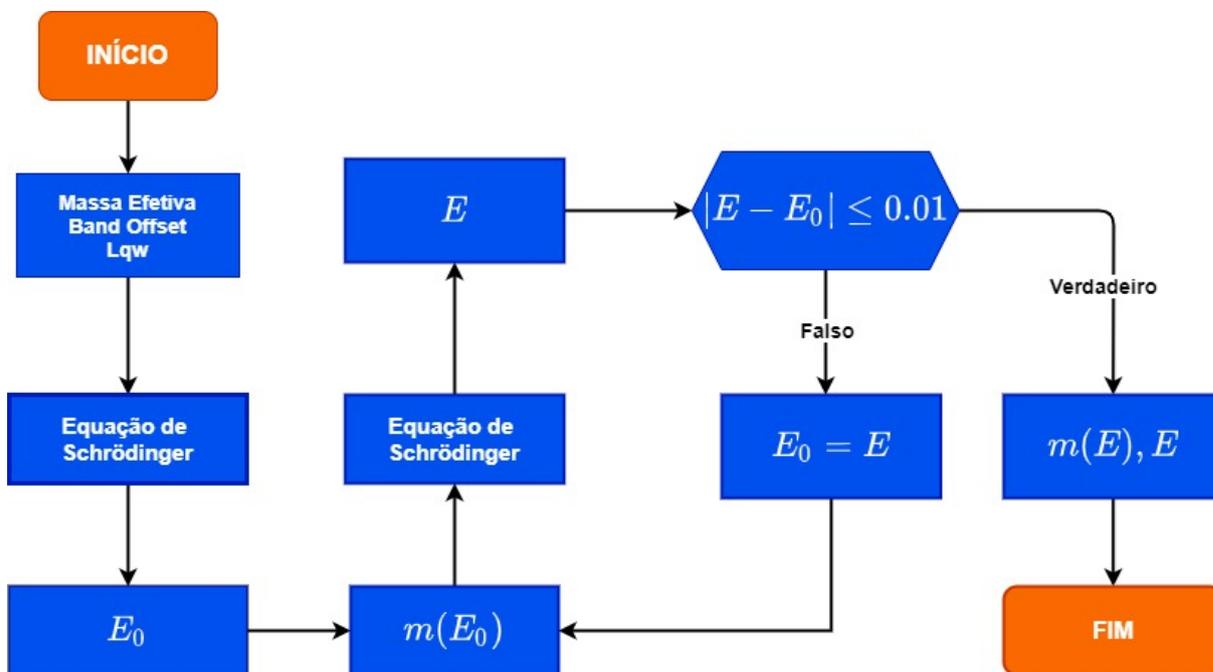
$$m_b^*(E) = m_b^* \left[1 - \frac{(V - E + E_c)}{E_b} \right], \quad (58)$$

onde E_b e E_w são os *band gaps* normalizados nas regiões da barreira e do poço quântico, respectivamente, que dependem do parâmetro de não parabolicidade γ . Podemos agora tratar o problema de forma desacoplada entre as bandas de valência e condução, mas para tanto há a necessidade da inclusão dessa normalização no *band gap*, como se o problema desacoplado fosse dado por bandas fictícias separadas por *gaps* de energia E_b e E_w .

Mesmo tendo separado o problema, introduzindo uma massa que depende da energia, tal problema pode ser resolvido iterativamente. Machado (2019). mostra uma forma de resolver a equação de Schrödinger Eq. (56), na aproximação de massa efetiva e bandas não parabólicas a partir de um método iterativo (MACHADO, 2019). O método é iniciado usando a massa efetiva, m^* e a descontinuidade da interface (*band offset*) na Eq. (56) e se obtém a energia do estado fundamental (E). Aplicando essa energia na Eq. (57), calcula-se a massa dependente da energia, $m(E)$, determinando uma correção para a massa efetiva. Com a massa atualizada, recalculamos a autoenergia (E_n) e este processo se repete até que o módulo da energia calculada no processo anterior (E_0) menos a do processo atual (E_n) esteja no limite imposto. Esse procedimento é

ilustrado na Fig. 9.

Figura 9 – Fluxograma do método iterativo para o cálculo na massa dependente da energia e da energia via Eq. de Schrödinger proposto por Machado (2019).



Fonte: Do autor.

Com a solução matricial pode-se obter os autovalores e autovetores no método iterativo. Entretanto esse processo não é algo simples e rápido, possuindo custo computacional elevado. Com isso, propomos a utilização de uma rede neural multi-camadas para a obtenção das autoenergias de sistemas heteroestruturados. Nesse sentido, a solução da Eq. (56) para diversas heteroestruturas, fornece resultados de autovalores que serão usados como parâmetros de saída para a rede neural. Uma vez bem treinada, ou seja, apresentando um erro baixo, ao passarmos os parâmetros de entrada¹⁰, a rede deve estar apta a retornar o espectro de autovalores para o sistema determinado por esses parâmetros.

¹⁰Parâmetros definidos na Sec. 8

6 INTELIGÊNCIA ARTIFICIAL

Inteligência artificial (IA) é uma área da ciência da computação voltada para o desenvolvimento de sistemas de computadores, que exibem as características que associamos à inteligência no comportamento humano, como percepção, processamento de linguagem natural, resolução e planejamento de problemas, aprendizado e adaptação e atuando no meio ambiente (TECUCI, 2012).

A IA possui diversos métodos para que os softwares consigam interpretar dados e obter uma saída, alguns desses métodos são: redes neurais artificiais, algoritmos genéticos, lógica *fuzzy*, solução e planejamento de problemas, raciocínio não monotônico, programação lógica, processamento de linguagem natural, visão computacional, entre outros (MELLIT; KALOGIROU, 2008). Como mencionado anteriormente neste trabalho utilizamos as redes neurais artificiais que serão detalhadas a seguir.

6.1 REDES NEURAIS ARTIFICIAIS

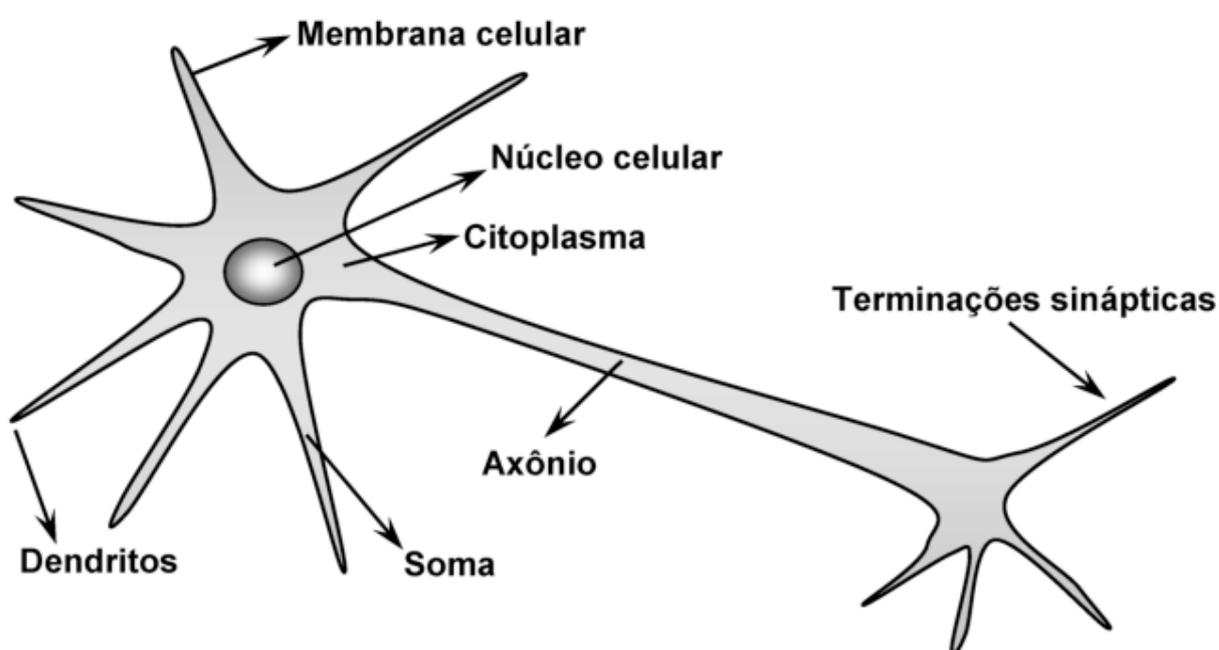
Para se desenvolver sistemas inteligentes é necessário a utilização de algoritmos para processar dados, aprender com eles e tomar decisões. Esse processo que ocorre na IA é denominado Aprendizado de Máquina (*ML* do inglês *Machine Learning*). Uma técnica de *ML* é o Aprendizado Profundo ou *Deep Learning (DL)*, que tem como fundamento as redes neurais artificiais (PACHECO; PEREIRA, 2018). Trata-se de modelo matemático inspirado em uma rede de neurônios biológicos como, por exemplo, do cérebro humano.

O cérebro, principal órgão do sistema nervoso central, é um órgão complexo constituído de mais de cem bilhões de neurônios, que são sua unidade básica (BEAR; CONNORS; PARADISO, 2002). As ligações entre esses bilhões de neurônios formam uma rede neural biológica, a qual torna possível para os seres vivos aprenderem. Estes neurônios possuem uma anatomia constituída de três diferentes partes, a soma ou corpo celular, os dendritos e o axônio (Fig. 10). Os neurônios, são células que têm como característica a transmissão de informações, num sistema de gatilho acionado a partir de limites eletroquímicos bem determinados, que acontece da seguinte forma. Os dendritos, que são vários prolongamentos finos que formam a árvore dendrital, captam os estímulos recebidos em um determinado período de tempo e os transmitem

ao corpo do neurônio, onde são processados (SILVA; SPATTI; FLAUZINO, 2010).

Quando tais estímulos atingirem determinado limite, o corpo da célula envia novo impulso que se propaga pelo axônio que é constituído por um único prolongamento e é transmitido às células vizinhas (FERNEDA, 2006; SILVA; SPATTI; FLAUZINO, 2010). Quando dois ou mais neurônios são interligados ocorre um processo de sinapse, que são conexões que viabilizam a transferência de estímulos do axônio de um neurônio para os dendritos de outros.

Figura 10 – Representação da anatomia de um neurônio biológico destacando sua principais partes.



Fonte: Ref. (SILVA; SPATTI; FLAUZINO, 2010).

6.2 NEURÔNIO MATEMÁTICO

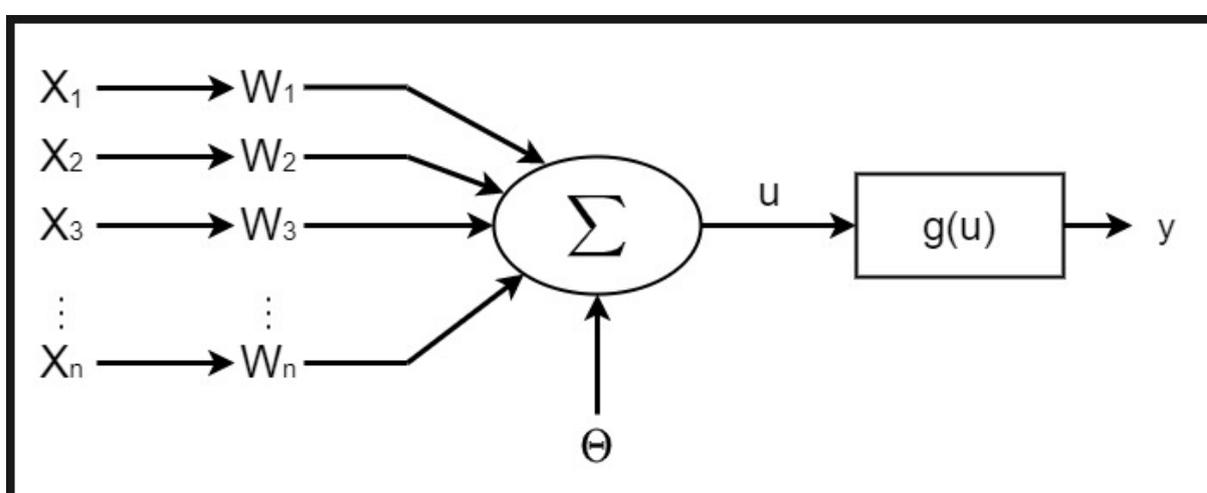
Inspirado no sistema nervoso dos seres vivos, o neurofisiologista Warren McCulloch e o matemático Walter Pitts, em 1943, no artigo “*A Logical Calculus of the Ideas Immanent in Nervous Activity*” (MCCULLOCH; PITTS, 1943) propõe um modelo matemático para o neurônio biológico utilizando a lógica proposicional devido ao caráter tudo ou nada (*all or none*) dos neurônios.

O neurônio matemático também chamado de neurônio artificial (NA), consiste em um modelo simplificado do biológico¹¹. Os impulsos elétricos recebidos pelos dendritos

¹¹O restante do capítulo tem como base a referência (SILVA; SPATTI; FLAUZINO, 2010).

no neurônio biológico são dados por sinais de entrada (*inputs*) representados pelo conjunto $(x_1, x_2, x_3, \dots, x_n)$. Os *inputs* são multiplicados por um conjunto de valores denominados pesos sinápticos, $(w_1, w_2, w_3, \dots, w_n)$, que determinam a importância que cada *input* terá para o resultado final. Em outras palavras, a modificação no valor de entrada, que leva ao valor da saída, se dá pelo valor do peso sináptico ao qual o valor de entrada é ponderado.

Figura 11 – Representação de um neurônio artificial onde as entradas são representadas por x_i , w_i são os pesos sinápticos, Θ é chamado de *bias*, $g(u)$ a função de ativação e y a saída do neurônio.



Fonte: Do autor.

A soma ou corpo da célula é representada por uma composição de dois módulos, o primeiro é uma função aditiva, somatório dos sinais de entrada multiplicados por seus respectivos pesos sinápticos. As entradas são ponderadas na soma e se alcançarem determinado patamar ou limiar, um novo impulso é gerado e passado adiante. Nos neurônios matemáticos esse patamar é dado por um valor (Θ) denominado *bias*. O *bias* possibilita que um neurônio apresente saída não nula ainda que todas as suas entradas sejam nulas (ACADEMY, 2019).

