

UNIVERSIDADE ESTADUAL PAULISTA
UNESP - Campus de São João da Boa Vista
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA (MESTRADO)

RÔMULO APARECIDO DE PAULA JUNIOR

**Otimização de moduladores Mach-Zehnder integrados em fotônica de silício
utilizando redes neurais artificiais**

São João da Boa Vista

2022

RÔMULO APARECIDO DE PAULA JUNIOR

**Otimização de moduladores Mach-Zehnder integrados em fotônica de silício
utilizando redes neurais artificiais**

Versão original

Dissertação apresentada ao Conselho de Curso do Programa de Pós-Graduação em Engenharia Elétrica (Mestrado) do Campus de São João da Boa Vista, Universidade Estadual Paulista, como parte dos requisitos para obtenção do diploma de mestrado em Engenharia Elétrica.

Área de concentração: Fotônica Integrada e Aprendizagem de Máquinas

Orientador: Prof. Dr. Ivan Aritz Aldaya Garde

Coorientadora: Dra. Yesica Raquel Rumaldo Bustamante

São João da Boa Vista

2022

P324o

Paula Junior, Rômulo Aparecido de

Otimização de moduladores Mach-Zehnder integrados em fotônica de silício utilizando redes neurais artificiais / Rômulo Aparecido de Paula Junior. -- São João da Boa Vista, 2022

150 p. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Engenharia, São João da Boa Vista

Orientador: Ivan Aritz Aldaya Garde

Coorientadora: Yesica Raquel Rumaldo Bustamante

1. Aprendizado do computador. 2. Moduladores (Eletrônica). 3. Interferência (Luz). 4. Comunicações ópticas. 5. Otimização combinatória. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Engenharia, São João da Boa Vista. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

**CERTIFICADO DE APROVAÇÃO**

TÍTULO DA DISSERTAÇÃO: Otimização de moduladores Mach-Zehnder integrados em fotônica de silício utilizando redes neurais artificiais

AUTOR: ROMULO APARECIDO DE PAULA JUNIOR

ORIENTADOR: IVAN ARITZ ALDAYA GARDE

COORIENTADORA: YESICA RAQUEL RUMALDO BUSTAMANTE

Aprovado como parte das exigências para obtenção do Título de Mestre em ENGENHARIA ELÉTRICA, área: Sistemas Eletrônicos pela Comissão Examinadora:

Prof. Dr. IVAN ARITZ ALDAYA GARDE (Participação Virtual)

Coordenadoria de Curso de Engenharia Eletronica e de Telecomunicacoes / Faculdade de Engenharia de Sao Joao da Boa Vista - UNESP

Prof. Dr. MARCELO LUÍS FRANCISCO ABBADE (Participação Virtual)

Coordenadoria de Curso de Engenharia Eletrônica e de Telecomunicações / Faculdade de Engenharia de São João da Boa Vista - UNESP

Pesquisador Pleno Dr. RIVAEEL STROBEL PENZE (Participação Virtual)

Sensoriamento Óptico / CPQD

Prof. Dr. LUCAS HEITZMANN GABRIELLI (Participação Virtual)

Faculdade de Engenharia Elétrica e Computação / UNICAMP

São João da Boa Vista, 02 de junho de 2022



Maria Luiza Serubi Barreto
Diretora Téc. Acadêmica

Dedico esta dissertação à minha família e, em especial, à minha mãe, por todo incentivo e ajuda dados durante a realização do trabalho. Dedico também aos meus orientadores e aos meus colegas da UNESP e da CPQD por todo o conhecimento que foi compartilhado.

Resumo

Uma das principais vantagens da integração fotônica baseada na plataforma de silício sobre isolante (*silicon-on-insulator*, SOI) é a sua compatibilidade com o processo de manufatura de metal óxido-semicondutor complementar, possibilitando a produção em massa e reduzindo drasticamente o custo dos dispositivos. Além do mais, o alto contraste de índice de refração nesta plataforma resulta em um forte confinamento da luz no meio guiado, o que permite a construção de dispositivos compactos. Estas propriedades tornam a plataforma de silício sobre isolante atraente no projeto de sistemas coerentes ópticos, tais quais vêm ganhando atenção para enlaces cada vez mais curtos, como é o caso interconexões de alta velocidade em *data-centers*. Entretanto, o projeto de moduladores de banda larga e de baixo consumo de potência ainda é um desafio nesta plataforma. Dentre as arquiteturas existentes, o modulador de Mach-Zehnder (*Mach-Zehnder Modulator*, MZM) se mostra um candidato atraente devido ao seu potencial de apresentar larga banda eletro-óptica, baixo consumo de potência e alta estabilidade térmica. Estas qualidades se destacam, em especial, em MZMs com deslocadores de fase baseados no efeito de dispersão de plasma, operando em polarização reversa de tensão. Entretanto, a partir de uma revisão bibliográfica a respeito de MZMs construídos em silício, podemos perceber que o desempenho dos mesmos está atingindo um platô. Por outro lado, até onde sabemos, a otimização de MZMs por algoritmos de inteligência artificial ainda não tem sido explorada na literatura.

Modelos de simulação eletromagnéticas apresentam um alto custo de tempo, enquanto que algoritmos de otimização como os evolutivos requerem numerosas inferências, ainda mais ao otimizar um dispositivo com múltiplos parâmetros de projeto, como é o caso do MZM. A partir de simulações realizadas em um entorno que integra os *softwares* CST, Lumerical (*Mode* e *Device*) e uma linguagem de alto nível, como o Python, simularam-se 10.000 configurações de MZM com parâmetros aleatórios. Os dados obtidos mediante estas simulações foram utilizadas para treinar uma rede neural que pode substituir o modelo eletromagnético de alta complexidade no cômputo da função de custo, de forma a acelerar algoritmos de otimização heurística. A simulação e a rede neural consideraram oito parâmetros de *design* no modelo, incluindo a largura do guia de onda, a largura das regiões dopadas, o comprimento do modulador e a tensão de polarização. No caso dos parâmetros de desempenho do modulador temos a banda eletro-óptica, a tensão de meia-onda e as perdas de inserção ópticas.

Os resultados mostraram que a rede neural foi capaz de emular o modelo de simulação, mostrando uma acurácia satisfatória na predição dos três parâmetros de desempenho do MZM. Ao utilizarmos a rede neural como forma de computar a figura de mérito dos MZMs, o custo de tempo é reduzido por um fator de 10^7 em relação ao modelo de simulação. Com esta demanda menor por recursos computacionais, é possível executarmos otimizações que consideram um elevado número de iterações e indivíduos nas populações. Além do mais, diferentes figuras de mérito podem ser utilizadas para dar ênfase na otimização de parâmetros de métrica distintos. Concluimos que esta configuração de otimização conjunta possibilita achar os limites da arquitetura em termos das métricas consideradas.

Palavras-chaves: Fotônica em Silício, Modulador de Mach-Zehnder, Otimização de *Design*, Aprendizado Profundo, Evolução Diferencial.

Abstract

One of the main advantages of silicon on insulator based integration is its compatibility with the complementary metal-oxide-semiconductor (CMOS) manufacturing process, which enables mass production and drastically reduces device costs. Furthermore, the high refractive index contrast inherent to this platform allows a high confinement of light in the guided medium, enabling the design of compact devices. These silicon-on-insulator (SOI) properties have attracted particular attention for the design of optical coherent systems, which have been getting more popular for shorter and shorter links, as the case of high-speed data-centers interconnections. However, the design of broadband and power-efficient modulators in this platform still poses a challenge. Amongst the existing architectures, the Mach-Zehnder Modulator (MZM) emerges as an attractive candidate, due to its potential for wide electro-optic bandwidth, low power consumption and high thermal stability. Such benefits stand out, in special, when it comes to Mach-Zehnder modulators composed of phase shifters based on the plasma dispersion effect, operating via reverse voltage polarization. However, from a literature review concerning MZMs made of silicon, we observed that the modulators performance is plateauing. On the other hand, to the best of our knowledge, MZM optimization via artificial intelligence algorithms is still unexplored in the literature.

Electromagnetic simulation models present a high computational complexity, while optimization algorithms such as genetic algorithm require several inferences, even more when a multi-parameter device such as the MZM is being optimized. The simulations were carried out in an wrapper comprising CST and Lumerical (Mode and Device) softwares, and the high-level programming language Python. From this wrapper, we obtained 10,000 MZM configurations of random parameters. The dataset was used to train an artificial neural network model that can replace the high complexity simulation model in the computation of the cost function, accelerating the convergence of heuristic optimization algorithms. The simulation and the neural network consider eight design parameters, including the waveguide width, the doped regions width, the modulator length and the bias voltage. For the modulator performance metrics we considered the electro-optic bandwidth, the half-wave voltage and the optics insertion loss.

The obtained results demonstrated that the network is able to emulate the simulation model with satisfactory accuracy for the three considered performance parameters. By utilizing the neural network for the MZMs evaluation phase of the heuristic algorithm, the time efficiency is increased by 10^7 times. By alleviating the burden on the hardware resources, it is possible to execute the optimization for high numbers of iterations and MZM samples. Furthermore, different figures of merit can be utilized to emphasize distinct performance parameters in the optimization. We concluded that this joint optimization configuration made it possible for us to explore the limits of the MZM architecture in terms of the considered metrics.

Keywords: Silicon Photonics, Mach-Zehnder Modulator, Design Optimization, Deep Learning, Differential Evolution.

Lista de figuras

- Figura 1 – Constelações representando os seguintes formatos de modulação digital: (a) OOK, (b) PAM-4, (c) BPSK, (d) 4-QAM, (e) 16-QAM e (f) 64-QAM. 27
- Figura 2 – Representação do transceptor integrado coerente por diagrama de blocos. No bloco superior temos a seção do transmissor e no bloco inferior temos a seção de recepção. 34
- Figura 3 – Diagrama esquemático ilustrando moduladores integrados baseados na plataforma de silício. (a) Modulador de eletro-absorção. (b) Moduladores de micro-anel. (c) Moduladores de Michelson. (d) Moduladores de Mach-Zehnder. 38
- Figura 4 – Estrutura de um MZM-IQ. 42
- Figura 5 – Esquemático de um interferômetro de Mach-Zehnder. 46
- Figura 6 – Saída normalizada em relação à entrada, tanto para a intensidade quanto para a amplitude complexa do campo elétrico, e defasagem do campo elétrico complexo de saída em termos da diferença de fase. 48
- Figura 7 – (a) Função transferência de um MZI em relação ao comprimento de onda. Os braços possuem um desbalanço de $100 \mu\text{m}$ em relação ao comprimento físico destes. (b) FSR, em relação ao comprimento de onda, do MZI considerado. 50
- Figura 8 – Configurações de alimentação. (a) *Single drive* (b) *dual drive* e (c) *series push-pull*. 53
- Figura 9 – Diferentes estruturas de diodos. (a) Estrutura de junção interdigital. (b) Junção PN lateral convencional e (c) junção PN lateral com regiões de dopagens intermediárias. (d) Estrutura PIPIN. (e) Estrutura PIN. . . . 58
- Figura 10 – Variação de (a) índice de refração e (b) atenuação em relação às variações de elétrons e de lacunas. 60
- Figura 11 – Circuito elétrico equivalente, considerando pequenos sinais, de um (a) diodo polarizado diretamente e (b) diodo polarizado reversamente. . . . 61