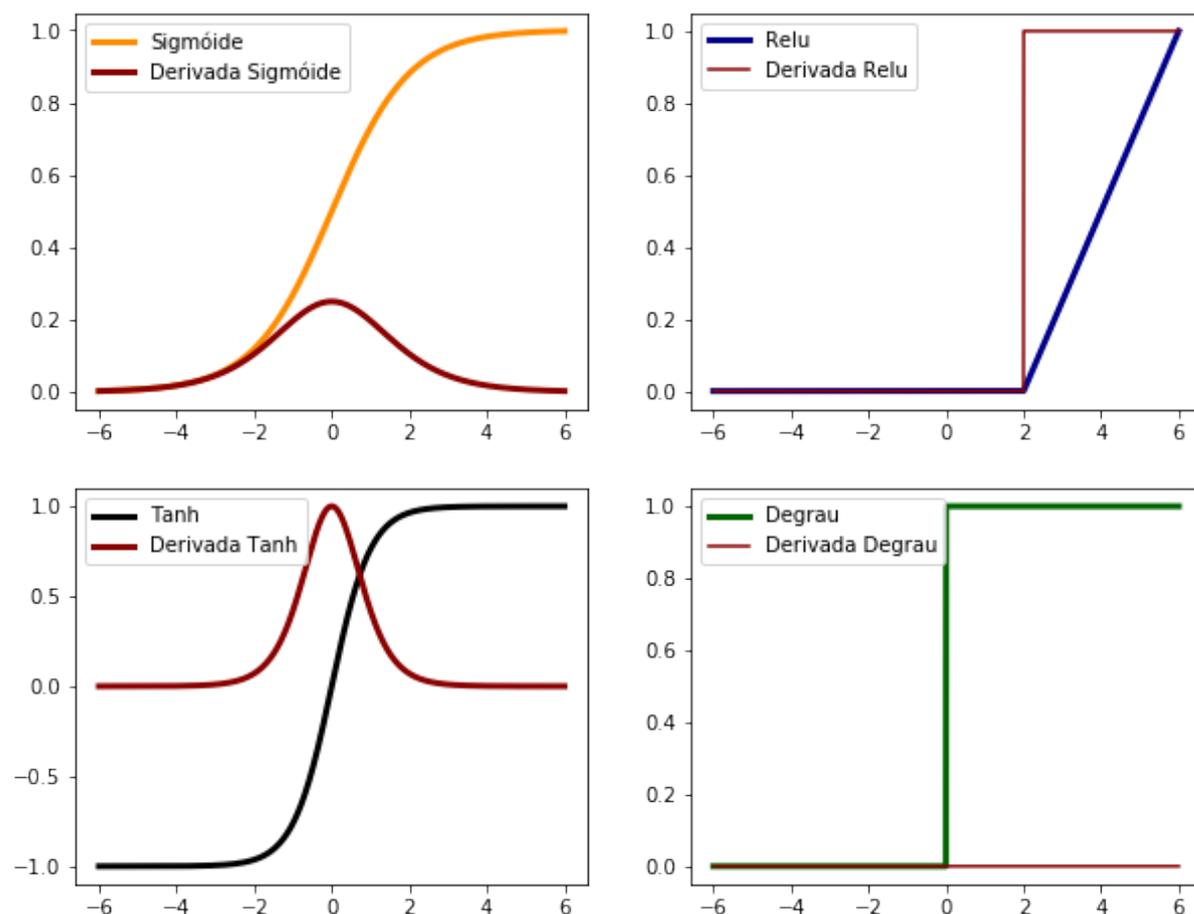
$$u = \sum_{i=1}^n w_i \times x_i - \Theta. \quad (59)$$

A função u é denominada potencial de ativação o qual é aplicado a uma função de ativação ($g(u)$) que definirá, com base nas entradas e pesos sinápticos, qual será a saída do neurônio. O axônio é aqui representado pela saída y (*output*), que é o

resultado da aplicação da função aditiva na função de ativação (ACADEMY, 2019), alguns exemplos dessas funções são mostrados na Fig. 12.

$$y = g(u) = g\left(\sum_{i=1}^n w_i \times x_i - \Theta\right). \quad (60)$$

Figura 12 – Exemplos de funções de ativação utilizadas em neurônios artificiais e suas respectivas derivadas.



Fonte: Do autor.

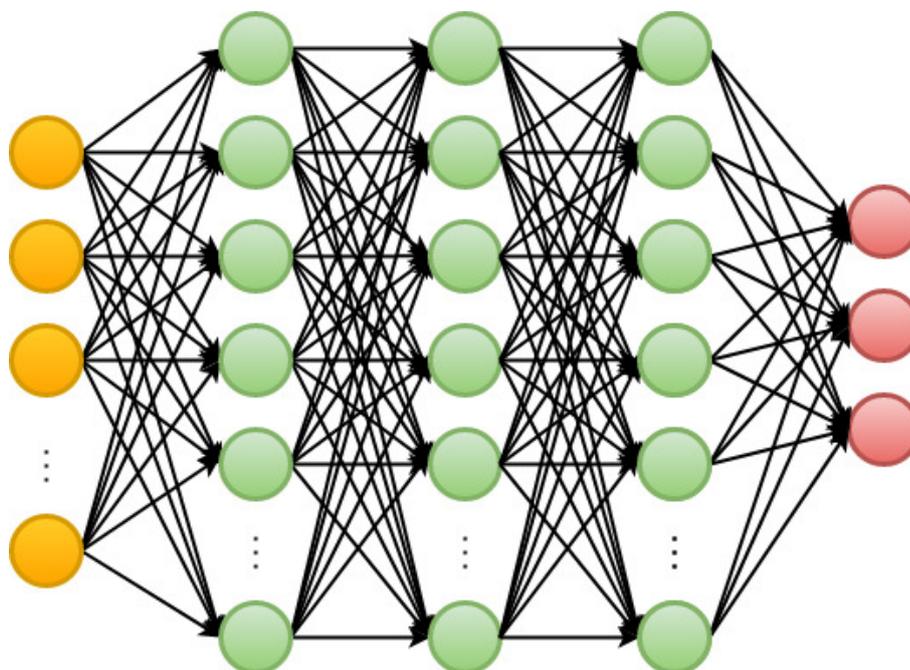
Em termos gerais e simples, um neurônio matemático é um componente que calcula a soma ponderada de vários *inputs*, aplica uma função de ativação e passa o resultado adiante. Quando vários neurônios estão conectados formam uma rede neural artificial.

6.3 ARQUITETURA DAS REDES NEURAIS ARTIFICIAIS

Existem diversas formas de organizar vários neurônios, caracterizando a arquitetura da rede. A arquitetura mais comum é a *perceptron* proposta por Frank Rosenblatt em 1958 (ROSENBLATT, 1958). Um *perceptron* é um neurônio matemático que recebe várias entradas e produz uma única saída binária. No presente trabalho utilizamos uma rede *perceptron* multicamadas (*MLP*, do inglês *multilayer perceptron*).

As *MLP* são arquiteturas com alimentação para frente (*feed forward*). Nesta arquitetura os inputs são conectados a várias camadas consecutivas de neurônios, denominadas camadas escondidas ou *hidden layers*, que são ligadas ao *output*. Se houver mais de uma camada oculta, estas são classificadas como redes neurais "profundas" (ou *deep neural network*).

Figura 13 – Diagrama de uma rede neural *perceptron* multicamadas profunda com 3 camadas ocultas.



Fonte: Do autor.

As *MLP* são empregadas a problemas relacionados à aproximação de funções, reconhecimento de padrões, otimização de sistemas, identificação e controle de processos etc. Nesta arquitetura o fluxo de sinais ocorre sempre para frente, como indicam as flechas na Fig. 13. A camada de entrada serve de *input* para a primeira *hidden layer*, que por sua vez será o *input* da próxima camada e assim por diante até chegarem no

output. Uma especificação muito importante das *MLP* é o fato que o *output* não precisa ser necessariamente apenas um neurônio como na *perceptron*, o *output* pode ter quantos neurônios forem necessários, número esse determinado por meio do aprendizado ou treinamento.

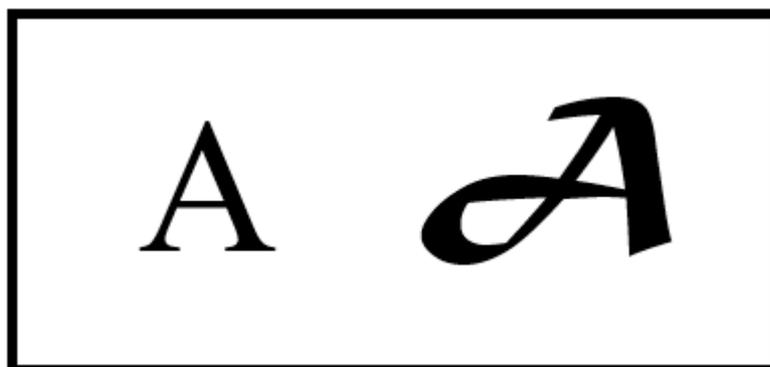
6.4 TREINAMENTO

Uma característica importante dos neurônios biológicos é a de adquirir conhecimento, aprender. Para o neurônio artificial aprender significa em encontrar os valores apropriados para os pesos sinápticos (w_i) e para os limiares de ativação (Θ). A capacidade humana de aprender ocorre a partir de experiências vividas anteriormente que são recebidas pelos neurônios processadas e distribuídas pelo cérebro (FERNEDA, 2006).

Já nas RNA existem alguns algoritmos de aprendizagem a fim de relacionar os *input(s)* com o(s) *output(s)*. Existem diversos tipos de treinamento como por exemplo, treinamento não-supervisionado, treinamento com reforço, treinamento supervisionado, entre outros. Para as *MLP* que são redes *feed forward* o tipo de treinamento é realizado efetivamente de forma supervisionada.

O treinamento supervisionado (TS) consiste em ter para cada conjunto de *inputs* seus respectivos *outputs*. Pode-se entender o TS como sendo uma aprendizagem onde se tem um professor que lhe passa uma tarefa (*inputs*) e verifica se você realizou tudo certo e alcançou o resultado esperado (*output*).

Figura 14 – Exemplos de dados de entrada para uma rede neural artificial.



Fonte: Do autor.

Supomos que a entrada para uma rede seja a imagem de uma letra (Fig. 14

esquerda), lembrando que podemos converter uma imagem em uma matriz, ou seja, se tem diversos valores de entrada, como se trata de um TS temos também que fornecer a saída como por exemplo "A".

Com base nestes dois valores, *input* e *output*, a rede se organiza ajustando os pesos e os limiares de modo que, se inserirmos como entrada a imagem no painel da direita da Fig. (14), por mais que não seja idêntica a figura com a qual foi treinada, a rede aprende alguns padrões e nos retorna como saída "A".

6.5 BACKPROPAGATION

O algoritmo utilizado para realizar o treinamento da redes *MLP* é chamado de algoritmo de retropropagação, ou do inglês *backpropagation* (*BP*). O processo de treinamento consiste em duas fases, a primeira chamada de propagação adiante (*forward*) onde as entradas da redes são propagadas camada a camada, da entrada até a saída, realizando previsões.

Na fase de *forward* os pesos sinápticos são gerados de maneira aleatória, a matriz de pesos será representada por W_{ji}^L onde os elementos da matriz conectam o j -ésimo neurônio da camada (L) ao i -ésimo neurônios da camada ($L - 1$), por exemplo para $L = 1$, faz referência a matriz que conecta a primeira camada escondida com a camada de entrada. A função que calcula as entradas ponderadas pelos pesos será dada por

$$I_j^L = \sum_{i=0}^n W_{ji}^L \times x_i. \quad (61)$$

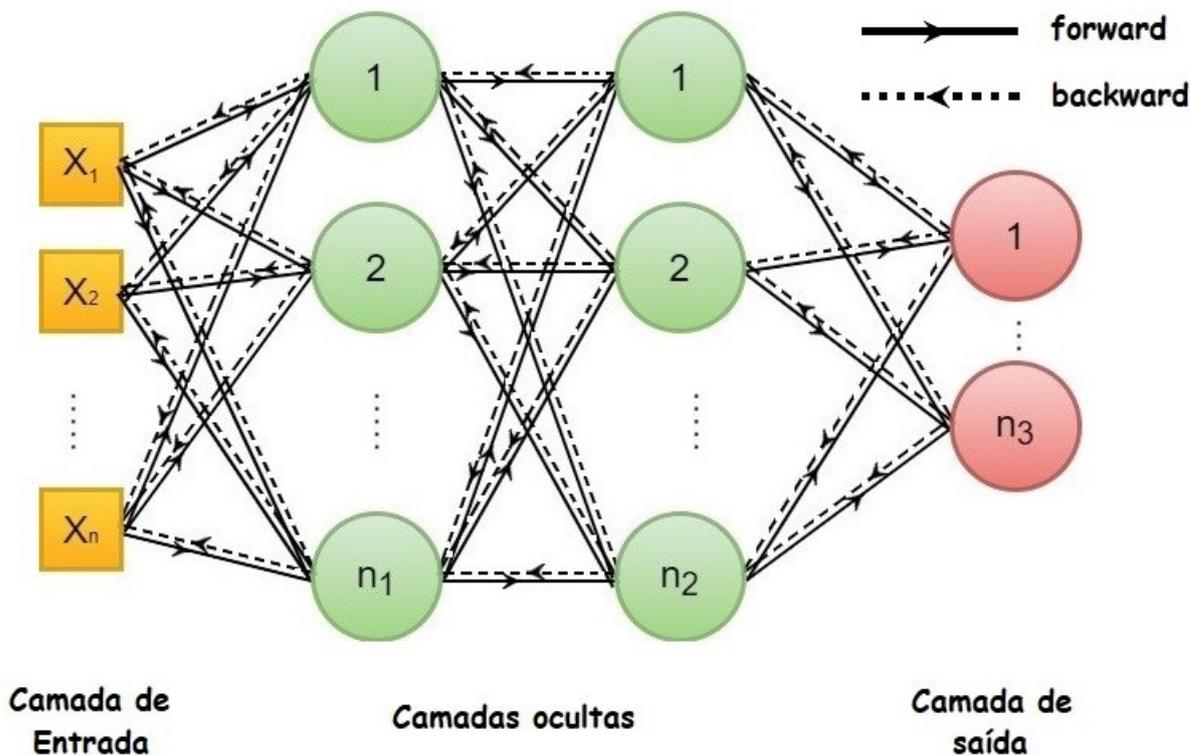
A saída de um neurônio, que é calculada quando a função soma é aplicada a uma função de ativação ($g(u)$), é

$$Y_j^L = g(I_j^L). \quad (62)$$

Ao realizar o processo descrito acima, e obter a saída Y , podemos calcular o erro das previsões. Lembrando que nas *MLP* estamos trabalhando com o treinamento supervisionado, temos em mãos as respectivas saídas desejadas (D_j). Utilizando o erro quadrático para calcular o desvio dos resultados, temos que o erro será

$$E(k) = \frac{1}{2} \sum_{j=1}^n (D_j(k) - Y_j^L(k))^2, \quad (63)$$

Figura 15 – Processo de treinamento *backpropagation* onde a linha contínua com fluxo da direita para esquerda representando a fase de *forward* onde os pesos são aleatoriamente calculados. Linha tracejada com fluxo da esquerda para direita representa a fase de *backward* onde ocorre a alteração dos pesos realizando a descida do gradiente.



Fonte: Do autor.

onde k é um determinado conjunto de amostra. Assumindo um conjunto de P amostras o erro quadrático global, ou o erro quadrático médio é dado por

$$E_M = \frac{1}{P} \sum_{k=1}^P E(k). \quad (64)$$

Para uma melhor compreensão do algoritmo de *backward*, este pode ser separado em duas partes, a primeira consiste no ajuste dos pesos que liga a última camada oculta e o *output* e a segunda consiste no ajuste entre as camadas ocultas propagando para a esquerda (Fig. 15) até chegar ao *input*.

6.5.1 Ajuste dos pesos da camada de saída

Para atualizar a matriz de pesos com a intenção de diminuir o erro global, iremos utilizar o erro médio calculado, Eq. (64), para nos mostrar em qual direção devemos

alterar os pesos. Para isso iremos calcular o gradiente dessa função

$$\nabla E = \frac{\partial E}{\partial W_{ji}^N}, \quad (65)$$

onde N representa a última matriz de pesos, a matriz que liga a última camada oculta com o *output*. Lembrando que $E = E(Y_j^N)$ é uma função composta, usando a propriedade da regra da cadeia, a Eq. (65) pode ser escrita como

$$\nabla E = \frac{\partial E}{\partial Y_j^N} \cdot \frac{\partial Y_j^N}{\partial I_j^N} \cdot \frac{\partial I_j^N}{\partial W_{ji}^N}. \quad (66)$$

Usando as definições (61), (62) e (63), derivando em ambos os lados com respeito a W_{ji}^N , I_i^N e Y_j^N , alcançamos as seguintes relações

$$\frac{\partial I_i^N}{\partial W_{ji}^N} = Y_i^{N-1} \quad (67)$$

$$\frac{\partial Y_j^N}{\partial I_i^N} = g'(I_j^N) \quad (68)$$

$$\frac{\partial E}{\partial Y_j^N} = -(D_j - Y_j^N) \quad (69)$$

Substituindo relações (67), (68) e (69) em (66), obtém-se

$$\frac{\partial E}{\partial W_{ji}^N} = -(D_j - Y_j^N) \cdot g'(I_j^N) \cdot Y_i^{N-1}. \quad (70)$$

A correção ΔW_{ji}^N aplicada a matriz de pesos é dada pela regra delta¹²

$$\Delta W_{ji}^N = -\eta \frac{\partial E}{\partial W_{ji}^N}, \quad (71)$$

sendo η é a taxa de aprendizagem do algoritmo *backpropagation*, e o sinal negativo indicando a descida do gradiente no espaço dos pesos, ou seja, buscando uma direção onde os pesos minimizam o valor do erro. Usando a Eq. (70) em (71) podemos escrever a correção da forma

¹²Também conhecida como método do gradiente descendente.

$$\Delta W_{ji}^N = \eta(D_j - Y_j^N) \cdot g'(I_j^N) \cdot Y_i^{N-1} \quad (72)$$

$$\Delta W_{ji}^N = \eta \cdot \delta_j^N \cdot Y_i^{N-1}, \quad (73)$$

onde δ_j^N é definido como o gradiente local em relação ao j-ésimo neurônio da camada de saída, dado por

$$\begin{aligned} \delta_j^N &= -\frac{\partial E}{\partial I_j^N} \\ &= -\frac{\partial E}{\partial Y_j^N} \cdot \frac{\partial Y_j^N}{\partial I_j^N} \\ &= (D_j - Y_j^N) \cdot g'(I_j^N). \end{aligned} \quad (74)$$

Como pode ser observado nas equações (73) e (74) para ajustar os pesos entre a última camada e a camada de saída é necessário ter o conhecimentos dos valores corretos D_j . Entretanto, para qualquer outra matriz de pesos que não seja a mencionada anteriormente, não temos as respostas esperadas de modo que não se pode aplicar o raciocínio realizado nesta seção.

6.5.2 Ajuste dos pesos das camadas intermediárias

Neste passo estamos no caminho da última camada oculta para sua antecessora, na Fig. 15 é o mesmo que caminhar da segunda camada oculta para a primeira. O fato de não termos um resposta, D , desejada nos leva ao problema de atribuição de crédito. Como todos os pesos da rede contribuem para o resultado, os erros eminentes das camadas ocultas percorrem toda a rede, assim o problema consiste em atribuir créditos ou culpa para o erro global da rede a cada decisão local. Ele é resolvido retropropagando-se os sinais de erro através da rede (HAYKIN, 2001).

$$\nabla E^{N-1} = \frac{\partial E}{\partial W_{ji}^{N-1}} \quad (75)$$

$$\nabla E^{N-1} = \frac{\partial E}{\partial Y_j^{N-1}} \cdot \frac{\partial Y_j^{N-1}}{\partial I_j^{N-1}} \cdot \frac{\partial I_j^{N-1}}{\partial W_{ji}^{N-1}} \quad (76)$$

analogamente as Eq. (67) e (68), temos

$$\frac{\partial I_i^{N-1}}{\partial W_{ji}^{N-1}} = Y_i^{N-2} \quad (77)$$

$$\frac{\partial Y_j^{N-1}}{\partial I_i^{N-1}} = g'(I_j^{N-1}). \quad (78)$$

Derivando a função erro, Eq. (63), em relação a Y_i^{N-2} e substituindo I_k^N pela sua definição dada por Eq. (61)

$$\frac{\partial E}{\partial Y_j^{N-1}} = \sum_{k=1}^N \frac{\partial E}{\partial I_k^N} \cdot \frac{\partial I_k^N}{\partial Y_j^{N-1}} \quad (79)$$

$$\frac{\partial E}{\partial Y_j^{N-1}} = \sum_{k=1}^N \frac{\partial E}{\partial I_k^N} \cdot \frac{\partial (\sum_{k=1}^N W_{kj}^N \times Y_j^{N-1})}{\partial Y_j^{N-1}} \quad (80)$$

$$\frac{\partial E}{\partial Y_j^{N-1}} = \sum_{k=1}^N \frac{\partial E}{\partial I_k^N} \cdot W_{kj}^N. \quad (81)$$

O fato da expressão acima possuir a matriz de pesos da última camada, os quais foram ajustados com base nos valores reais das saídas, este é a essência do *back-propagation* a utilização de pesos já ajustadas para melhoras dos pesos das camada anteriores. O gradiente local para a camada $N - 1$ é dado por:

$$\frac{\partial E}{\partial Y_j^{N-1}} = \sum_{k=1}^N \frac{\partial E}{\partial Y_j^N} \cdot \frac{\partial Y_j^N}{\partial I_j^N} \cdot W_{kj}^N \quad (82)$$

Usando as definições (68) e (69) a Eq. (82) toma a forma:

$$\frac{\partial E}{\partial Y_j^{N-1}} = \sum_{k=1}^N -(D_j - Y_j^N) \cdot g'(I_j^N) \cdot W_{kj}^N \quad (83)$$

$$\frac{\partial E}{\partial Y_j^{N-1}} = - \sum_{k=1}^N \delta_k^N \cdot W_{kj}^N. \quad (84)$$

Finalmente com as equações em mãos podemos calcular o gradiente do erro para a camada $N - 1$, que fica

$$\frac{\partial E}{\partial W_{ji}^{N-1}} = - \left(\sum_{k=1}^N \delta_k^N \cdot W_{kj}^N \right) \cdot g'(I_j^{N-1}) \cdot Y_i^{N-2} \quad (85)$$

e a correção nos pesos da camada $N - 1$ se dá por:

$$\Delta W_{ji}^{N-1} = -\eta \frac{\partial E}{\partial W_{ji}^{N-1}} \quad (86)$$

$$\Delta W_{ji}^{N-1} = \eta \cdot \delta_j^{N-1} \cdot Y_i^{N-2} \quad (87)$$

onde δ_j^{N-1} é o gradiente local em relação ao j -ésimo neurônio da camada $N - 1$ dado por:

$$\delta_j^{N-1} = - \left(\sum_{k=1}^N \delta_k^N \cdot W_{kj}^N \right) \cdot g'(I_j^{N-1}) \quad (88)$$

para o ajuste dos pesos das camadas anteriores se realiza os processos descritos acima, chegando aos pesos da camada de entrada onde se encerra o primeiro ciclo deste processo, que é repetido várias vezes para um ajuste fino dos pesos.

7 METODOLOGIA

Apoiado nas aproximações mostrados anteriormente, nosso objetivo foi verificar se as redes neurais artificiais são capazes de, a partir de um conjunto de dados (*dataset*) obter as energias com o efeito da não-parabolicidade da banda de condução para heteroestruturas semicondutoras com perfil de poço quântico.

7.1 SOLUÇÃO MATRICIAL DA EQUAÇÃO DE SCHRÖDINGER

Como mencionado anteriormente para se realizar o treinamento de uma rede neural é necessário possuir os parâmetros de saída, que no nosso caso é a energia não-parabólica. A obtenção destas são dadas pelo método iterativo mostrado na Fig. 9. Durante este processo é necessário resolver a equação de Schrödinger, lembrando-a,

$$-\frac{\hbar^2}{2} \frac{d}{dz} \left[\frac{1}{m^*(E)} \frac{d\psi_c}{dz} \right] + V(z)\psi_c = (E - E_c)\psi_c \quad (89)$$

que aplicando as propriedades do cálculo diferencial e assumindo que a energia do fundo da banda de condução seja zero, $E_c = 0$, pode ser reescrita como

$$-\frac{\hbar^2}{2} \left[\frac{1}{m^*(E)} \frac{d^2\psi_c}{dz^2} + \left(\frac{d}{dz} \frac{1}{m^*(E)} \right) \frac{d}{dz} \right] + V(z)\psi_c = E\psi_c \quad (90)$$

Uma forma de resolver a equação acima é de maneira matricial como demonstra Machado (2019). Numericamente a variável posição é discreta, de modo que $x \rightarrow z_i = i \times \Delta z$ com $i = \dots, -2, -1, 0, 1, 2, \dots$ e Δz é a distância entre os pontos da grade, assim uma função $f(z)$ se torna $f(z_i)$ ou simplesmente f_i . A derivada de f_i é calculada aplicando uma matriz de diferenciação (D), de modo que $Df_i = w_i$ onde w_i é a derivada de f no ponto i

Foi utilizado para a solução da Eq. (90) as matrizes de Chebyshev que são construídas a partir da interpolação da função em termos dos polinômios de Lagrange, empregados em uma grade não uniforme no intervalo $[-1, 1]$ dados por (TAHER; MALEK; MOMENI-MASULEH, 2013):

$$z_k = \cos \frac{k\pi}{N}, \quad k = 0, 1, \dots, N \quad (91)$$

Assim a matriz de diferenciação usando os pontos de Chebyshev é dada por

$$\bar{D}_{ij} = \begin{cases} \frac{2N^2+1}{6} & i = j = 0 \\ -\frac{2N^2+1}{6} & i = j = N \\ -\frac{x_j}{2(1-x_j^2)} & i = j \neq 0, N \\ \frac{\gamma_i (-1)^{(i+j)}}{\gamma_j (x_i - x_j)} & i \neq j \end{cases}$$

onde $\gamma_i = 1$, exceto para $\gamma_0 = \gamma_N = 2$. Assim a Eq. (90) é reescrita como

$$(\bar{D}^2 + M' + \mathbf{IV})|\psi\rangle = E|\psi\rangle \quad (92)$$

onde

$$M' = \left(\frac{d}{dz} \frac{1}{m^*(E)} \right) \frac{d}{dz}, \quad (93)$$

em que os autovalores E e os autovetores $|\psi\rangle$ são obtidos diagonalizando o Hamiltoniano.

7.2 REDES NEURAS ARTIFICIAIS

Na construção das redes neurais artificiais alguns cuidados devem ser tomados, um deles é realizar a normalização dos dados. Esta etapa é muito importante, pois se o conjunto de teste possuir dados fora do limite dos dados utilizados para o treinamento, a rede pode não retornar valores dentro do esperado. A normalização pode ser feita utilizando o desvio padrão e a média dos dados conforme a relação

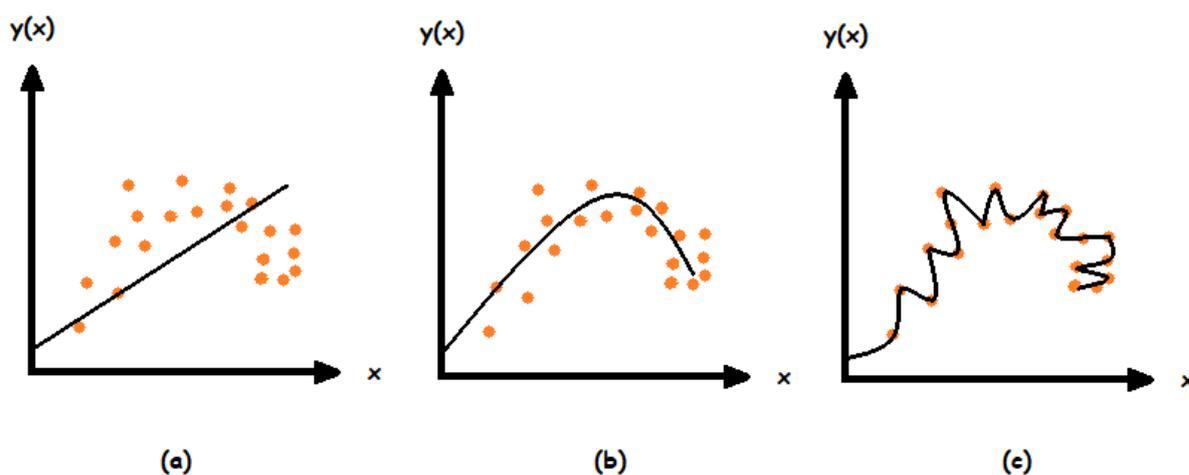
$$x_{norm} = \frac{x - x_{med}}{x_{des}}, \quad (94)$$

onde x_{med} é a média dos valores de entrada e x_{des} é o desvio padrão. Esse tipo de normalização garante uma média aproximadamente igual a zero e um desvio padrão unitário, facilitando o uso de funções de ativação binárias como a função sigmoide, por exemplo.

Outro cuidado que se deve ter é com o aumento excessivo das camadas e de neurônios da rede neural. Nesse caso, a generalização pode se tornar inapropriada. Pode ocorrer uma memorização excessiva dos parâmetros de entrada, problema

conhecido como *overfitting* ou sobreajuste (Fig. 16 (c)). Por outro lado, caso a rede possua uma quantidade pequena de neurônios pode ocorrer o processo de *underfitting* ou subajuste (Fig. 16 (a)), no qual a rede não consegue nenhuma generalização, ocorrendo um erro alto, tanto na fase de treinamento quanto na fase de teste.

Figura 16 – Exemplos de curvas de aprendizagem de uma rede neural artificial onde ocorreu (a) *Underfitting* onde a rede não encontra uma generalização para os dados. (b) Boa aprendizagem onde a rede não encontra um padrão para os dados. (c) *Overfitting* quando a rede "decora" certo conjunto de dados mas não encontra uma generalização para estes.



Fonte: Do autor.

Uma forma de driblar o *overfitting* é ter um conjunto de dados grande ou utilizando técnicas de regularização (GÉRON, 2019), como Regularização L1 e L2, *Dropout*, *Data Augmentation*, Regularização Max-Norm e *Early Stopping*. Na técnica *Early Stopping*, durante todo o processo de aprendizagem a rede é constantemente checada pela aplicação de um subconjunto de teste, sendo finalizado quando começa a haver elevação do erro quadrático médio entre épocas sucessivas.

Uma técnica muito utilizada é o *Dropout*, em cada etapa do treinamento, cada neurônio tem uma probabilidade de ser temporariamente "descartado", ou seja, durante esta etapa do treinamento ele será totalmente ignorado, mas poderá estar ativo na próxima.

Em relação a quantidade de camadas e de neurônios em cada camada, o uso de uma única camada escondida é suficiente, já que essa estrutura é capaz de aproximar qualquer equação não-linear. Duas camadas escondidas já são capazes de representar qualquer relação entre os dados, mesmo aquelas que não podem ser representadas por

equações. Mais do que duas camadas escondidas só são necessárias em problemas ainda mais complexos (CECCON, 2020).

Quanto ao número de neurônios por camada Heaton (2008) destaca que não há fórmula exata, sendo a melhor maneira a tentativa e erro, contudo propõe três passos que podem ser o ponto de partida inicial que são (HEATON, 2008):

- a) número de neurônios ocultos deve estar entre o tamanho da camada de entrada e o tamanho da camada de saída. ($n_{inp} \leq n_{oc} \leq n_{out}$)
- b) O número de neurônios ocultos deve ser $2/3$ do tamanho da camada de entrada, mais o tamanho da camada de saída. ($n_{oc} = 2/3(n_{inp} + n_{out})$)
- c) O número de neurônios ocultos deve ser menor que o dobro do tamanho da camada de entrada. ($n_{oc} < 2n_{in}$)

Baseados nesses passos, foram construídas redes neurais de diferentes arquiteturas e testadas utilizando um *dataset* com parâmetros de diversas heteroestruturas, os quais são mostrados a seguir.

8 RESULTADOS

Como mencionado no estudo das heteroestruturas, são necessárias diversas aproximações que foram apresentadas nas seções anteriores. Sendo assim o objetivo central desta dissertação foi analisar se o uso das redes neurais artificiais são capazes de, após um treinamento supervisionado, generalizar os dados de heteroestruturas semicondutoras e realizar previsões dos níveis eletrônicos que surgem na banda de condução dessas heteroestruturas. Além disso, foram determinados quais os efeitos sobre estes estados ao levar em conta o efeito de bandas não-parabólicas para o estado fundamental do sistema.

Entretanto para que aconteça o treinamento supervisionado, via *backpropagation*, é necessário possuir as saídas para um determinado conjunto de dados. Para o problema aqui proposto, a obtenção destas respostas se faz resolvendo a equação de Schrödinger com a massa dependente da posição (Eq. 90), que foi solucionada com a utilização das matrizes de Chebyshev e o método iterativo conforme Machado (2019), baseado no modelo de bandas desacopladas de Nelson, Miller e Kleinman (1987).