Figura 12 – Seção transversal de um guia do tipo <i>rib</i> com três níveis de dopagem. No centro do guia óptico há a formação de uma junção PN. A figura também mostra uma simplificação de um circuito elétrico equivalente à junção.	62
Figura 13 – (a) Variação do índice efetivo e das perdas ao aplicarmos uma tensão reversamente polarizada e (b) variação da fase do sinal para os mesmos valores de tensão.	64
Figura 14 – Seção transversal de uma linha CPS simétrico com única camada de substrato.	68
Figura 15 – Perspectiva isométrica e seção transversal da conexão do TWE com os guias de ondas do MZM.	70
Figura 16 – Circuito equivalente RC paralelo da linha de transmissão com a junção PN como terminação. A janela superior mostra os elementos da junção PN antes de tomarmos o circuito RC paralelo equivalente. Figura adaptada de (POZAR, 2005).	71
Figura 17 – Comparação entre três modelos de regressão, em que um deles apresenta <i>underfitting</i> , o outro apresenta <i>overfitting</i> , e o último demonstra um mapeamento adequado. (a) Conjunto de dados, de teste e de treinamento, e as curvas de mapeamento obtidas pelos modelos de regressão. (b) Erro do conjunto de treinamento e de validação em relação às iterações de treinamento.	81
Figura 18 – (a) Uma distribuição gaussiana com média 10 e desvio padrão de 4 e (b) a mesma distribuição após a estandardização.	83
Figura 19 – Função de ativação unidade linear retificada (ReLU).	83
Figura 20 – Uma ANN totalmente conectada composta de L_c camadas com quantidades arbitrárias de unidades neurais. Tem-se n neurônios de entrada K neurônios de saída.	85
Figura 21 – Diagrama de blocos do algoritmo de DE.	97
Figura 22 – Diagrama de blocos do nosso esquema de otimização conjunta de ANNs com algoritmos de otimização.	98

Figura 23 – Ilustração do MZM a ser implementado. A parte superior representa uma seção do modulador e dos guias de transmissão, enquanto a parte do meio representa a seção transversal do modulador completo e abaixo temos a seção transversal de um dos braços do modulador. Figura adaptada de (MOTTA <i>et al.</i> , 2017).	101
Figura 24 – Diagrama de blocos da geração de dados a partir do modelo de simulação. Os blocos com contornos em preto indicam a utilização da linguagem de programação de alto nível, enquanto que os blocos com contornos vermelhos e azuis indicam, respectivamente, simulações ópticas e elétricas.	102
Figura 25 – <i>Design</i> de um CPS com extensões em formato de T (MOTTA <i>et al.</i> , 2017).	103
Figura 26 – Ilustração da ANN utilizada no trabalho. As entradas são os parâmetros de <i>design</i> e a saída são os parâmetros de métricas referentes ao MZM. (a) ANN arbitrária aplicada ao nosso problema e (b) ANN utilizada na otimização aqui presente.	107
Figura 27 – Curva de aprendizado para o caso da ANN que demonstra 4 camadas escondidas. A figura compara a otimização da ANN ao utilizar BN, DO, RES, todas combinações de 2 destes e o modelo completo. As curvas de aprendizado são em relação ao (a) conjunto de treinamento e (b) conjunto de testes.	112
Figura 28 – Curvas de aprendizado comparando as ANNs de 3 a 6 camadas escondidas, tanto para o caso da ANN base quanto para a ANN completa. Curvas de aprendizado em relação ao (a) conjunto de treinamento e (b) conjunto de testes.	113
Figura 29 – Decaimento da taxa de aprendizado em relação às épocas do treinamento da ANN.	113
Figura 30 – Sobreposição de um subconjunto dos valores de teste simulados e previstos pelo modelo baseado em ANN para (a) BW_{EO} , (b) IL e (c) V_{π} . Relação entre os valores simulados e previstos, considerando todo o conjunto de testes para (d) BW_{EO} , (e) IL e (f) V_{π} . A linha tracejada indica os valores ideais.	114

Figura 31 – Métricas de desempenho do MZM do conjunto de treinamento da ANN e as métricas obtidas por 100 execuções do algoritmo DE. Os eixos são dados pela BW_{EO} e pelo V_{π} , e a intensidade de cor representa os valores de IL.	116
Figura 32 – FOM do melhor indivíduo da geração considerando 100 inicializações aleatórias.	117
Figura 33 – Otimização por DE de apenas duas métricas de desempenho do MZM, enquanto o outro é limitado para determinado valor de limiar. (a) Métricas obtidas ao limitar o BW_{EO} , (b) métricas obtidas ao limitar o IL e (c) métricas obtidas ao limitar o V_{π}	118
Figura 34 – Sobreposição de um subconjunto dos valores de métricas relativas aos MZMs otimizados pelos algoritmos heurísticos, previstos pela ANN e pelo modelo de simulação eletromagnética, para (a) BW_{EO} , (b) IL e (c) V_{π} . Relação entre os valores simulados e previstos, considerando os 400 MZM otimizados, para (d) BW_{EO} , (e) IL e (f) V_{π} . A linha tracejada indica a relação ideal.	120
Figura 35 – Seção infinitesimal de uma linha de transmissão, com a tensão e a corrente propagante em função do tempo e os parâmetros distribuídos ilustrados. Figura adaptada de (POZAR, 2005).	137
Figura 36 – Linha de transmissão representada como uma rede de duas portas. . .	139
Figura 37 – Diagrama de fluxo do algoritmo genético. A caixa em azul se refere ao algoritmo de evolução diferencial.	142
Figura 38 – Esquemático do cruzamento entre dois cromossomos.	143
Figura 39 – Diagrama de fluxo do algoritmo de otimização por enxames de partículas.	145
Figura 40 – Diagrama de fluxo do algoritmo de DA.	146
Figura 41 – Curva de convergência da otimização pelos algoritmos heurísticos, considerando 100 inicializações aleatórias para cada um destes.	148
Figura 42 – Métricas de desempenho dos MZMs obtidos pela otimização via algoritmos heurísticos para 100 inicializações aleatórias. (a) Conjunto de dados do treinamento da ANN e os MZMs otimizados pelo GA, PSO e DA. (b) Resultados da otimização por GA, por (b) PSO e por (d) DA.	150

Lista de tabelas

- Tabela 1 – Comparativo entre MZMs em silício reportados em diversos trabalhos, em ordem cronológica de publicação. A tabela inclui a BW_{EO} , a V_{π} , a IL, figura de mérito definida e a modulação utilizada. 43
- Tabela 2 – Resultados de simulação obtidos para diferentes estruturas de diodos. Tais resultados incluem as métricas BW_{EO} , IL e V_{π} obtidos em (PEREZ-GALACHO *et al.*, 2015) 59
- Tabela 3 – Parâmetros variáveis do MZM no modelo de simulação, com seus respectivos alcances de valores. 101
- Tabela 4 – Número de camadas escondidas por configuração de ANN, número de neurônios pertencentes a cada camada e configurações utilizadas. . . . 108
- Tabela 5 – FOMs utilizados durante a otimização. A tabela inclui os FOMs que foram utilizados ao limitar alguma das métricas, de forma que seu valor não seja ultrapassado 110
- Tabela 6 – Valores de MSE, MAE e r_p para as métricas BW_{EO} , IL e V_{π} e para as amostras em que $V_{\pi} < 10$ V. Não calculamos o r_p para o caso em que $V_{\pi} < 10$ V pois a correlação de Pearson para conjunto de dados com diferentes amostras pode não ser muito precisa. 115
- Tabela 7 – Avaliação das métricas obtidas pelos 400 MZMs otimizados pelos algoritmos heurísticos. A tabela inclui o MAE, MSE e r_p ao relacionar as métricas obtidas pelos diferentes métodos. A tabela também inclui a média das métricas para cada método, assim como o erro percentual entre estas. 120
- Tabela 8 – Média, desvio padrão e valor máximo para a FOM obtida pelos algoritmos heurísticos, considerando 100 inicializações aleatórias. A tabela também mostra os valores mínimos e máximos obtidos para cada um dos parâmetros de métrica do MZM. 149

Lista de abreviaturas e siglas

ADAM	<i>Adaptive Moment Estimation</i> - Estimativa Adaptativa de Momento
ADC	<i>Analog to Digital Converter</i> - Conversor Analógico para Digital
ANN	<i>Artificial Neural Network</i> - Rede Neural Artificial
BN	<i>Batch Normalization</i> - Normalização de <i>Batch</i>
BPSK	<i>Binary Phase Shift Keying</i> - Chaveamento Binário de Fase
BP	<i>Backward Propagation</i> - Retropropagação
CMOS	<i>Complementary Metal Oxide Semiconductor</i> - Metal Óxido-Semicondutor Complementar
CPS	<i>Coplanar Stripline</i> - Guia de Onda de Tira Coplanar
CPU	<i>Central Processing Unit</i> - Unidade Central de Processamento
CPW	<i>Coplanar Waveguide</i> - Guia de Onda Coplanar
CSA	<i>Classical Simulated Annealing</i> - Recozimento Clássico
CW	<i>Continuous Wave</i> - Onda Contínua
DA	<i>Dual Annealing</i> - Recozimento Duplo
DE	<i>Differential Evolution</i> - Evolução Diferencial
DL	<i>Deep Learning</i> - Aprendizado Profundo
DO	<i>Dropout</i>
DSP	<i>Digital Signal Processing</i> - Processamento de Sinais Digitais
DV	Distribuição de Visitação
FCA	<i>Free Carrier Absorption</i> - Absorção de Portadores Livres
FC	<i>Fully Connected</i> - Totalmente Conectada
FEEM	<i>Finite Element Eigenmode</i> - Automodo de Elemento Finito

FOM	<i>Figure of Merit</i> - Figura de Mérito
FP	<i>Feedforward Propagation</i> - Propagação Avante
FSA	<i>Fast Simulated Annealing</i> - Recozimento Rápido
GA	<i>Genetic Algorithm</i> - Algoritmo Genético
GD	<i>Gradient Descent</i> - Gradiente Descendente
GSA	<i>Generalized Simulated Annealing</i> - Recozimento Generalizado
GaAs	<i>Gallium Arsenide</i> - Arsenieto de Gálio
IA	Inteligência Artificial
IF	<i>Intermediate Frequency</i> - Frequência Intermediária
IoT	<i>Internet of Things</i> - Internet das Coisas
IP	<i>Internet Protocol</i> - Protocolo de Internet
IM-DD	<i>Intensity Modulation - Direct Detection</i> - Modulação de Intensidade e Detecção Direta
ISI	<i>Intersymbol Interference</i> - Interferência Intersimbólica
InP	<i>Indium Phosphide</i> - Fosfeto de Índio
LD	<i>Laser Diode</i> - Diodo Laser
LO	<i>Local Oscillator</i> - Oscilador Local
LiNbO ₃	<i>Lithium Niobate</i> - Niobato de Lítio
MAE	<i>Mean Absolute Error</i> - Erro Absoluto Médio
ML	<i>Machine Learning</i> - Aprendizado de Máquinas
MLP	<i>Multilayer Perceptron</i> - Percéptron Multi-camadas
MMI	<i>Multimode Interferometer</i> - Interferômetro Multi-Modo
MSE	<i>Mean Squared Error</i> - Erro Quadrático Médio

MZI	<i>Mach-Zehnder Interferometer</i> - Interferômetro de Mach-Zehnder
MZM	<i>Mach-Zehnder Modulator</i> - Modulador de Mach-Zehnder
OIF	<i>Optical Internetworking Forum</i> - Fórum de Interconexão Óptica
OOK	<i>On-Off Keying</i> - Chaveamento Liga Desliga
OPLL	<i>Optical Phase-Locked Loop</i> - Laço de Sincronismo de Fase Óptica
PAM	<i>Pulse Amplitude Modulation</i> - Modulação de Pulso em Amplitude
PDE	<i>Plasma Dispersion Effect</i> - Efeito de Dispersão de Plasma
PSK	<i>Phase Shift Keying</i> - Chaveamento por Deslocamento de Fase
PSO	<i>Particle Swarm Optimization</i> - Otimização por Enxame de Partículas
PSR	<i>Polarization Splitter and Rotator</i> - Divisor e Rotor de Polarização
PS	<i>Power Splitter</i> - Divisor de Potência
QAM	<i>Quadrature Amplitude Modulation</i> - Modulação de Amplitude em Quadratura
QCSE	<i>Quantum-Confined Stark Effect</i> - Efeito Stark Confinado por Quantum
QPSK	<i>Quadrature Phase Shift Keying</i> - Chaveamento por Deslocamento de Fase em Quadratura
RAM	<i>Random Access Memory</i> - Memória de Acesso Aleatória
RES	<i>Residual Connections</i> - Conexões Residuais
ReLU	<i>Rectified Linear Unit</i> - Unidade Linear Retificada
SGD	<i>Stochastic Gradient Descent</i> - Gradiente Descendente Estocástico
SOI	<i>Silicon-on-Insulator</i> - Silício em Isolante
SSP	<i>Series Push-Pull</i>
Si	<i>Silicon</i> - Silício
TEM	<i>Transverse Electromagnetic</i> - Transversal Electromagnético

TE	<i>Transverse Electric</i> - Transversal Elétrico
TIA	<i>Transimpedance Amplifiers</i> - Amplificadores de Transimpedância
TM	<i>Transverse Magnetic</i> - Transversal Magnético
TW	<i>Traveling Wave</i> - Onda Viajante
TWE	<i>Traveling Wave Electrode</i> - Eletrodo de Onda Viajante
WDM	<i>Wavelength Division Multiplexing</i> - Multiplexação por Divisão de Comprimento de Onda
Z-score	<i>Standard Score</i> - Escore Padronizado

Lista de símbolos

E_{out}	Amplitude do campo elétrico na saída do combinador do interferômetro de Mach-Zehnder
E_{oi}	Amplitude do campo elétrico na saída do braço i do interferômetro de Mach-Zehnder
E_i	Amplitude do campo elétrico no braço i do interferômetro de Mach-Zehnder
E_{in}	Amplitude do campo elétrico na entrada do divisor de potência do interferômetro de Mach-Zehnder
a_k	Ativação do k -ésimo neurônio artificial
V_j	Barreira de potencial da junção PN
\mathbf{E}	Campo elétrico
\mathbf{E}_T	Campo elétrico transversal
\mathbf{H}	Campo magnético
\mathbf{H}_T	Campo magnético transversal
C_j	Capacitância da junção PN
c_d	Capacitância de difusão
C	Capacitância distribuída da linha de transmissão
C_{tpn}	Capacitância equivalente da linha de transmissão carregada
q	Carga elementar
κ	Coefficiente de absorção
α	Coefficiente de atenuação
α_{RF}	Coefficiente de atenuação de RF

α_i	Coefficiente de atenuação do i -ésimo braço do interferômetro de Mach-Zehnder
r_p	Coefficiente de correlação de Pearson
Γ	Coefficiente de reflexão
Γ_S	Coefficiente de reflexão na fonte
Γ_T	Coefficiente de reflexão na terminação
$\beta^{(1)}$ e $\beta^{(2)}$	Coefficientes de primeiro e de segundo momento no otimizador ADAM
λ	Comprimento de onda
λ_{RF}	Comprimento de onda de RF
L_{ps}	Comprimento do deslocador de fase
L_i	Comprimento do i -ésimo braço do interferômetro de Mach-Zehnder
n_i	Concentração intrínseca de portadores
G_d	Condutância da junção PN
G	Condutância distribuída da linha de transmissão
S	Conjunto de dados para treinamento da rede neural
K_B	Constante de Boltzmann
β_i	Constante de fase do i -ésimo braço do interferômetro de Mach-Zehnder
i	Corrente elétrica
$\Delta\phi$	Defasamento entre os braços do interferômetro de Mach-Zehnder
ΔN	Densidade em excesso de elétrons
ΔN^{++}	Densidade em excesso de elétrons na região com dopagem forte
ΔN^+	Densidade em excesso de elétrons na região com dopagem intermediária
ΔP	Densidade em excesso de lacunas

ΔP^{++}	Densidade em excesso de lacunas na região com dopagem forte
ΔP^+	Densidade em excesso de lacunas na região com dopagem intermediária
ΔT	Diferença de temperatura
σ_{BN}	Desvio padrão considerando as amostras de um <i>mini-batch</i>
g_{strip}	Espaçamento entre as tiras do guia de onda de tira coplanar
g_e	Espaçamento efetivo
t	Espaçamento entre eletrodos
g	Espaçamento entre extensões T do eletrodo
t_{RIB}	Espessura do guia de onda RIB
FSR	<i>Free Spectral Range</i> - Faixa Espectral Livre
ϕ_i	Fase acumulada pelo i -ésimo braço do interferômetro de Mach-Zehnder
ϵ_{ADAM}	Fator de correção do otimizador Adam
Q	Fator de qualidade
k_i	Fração da potência sendo guiada para o i -ésimo braço do interferômetro de Mach-Zehnder
ω	Frequência angular
h	Função de ativação da rede neural
J	Função de custo
T_m	Função de transferência do Mach-Zehnder
$T_m^{(f)}$	Função de transferência do Mach-Zehnder em relação ao campo elétrico complexo
δ	Gradiente local
Z	Impedância
Z_0	Impedância característica da linha de transmissão

Z_{0-cps}	Impedância característica do guia de ondas CPS
Z_s	Impedância da fonte
Z_t	Impedância do terminal da linha de transmissão
n_{g-eff}	Índice de grupo efetivo
n	Índice de refração
\bar{n}	Índice de refração complexo
n_c	Índice de refração do núcleo
n_{eff}	Índice de refração efetivo
n_{eff-mw}	Índice de refração efetivo de RF
$n_{eff,int}$	Índice de refração efetivo intrínseco
L	Indutância distribuída da linha de transmissão
IL	<i>Insertion loss</i> - Perdas por Inserção
I_{out}	Intensidade na saída do interferômetro de Mach-Zehnder
I_{in}	Intensidade óptica da onda de entrada do interferômetro de Mach-Zehnder
W_i	Largura da camada de dopagem do i -ésimo bloco
r	Largura da conexão entre o eletrodo e sua extensão T
W_d	Largura da região de depleção
W	Largura das extensões T do eletrodo
BW_{EO}	Largura de banda eletro-óptica
W_e	Largura de <i>strip</i> efetiva
$\mathcal{H}^{(l)}$	Mapeamento realizado pela l -ésima camada de uma rede neural artificial
μ_{BN}	Média em relação às amostras de um mini-batch

n_{inputs}	Número de entradas numa unidade neural
OMA	<i>Optical Amplitude Modulation</i> - Amplitude de Modulação Óptica
λ_{reg}	Parâmetro de regularização
γ_{BN}	Parâmetro do BN que escalona a distribuição de amostras
β_{BN}	Parâmetro do BN que desloca a distribuição de amostras
S	Parâmetros S (Matriz)
α_{mod}	Perdas de modulação
μ	Permeabilidade magnética do meio
ϵ	Permissividade do meio
ϵ_0	Permissividade do vácuo
ϵ_s	Permissividade elétrica relativa do silício
ϵ_{SiO_2}	Permissividade elétrica relativa da sílica
ϵ_r	Permissividade relativa
P_{sd}	Potência dissipada utilizando um único <i>driver</i> em um Modulador de Mach-Zehnder
P_{dd}	Potência dissipada utilizando dois <i>drivers</i> em um Modulador de Mach-Zehnder
u	Potencial de ativação da unidade neural
Φ	Potencial eletrostático escalar
P_{dyn}	Potência dissipada do modulador de Mach-Zehnder operando de forma dinâmica
P_H	Potência do maior símbolo de modulação
P_L	Potência do menor símbolo de modulação
\mathbf{m}_t e \mathbf{v}_t	Primeiro e segundo momentos de gradiente

n	Quantidade de atributos de entrada de uma amostra
k	Quantidade de atributos de saída de uma amostra
L_c	Quantidade de camadas escondidas na rede neural artificial
s_l	Quantidade de unidades neurais da l -ésima camada
ER	<i>Extinction Ratio</i> - Razão de Extinção
k_e	Razão de aspecto efetiva entre espaçamento e largura dos substratos
$k^{(1)}$	Razão de aspecto entre espaçamento e largura dos substratos
R_S	Resistência da fonte
R_j	Resistência da junção PN
R_{si}	Resistência de folha do i -ésimo bloco
R_T	Resistência de terminação
R	Resistência distribuída da linha de transmissão
R_{tpn}	Resistência equivalente da linha de transmissão carregada
r_b	Resistências da junção PN
T	Temperatura
V_{bias}	Tensão DC aplicada no modulador de Mach-Zehnder
V_π	Tensão de meia-onda
V	Tensão externa aplicada
V_{avg}	Tensão média entre eletrodos
b	Termo de <i>bias</i> da rede neural
$L2$	Termo de custo de regularização
r	Termo de probabilidade de DO
TPP	<i>Transmission Power Penalty</i> - Penalidade de Potência de Transmissão

Δn	Varição de índice de refração
Δn_{eff}	Varição de índice de refração efetivo
$\Delta\alpha$	Varição de coeficiente de atenuação
c	Velocidade da luz no vácuo
\mathbf{w}	Vetor de pesos de uma camada da rede neural
\mathbf{v}	Vetor doador
\mathbf{u}	Vetor tentativa

Sumário

1	Introdução	25
1.1	<i>Fotônica Integrada para Sistemas de Comunicações Ópticas</i>	29
1.2	<i>Moduladores Ópticos Integrados</i>	34
1.3	<i>Justificativa</i>	43
1.4	<i>Contribuições do Trabalho</i>	44
1.5	<i>Organização do Documento</i>	45
2	Fundamentos de Moduladores de Mach-Zehnder	46
2.1	<i>Interferômetros de Mach-Zehnder</i>	46
2.2	<i>Moduladores de Mach-Zehnder</i>	50
2.2.1	Configurações de Alimentação	52
2.2.2	Figuras de Mérito de um MZM	54
2.3	<i>Deslocadores de Fase</i>	57
2.3.1	Estruturas de Diodos	57
2.3.2	Efeito de Dispersão de Plasma	59
2.3.3	Polarização Direta e Reversa da Junção PN	61
2.3.4	Resposta Óptica	63
2.4	<i>Eletrodos de Ondas Viajantes</i>	66
2.4.1	Guia de Ondas de Tiras Coplanares sem Carga	67
2.4.2	Linha de Transmissão com Carga	70
2.4.3	Banda Eletro-Óptica Considerando os Parâmetros da Linha de Transmissão	73
3	Fundamentos de Inteligência Artificial	77
3.1	<i>Princípios de Aprendizagem de Máquinas e Aprendizagem Profunda</i>	77
3.1.1	Aprendizagem de Máquinas	78
3.1.2	Funcionamento de Redes Neurais Artificiais	82
3.1.3	Aprimoramento de Redes Neurais Artificiais	88
3.1.4	Métricas Do Mapeamento da ANN	94
3.2	<i>Evolução Diferencial</i>	95
3.3	<i>Otimização Integrando Algoritmos Heurísticos e Redes Neurais Artificiais</i>	97

4	Método de Otimização	100
4.1	<i>Geração de Dados</i>	100
4.1.1	Simulação da Seção Elétrica	102
4.1.2	Simulação da Seção Óptica	104
4.2	<i>Otimização do Modulador</i>	106
4.2.1	Mapeamento por Rede Neural	106
4.2.2	Otimização por Evolução Diferencial	109
4.2.3	Métricas de Otimização	109
5	Resultados	111
5.1	<i>Resultados da Rede Neural Artificial</i>	111
5.1.1	Curvas de Aprendizado	111
5.1.2	Desempenho da Rede Neural Artificial	113
5.2	<i>Resultados de Otimização Utilizando Evolução Diferencial</i>	115
5.3	<i>Validação dos Moduladores Otimizados com o Modelo de Simulação</i>	119
6	Conclusões e Trabalho Futuro	121
	Referências	122
	 APÊNDICES	 136
	Apêndice A – Análise de Parâmetros Distribuídos	137
A.1	<i>Parâmetros S e de Transmissão</i>	139
	Apêndice B – Outros Algoritmos Heurísticos	142
B.1	<i>Funcionamento dos Algoritmos Heurísticos</i>	142
B.1.1	Algoritmo Genético	142
B.1.2	Otimização por Enxame de Partículas	144
B.1.3	Dual Annealing	146
B.2	<i>Metodologia dos Algoritmos Heurísticos</i>	148
B.3	<i>Resultados dos Algoritmos Heurísticos</i>	148