Inicialmente a Eq. (90) foi solucionada para o sistema gerado pela heteroestrutura de arseneto de gálio-alumínio ($\text{Al}_x\text{Ga}_{1-x}\text{As}$) e arseneto de gálio (GaAs). Para solução da Eq. (90) foram necessários dois parâmetros iniciais que são o *band offset*, que define a profundidade do poço, e a massa efetiva, que são dados por (NELSON; MILLER; KLEINMAN, 1987),

$$V = 0,6 \times (1,425x - 0,9x^2 + 1,1x^3). \quad (95)$$

$$m = 0,0665 \times m_0, \quad (96)$$

Uma vez que a profundidade do poço depende da concentração, está foi escolhida como sendo 0.37. Ao comparar os resultados obtidos com os de Nelson, Miller e Kleinman (1987) para o estado fundamental ($n = 1$) podemos dizer que o método utilizado retornou valores aceitáveis, como mostra a Tabela 1 para as energias parabólicas.

Ao analisar as energias não-parabólicas se observou a mesma tendência, do trabalho de referência, de que a energia do estado fundamental sofre um aumento em relação àquela calculada para sistemas parabólicos, como mostra a Tabela 2.

A diferença existente entre os resultados apresentados na Tab. 1, ao compararmos

Tabela 1 – Comparação entre as energias parabólicas dadas por Nelson, Miller e Kleinman (1987) e usando as matrizes de Chebyshev para diversas larguras de poços quânticos.

Lqw (Å)	Nelson (meV)	Chebyshev (meV)
5	265,25	266,70
20	179,89	178,65
50	79,7	79,56
100	31,60	32,01
200	10,36	10,67

Fonte: Do autor.

Tabela 2 – Comparação entre as energias não-parabólicas dadas por Nelson, Miller e Kleinman (1987) e usando as matrizes de Chebyshev para diversas larguras de poços quânticos.

Lqw (Å)	Nelson (meV)	Chebyshev (meV)
5	265,27	266,90
20	180,90	180,51
50	81,17	81,41
100	32,30	32,80
200	10,55	10,87

Fonte: Do autor.

os métodos, pode ser associada a erros numéricos inerentes ao método de Chebyshev, método escolhido por conta da possibilidade de se resolver um grande número de estruturas de modo sequencial. Para os resultados da Tab. 2, além dos erros numéricos associados ao método, temos a massa dependente da energia que, para Nelson, Miller e Kleinman (1987) é calculada de maneira analítica, já no modelo aqui utilizado é calculada de maneira iterativa.

Com o modelo validado, o próximo passo foi construir um conjunto de dados suficientes para realizar o treinamento da rede.

8.1 CONSTRUÇÃO DO DATASET

Um das formas de evitar que a rede neural caia em *overfitting*, além das técnicas de validação cruzada, é possuir um *dataset* grande o suficiente. Para isso foram usadas juntamente com a heteroestrutura de $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ as estruturas InN/AlN , InN/GaN , AlN/GaN , $\text{Ga}_x\text{In}_{1-x}\text{As}/\text{Al}_{0,41}\text{Ga}_{0,59}\text{As}$ e $\text{Ga}_{0,47}\text{In}_{0,53}\text{As}/\text{InP}$ e, para cada estrutura, diversas configurações foram obtidas variando parâmetros como largura do poço

e concentração, que está ligada diretamente com o *band offset* de algumas delas.

Para as heteroestruturas InN/AlN, InN/GaN, AlN/GaN e $\text{Ga}_{0,47}\text{In}_{0,53}\text{As}/\text{InP}$ o *bandoffset* é fixo como mostra a Tab. 3, de modo que o único parâmetro variável que altera a configuração da estrutura é a largura do poço. Assim foram modeladas estruturas que tinham uma largura de poço de 5 até 200 Å com um passo de 1 Å.

Tabela 3 – Heteroestruturas utilizadas para a construção *dataset* e seus respectivos *band offset*

Heteroestrutura	Band offset (eV)
$\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$	$0,6(1,425x - 0,9x^2 + 1,1x^3)$
$\text{Ga}_x\text{In}_{1-x}\text{As}/\text{Al}_{0,41}\text{Ga}_{0,59}\text{As}$	$0,31 + 1,5(1-x) - 0,4(1-x)^2$
AlN/GaN	2
InN/Ga	0,45
InN/AlN	2,7
$\text{Ga}_{0,47}\text{In}_{0,53}\text{As}/\text{InP}$	0,22

Fonte: Do autor.

Entretanto para $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ e $\text{Ga}_x\text{In}_{1-x}\text{As}/\text{Al}_{0,41}\text{Ga}_{0,59}\text{As}$, além de variar a largura do poço, também foi possível alterar o *bandoffset*, já que este depende da concentração do material (Tab. 3). Assim, para cada poço criado para $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ a concentração foi de 0,1 até 0,45 variando de 0,01. Para $\text{Ga}_x\text{In}_{1-x}\text{As}/\text{Al}_{0,41}\text{Ga}_{0,59}\text{As}$ cada poço teve a concentração de 0 a 0,45, com um passo de 0,05. Com estas pequenas variações parâmetros foram geradas 8.591 estruturas, as quais foram usados no treinamento da RNA.

Para todas as estruturas descritas anteriormente foram obtidas as energias levando em consideração os efeitos da não-parabolicidade e também as energias sem considerar tais efeitos. Com as energias, que foram utilizadas como resposta para a rede, o passo seguinte foi determinar quais seriam os parâmetros de entrada. Foram adicionados os parâmetros estruturais tanto do material da barreira¹³ quanto do material do poço¹⁴, que são as massas efetivas (M_w e M_b), os *band gaps* (E_{g_w} e E_{g_b}), parâmetros de rede (A_w e A_b), constante dielétrica estática (E_{w_s} e E_{b_s}) e para altas frequências ($E_{w_{hf}}$ e $E_{b_{hf}}$). Adicionalmente foram consideradas também as características da heteroestrutura que são a concentração (con), largura do poço quântico (L_{qw}), o *Band offset* (E_{qw}), os cinco primeiros níveis de energia parabólica (E_{i_SNP})¹⁶ e os

¹³representando pelo índice b

¹⁴representando pelo índice w

¹⁶O índice i representa o nível energético calculado sem a aproximação de banda não-parabólica.

cinco primeiros níveis de energia de um poço quântico infinito (E_{inf})¹⁵ parâmetros estes obtido em (LEVINSHTEIN, 1997; LEVINSHTEIN et al., 1999). Por fim foram adicionadas ao *dataset* as energias com não-parabolicidade (E_{CNP}), usando estes parâmetros para as 8.591 heteroestruturas o *dataset* ficou com 223.366 dados.

8.2 APLICAÇÃO DA REDES NEURAIS ARTIFICIAIS

Com o *dataset* construído a próxima etapa foi encontrar qual a topologia que melhor generalizaria os dados. Usando os modelos propostos por Heaton (2008), foram construídas redes neurais com uma, duas e três camadas ocultas. Relembrando, no primeiro modelo é proposto que o número de neurônios ocultos deve estar entre o tamanho da camada de entrada e o tamanho da camada de saída. O segundo modelo diz que o número de neurônios ocultos deve ser 2/3 do tamanho da camada de entrada mais o tamanho da camada de saída e o último modelo determina que o número de neurônios ocultos deve ser menor que o dobro do tamanho da camada de entrada. Para realizar todo o processo de construção da redes, aplicação do algoritmo *backpropagation*, validação dos resultados foram utilizadas a biblioteca de código aberto TensorFlow e a biblioteca keras (BRAIN, 2015; CHOLLET, 2015).

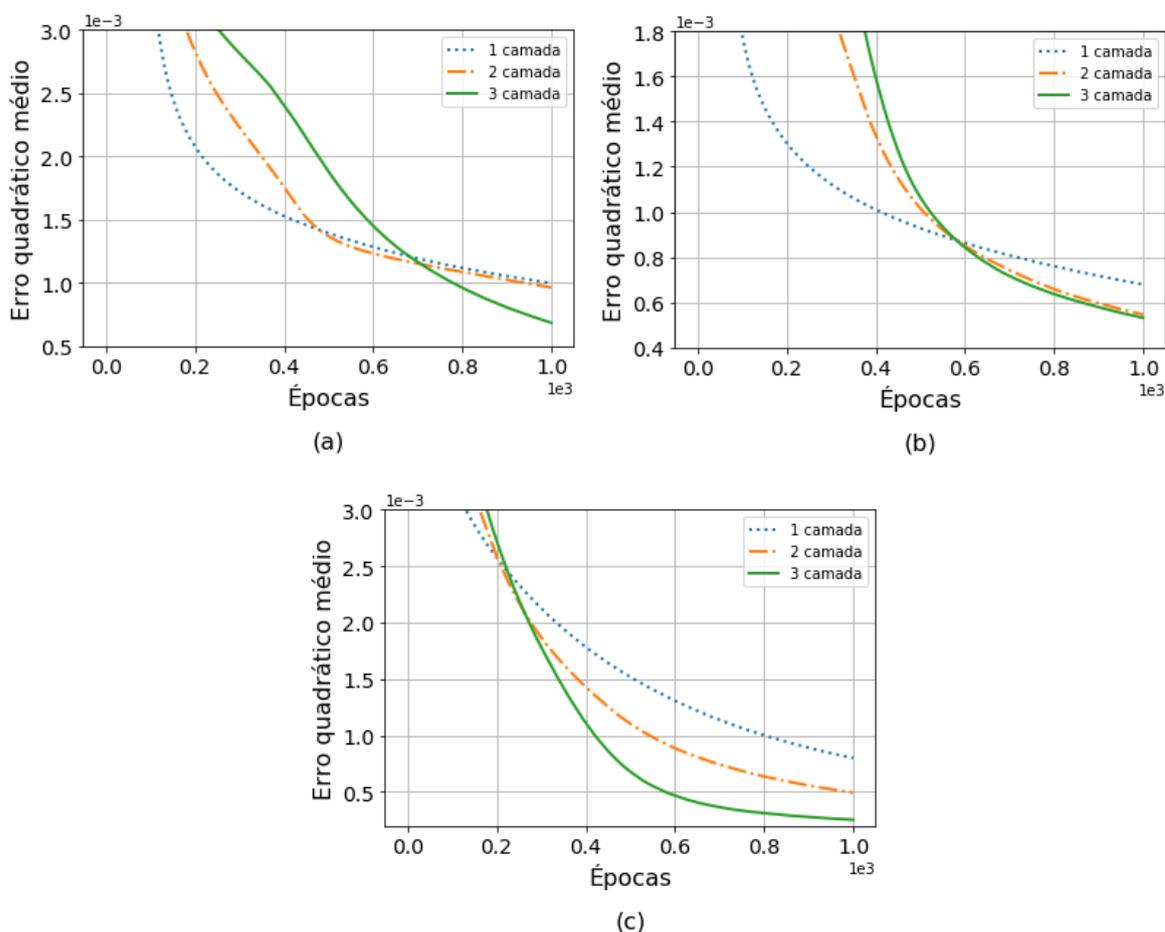
A rede ficou limitada a 1.000 épocas inicialmente, para o treinamento utilizou-se 80% dos dados e o restante foi utilizado para fazer previsões com o modelo treinado. Durante o treinamento em cada época 20% dos 80% dos dados foram utilizados para validação. A Fig. 17, mostra as curvas do erro quadrático médio (*MSE*, do inglês *mean square error*) durante todo o treinamento. Foram utilizados todos os parâmetros mencionados na sec. 8.1, totalizando 25 parâmetros de entrada.

Na Fig. 17(a) foi apresentado o resultado do treinamento usando o primeiro método de Heaton (2008), para três redes com uma, duas e três camadas ocultas. A rede com apenas uma camada (linha pontilhada azul) era composta por 13 neurônios, já a rede com duas camadas (linha tracejada com pontos laranja) foi construída com 13 neurônios na primeira e 7 na segunda camada e a rede com três camadas ocultas (linha cheia verde) possuía 13 neurônios na primeira, 7 na segunda e 4 na terceira camada. Na Fig. 17 (b), utilizando o segundo método, a quantidade de neurônios para

¹⁵Parâmetro que foi simulado usando a massa do elétron igual a m_w , de modo que para uma mesma largura de poço cada estrutura possui seu espectro de energia particular, sendo mais um dado característico da heteroestrutura, aumentando o *dataset*.

a rede com uma camada era 18, com duas camadas 18 e 13 e com três camadas 18, 13 e 10 neurônios. Por fim a Fig. 17 (c) possui na rede com uma camada 49 neurônios, com duas camadas 49 e 97 neurônios e com três camadas 49, 97 e 196 neurônios nas camadas ocultas.

Figura 17 – Curvas de erro quadrático médio de rede neurais com uma, duas e três camadas seguindo o método de Heaton (2008). (a) Método a): $n_{inp} \leq n_{oc} \leq n_{out}$ (b) Método b): $n_{oc} = 2/3 (n_{inp} + n_{out})$ (c) Método c): $n < 2n_{in}$.



Fonte: Do autor.

A Tab. 4 mostra um resumo de como foram construídas as redes treinadas apresentadas na Fig. 17, mostrando a quantidade de camadas, neurônios e as funções de ativação utilizadas em cada uma.

Em todos os casos, quanto maior o número de camadas ocultas ao final do treinamento, menor foi o *MSE* associado a rede e, dentro das três arquiteturas utilizadas, a que apresentou menor erro foi o modelo apresentado na Fig. 17 (c) com um *MSE* igual a $3,4811 \times 10^{-4}$ e um erro absoluto médio (*MAE*, do inglês *mean absolute error*) igual a

Tabela 4 – Topologia das redes treinadas para os três modelos de Heaton (2008)

FIGURA	MODELO	ENTRADA	CAMADA 1		CAMADA 2		CAMADA 3	
			Neur.	Ativ.	Neur.	Ativ.	Neur.	Ativ.
17 (a)	"1 camada"	25	13	Sigmoide	-	-	-	-
	"2 camadas"		13	Sigmoide	7	ReLu	-	-
	"3 camadas"		13	Sigmoide	7	ReLu	4	ReLu
17 (b)	"1 camada"	25	18	Sigmoide	-	-	-	-
	"2 camadas"		18	Sigmoide	13	ReLu	-	-
	"3 camadas"		18	Sigmoide	13	ReLu	10	ReLu
17 (c)	"1 camada"	25	49	Sigmoide	-	-	-	-
	"2 camadas"		49	Sigmoide	97	ReLu	-	-
	"3 camadas"		49	Sigmoide	97	ReLu	196	ReLu

Fonte: Do autor.

0,0077.

Para o treinamento das redes apresentadas na Fig. 17 foi utilizado o método do gradiente descendente durante o processo de *backpropagation*, contudo existem algoritmos de otimização criados para melhorar o treinamento e a velocidade de processamento (RUDER, 2016). Com a intenção de verificar qual algoritmo de otimização levaria a rede ao menor erro, foram construídas oito redes com diferentes otimizadores além da rede com gradiente descendente. O algoritmo e os *MSE* e *MAE* resultantes foram apresentados na Tab. 5. Como se pode notar, o otimizador que apresentou menor *MSE* foi o Nadam, sendo este o escolhido para o treinamento das redes neurais apresentadas adiante.

Tabela 5 – Comparação entre os valores de *MSE* e *MAE* de vários otimizadores.