1 Introdução

A exigência por sistemas de telecomunicações cada vez com capacidades superiores de transmissão se mostra crescente, principalmente quando se trata de sistemas ópticos, amplamente utilizados devido à disposição de uma larga banda (ZHOU; LIU; URATA, 2017). Tais sistemas devem dispor de capacidade suficiente para atender as demandas impostas por serviços como os de *streaming*, computação em nuvem e computação distribuída (THRASKIAS *et al.*, 2018). De acordo com a Cisco Systems, há previsões de que o número de dispositivos conectados por protocolo de Internet (*Internet Protocol*, IP) será de 29,3 bilhões em 2023, em relação aos 18,4 bilhões em 2018, o que representa um aumento de quase 60% em apenas 5 anos (CISCO, 2022). Também é de se esperar que as taxas médias de Internet fixa aumentem de 45,9 Mbps para 110,4 Mbps. Outra pesquisa indica que, em 2010, haviam menos de 2 bilhões de usuários de Internet ao redor do mundo (DATA, 2022). Este número mais do que dobrou até 2018, em que 4,2 bilhões de pessoas possuíam acesso à Internet e, em 2021, o número de usuários subiu para 5,1. Já no caso do Brasil temos que em 2008 haviam 67 milhões de usuários de Internet e, em 2019, o país apresentou 149 milhões de usuários. Outra fonte mais recente mostra que em 2021 Brasil apresentou quase 162 milhões de usuários de Internet, o que representa 76% da população do país (STATISTA, 2022). Outro campo que se espera que cause um significativo aumento da capacidade que os operadores devem oferecer é a Internet das coisas (*Internet of Things*, IoT). A tecnologia de IoT se encontra em um alto grau de amadurecimento, e esta tem como proposta a interconexão de diversos objetos cotidianos por meio da rede. IoT abrange comunicações de máquina para máquina, de humanos para máquinas e de humanos para humanos. De acordo com (IOT-ANALYTICS, 2022), em 2021 o número de dispositivos conectados em IoT eram de 12,2 bilhões, e até 2025 haverão 27 bilhões de dispositivos. As previsões são que, em 2025, 152.000 dispositivos de IoT serão conectados, por minuto, na Internet (DATAPROT, 2022).

É de se esperar que este excessivo número de dispositivos conectados resulte num acréscimo do congestionamento da rede de transferência de dados em centros de processamento de dados (*data-centers*). Para o atendimento de tais demandas, os *data-centers* apresentam um aumento tanto em relação à taxa de dados de seus enlaces, quanto

na distância que estes dados devem percorrer, o que se dá devido à ampliação física de tais centros (BENNER, 2012).

A transferência de dados em um *data-center* pode ocorrer de um computador para outro, de uma unidade de processamento central (*Central Processing Unit*, CPU) para outra CPU ou para uma memória, de uma memória para uma CPU e também devemos levar em consideração a entrada e a saída de dados do centro para a rede Internet. Deve-se dar atenção especial ao projeto de telecomunicações em *data-centers*, pois se este apresentar um desempenho sub-ótimo, resultará em um gargalo na funcionalidade dos servidores. Não obstante, um trabalho de 2012 indica que a transferência de dados entre *chips* apresenta um custo de energia até 200 vezes maior do que o custo para processamento desses dados (BENNER, 2012). Estes fatos justificam o estudo de transferência de dados cada vez mais eficiente em termos de desempenho e consumo de energia em centros de processamentos.

Em um *data-center*, a transferência de dados pode se dar por meios guiados ou por ar livre. Meios sem fio são convenientes por possuírem dispositivos mais simples e compactos. Apesar de haver estudos intensivos em novos métodos e novas frequências de operação (VARDHAN *et al.*, 2010), estes ainda apresentam limitações em largura de banda e distância e podem apresentar interferência entre canais (*crosstalk*) na ausência de esquemas eficientes de acesso múltiplo ao meio. Meios sem fio são adequados em enlaces de relativamente curto alcance, e também que requeiram baixa largura de banda. Por outro lado, interconexões guiadas podem ser implementadas a partir de guias elétricos ou guias ópticos, cada qual com suas respectivas características. Canais ópticos se destacam em relação a canais elétricos devido ao maior produto entre largura de banda e comprimento disponível para transmissão de informação, mesmo sem utilizar multiplexação por comprimento de onda (*Wavelength Division Multiplexing*, WDM) (STUCCHI *et al.*, 2013). Além disso, enlaces ópticos possuem menores perdas e são mais resilientes à interferência entre canais, o que possibilita uma maior densidade de interconexões (MILLER, 2009). Levando em consideração que, geralmente, os dados são gerados e processados no domínio elétrico, a implementação de interconexões ópticas impõe duas etapas adicionais: a conversão do domínio elétrico para domínio óptico e do domínio óptico para o domínio elétrico. Estas etapas adicionais elevam o custo total do sistema, por isto canais elétricos ainda são preferíveis em casos de curta distância e em enlaces que não requerem requisitos rigorosos

de largura de banda. O desenvolvimento de conversores eletro-ópticos de baixo custo pode mudar este cenário.

Para o atingimento de taxas na ordem de Tbps, sistemas coerentes ópticos têm sido introduzidos. Além da já tradicionalmente utilizada modulação por intensidade com detecção direta (*Intensity Modulation Direct Detection*, IM-DD), como o chaveamento liga-desliga (*On-Off Keying*, OOK), os sistemas coerentes ópticos permitem a codificação de informação digital por modulação de amplitude, fase e também multiplexação de estados de polarização ortogonais (AGRAWAL, 2014). Alguns dos formatos de modulação que sistemas coerentes suportam são o chaveamento por deslocamento de fase (*Phase Shift Keying*, PSK), chaveamento por deslocamento de fase em quadratura (*Quadrature Phase Shift Keying*, QPSK) e modulação de amplitude em quadratura (*Quadrature Amplitude Modulation*, QAM).

Estes formatos avançados de modulação, tais quais apresentando os diagramas de constelação ilustrados pela Figura 1, são capazes de codificar informação tanto na amplitude quanto na fase da onda portadora. A Figura 1(a) ilustra o caso do OOK, em que há apenas dois níveis de intensidade e a fase não é relevante para a informação. A Figura 1(b) ilustra a modulação de pulso em amplitude (*Pulse Amplitude Modulation*, PAM) com quatro níveis (PAM-4). Neste tipo de modulação, a informação é codificada

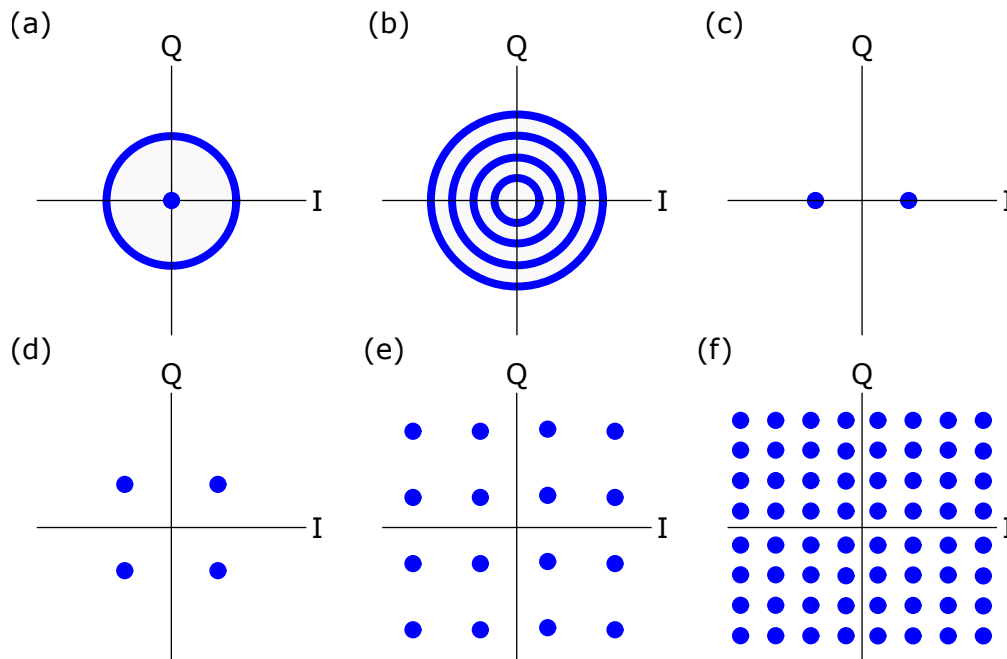


Figura 1 – Constelações representando os seguintes formatos de modulação digital: (a) OOK, (b) PAM-4, (c) BPSK, (d) 4-QAM, (e) 16-QAM e (f) 64-QAM.

apenas na amplitude, porém há múltiplos níveis, permitindo uma maior eficiência espectral. A Figura 1(c) esquematiza a modulação por chaveamento binário de fase (*Binary Phase Shift Keying*, BPSK), em que há dois símbolos de informação, sendo a diferença na representação destes um defasamento de 180° . Já as Figuras 1(d-f) ilustram o diagrama de constelação para o formato de modulação QAM com, respectivamente, 4, 16 e 64 níveis. No formato de modulação QAM, a informação é codificada tanto na amplitude quanto na fase da portadora, por isto este apresenta a maior eficiência espectral entre os formatos discutidos anteriormente. Devemos nos atentar ao fato de que quanto mais símbolos representados por um formato de modulação, menor vai ser a distância entre estes no espaço do diagrama de constelação (considerando a mesma potência máxima de modulação). Isto significa que o sistema é mais sensível a ruído e distorções, podendo ocasionar uma detecção errônea do nível de amplitude ou da fase. Tais esquemas de que combinam modulação de fase e de amplitude já eram presentes em sistemas coerentes de rádio frequência e de micro-ondas desde os anos 60 (CAMPOPIANO; GLAZER, 1962). Apesar de já terem sido alvo de estudos em sistemas ópticos nos anos 80 (KAZOVSKY, 1989), foi apenas a partir de 2000, com o avanço do processamento digital de sinais (*Digital Signal Processing*, DSP), que se fez possível sua aplicação comercialmente (LI, 2009).

Na seção do transmissor do sistema coerente, a modulação é realizada codificando informação no campo elétrico, com a liberdade de modular sua fase além de sua intensidade, de forma independente. Em configurações QAM, por exemplo, são combinados sinais modulados em amplitude, tanto para o eixo em fase quanto para o eixo em quadratura (tais eixos são defasados em 90°). Como consequência da modulação de fase, a transformada de Fourier do campo modulado não apresenta simetria hermitiana em torno a sua frequência central. Já na seção de recepção do sistema coerente, o receptor deve ser capaz de decodificar as amplitudes nos eixos em fase e em quadratura. Para este fim, emprega-se uma rede híbrida 2×4 de 90° e um oscilador local (*Local Oscillator*, LO), ou seja, um laser de onda contínua (*Continuous Wave*, CW). Com a combinação da rede híbrida com o LO, discrimina-se os componentes ortogonais e recupera-se a informação em amplitude e em fase (KIKUCHI, 2016). Maiores detalhes sobre esse processo serão discutidos na Seção 1.1.

Podemos classificar a detecção do sinal como sendo de três tipos, sendo elas a homódina, heteródina e intradina. A maior limitação da detecção homódina é a necessidade de sincronia de frequência e fase do LO com o sinal. Para isto, é necessária uma malha de captura de fase óptica (*Optical Phase-Locked Loop*, OPLL), o que aumenta

significativamente a complexidade do sistema (NORIMATSU; IWASHITA; SATO, 1990). Já no caso da detecção heteródina, requer-se que a diferença entre a frequência central do sinal e do LO seja maior do que a banda que o sinal ocupa. Essa diferença de frequências, denominada frequência intermediária (*Intermediate Frequency*, IF), apresenta o fato de que requer de uma banda de detecção excessivamente alta (KIKUCHI, 2016). De fato, a detecção heteródina requer uma banda de até 5 vezes maior largura do que a banda requerida pela detecção homódina, de acordo com (KAZOVSKY, 1985), o que impõe um desafio tanto na foto-detecção quanto no processamento dos dados. No caso de detecção intradina, devido a implementação mais simples, esta representa uma solução viável para receptores em sistemas coerentes ópticos. Neste tipo de detecção, o LO não precisa estar completamente em sincronia com o sinal. Diferenças de frequência e fase são compensadas via algoritmos de DSP (DERR, 1991).

A diversidade de amplitude e fase resulta em uma melhoria na eficiência espectral dos canais de transmissão, permitindo um aumento significativo na capacidade total do sistema. Sistemas coerentes digitais devem fazer uso de um DSP competente, o qual deve ser capaz de compensar efeitos lineares e não lineares (LI, 2009). Sendo assim, a compatibilidade com transmissão e recepção coerente é fundamental para um transceptor competitivo atual de alto desempenho. A integração de tais sistemas ópticos de alto desempenho se faz conveniente devido sua compactidade e seu custo reduzido, sobretudo quando a manufatura é realizada em larga escala. Na próxima seção iremos discutir o conceito de fotônica integrada, e daremos um exemplo de implementação de um transceptor coerente digital em *chip*.

1.1 Fotônica Integrada para Sistemas de Comunicações Ópticas

Fotônica integrada se refere à dispositivos e circuitos ópticos integrados em um *chip*, em que estes componentes apresentam funções que permitem a manipulação de ondas de luz (radiação eletromagnética nas faixas de comprimentos de ondas de 500 nm até 2000 nm) (COLDREN; CORZINE; MASANOVIC, 2012; SIEW *et al.*, 2021a). Na área de telecomunicações, há diversas plataformas para a implementação de fotônica integrada, sendo algumas das plataformas mais comuns o niobato de lítio (LiNbO_3), o fosfeto de índio (InP), o arsenieto de gálio (GaAs) e o silício (Si) (SIEW *et al.*, 2021b). Como é de se

esperar, cada uma destas plataformas apresentam vantagens e desvantagens únicas. No desenvolvimento de projetos de fotônica integrada deve-se levar em consideração tanto o desempenho que a plataforma de fabricação oferece para uma função de interesse, quanto a disponibilidade de manufatura da *foundry* ou do laboratório.

A escolha da plataforma de silício para o projeto de dispositivos fotônicos é justificada, principalmente, devido a compatibilidade desta plataforma com o processo de manufatura de metal óxido-semicondutor complementar (*Complementary Metal-Oxide Semiconductor*, CMOS) (SIEW *et al.*, 2021a). Sendo um processo amadurecido, CMOS já vem sendo utilizado há décadas para produção de dispositivos eletrônicos, estes operando na faixa de frequências de RF e de micro-ondas (CHROSTOWSKI; HOCHBERG, 2015). Fotônica integrada em silício é alvo de estudo desde 1985, com a premissa de produzir guias de ondas em silício sobre isolante (*Silicon-On-Insulator*, SOI) (REED; HEADLEY; PNG, 2005). A característica do SOI é que, neste caso, a junção do silício é construído acima de um isolante elétrico. Sistemas fabricados em SOI encontram aplicações em microeletrônica devido à algumas propriedades elétricas desta plataforma como, por exemplo, a redução da capacitância parasita nos dispositivos (CELLER; CRISTOLOVEANU, 2003). No caso de aplicações em fotônica integrada, o guia de onda é composto de silício e envolvido por dióxido de silício (sílica). Devido ao alto contraste de índice de refração entre o silício (índice de refração de $n = 3,47$, em 1550 nm) e a sílica ($n = 1,444$ em 1550 nm), há um alto grau de confinamento dos modos propagantes no guia, o qual permite a implementação de dispositivos compactos.

Apesar de seus benefícios, a plataforma SOI ainda apresenta alguns desafios principalmente na fabricação de alguns dispositivos ativos como fontes de luz, fotodetectores e moduladores (THOMSON *et al.*, 2016). A emissão de luz em componentes de transmissão óptica se dá a partir de um bombeio elétrico, o qual é responsável por elevar o estado de energia dos átomos do material semicondutor. A excitação de um átomo causa uma elevação dos elétrons pertencentes à banda de valência para a banda de condução, gerando lacunas na banda de valência. Quando um elétron que se encontra na banda de condução se recombina com uma lacuna pertencente à banda de valência, no caso de recombinação radiativa, emite-se um ou vários fóton tal que a energia total é proporcional à diferença de energia entre ambas as bandas (AGRAWAL, 2014). O material composto de silício é demasiadamente ineficiente para emitir luz devido à sua característica de *band gap* indireto. Ou seja, em silício, o mínimo de energia pertencente à banda de condução se encontra em

diferentes níveis de momento de cristal em relação ao máximo de energia que a banda de valência possui. A solução mais tradicional para implementação de *lasers* em SOI é por fontes externas, apesar de haver soluções mais eficientes em estudo como, por exemplo, a integração monolítica de um material com um ganho óptico elevado, como germânio ou algum material semi-condutor composto de III-V (CHROSTOWSKI; HOCHBERG, 2015). No caso de fotodetectores, tem-se o fato de que silício é praticamente transparente para faixas de comprimento de onda maiores que 1100 nm, o que implica uma baixa eficiência ao operar nas bandas utilizadas para telecomunicações (THOMSON *et al.*, 2016). Há algumas técnicas para aumentar a absorção de fótons em silício, como implantação iônica, porém esta ainda não apresenta grande eficiência, dado que ainda é requerida altas tensões reversas de operação (DOYLEND; JESSOP; KNIGHTS, 2010). Outra solução é integração híbrida com materiais III-V (CHANG *et al.*, 2010), o que resulta em dispositivos eficientes porém de alta complexidade de fabricação. O método que mais se destaca é o de crescimento epitaxial de germânio, que resulta em uma alta eficiência e permite integração monolítica com a tecnologia CMOS. A estrutura de moduladores em SOI será descrita em detalhes na Seção 1.2.

Podemos classificar a integração de dispositivos fotônicos em duas categorias. Por um lado, a integração monolítica se refere ao caso em que todos componentes de um circuito fotônico são integrados em um mesmo substrato. Já a integração híbrida é o tipo em que diferentes componentes são fabricados em diferentes substratos, e então combinados de forma mecânica. Apesar da integração híbrida apresentar funcionalidades inéditas, sua integração é susceptível à falhas devido às conexões entre os diferentes elementos, apresentando sensibilidade ao desalinhamento por expansões térmicas e vibrações (HUNSPERGER, 2009). Com integração monolítica em CMOS é possível sua produção em massa, o que reduz significativamente o preço da fabricação de dispositivos.

Podemos citar um exemplo de circuito fotônico, o transceptor, tal qual faz utilização de diversos dispositivos integrados, tanto por integração monolítica como por integração híbrida. Transceptores são componentes que combinam a capacidade de inserir (transmissão) e extrair (recepção) informação em um ou diversos canais. Um transceptor óptico coerente é composto por diversos subsistemas mais simples, os quais são responsáveis pela geração e detecção do sinal. A Figura 2 representa o diagrama de blocos de um transceptor integrado coerente. O transceptor ilustrado possui diversidade de fase e de polarização, ou seja, além de ser capaz de codificar e decodificar informação em fase e em quadratura, o mesmo

é capaz de utilizar ambas polarizações ortogonais para transmissão de informações em paralelo. Tal técnica é denominada multiplexação de polarização (IVANOVICH *et al.*, 2018). Em fotônica integrada, geralmente, ambas as polarizações a serem manipuladas pelo sistema são denominadas modos transversal elétrico (*Transverse Electric*, TE) fundamental e transversal magnético (*Transverse Magnetic*, TM) fundamental. O modo TE consiste do campo elétrico estar oscilando na direção paralela ao plano do substrato, e utilizamos a notação que o estado de polarização está em X . O modo TM, por sua vez, oscila na direção ortogonal ao plano do substrato, então referimos que o estado de polarização deste está em Y . É válido observar que, devido ao fato do guia de ondas ser formado por dielétricos, e não por metal, não há modos puramente TE ou TM (WENG; LIN; WU, 2021). Entretanto, por simplicidade, iremos nos referir aos modos quasi-TE e quasi-TM como modos TE e TM. Na seção do transmissor, o bloco de DSP é encarregado da geração de uma certa sequência binária, o seu mapeamento para o formato de modulação desejado, a formatação do pulso e pré-compensação de impedimentos de transmissão. Os pulsos elétricos gerados são então transmitidos para os *drivers*, tais quais possibilitam o controle dos níveis de tensão requeridos para uma modulação eficiente. Do ponto de vista do domínio óptico, gera-se uma onda de luz de onda contínua a partir do diodo laser (*Laser Diode*, LD).

Para o caso do transmissor com diversidade de fase e polarização, a onda luminosa tem sua potência dividida em quatro ramos, idealmente com potências idênticas, a partir de duas sequências de divisores de potência (*Power Splitter*, PS), o que totaliza 3 unidades de PSs. Destes quatro pares, cada um dos ramos é inserido individualmente em um modulador de amplitude. Este tipo de modulador é capaz de modular tanto a intensidade quanto a fase do sinal. Quatro moduladores são utilizados para a modulação do sinal óptico, sendo um para representação da amplitude em fase e outro para amplitude em quadratura, este padrão se repete tanto para o estado de polarização na direção horizontal (em X) quanto para o estado de polarização vertical (em Y). Os defasadores de 90° são integrados na saída de apenas um modulador de cada par, com a finalidade de associar os símbolos no eixo de quadratura. Após a defasagem, as componentes em fase e em quadratura são combinadas por meio de um PS. Ambos os sinais modulados em quadratura apresentam um estado de polarização TE. Estes sinais são então processados por um divisor e rotor de polarização (*Polarization Splitter and Rotator*, PSR). O PSR primeiramente rotaciona o estado de polarização de um dos sinais, de forma a manter o primeiro em TE fundamental

e convertendo o segundo para TM fundamental. O PSR então combina os feixes de ambas polarizações para gerar o sinal multiplexado em polarização para ser transmitido.

Na seção de receptor ilustrado na Figura 2, as duas polarizações da onda incidente são demultiplexadas, e posteriormente a TM é convertida a TE através de um PSR. O *crossing* é um dispositivo comum em roteamento de sinais ópticos, pois este permite o cruzamento de dois guias de ondas evitando interferências. Um *crossing* é utilizado para rotear os sinais para as entradas das portas híbridas de 90° . A porta híbrida de 90° , por sua vez, mistura o sinal de entrada com o sinal do oscilador local em quatro componentes. Destas componentes, duas são relacionadas ao eixo em fase, defasadas em π , e as outras duas relacionadas ao eixo em quadratura, também com um defasamento de π entre estas (GUAN *et al.*, 2017). A porta híbrida resulta em quatro sinais de luz que serão inseridos em dois pares de fotodiodos. Após a fotodetecção, os amplificadores diferenciais são responsáveis por combinar as amplitudes de valores positivos e negativos (defasadas em 180°), tanto para os eixos em fase quanto em quadratura. O banco de amplificadores de transimpedância (*Transimpedance Amplifiers*, TIAs) possibilita a conversão de corrente elétrica para tensão elétrica enquanto amplifica o sinal. O sinal elétrico é amostrado por um conversor analógico para digital (*Analog to Digital Converter*, ADC) antes de ser processado pelo bloco de DSP do receptor. Efeitos lineares e não lineares, que apresentam impedimentos na detecção do sinal, são compensados pelo bloco de DSP. Este último bloco também recupera a informação transmitida.