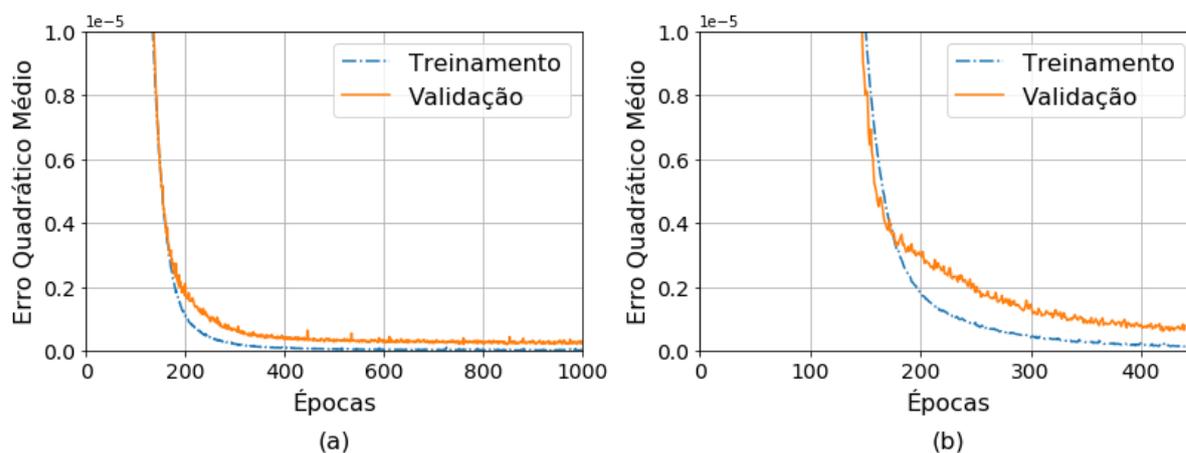
OTIMIZADOR	<i>MSE</i>	<i>MAE</i>
Adadelta	0,0079	0,0469
Adagrad	0,0061	0,0443
Adam	$1,0456 \times 10^{-4}$	0,0028
Adamax	$9,71480 \times 10^{-5}$	0,003
RMSprop	$5,3031 \times 10^{-5}$	0,0019
Ftrl	0,0092	0,0533
Nadam	$1,984 \times 10^{-5}$	0,0029
SGD	0,0049	0,0411

Fonte: Do autor.

A Fig. 18 (a) mostra a curva do *MSE* para o treinamento e para a validação, que utiliza 20 dos 80% de parâmetros separados inicialmente. Durante o treinamento,

nota-se que a curva de treinamento (linha descontinua azul) tem a mesma tendência da curva de validação (linha contínua laranja). Observa-se que por volta de 400 épocas a curva deixa de cair estabilizando em uma tendência linear, não sofrendo grandes variações. Tal fato pode colocar a rede em *overfitting* e, para evitá-lo foi aplicado a técnica de *Early Stopping* durante o treinamento. Ao se utilizar tal técnica, por meio da biblioteca Keras, a própria rede de maneira autônoma escolhe um ponto onde a métrica com a qual rede esta sendo monitorada, neste caso *MSE*, encontra o melhor ponto para interromper o treinamento. A Fig. 18 (b) mostra a curva de *MSE* com *Early Stopping*, que apresentou um *MSE* para o conjunto de dados para teste igual a $1,9535 \times 10^{-4}$.

Figura 18 – Curva de *MSE* para a rede com a melhor arquitetura utilizando o otimizador NADAM. (a) Sem aplicação da técnica *Early Stopping*. (b) Com *Early Stopping*.



Fonte: Do autor.

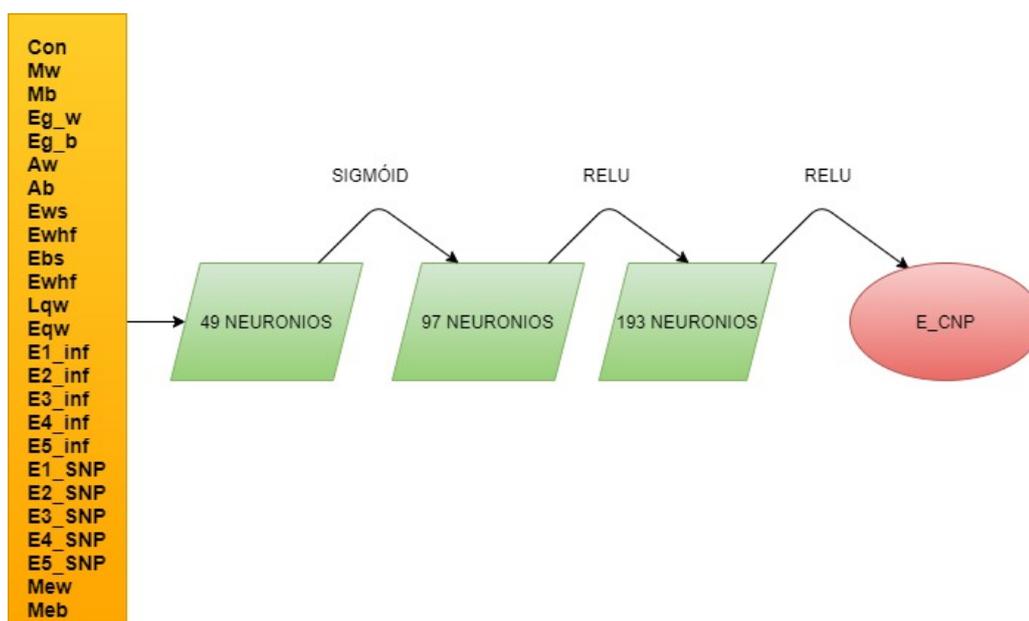
Por fim, com a topologia da rede escolhida, esta foi aplicada para previsão das energias e posteriormente foram realizadas modificações no *dataset* para testar todo potencial das redes neurais aplicadas à solução do problema estudado.

8.3 PREVISÃO DAS ENERGIAS NÃO-PARABÓLICAS

Com a melhor arquitetura encontrada verificou-se quão eficiente foi a rede em generalizar os parâmetros de entrada e encontrar as saídas desejadas, neste caso, a energia com efeito de banda não-parabólica. Inicialmente o *dataset* utilizado foi o mais amplo construído, contento os parâmetros estruturais e os obtidos via Eq. de Schrödinger como representado na Fig. 19.

Para este *dataset*, a rede apresentou um erro quadrático médio igual a $6,7779 \times 10^{-5}$

Figura 19 – Esquema da rede neural utilizando o mais amplo *dataset*.



Fonte: Do autor.

e um erro absoluto médio igual a 0,0038. A Tab. 6 mostra uma comparação entre os alguns valores de energia previstos pela rede à partir dos valores simulados usando o método iterativo matricial que tem como base os polinômios de Chebyshev, para os cinco primeiros autoestados.

Tabela 6 – Comparação entre as energias não-parabólicas calculadas via matriz de diferenciação e previstas pelas redes neurais artificiais para algumas larguras de poços quânticos.

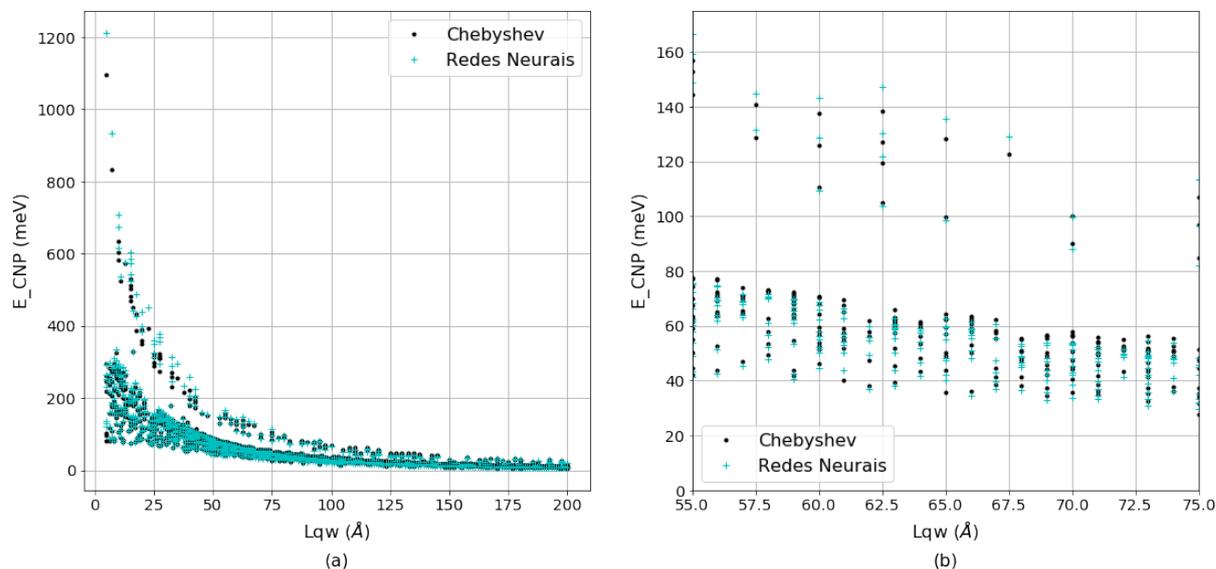
Lqw (Å)	Redes Neurais (meV)	Chebyshev (meV)
5	1.212,9	1.096,3
7	934,8	832,2
10	616,2	582,0
11	536,1	523,2
16	328,7	329,6

Fonte: Do autor.

Já a Fig. 20 (a) mostra a comparação dos 20% dos dados que foram retirados aleatoriamente do conjunto total de dados antes do treinamento para fazer as previsões com o modelo treinado e a Fig. 20 (b) apresenta um recorte dos dados centrais para melhor visualização.

Utilizando-se todos os parâmetros de entrada, após o treinamento, foi possível notar

Figura 20 – Comparação dos 20% de dados de teste previstos pela rede neural com as respectivas saídas calculadas via matriz de diferenciação. (a) Todo o conjunto de teste. (b) Recorte dos dados centrais da Fig. (a).



Fonte: Do autor.

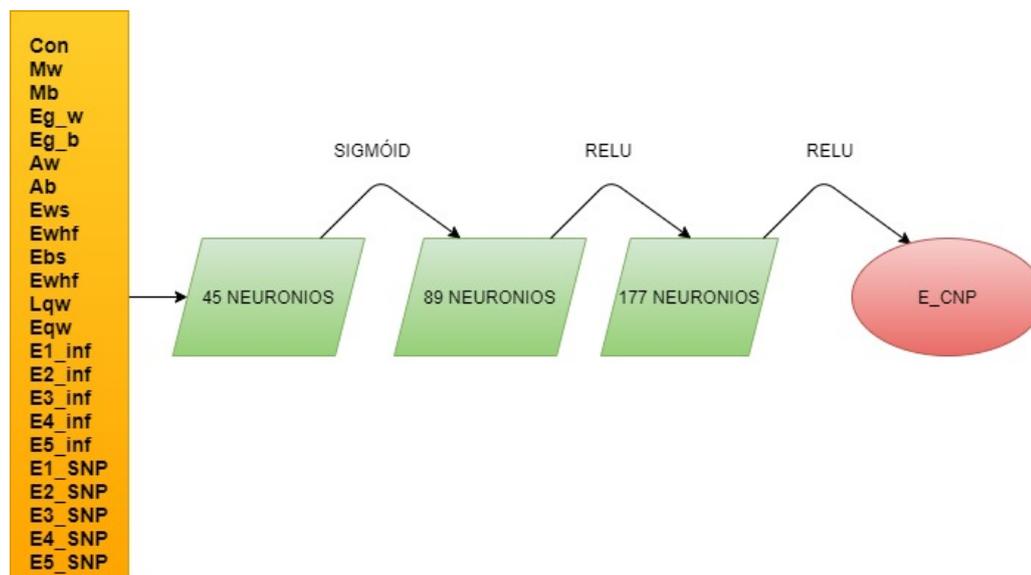
que a rede conseguiu encontrar um padrão e fazer boas aproximações dos valores de energia. Entretanto, dentre os parâmetros de *input* se levou em consideração a massa dependente da energia, que é a principal característica retornada no modelo de Machado (2019) e Nelson, Miller e Kleinman (1987). Essa massa foi obtida da solução da equação de Schrödinger e está relacionada ao parâmetro de não-parabolicidade que acopla as bandas de condução e valência. Assim, um conjunto de dados mais genérico, sem este parâmetro, seria de maior interesse.

8.3.1 *Dataset* com remoção das massas dependentes da energia

Removendo as massas dependentes da energia do grupo de parâmetros de entrada e fazendo novamente o treinamento, tem-se agora como entrada 23 parâmetros, o que implica em uma alteração na quantidade de neurônios. As camadas passaram a conter 45, 89 e 177, respectivamente, e a camada de saída ainda com apenas um neurônio prevendo a energia não-parabólica, como mostra a Fig. 21.

Com esta variação dos parâmetros de entrada, houve alterações nos erros, onde o erro quadrático médio passou a ser $1,9467 \times 10^{-5}$ e o erro absoluto médio tornou-se 0,0023, um decréscimo de 71,3% para o *MSE* em relação ao treinamento realizado

Figura 21 – Esquema da rede neural utilizado com a primeira variação do *dataset*.



Fonte: Do autor.

incluindo-se as massas dependentes da energia. Entretanto, considerando que foram retirados do *dataset* dois parâmetros, que leva a uma perda de 13.746 dados no *dataset*, tal resultado mostra que as massas não são relevantes para o treinamento, pelo contrário podem estar servindo como ruído e a rede estaria tentando aprendê-las. Algumas previsões são mostradas na Tab. 7 e a Fig. 22 mostra todos os dados utilizados no teste, em função da largura dos poços quânticos.

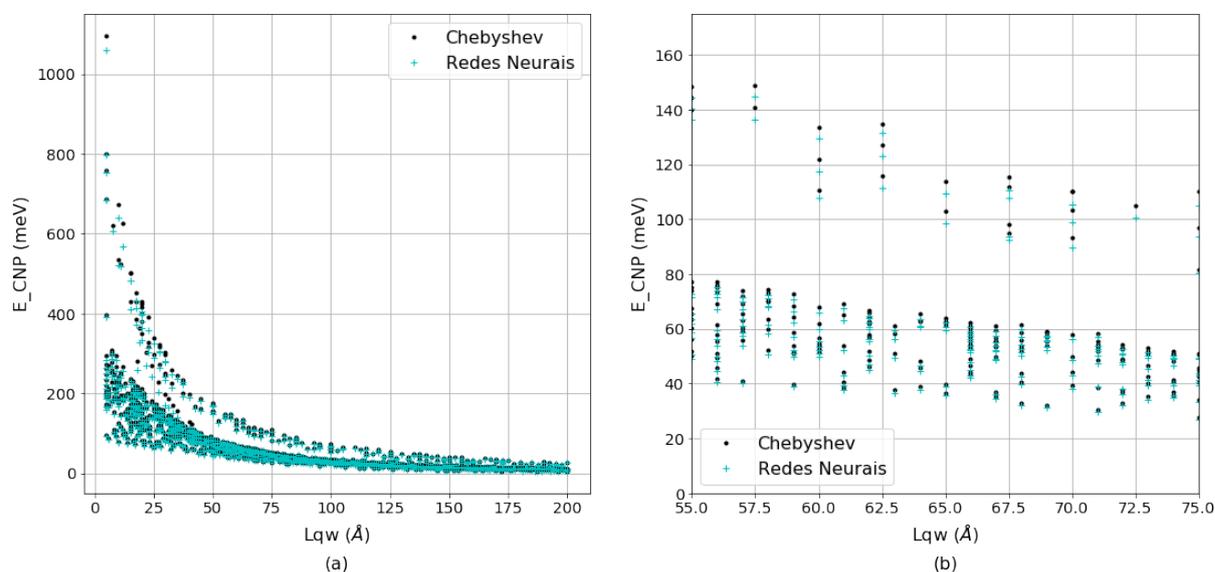
Tabela 7 – Comparação entre as energias não-parabólicas calculadas via matriz de diferenciação e previstas pelas redes neurais artificiais com diminuição do *dataset*.

Lqw (Å)	Redes Neurais (meV)	Chebyshev (meV)
5	1.060,8	1.096,3
11	519,3	523,2
18	259,6	281,2
27	143,3	155,6
33	103,4	113,4

Fonte: Do autor.

Para a construção dos *datasets* utilizados anteriormente foi necessário resolver a equação de Schrödinger para todas as estruturas analisadas. Tal solução requer um esforço computacional considerável, de modo que para a construção do *dataset*

Figura 22 – Comparação dos 20% de dados de teste previstos pela rede neural com as respectivas saídas para primeira redução do *dataset* (a) Todo o conjunto de teste. (b) Recorte dos dados centrais da Fig. (a).



Fonte: Do autor.

foram gastos cerca de seis semanas¹⁷ de processamento somente para esse passo. Se ainda levar em conta que problemas externos que podem acabar interrompendo o processo, uma escolha prudente de entradas não deveria levar em conta as massas dependentes da energia e os cinco estados de energia parabólica.