Como é perceptível pela Figura 2, um transceptor óptico coerente é composto de diversos dispositivos. O funcionamento conjunto destes dispositivos irá definir as características do transceptor, como o tamanho, desempenho e consumo de energia. Um dos grupos de fabricantes que desenvolve estas especificações é a OIF (*Optical Internetworking Forum*) (OIF, 2021). Estas especificações incluem requisitos elétricos e eletro-ópticos, requisitos mecânicos e características operacionais do transceptor e dos dispositivos individuais que o compõe. O projeto de moduladores, dispositivo fundamental para transceptores, ainda apresenta alguns desafios, principalmente se tratando de sua integração na plataforma de silício, a qual está sujeita a certas limitações. Na Seção 1.2 discutiremos algumas das principais arquiteturas e plataformas de moduladores ópticos integrados.

1.2 Moduladores Ópticos Integrados

Como mencionado, moduladores eletro-ópticos são responsáveis pelo mapeamento da informação contida na onda de micro-ondas para a onda de luz. O método pioneiro e mais simples de modulação da onda óptica é denominado modulação direta. Desta maneira, a fonte de luz é simplesmente ligada (maior nível de intensidade) e desligada (menor nível de intensidade), o que limita a quantidade de *bits* para apenas dois por cada símbolo transmitido. A configuração de modulação direta é, simultaneamente, mais compacta e menos custosa. Entretanto, além das limitações em relação à baixa eficiência espectral obtida, IM-DD introduz gorjeio (*chirp*) no sinal, o que ocasiona distorções durante a

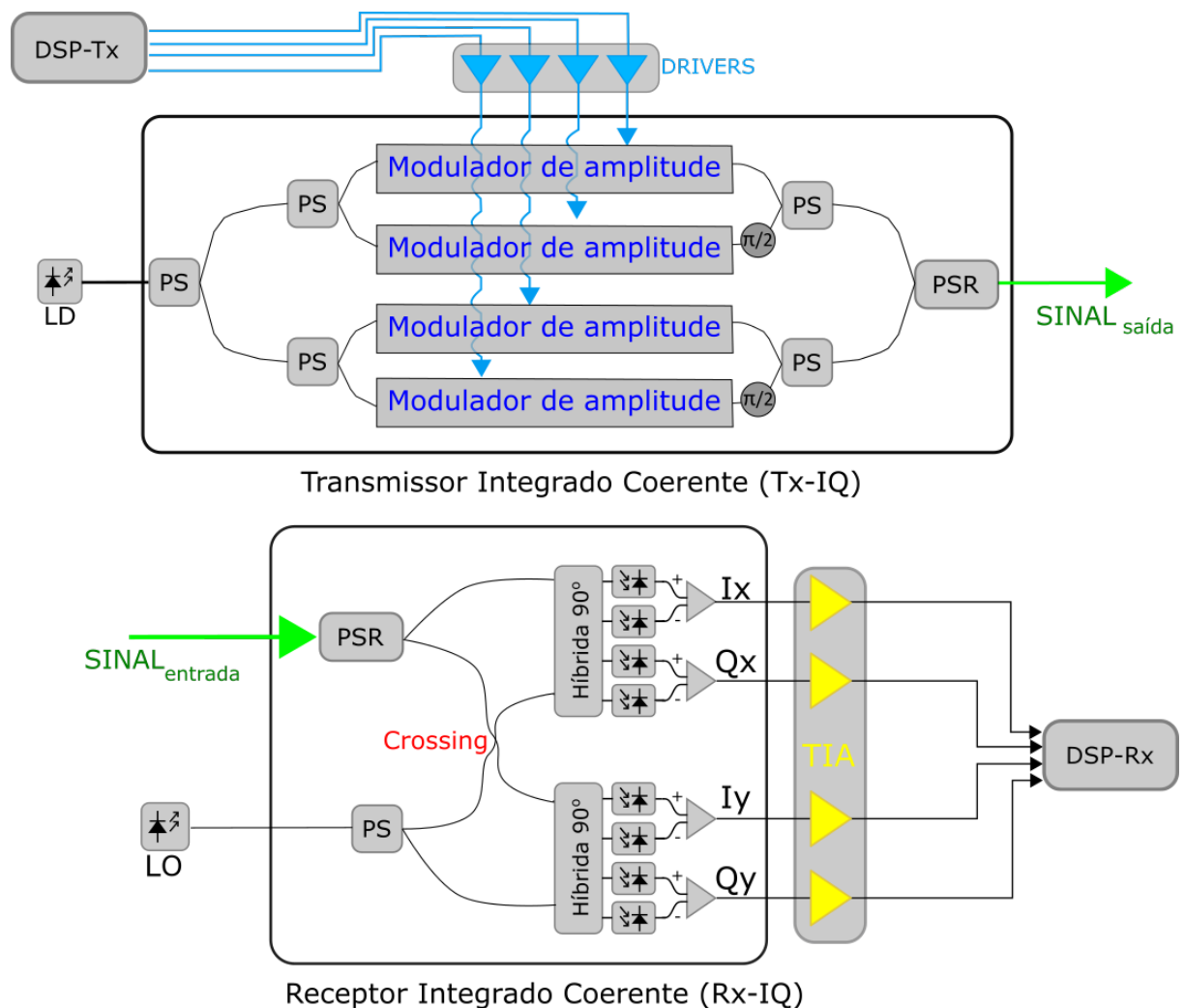


Figura 2 – Representação do transceptor integrado coerente por diagrama de blocos. No bloco superior temos a seção do transmissor e no bloco inferior temos a seção de recepção.

transmissão pela fibra (AGRAWAL, 2014). Em sistemas de alto desempenho é necessária a implementação de modulação externa, assim como ilustrada na Figura 2. Além do mais, por modulação externa, somos capazes de anular o gorjeio ou até mesmo ajustar o seu valor para a reversão de algumas distorções como a dispersão cromática (FIGUEIREDO; SOUZA; CHIUCHIARELLI, 2017). Atualmente os principais desafios na modulação óptica, especialmente na plataforma de silício, se encontram no desenvolvimento de moduladores que apresentam baixo gorjeio, alta eficiência de modulação, larga banda eletro-óptica (*Electro-Optic Bandwidth, BW_{EO}*) e baixas perdas (RAHIM *et al.*, 2021).

Em silício, a modulação (externa) é realizada por meio de mudanças no índice de refração complexo (\bar{n}) do material, tal qual é dado pela seguinte equação:

$$\bar{n} = n + j \cdot \kappa, \quad (1)$$

em que n é a componente real do índice de refração e κ é o coeficiente de absorção. A atenuação α pode ser obtida a partir do coeficiente de absorção, e esta relação é dada por:

$$\alpha = \frac{4\pi\kappa}{\lambda}, \quad (2)$$

em que λ é o comprimento de onda de operação no meio.

Uma forma de realizar variações do índice de refração é a partir do efeito termo-óptico, que modula o índice de refração por meio de variações na temperatura do material. A diferença de temperatura é induzida por aquecedores térmicos controlados, geralmente por uma corrente externa aplicada. Entretanto, principalmente na plataforma de silício, este efeito apresenta uma dinâmica lenta, o que limita a utilização de tal efeito apenas para chaveamentos, comutações e também para sintonização de moduladores (COCORULLO; RENDINA, 1992). No caso de moduladores em silício de alto desempenho a variação do índice de refração é resultante da aplicação de um campo elétrico externo (efeito eletro-óptico). Mudanças na componente real de \bar{n} resulta em mudanças na velocidade de propagação da onda óptica (eletro-refração), assim como mudanças na componente imaginária de \bar{n} resulta em mudanças na absorção do material (eletro-absorção) (REED *et al.*, 2010).

Alguns dos efeitos eletro-ópticos mais comumente explorados são os efeito Pockels, efeito Kerr e os efeito de eletro-absorção Franz-Keldysh e o Efeito Stark de confinamento quântico (*Quantum-Confined Stark Effect, QCSE*) (REED *et al.*, 2010). O efeito Poc-

kels ocasiona uma variação no índice de refração de forma linearmente proporcional ao módulo do campo elétrico externo aplicado (AGRAWAL, 2014). Entretanto, em um cristal centrossimétrico como o silício, o efeito Pockels apresenta uma eficiência relativamente baixa (REED *et al.*, 2010). Há algumas técnicas de dopagem que quebram a simetria do silício com a introdução de materiais como titanato de bário (OLIVARES *et al.*, 2019) ou nitrato de silício (CHMIELAK *et al.*, 2011), o que permite um maior aproveitamento do efeito Pockels. Entretanto, tais técnicas aumentam significativamente a complexidade de fabricação, impedindo a integração com o processo CMOS. O efeito Kerr, por sua vez, é similar ao efeito Pockels, porém a variação do índice de refração é proporcional ao quadrado do módulo do campo elétrico aplicado (efeito não linear) (AGRAWAL, 2014). No silício, o efeito Kerr pode ser excitado, porém não de forma intensa o suficiente para sua aplicação em moduladores de alta velocidade (SOREF; BENNETT, 1987). Por outro lado, o efeito de Franz-Keldysh é resultante de uma perturbação nas bandas de energia causada pela aplicação de um campo elétrico externo no meio. Tal perturbação resulta em um tunelamento quântico entre as bandas de valência e de condução (SOREF; BENNETT, 1987). Apesar deste efeito, em silício, apresentar tanto eletro-absorção quanto eletro-refração, a variação de índice de refração e de absorção não é intensa o suficiente para a obtenção de índices de modulação satisfatórias. Para fins de comparação, em (SOREF; BENNETT, 1987) mediu-se experimentalmente as variações de índice de refração sob efeito de um campo elétrico externo com amplitude de $1 \cdot 10^5$ V/cm. Com o efeito Kerr, obteve-se uma variação de índice de $\|\Delta n\| = 1 \cdot 10^{-6}$ em um comprimento de onda de $\lambda = 1,3 \mu\text{m}$. Com o efeito Franz-Keldysh, obteve-se um $\|\Delta n\| = 1,3 \cdot 10^{-6}$ em um comprimento de onda de $\lambda = 1,07 \mu\text{m}$ (ponto em que ocorre o máximo de variação de índice de refração).

Quando se trata de silício, o método mais eficiente para variação do índice de refração é a partir da utilização do efeito de dispersão de plasma (*Plasma Dispersion Effect*, PDE). O PDE se baseia no fato de que o índice de refração de um material depende da densidade de portadores livres no mesmo. Desta forma, pode-se variar o índice de refração ao injetar ou ao extrair portadores do material. Tais variações de portadores também causam uma variação na absorção por portadores livres (*Free Carriers Absorption*, FCA) (SOREF; BENNETT, 1987). Ambos os efeitos PDE e FCA, podem ser quantificados aproximadamente pelo modelo de Drude (TAKENAKA; TAKAGI, 2012). Este modelo nos diz que a variação de índice de refração e de absorção não dependem apenas da quantidade de portadores livres, mas são também dependentes da variação da frequência de plasma no

meio. A variação do índice de refração é inversamente proporcional às massas efetivas de condução dos elétrons e das lacunas. A variação da absorção, além de ser dependente do inverso das massas efetivas de condução dos elétrons e das lacunas, também é inversamente proporcional às mobilidades de elétrons e de lacunas. No presente trabalho, utilizaremos o modelo empírico de Soref e Bennett (SOREF; BENNETT, 1987) para modelar o PDE, e mais detalhes sobre este serão discutidos na Subseção 2.3.2.

Enquanto os efeitos Pockels e Kerr produzem variações de índice de refração na ordem de $1 \cdot 10^{-6}$, com o PDE é possível obtermos um $\|\Delta n\| = 1,5 \cdot 10^{-3}$ com uma injeção ou depleção de portadores de 10^{18} portadores/cm³ (SOREF; BENNETT, 1987). Este fato mostra o motivo do PDE ser o método mais utilizado para moduladores de alto desempenho em silício, e isto justifica sua utilização no desenvolvimento do modulador no presente trabalho.

Podemos prever o funcionamento do modulador a partir de alguns parâmetros de desempenho do mesmo, os quais serão discutidos, em sua maioria, no Capítulo 2. Em relação à presente seção, limitaremos-nos em definir apenas três das mais importantes, sendo elas a BW_{EO} , a tensão de meia-onda (V_π) e as perdas de inserção (*Insertion Loss*, IL). Definimos a BW_{EO} como sendo o valor de frequência em que o espectro de transmissão do modulador, do ponto de vista da resposta óptica resultante de uma entrada elétrica, decai pela metade (-3 dB), considerando uma análise de pequenos sinais. A V_π é definida como a tensão necessária que se deve aplicar no modulador para causar uma mudança de fase de 180° (π) na portadora óptica. A IL contabiliza a atenuação óptica causada pela inserção do dispositivo no sistema fotônico, levando em consideração absorção, reflexão, espalhamento, entre outras (FENG *et al.*, 2012). O equacionamento destes parâmetros, assim como uma explicação mais aprofundada, será discutido no Capítulo 2.

Diversas estruturas têm sido propostas e analisadas para a modulação do sinal óptico, tais quais baseadas nos efeitos de eletro-absorção e eletro-refração descritos acima. Podemos classificar os tipos mais comuns de moduladores de acordo com seu funcionamento, sendo estes baseados em eletro-absorção, em ressonância e em interferometria. Nos próximos parágrafos será realizada uma discussão a respeito das diferenças mais relevantes entre estes dispositivos. É válido ressaltar que as características destacadas se referem a moduladores fabricados na plataforma de silício, pois estes que serão tratados no presente trabalho.

Moduladores baseados em eletro-absorção são, como o próprio nome alega, moduladores de intensidade que operam ao variar a absorção do meio propagante. Dois dos efeitos

mais utilizados em moduladores de eletro-absorção são o efeito de Franz-Keldysh e QCSE. O QCSE apresenta uma variação mais abrupta na absorção em relação ao Franz-Keldysh, o que significa uma maior eficiência de modulação (MILLER, 2009). Entretanto, além de ser necessária a integração de um segundo material em conjunto com o silício para excitação do QCSE, ainda a fabricação de moduladores baseados neste efeito é demasiadamente complexa.

Moduladores de eletro-absorção são capazes de atingir altos valores de BW_{EO} e de eficiência de modulação, porém possuem a desvantagem de causar alta IL e introduzir gorjeio significativo no sinal. Um estudo dimensionou e fabricou um modulador baseado em eletro-absorção formado por silício integrado com germânio em heteroestrutura, assim como ilustrado na Figura 3(a). Tal modulador utiliza o efeito de Franz-Keldysh como base para sua operação. Os resultados obtidos foram de 56 GHz de BW_{EO} , 4 V de V_{π}^1 e 10,6 dB de perdas (MASTRONARDI *et al.*, 2018). Apesar de apresentarem uma larga BW_{EO} e alta eficiência de modulação, verifica-se que este tipo de moduladores introduz altas perdas ópticas. Além do mais, tal estrutura não permite a modulação de fase do sinal, apenas da intensidade, e sua dificuldade de fabricação na plataforma de silício limita sua integração.

Moduladores baseados em ressonância geralmente são construídos a partir de micro-anéis ressonantes, como ilustrado na Figura 3(b). Micro-anéis são estruturas fechadas, geralmente circulares, que são dispostas ao lado do guia de onda. Devido à proximidade

¹ Neste caso, a V_{π} refere-se à tensão necessária para excursionarmos entre o mínimo e o máximo de intensidade do sinal modulado.

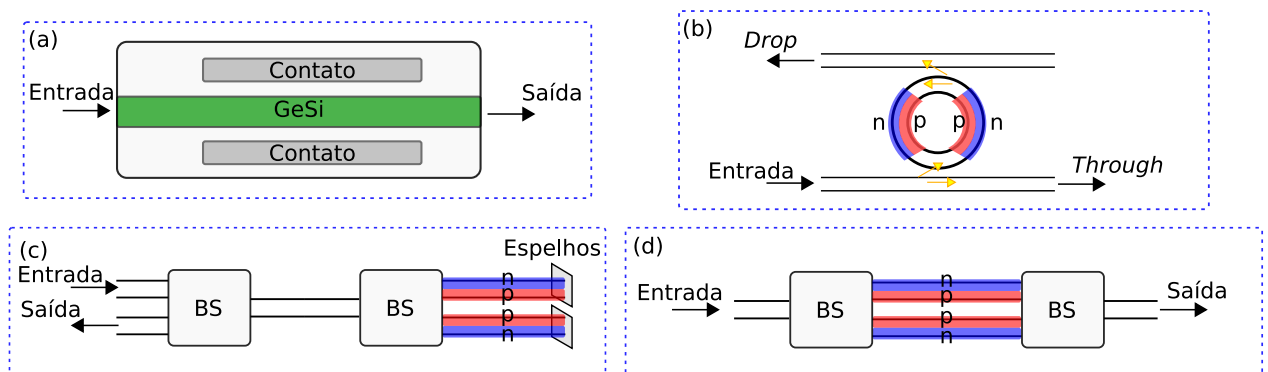


Figura 3 – Diagrama esquemático ilustrando moduladores integrados baseados na plataforma de silício. (a) Modulador de eletro-absorção. (b) Moduladores de micro-anel. (c) Moduladores de Michelson. (d) Moduladores de Mach-Zehnder.

com o guia, ocorre acoplamento da luz para a estrutura em anel, tal qual é capaz de confinar luz, dado que as condições de ressonância sejam obedecidas. De forma geral, dentro das condições de ressonância, toda a energia óptica, idealmente, é confinada dentro do anel, devido a interferências construtivas de sucessivas voltas dadas pela luz. Fora das condições de ressonância, a luz dentro do anel sofre interferência destrutiva a cada volta dada, fazendo com que a luz não seja armazenada na estrutura. Como resultado disto, um ressoador em anel possui alta seletividade em comprimento de onda, funcionando como um filtro óptico de banda estreita (FENG *et al.*, 2012). Operando próximo do comprimento de onda de ressonância, há uma alta sensibilidade à fase da portadora na transmissão do micro-anel. Explorando esse efeito, é possível a fabricação de moduladores eficientes a partir da exploração do efeito de dispersão de plasma no guia que compõe o anel (CHROSTOWSKI; HOCHBERG, 2015). É válido observar que, apesar do funcionamento destes dispositivos ser também baseado em interferência, por conveniência, os classificamos como moduladores ressonantes, pois estes são compostos de cavidades ressonantes. Na fabricação de moduladores de micro-anel utiliza-se uma junção PN, de forma que a região de depleção se sobrepõe ao modo óptico no interior do micro-anel. Alterações na tensão aplicada à junção PN causam variações do índice de refração, resultando em um deslocamento do espectro de transmissão do modulador (FENG *et al.*, 2012). A Figura 3(b) ilustra o esquemático de um modulador de micro-anel do tipo *add-drop*, ou seja, há um guia de onda adicional que serve para extrair luz do anel. Em um trabalho recente, em que a otimização da junção PN é realizada, o modulador baseado em micro-anel foi capaz de obter resultados que apresentam uma BW_{EO} de mais de 67 GHz, porém com uma V_{π} estimada superior a 100 V (ZHANG *et al.*, 2020). Um outro estudo obteve uma BW_{EO} de 30 GHz, uma V_{π} de apenas 2 V e uma perda de inserção de 7 dB (XUAN *et al.*, 2014). As maiores vantagens dos moduladores baseados em micro-anel ressonantes em relação aos moduladores de Mach-Zehnder são: o tamanho compacto, menor consumo de energia e, geralmente, uma melhor eficiência de modulação. Entretanto, os moduladores em micro-anel possuem algumas características negativas que evitam a sua adoção ao considerar aplicações em sistemas de baixo custo e alto desempenho, sendo estas uma maior complexidade de fabricação, um compromisso entre BW_{EO} e eficiência de modulação geralmente menor, uma alta sensibilidade térmica e uma introdução significativa de gorjeio no sinal (FENG *et al.*, 2012; CHROSTOWSKI; HOCHBERG, 2015).

Para entendimento dos moduladores baseados em interferometria, devemos primeiramente introduzir o dispositivo passivo que serve como base para o *design* de tais moduladores. Um interferômetro de Mach-Zehnder (*Mach-Zehnder Interferometer*, MZI) consiste de um dispositivo que faz proveito do fenômeno de interferência entre ondas eletromagnéticas. Primeiramente, a onda provinda de um guia tem sua potência dividida entre dois ramos, os quais são denominados braços do MZI. A divisão de potência pode ser realizada por qualquer PS, como um acoplador direcional, *Y-branch* ou um interferômetro multi-modo (*Multimode Interferometer*, MMI). Após a propagação pelo dispositivo, as ondas providas de ambos os braços são recombinaadas (CHROSTOWSKI; HOCHBERG, 2015). Considerando uma condição ideal, em que o PS tem uma razão de divisão perfeitamente balanceada e em que não há perdas na propagação através do MZI, a intensidade da onda de saída será dependente apenas do desfasamento causado nos braços. Quando ambas ondas estão em fase, haverá interferência construtiva e a potência de saída será a mesma da entrada. Quando há uma defasagem de π entre as ondas, haverá interferência destrutiva, o que anulará a onda de luz.

Moduladores de Michelson são estruturas que têm sua operação baseada no efeito de interferência do MZI. A particularidade destes se dá no fato de que a modulação é explorada com o dobro de eficiência para um dado comprimento. Isto ocorre pois um espelho é posicionado na extremidade de cada braço. Como ilustrado na Figura 3(c), o feixe de luz introduzido no dispositivo é ramificado em dois braços e então recombinaados. A variação de fase realizada em ambos os braços será convertida em modulação de amplitude como consequência do efeito de interferência. Após a propagação pelos braços, o sinal é então recombinaado e extraído do dispositivo. O comprimento de cada braço modulante pode ser reduzido pela metade, em relação ao MZM, enquanto a eficiência é mantida a mesma, apresentando então um tamanho mais compacto. Apesar desta qualidade ser conveniente para o nosso objetivo, moduladores de Michelson, em geral, possuem maiores perdas ópticas e menores bandas eletro-ópticas comparados com MZMs. Podemos citar o desempenho obtido por alguns trabalhos. Em (PATEL *et al.*, 2014), a V_π atingida foi de 4 V, porém a banda de 23,5 GHz e as perdas de 8,8 dB não são capazes de atender certas demandas. Em outro trabalho, em que se obteve uma eficiência de modulação ainda superior, foi obtido uma V_π de 3 V, porém apenas 12 GHz de banda foram mensurados, contabilizando com uma perda de 8 dB (Li *et al.*, 2013).