Sendo assim testou se um *dataset* que possui como dados de entrada apenas parâmetros estruturais e as energias para aproximação de poço infinito, dados que são bem conhecidos na literatura.

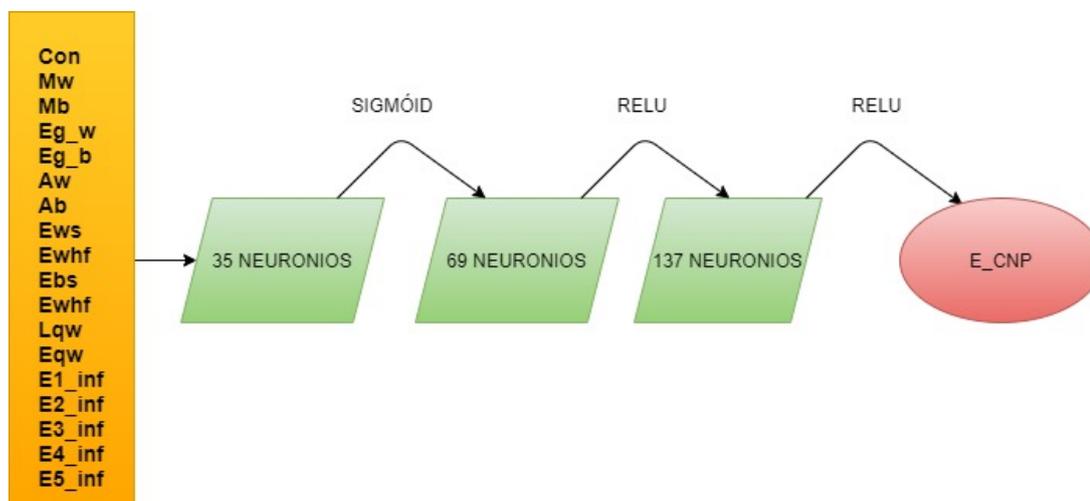
8.3.2 *Dataset* sem parâmetros obtidos via equação de Schrödinger

Com a remoção das energias parabólicas o novo conjunto de entrada continha 28 parâmetros, assim a rede utilizada para o treinamento sofreu modificações em suas camadas ocultas tendo 35, 69 e 137 (Fig. 23).

Esta rede teve como resposta um erro quadrático de $2,7813 \times 10^{-5}$ e o absoluto igual a 0,0020 que, pela mesma ideia mostrada anteriormente, já era esperado um aumento nos erros devido à diminuição do *dataset*. Contudo, o erro ainda continua menor que a do *dataset* utilizado para realizar o primeiro teste, confirmando nossa suspeita de que as massas com dependência da energia são os parâmetros menos importantes para o

¹⁷Em uma workstation com 20 cores, 68Gb ram - modelo E5-2630V4

Figura 23 – Esquema de rede neural com *dataset* sem as energias parabólicas parâmetros obtidos via Schrödinger.



Fonte: Do autor.

treinamento. Na Tab. 8 é mostrado uma parte dos dados e sua comparação com os resultados simulados por Chebyshev. A Fig. 24 mostra todos os dados.

Tabela 8 – Comparação entre as energias não-parabólicas calculadas via matriz de diferenciação e previstas pelas redes neurais artificiais com a segunda diminuição do *dataset*.

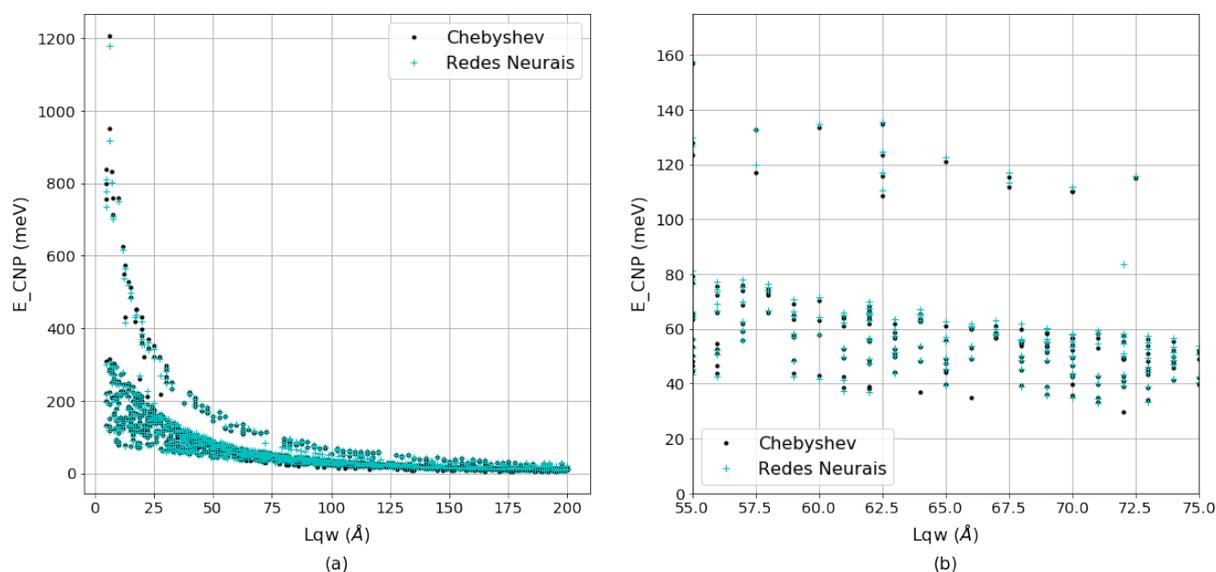
Lqw (Å)	Redes Neurais (meV)	Chebyshev (meV)
6	982,5	951,3
11	524,3	523,2
14	392,4	391,6
15	358,0	358,6
20	236,9	242,6

Fonte: Do autor.

8.3.3 *Dataset* apenas com parâmetros estruturais

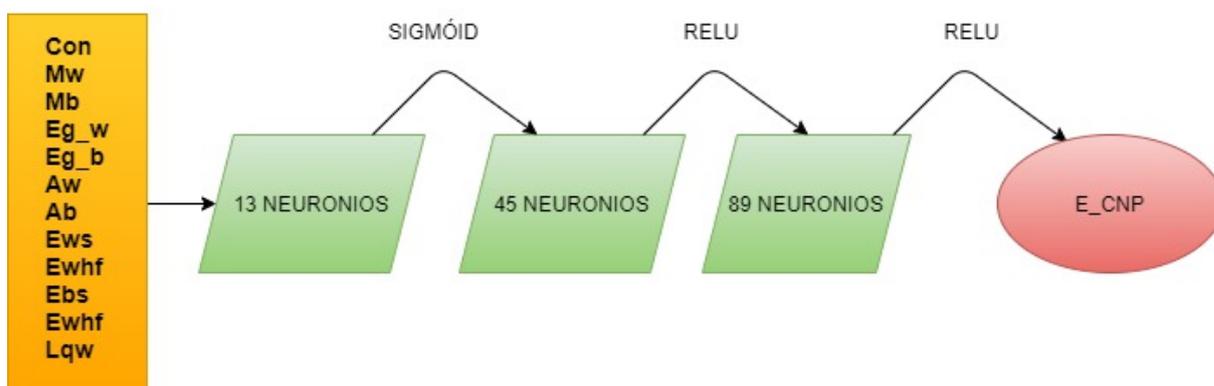
Com a intenção de simplificar mais uma vez o *dataset*, retirou-se as energias para o poço infinito restando apenas características estruturais da heteroestrutura e dos materiais que a compõe. Neste *dataset* foi retirado também o *band offset*, que é um parâmetro estrutural, porém está diretamente ligado à concentração (Tab. 3), que foi mantida. Assim temos um total de 12 parâmetros de entrada, levando a uma rede com 23, 45 e 89 neurônios em cada camada oculta, como mostra a Fig. 25.

Figura 24 – Comparação dos 20% de dados de teste previstos pela rede neural com as respectivas saídas para o *dataset* sem parâmetros obtidos via método iterativo. (a) Todo o conjunto de teste. (b) Recorte dos dados centrais da Fig. (a).



Fonte: Do autor.

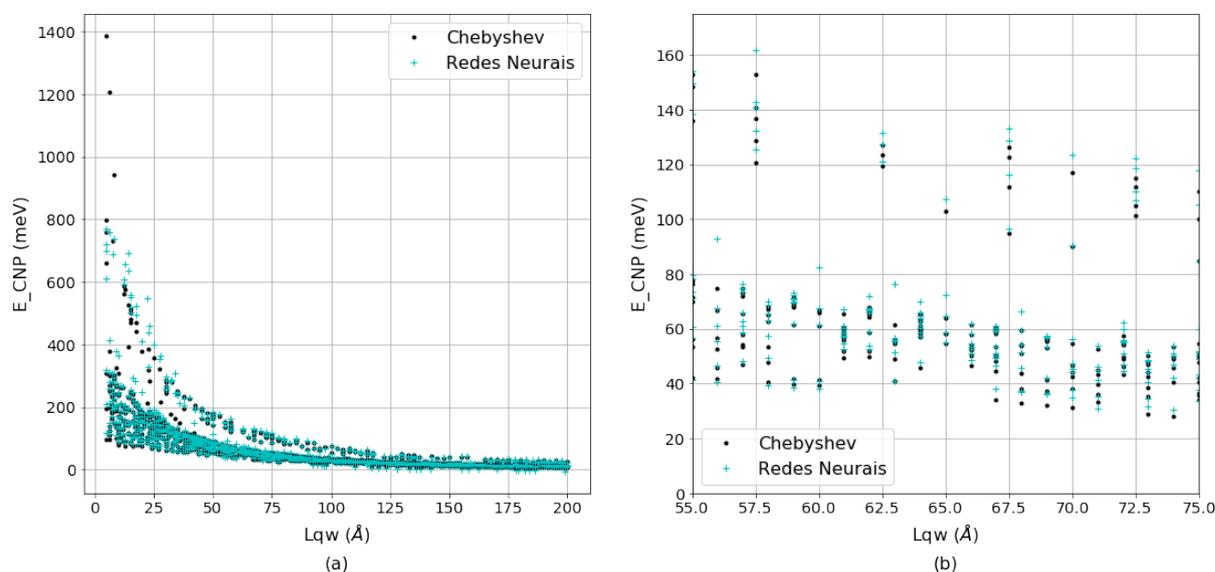
Figura 25 – Esquema da rede neural utilizado o *dataset* com apenas parâmetros estruturais.



Fonte: Do autor.

Esta rede foi a que apresentou maior MSE sendo este igual a $6,0478 \times 10^{-4}$ e 0,0020 para o MAE , todavia se levarmos em consideração a grande quantidade de dados retirados, tal valor pode ser considerado aceitável. A Fig. 26 mostra as previsões realizadas por esta rede e a Tab. 9 um pouco dos dados.

Figura 26 – Comparação das energias não-parabólicas para a rede neural com o *dataset* com apenas parâmetros estruturais. (a) Todo conjunto de teste. (b) Recorte dos dados centrais da Fig (a).



Fonte: Do autor.

Tabela 9 – Comparação entre as energias não-parabólicas usando conjunto de entrada com apenas dados estruturais.

Lqw (Å)	Redes Neurais (meV)	Chebyshev (meV)
169	14,2	14,4
110	22,1	23,0
106	20,5	21,3
174	11,9	11,9
100	50,8	56,0

Fonte: Do autor.

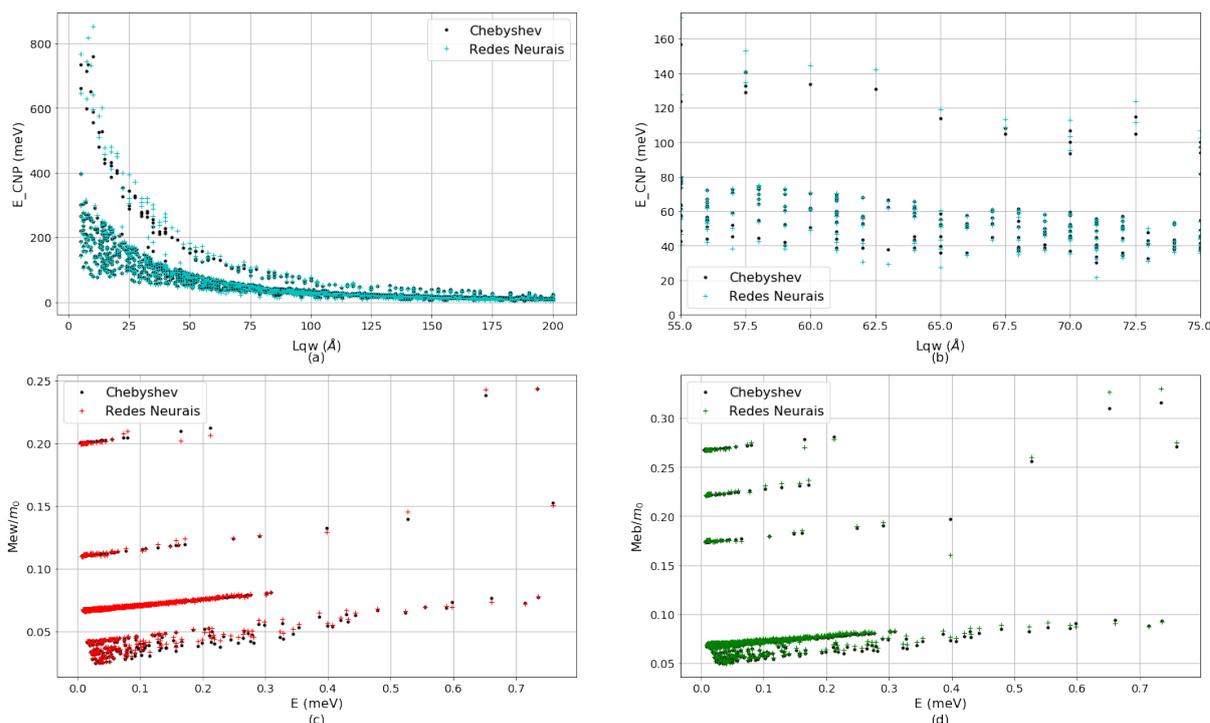
8.4 PREVISÃO DA ENERGIA NÃO-PARABÓLICA E DAS MASSAS DEPENDENTES DA ENERGIA

Uma vez que as redes neurais foram capazes de prever de maneira satisfatória a energia com não parabolicidade, levantou-se a questão se esta teria a capacidade de prever outros parâmetros. Foram escritas redes que possuem os mesmos parâmetros de entrada das Fig. 21, 23 e 25, porém tendo como saída, além das energias não-parabólicas (E_{CNP}), também a massa dependente da energia da barreira (M_{eb}) e do poço (M_{ew}).

A rede neural artificial para o primeiro *dataset* (Fig. 21) obteve um *MSE* igual a

$1,8894 \times 10^{-5}$ e $0,0010$ para os erros quadrático e absoluto respectivamente. A Fig. 27 mostra as previsões e os valores esperados tanto para a energia não-parabólica (Fig. 27(a) e (b)) quanto para as massas dependentes da energia (Fig. 27(c) e (d)).

Figura 27 – Comparação dos 20% de dados entre as previsões da rede neural com sua respectiva saída. (a) Energias não-parabólicas. (b) *Zoom* da Fig. (a). (c) Massa no poço. (d) Massa na barreira.

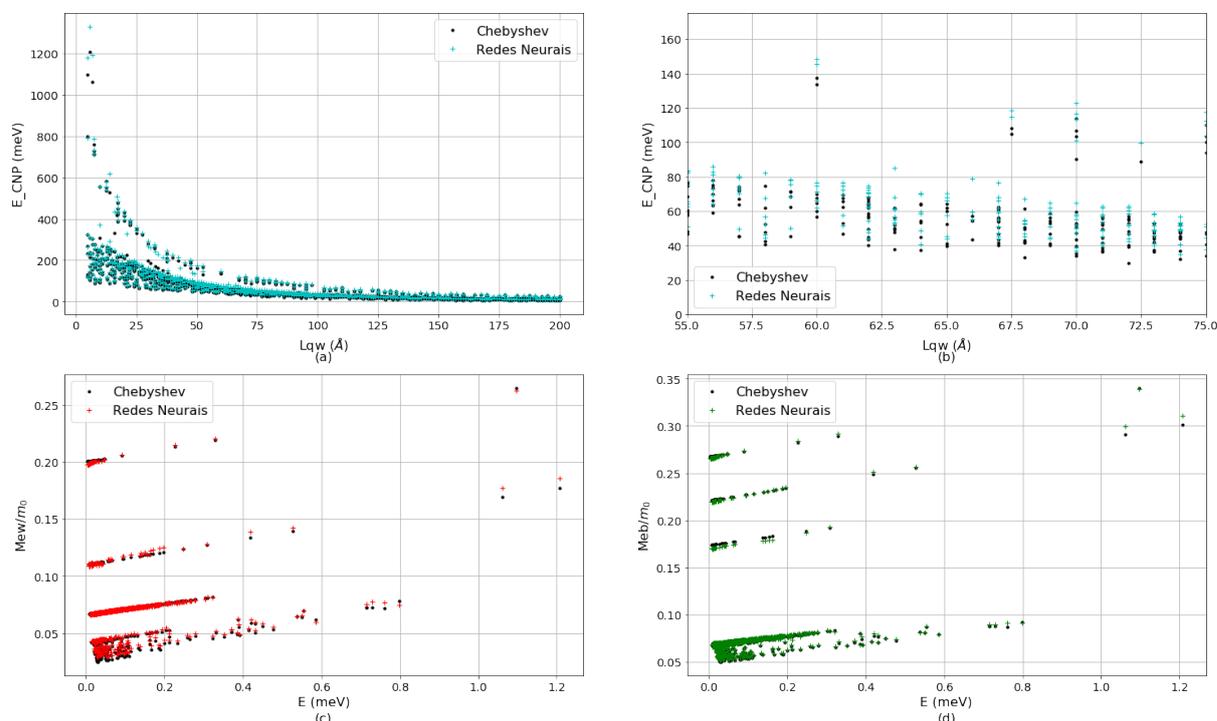


Fonte: Do autor.

Levando em conta que agora a rede está prevendo três parâmetros, pode-se dizer que os erros são aceitáveis. Pelas Fig. 27(a) e (b) é possível ver que a rede conseguiu encontrar um padrão para as energias. E para as massas ocorre o mesmo, onde pode ser notado nas Fig. 27(c) e (d) a classificação em conjuntos de dados próximos, onde cada conjunto representa uma heteroestrutura, mostrando que o modelo não está limitado a apenas uma heteroestrutura.

Como feito anteriormente, realizou-se reduções nos parâmetros de entrada. A primeira redução ocorreu retirando dos dados aqueles que necessitam da solução da Eq. Schrödinger aplicada às heteroestruturas para sua obtenção. A Fig. 28 mostra as previsões realizadas pela rede, apresentando um MSE igual a $3,3118 \times 10^{-5}$ e $0,0020$ para o MAE .

Figura 28 – Comparação entre as previsões da rede neural com sua respectiva saída para o *dataset* sem parâmetros via Schrödinger. (a) Energias não-parabólicas. (b) Zoom da Fig. (a). (c) Massa no poço. (d) Massa na barreira.



Fonte: Do autor.

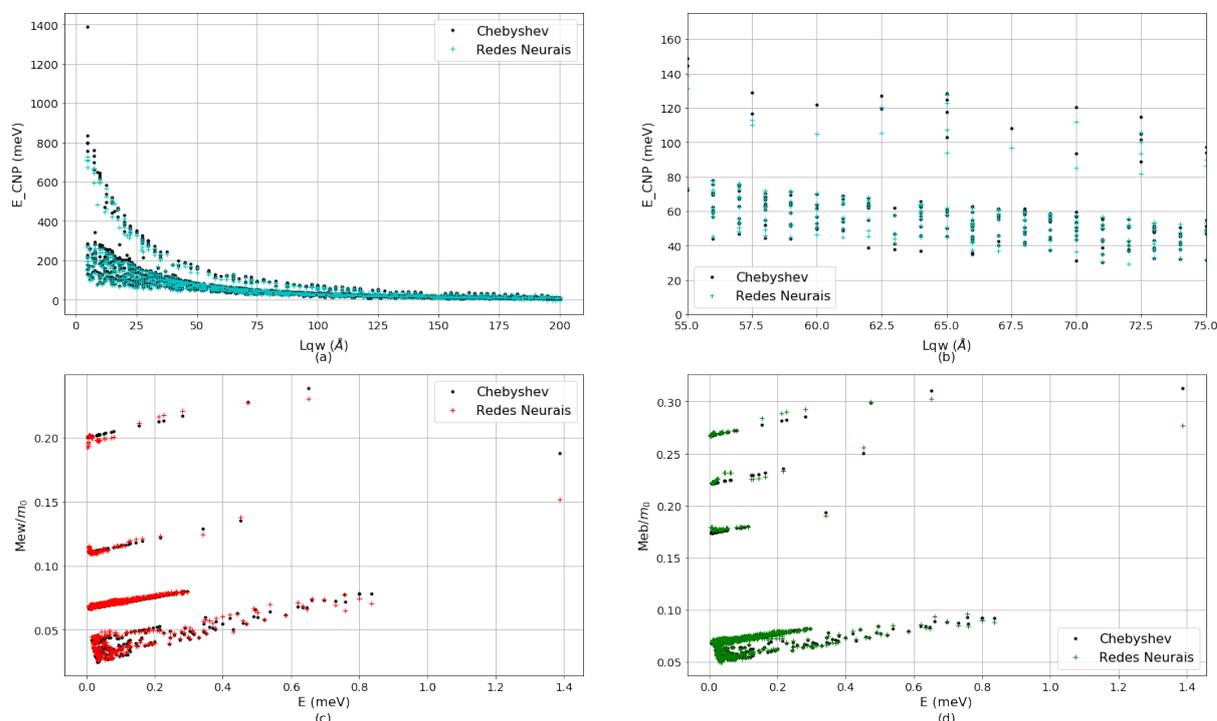
Já para a outra diminuição do conjunto de entrada, mantendo apenas parâmetros estruturais, a resposta desta nova rede foi de $7,1200 \times 10^{-5}$ para o *MSE* e de 0,0021 para o *MAE*. As previsões são mostradas na Fig. 29.

Nesse caso, onde tentamos prever além das energias também as massas, a rede neural que possui a maior quantidade de dados de entrada foi a que apresentou menor erro quadrático. O *MSE* aumentou a medida que se retirou parâmetros de entrada. Tal comportamento era esperado pois para encontrar a massa dependente da energia com o modelo iterativo partimos da energia parabólica, por isso o *dataset* que leva em conta tais energias mostrou melhor desempenho nas previsões das massas.

8.5 PREVISÃO DA ENERGIA NÃO-PARABÓLICA, ENERGIA PARABÓLICA DO ESTADO FUNDAMENTAL E DAS MASSAS DEPENDENTES DA ENERGIA

Com a intenção de testar mais a fundo a rede, realizou uma última arquitetura onde se utilizou a rede para prever as energias não-parabólicas, as massas na barreira, no poço e também a energia do nível fundamental do sistema sem não-parabolicidade.

Figura 29 – Comparação entre as previsões da rede neural com sua respectiva saída para o *dataset* apenas com dados estruturais. (a) Energias não-parabólicas. (b) Zoom da Fig. (a). (c) Massa no poço. (d) Massa na barreira



Fonte: Do autor.

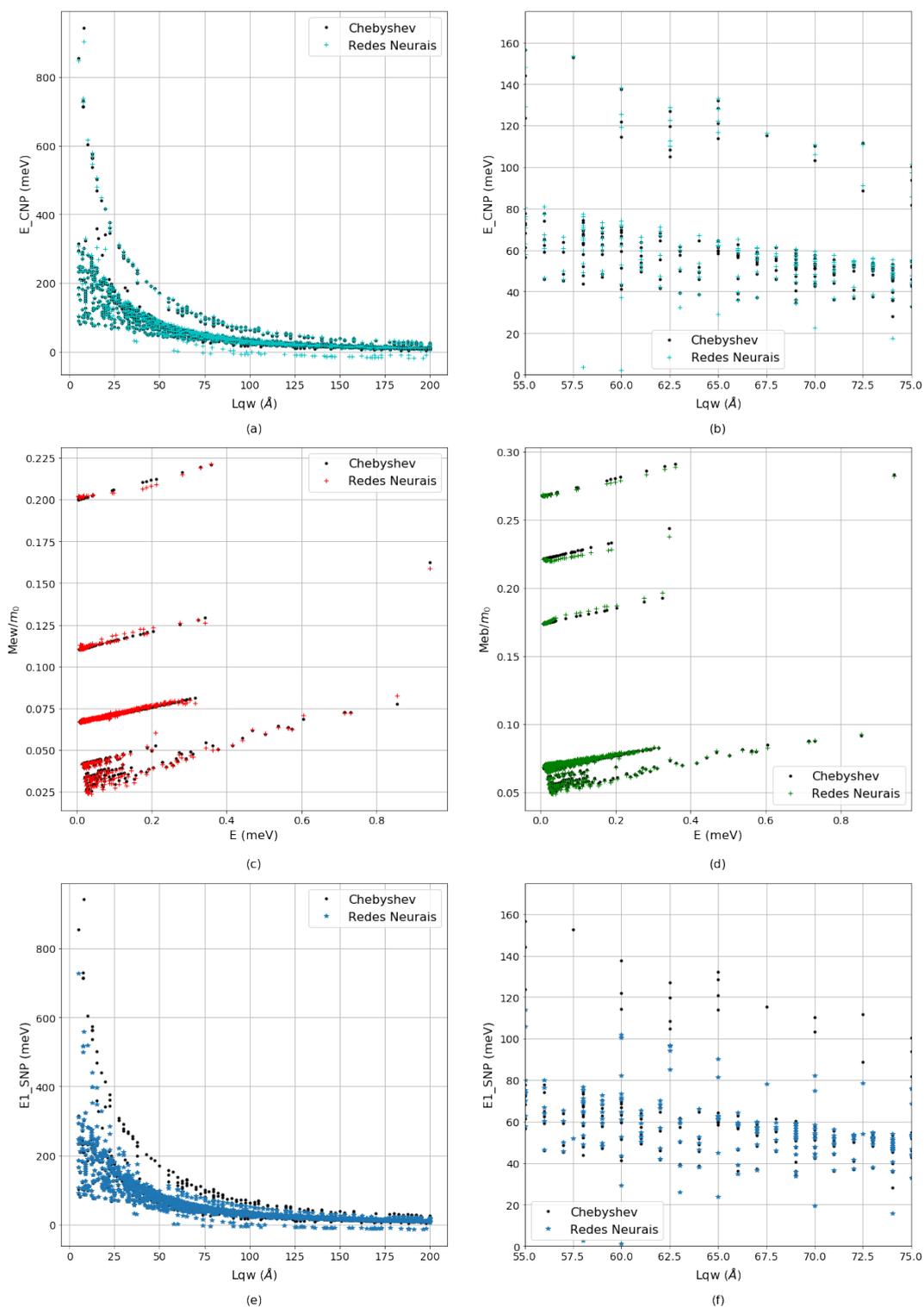
Para tanto, foram utilizados dois conjuntos de entrada, o primeiro contendo os parâmetros estruturais e as energias do poço infinito e o último contendo apenas parâmetros estruturais.

Os valores de MSE e MAE para o primeiro modelo foram iguais a $3,5182 \times 10^{-5}$ e 0,0016, respectivamente, e para o segundo modelo foram de $1,2991 \times 10^{-4}$ e 0,0024. Houve um pequeno aumento no erro quando ocorreu a diminuição do *dataset*. Porém, lembrando que estamos prevendo um grupo maior de parâmetros, a rede se comportou de uma ótima maneira, retornando bons resultados que são apresentados nas Fig. 30 e 31.

Baseados nos dados apresentados, podemos concluir que as redes neurais mostraram um excelente desempenho na previsão das energia de heteroestruturas utilizando como parâmetros de entrada apenas dados estruturais. Tal método se mostrou também interessante na previsão de outras características além da energia como a previsão das massas dependentes da energia.

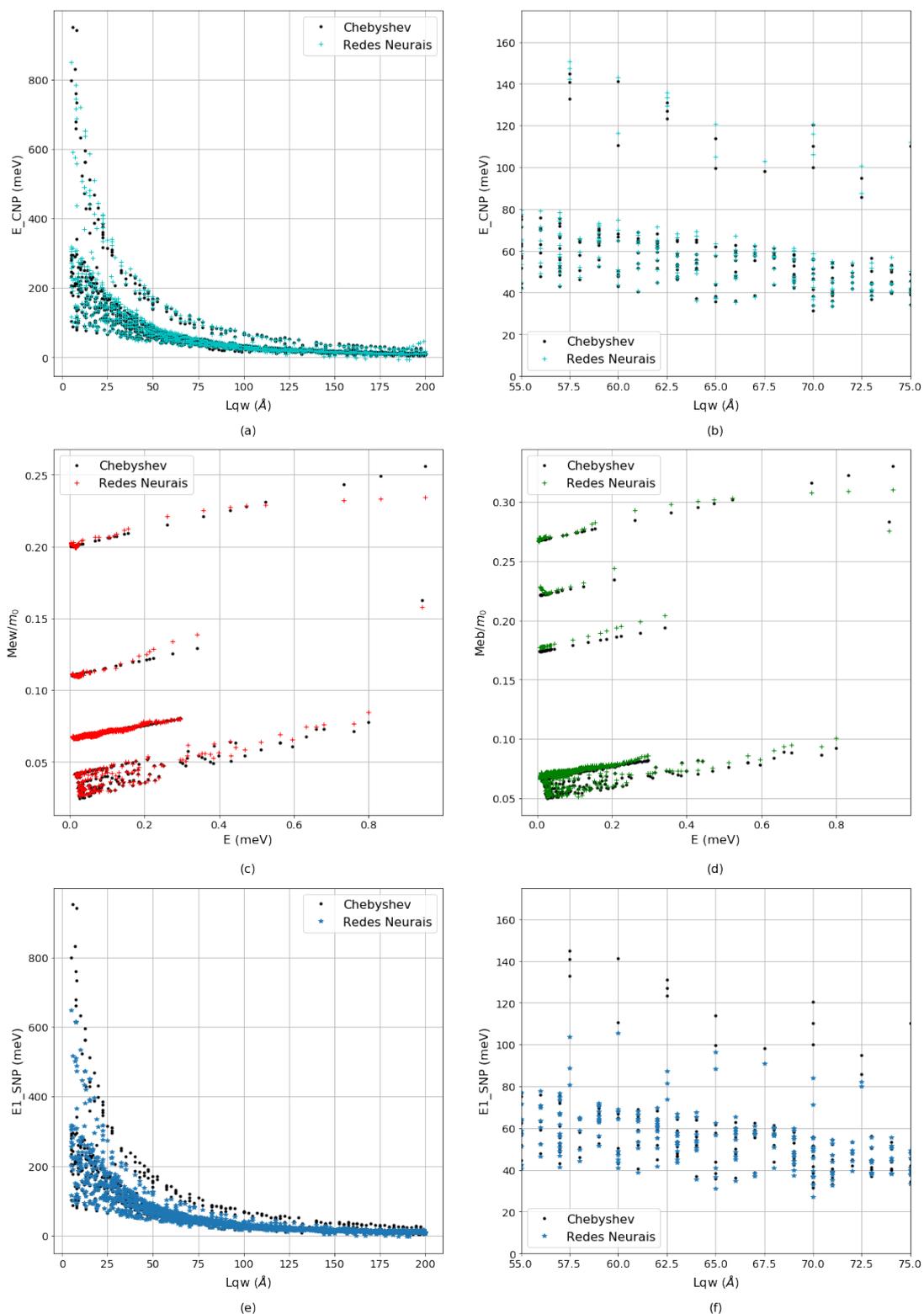
Considerando que a obtenção destes parâmetros via equação de Schrödinger

Figura 30 – Comparação entre as previsões da rede neural com sua respectiva saída para o *dataset* com dados estruturais e energias de poço infinito. (a) Energias não-parabólicas. (b) *Zoom* da Fig. (a). (c) Massa no poço. (d) Massa na barreira. (e) Energia parabólica. (f) *Zoom* da Fig. (e)



Fonte: Do autor.

Figura 31 – Comparação entre as previsões da rede neural com sua respectiva saída para o *dataset* com dados estruturais. (a) Energias não-parabólicas. (b) *Zoom* da Fig. (a). (c) Massa no poço. (d) *Zoom* da Fig. (c). (e) Energia parabólica. (f) *Zoom* da Fig. (e)



Fonte: Do autor.

requer um esforço computacional considerável, no caso deste trabalho cerca de seis semanas as redes neurais podem se tornar uma alternativa mais que interessante, uma vez que a obtenção destes dados para diversas heteroestruturas leva minutos¹⁵. É possível sempre que se obtiver parâmetros novos adicioná-los ao *dataset* refinando ainda mais os dados e aumentando seu alcance para outras heteroestruturas.

¹⁵processador Core i7 8550U (4 núcleos físicos + 8 Threads) e placa dedicada AMD Radeon 520 com 2 GB do padrão gDDR5

9 CONCLUSÃO

Dispositivos heteroestruturados são de grande aplicabilidade, entretanto a obtenção de suas propriedades físicas é feita numericamente resolvendo a equação de Schrödinger. Devido ao crescimento preciso das heteroestruturas a diferença entre os *gaps* de energia dos materiais que compõe a estrutura, gera na banda de condução, um perfil de potencial igual a um poço quântico. Entretanto o problema não é apenas resolver um poço quântico finito. Uma vez que o acoplamento entre as bandas de valência e condução leva a uma alteração da energia, o problema requer a resolução do modelo de bandas não-parabólicas.

Para contornar este problema de bandas não-parabólicas foi utilizado o método iterativo para a solução da equação de Schrödinger conforme Machado (2019), que tem como base o modelo empírico de Nelson, Miller e Kleinman (1987). O modelo desacopla as bandas de valência e condução por meio de uma renormalização do *gap* e a massa do elétron é substituída por uma massa efetiva, que depende da energia da banda.

Contudo devido à complexidade da Eq. de Schrödinger para um cristal, diversas aproximações além da proposta por Nelson, Miller e Kleinman (1987) são necessárias para a solução do problema, que leva a um gasto computacional e tempo necessário relativamente altos.

Neste contexto, este trabalho teve como foco apresentar as RNA como uma alternativa a esse processo. As RNA utiliza, um conjunto de dados para treinamento a partir do qual o algoritmo encontra generalizações e consegue realizar previsões eficientes com um tempo de processamento pequeno. Assim uma rede neural foi construída e treinada afim de realizar previsões para o nível fundamental da banda de condução levando em conta a não-parabolicidade da banda.

Para realizar o treinamento da rede foi necessário resolver a equação de Schrödinger para diversas heteroestruturas, processo que levou cerca de seis semanas. Com os dados obtidos foi possível construir um robusto *dataset* contendo 223.366 dados. Dados estes que são estruturais além da energia e das massas dependentes da energia dos materiais da heteroestrutura.

Com o conjunto de dados foi construída uma rede tipo *perceptron* multicamadas, contendo três camadas ocultas. O número de neurônios em cada camada seguiu a

proposta por Heaton (2008). Após realizar o treinamento da rede esta retornou um erro quadrático médio igual $3,4811 \times 10^{-4}$ onde foi utilizado o método *backpropagation* para o treinamento e a técnica de *Early Stopping* para evitar *overfitting*.

Para esta arquitetura foram realizadas diversas modificações no *dataset* a fim de testar e verificar o poder de generalização da rede. Tal processo mostrou que ao diminuir o *dataset* ocorre um aumento no erro, entretanto a rede ainda consegue generalizar bem as propriedades da estrutura. Sendo esta afirmação válida para um conjunto de dados contendo apenas parâmetros estruturais, não necessitando a resolução prévia da Eq. de Schrödinger para as heteroestruturas, propiciando uma queda no tempo de preparo do *dataset* Sendo ainda possível prever outras características além de apenas a energia.

O presente trabalho abre caminho para trabalhos futuros nos quais a simulação para a banda de valência, bem como de outras propriedades como energias de transição e de absorção podem ser realizadas. Considerando o *dataset* apenas com parâmetros estruturais, ao menos em tese, acreditamos que a rede seja capaz de generalizar propriedades dos buracos na banda de valência, por exemplo. Esperamos ser possível adicionar parâmetros diversos ao *dataset*, levando a uma melhor adequação dos dados pela rede, parâmetros que podem ser, por exemplo efeitos de campos externos.

Sendo assim, as RNA se mostraram bastante promissora no estudo de heteroestrutura semicondutoras, reduzindo o tempo de processamento de semanas para minutos. Vale ressaltar que todo o *dataset* aqui utilizado foi construído utilizando as energias via solução da equação de Schrödinger, uma vez que o *dataset* para o treinamento seja construído a partir de dados experimentais todos os erros numéricos são descartados restando apenas o erro decorrente do processamento pela rede e eventuais erros experimentais.

Por fim, acreditamos que as redes neurais munidas com um bom *dataset* são capazes de simular de maneira satisfatória propriedades óticas e transporte, e não somente de semicondutores mas de outras classes de materiais.

REFERÊNCIAS

- ACADEMY, D. S. **Deep Learning Book**. 2019. Disponível em: <http://deeplearningbook.com.br/>. Acesso em: 2 dez. 2020
- ALESSIO, R. de C. P. V. **Implantação de um espectrômetro de estados metaestáveis de alvos gasosos**. Dissertação (Mestrado) — Universidade Federal de Juiz de Fora, Juiz de Fora, MG, 2009.
- ALHARBI, F. An explicit fdm calculation of nonparabolicity effects in energy states of quantum wells. **Optical and Quantum Electronics**, v. 40, p. 551–559, 2008.
- ALPAYDIN, E. **Introduction to Machine Learning**. 1. ed. Cambridge: MIT Press, 2004.
- ANGHINONI, B. **Estudo teórico da resistividade longitudinal de um poço quântico duplo com a presença do efeito Rashba**. Dissertação (Mestrado) — Universidade de São Paulo, São Paulo, SP, 2012.
- ASHCROFT, N.; MERMIN, N.; OLIVEIRA, M. D. **Física Do Estado Sólido**. 1. ed. São Paulo: Cengage Learning, 2010.
- BARR, A.; FEIGENBAUM, E. A. **The handbook of artificial intelligence**. Los Altos, CA: Morgan Kaufmann, 1981.
- BASTARD, G. **Wave Mechanics Applied to Semiconductor Heterostructures**. Les Editions de Physique, 1988. (Monographies de physique).
- BEAR, M.; CONNORS, B.; PARADISO, M. Neurociências: **Desvendando o Sistema Nervoso**. Artmed Editora, 2002.
- BEZERRA, A. T. **Estudos de Efeitos de Spin em Diodos de Tunelamento Ressonante do Tipo-P**. Dissertação (Mestrado) — Universidade Federal de São Carlos, São Carlos, SP, 2010.
- BRAIN, G. 2015. **TensorFlow**. 2015. Disponível em: <https://www.tensorflow.org/?hl=pt-br>. Acesso em: 2 dez. 2020
- CECCON, D. **Quantas camadas escondidas e quantos neurônios incluir numa rede neural artificial?** 2020. Disponível em: <https://iaexpert.academy/2020/05/04/quantas-camadas-escondidas-e-quantos-neuronios-incluir-numa-rede-neural-artificial/>. Acesso em: 2 dez. 2020
-
- CHIQUITO, A. J. Cálculo da distribuição de cargas na banda de condução em junções semicondutoras: um exercício de física (computacional). **Rev. Bras. Ensino Fís.**, São Paulo, v. 21, n. 4, p. 514–523, dez. 1999.

CHIQUITO, A. J.; LANCIOTTI, J. F. O transistor, 50 anos. **Rev. Bras. Ensino Fís.**, São Paulo, v. 20, n. 4, p. 309–314, dez. 1998.

CUSTODIO, R. Uma forma simplificada de deduzir as equações de hartree e hartree-fock. **Quím. Nova**, scielo, São Paulo, v. 38, n. 7, p. 995 – 1001, 08 2015.

EISBERG, R.; RESNICK, R. **Física Quântica**: átomos, moléculas, sólidos, núcleos e partículas. Rio de Janeiro: ELSEVIER EDITORA, 1979.

EISENSTEIN, J. P. et al. Effect of inversion symmetry on the band structure of semiconductor heterostructures. **Phys. Rev. Lett.**, American Physical Society, v. 53, p. 2579–2582, Dec 1984.

FERNEDA, E. Redes neurais e sua aplicação em sistemas de recuperação de informação. **Ciência da Informação**, scielo, v. 35, p. 25 – 30, 04 2006.

GEORG, H. de C. **Efeitos do meio em propriedades conformacionais e eletrônicas de moléculas**. Tese (Doutorado) — Universidade de São Paulo, São Paulo, SP, 2006.

GÉRON, A. **Mãos à Obra: Aprendizado de Máquina com Scikit-Learn & TensorFlow**. Alta Books, 2019.

GRIFFITHS, D. **Mecânica quântica**. 2. ed. São Paulo: PRENTICE HALL BRASIL, 2011.

GROSSO, G.; PARRAVICINI, G. **Solid State Physics**. 2. ed. Elsevier Science, 2013.

HAYKIN, S. **Redes Neurais**: Princípios e Prática. Bookman, 2001.

HEATON, J. **Introduction to Neural Networks with Java**. Heaton Research, 2008.

CHOLLET, F. **Keras**. 2015. Disponível em: <https://keras.io/>. Acesso em: 2 dez. 2020

KITTEL, C. **Introdução à Física do Estado Sólido**. 8. ed. Rio de Janeiro: LTC, 2006.

LAMAS, T. E. **Epitaxia por feixe molecular de camadas dopadas do tipo p para construção de dispositivos optoeletrônicos**. Tese (Doutorado) — Instituto de Física, Universidade de São Paulo, São Paulo, SP, 2004.

LEAVITT, R. P. Empirical two-band model for quantum wells and superlattices in an electric field. **Phys. Rev. B**, American Physical Society, v. 44, p. 11270–11280, Nov 1991.

LEVINE, I. **Quantum Chemistry**. 7. ed. Pearson, 2013. (Pearson advanced chemistry series).

LEVINSHTEIN, M. **Handbook Series on Semiconductor Parameters**. World

Scientific, 1997.

LEVINSHTEIN, M. et al. **Handbook Series on Semiconductor Parameters: Ternary and quaternary III-V compounds.** World Scientific Publishing Company, 1999.

LIMA, J. P. H. **Redes Neurais Aplicadas à otimização de processos de deposição de filmes finos poliméricos.** Dissertação (Mestrado) — Escola Politécnica, Universidade de São Paulo, São Paulo, SP, 2006.

MACHADO, G. M. **Recuperando o uso da aproximação de massa efetiva em perovskitas híbridas de baixa dimensionalidade por meio da aproximação de bandas não parabólicas.** Dissertação (Mestrado) — Universidade Federal de Alfenas, Alfenas, MG, 2019.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, v. 5, p. 115 – 133, 1943.

MEDEIROS, V. C. de. **Estudo da dissociação das moléculas da série $CF_{3n}H_nCL(n = 0 - 3)$ nos estados fundamental e excitados $n - \sigma^*$.** Dissertação (Mestrado) — Universidade Federal da Paraíba, João Pessoa, PB, 2011.

MELLIT, A.; KALOGIROU, S. A. Artificial intelligence techniques for photovoltaic applications: A review. **Progress in Energy and Combustion Science**, v. 34, n. 5, p. 574 – 632, 2008.

NELSON, D. F.; MILLER, R. C.; KLEINMAN, D. A. Band nonparabolicity effects in semiconductor quantum wells. **Phys. Rev. B**, American Physical Society, v. 35, p. 7770–7773, May 1987.

OLIVEIRA, H. L. **Estudo de Excitação Eletrônica em Moléculas de Camada Aberta.** Tese (Doutorado) — Universidade de São Paulo, Florianópolis, SC, 2008.

PACHECO, C. A. R.; PEREIRA, N. S. Deep learning conceitos e utilização nas diversas Áreas do conhecimento. **Revista Ada Lovelace**, v. 2, p. 34–49, dez. 2018.

PACOBAYHA, L. H. **Propriedades de Transporte de Spins Polarizados em Heteroestruturas Semicondutoras.** Tese (Doutorado) — Universidade Federal de São Carlos, São Carlos, SP, 2008.

PERSSON, A.; COHEN, R. M. Reformulated hamiltonian for nonparabolic bands in semiconductor quantum wells. **Phys. Rev. B**, American Physical Society, v. 38, p. 5568–5575, Sep 1988.

ROOS, O. von. Position-dependent effective masses in semiconductor theory. **Phys. Rev. B**, American Physical Society, v. 27, p. 7547–7552, Jun 1983.

ROSENBLATT, F. The perceptron: a probabilistic model for information storage and retrieval in the brain. **Psychological Review**, v. 65, p. 368 – 408, 1958.

ROUL, B. et al. Binary group iii-nitride based heterostructures: Band offsets and transport properties. **Journal of Physics D Applied Physics**, 09 2015.

RUDER, S. **An overview of gradient descent optimization algorithms**. CoRR, abs/1609.04747, 2016. Disponível em: <http://arxiv.org/abs/1609.04747>. Acesso em: 2 dez. 2020

SANTOS, W. F. da S. **Predição da Dinâmica de Lasers de Pontos Quânticos Semicondutores Utilizando Redes Neurais Artificiais**. Dissertação (Mestrado) — Universidade Federal da Bahia, Salvador, BA, 2018.

SERWAY, R. A.; JEWETT, J. J. W. **Physics for Scientists and Engineers with Modern Physics**. 9. ed. Boston: CENGAGE Learning, 2013.

SILVA, A. S. F. et al. Inteligência computacional aplicada a síntese de semicondutores quaternários InAlGaAs. In: BRAGA, A. d. P.; Bastos Filho, C. J. A. (Ed.). **Anais do 11 Congresso Brasileiro de Inteligência Computacional**. Porto de Galinhas, PE: SBIC, 2013. p. 1–6.

SILVA, I. D.; SPATTI, D.; FLAUZINO, R. **Redes Neurais Artificiais Para Engenharia e Ciências Aplicadas**. Fundamentos Teóricos e Aspectos Práticos. ARTLIBER, 2010.

SILVA, J. da Costa e. **Confinamento Quântico em Hetero-estruturas Semicondutoras de Baixa Dimensionalidade**. Tese (Doutorado) — Universidade Federal do Ceará, Fortaleza, CE, 2010.

SINGH, J. **Electronic and Optoelectronic Properties of Semiconductor Structures**. Cambridge: Cambridge University Press, 2003.

TAHER, A. S.; MALEK, A.; MOMENI-MASULEH, S. Chebyshev differentiation matrices for efficient computation of the eigenvalues of fourth-order sturm–liouville problems. **Applied Mathematical Modelling**, v. 37, n. 7, p. 4634 – 4642, 2013.

TECUCI, G. Artificial intelligence. **WIREs Computational Statistics**, v. 4, n. 2, p. 168–180, 2012.

TIPLER, P. A.; LLEWELLYN, R. A. **Física Moderna**. 6. ed. Rio de Janeiro: LTC, 2010.

VIANNA, J.; FAZZIO, A.; CANUTO, S. **Teoria quântica de moléculas e sólidos: simulação computacional**. Livraria da Física, 2004.

YU, P.; CARDONA, M. **Fundamentals of Semiconductors: Physics and Materials Properties**. 4. ed. Springer Berlin Heidelberg, 2010. (Graduate Texts in Physics).