Um modulador de Mach-Zehnder (*Mach-Zehnder Modulator*, MZM), ilustrado pela Figura 3(d), opera a partir da variação do índice de refração de um ou de ambos os braços do MZI. A possibilidade de modificar a velocidade da luz nos braços do MZM nos permite manipular a condição de interferência, ou seja, MZM converte modulações de fase em modulações de amplitude. Os maiores inconvenientes apresentados por MZMs é que estes são menos compactos em relação aos outros tipos de moduladores, e estes também possuem um elevado consumo de energia. Porém, a utilização de MZMs no presente trabalho é justificada por algumas características relevantes que estes apresentam. Uma das maiores vantagens ao utilizar MZMs é a larga BW_{EO} que este é capaz de alcançar. Além disso, MZMs são de relativamente simples fabricação, e também apresentam uma alta tolerância às eventuais variações de temperatura, dispensando a utilização de reguladores térmicos. O Capítulo 2 é dedicado a discutir, detalhadamente, a respeito de MZMs fabricados na plataforma de silício.

Uma grande vantagem de moduladores baseados em interferometria é que estes nos permite modular tanto a amplitude quanto a fase do sinal, propriedade importante para implementação de formatos de modulação multi-níveis de amplitude e de fase, como o QAM. Os moduladores que integram o transceptor com diversidade em fase e em polarização representado na Figura 2 devem ser capazes de modular a fase do sinal simultaneamente com sua amplitude. Na Seção 2.2 mostramos como um MZM é capaz de agir como um modulador tanto de amplitude como de fase.

A Figura 4 ilustra o diagrama de um MZM de estrutura IQ. Podemos interpretar tal estrutura como sendo um MZM composto de mais dois MZMs internos, sendo estes dispostos um em cada braço do MZM externo. Como exemplo de seu funcionamento, o formato de modulação 4-QAM ilustrado na Figura 1(d) pode ser facilmente gerado pela estrutura IQ. Primeiramente, definimos a resposta de cada modulador como sendo de -1 e 1 . E então, considerando um deslocamento de fase de π no braço inferior da Figura 4, temos que as quatro possíveis combinações na saída do MZM composto formam um diagrama de constelação com os valores $\pm(1 \pm j)/\sqrt{2}$. O mesmo conceito pode se estender para os formatos de modulação mais complexos, como os da Figura 1(e-f), ao considerar mais níveis de amplitude em cada modulador interno.

Ao se tratar do desenvolvimento de moduladores eletro-ópticos de alto desempenho, em específico o MZM, é válido também destacarmos os atributos negativos e positivos apresentados por cada plataforma. MZMs desenvolvidos nas plataformas de LiNbO_3 e InP

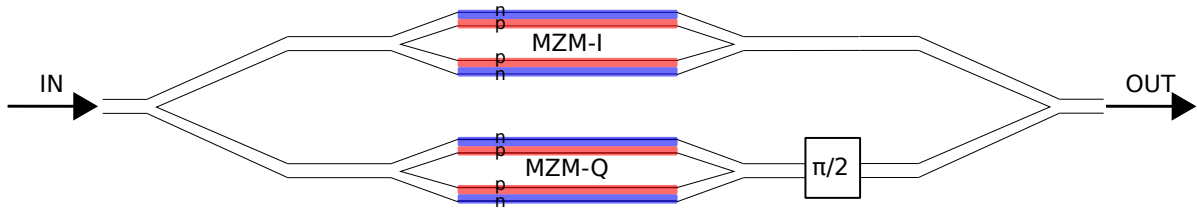


Figura 4 – Estrutura de um MZM-IQ.

possuem algumas figuras de mérito superiores que os construídos em silício. Comercialmente, MZMs baseados em LiNbO_3 e InP apresentam uma V_π entre 1,5 V e 4 V, IL entre 3 dB e 5 dB e BW_{EO} superiores a 35 GHz (JACQUES *et al.*, 2018). MZMs comerciais construídos na plataforma de silício apresentam uma V_π de 5 V a 6 V, IL superiores a 6 dB e larguras de banda entre 30 GHz e 35 GHz (JACQUES *et al.*, 2018). Alguns dos atributos positivos ao fabricar moduladores na plataforma de silício, além do processo de fabricação ser mais amadurecido, é o fato deste possuir uma menor sensibilidade a variações térmicas em relação à plataforma de InP , e de também oferecer dispositivos mais compactos em geral (WENG; LIN; WU, 2021). Uma desvantagem de moduladores em silício e também em InP em relação a moduladores baseados em LiNbO_3 é o comportamento não linear da variação do índice de refração em relação à variação de tensão aplicada no material. Tal não linearidade acrescenta gorjeio no sinal, o que limita a faixa de operação do modulador. Esta não linearidade pode ser parcialmente compensada por algoritmos de DSP (GUTIERREZ *et al.*, 2013). Por outro lado, no caso de LiNbO_3 , *drivers* de alta potência são requeridos (WANG *et al.*, 2018). Uma outra plataforma que já foi também explorada para fabricação de moduladores integrados é a baseada em GaAs (GRIFFIN *et al.*, 2003; STAMPOULIDIS *et al.*, 2013), atingindo uma banda de 22,5 GHz e uma V_π de 3V em alguns experimentos (STAMPOULIDIS *et al.*, 2013). O ponto que apresenta maior justificativa da escolha do SOI como plataforma para construção do MZM é sua compatibilidade com o processo de manufatura CMOS.

A Tabela 1.2 compila diversos trabalhos que foram realizados entre 2008 e 2021 com o intuito de otimizar MZMs na plataforma de silício. Em outras palavras, tais trabalhos representam o estado da arte em MZMs alcançados em sequência cronológica de ano de publicação. Os trabalhos citados, em sua maioria, otimizaram parâmetros físicos de forma a superar eventuais limitações. Os parâmetros a serem comparados na Tabela 1.2 são a BW_{EO} , a V_π e a IL. Percebe-se que a BW_{EO} apresenta grandes melhorias conforme a data de publicação é mais recente. Como o silício é uma plataforma com efeitos eletro-ópticos

limitados, o maior desafio é manter uma V_π baixa enquanto se estende a banda. Também podemos observar que há um compromisso entre a banda e a V_π . Ao variarmos alguns parâmetros de forma a ampliarmos a BW_{EO} do MZM, podemos observar também um aumento na V_π . Como podemos perceber, o trabalho (LIU *et al.*, 2008), por exemplo, apresenta uma banda relativamente alta para época em que foi publicado, porém possui uma V_π proibitivamente alta na prática. Como indicado na Tabela 1.2, o MZM apresentado em (JACQUES *et al.*, 2020b) foi desenvolvido para operar na banda O, enquanto que os demais operam na banda C. Em nosso trabalho, demos preferência para operação na banda C, pois esta é a mais utilizada em telecomunicações hoje em dia, apresentando a menor perda de propagação em fibras entre as diferentes bandas (TAMURA *et al.*, 2017). Na Tabela 1.2, o termo L_{ps} se refere ao comprimento físico do MZM.

Tabela 1 – Comparativo entre MZMs em silício reportados em diversos trabalhos, em ordem cronológica de publicação. A tabela inclui a BW_{EO} , a V_π , a IL, figura de mérito definida e a modulação utilizada.

Referência	BW_{EO} [GHz]	V_π [V]	IL [dB]	Modulação
(LIU <i>et al.</i> , 2008)	20	13,33	6,4	OOK
(PARK <i>et al.</i> , 2009)	7	12	6,6	OOK
(WATTS <i>et al.</i> , 2010)	8	5	3,6	OOK
(CHEN; PETERS; BOWERS, 2011)	27	4,8	3	OOK
(BAEHR-JONES <i>et al.</i> , 2012)	10	6	8,7	OOK
(STRESHINSKY <i>et al.</i> , 2013)	18,5	1,5	5,5	OOK
(XU <i>et al.</i> , 2014)	35	5,3	3,3	OOK
(PATEL <i>et al.</i> , 2015)	41	11	3,8	PAM-4
(ZHOU <i>et al.</i> , 2016)	33	4,95	9,5	BPSK
(PANTOUVAKI <i>et al.</i> , 2017)	27	2,5	2,2	OOK
(ZHENG <i>et al.</i> , 2018)	30	2,5	2	PAM-4
(SEPEHRIAN <i>et al.</i> , 2019)	36	7,3	6,8	16-QAM
(SAMANI <i>et al.</i> , 2019)	40	7,4	6,3	PAM-4
(JAFARI <i>et al.</i> , 2019)	40	$1,5 \text{ V}\cdot\text{cm}$ ($L_{ps} = 56 \mu\text{m}$)	1	OOK
(DENIEL <i>et al.</i> , 2019)	18	4	5	PAM-4
(HINAKURA; ARAI; BABA, 2019)	30	5,2	6	OOK
(ZHOU <i>et al.</i> , 2019)	54 [6 dB] & 37 [3 dB]	7,3	6,3	16/64-QAM
(JACQUES <i>et al.</i> , 2020b) (Banda O)	47	5,4	5,7	PAM-8
(ZHOU <i>et al.</i> , 2020)	50 [6 dB]	6,3	5,6	16/64-QAM
(JACQUES <i>et al.</i> , 2020a)	45	10	3,4	PAM-8
(ZHANG <i>et al.</i> , 2021)	40	≈ 14 ($2,1 \text{ V}\cdot\text{cm}$, $L_{ps} = 1,5 \text{ mm}$)	6,125	SSB PAM-4
(MOSHAEV; LEIBIN; MALKA, 2021)	34,11	1,63	1,45	PAM-4
(JAFARI; SHI; LAROCHELLE, 2021)	17,6	2,33	17	OOK

1.3 Justificativa

Como citado anteriormente, a plataforma de silício apresenta a maior integrabilidade de MZMs em sistemas fotônicos práticos. Entretanto, como podemos perceber pelos trabalhos atuais, o desempenho dos moduladores nesta plataforma estão atingindo um platô. Por um lado, MZMs possuem diversos parâmetros geométricos e de operação que

influenciam no desempenho final do dispositivo, o que dificulta a otimização conjunta de todos estes fatores. Por outro lado, simulações eletromagnéticas possuem uma complexidade computacional elevada, o que dificulta ainda mais o *design* do modulador. E além disso, o compromisso entre eficiência de modulação, BW_{EO} e IL impõe um desafio adicional na otimização.

Algoritmos heurísticos são eficientes na busca de uma solução para o problema de otimização de *design* de dispositivos como o MZM. Porém, levando em conta que estes ainda requerem uma alta quantidade de iterações e que modelos de simulação eletromagnéticas apresentam um alto custo computacional, a otimização de diversos parâmetros pode ser difícil. Após uma pesquisa bibliográfica, não encontramos nenhum trabalho que otimizou um MZM na plataforma de SOI utilizando algoritmos heurísticos. No melhor do nosso conhecimento, esta é a primeira vez que se otimiza um MZM utilizando ANNs.

1.4 Contribuições do Trabalho

Neste trabalho apresentamos um método de otimização de um MZM integrado em fotônica de silício. Este método é baseado em redes neurais e otimização heurística, e tem como finalidade encontrar o limite possível de desempenho destes dispositivos. Primeiramente, desenvolvemos uma rede neural artificial (*Artificial Neural Network*, ANN) profunda de forma a modelar o mapeamento realizado pelo simulador eletromagnético. Consideramos oito parâmetros de *design*, sendo seis destes relacionados às dimensões transversais da junção PN, um deles é o comprimento do modulador e o último é a tensão de polarização aplicada. Já em relação aos parâmetros de desempenho, consideramos a BW_{EO} , a IL e a V_{π} . O modelo treinado nos permitirá inferir as métricas de desempenho do modulador a partir dos parâmetros de *design* muito mais rapidamente que o modelo de simulação. Em outras palavras, idealmente, a ANN tem a mesma função que o simulador eletromagnético, porém com a possibilidade de obtenção de resultados com um intervalo de tempo reduzido em diversas ordens de magnitude. O único custo adicional é o da obtenção do conjunto de dados de treinamento, entretanto este representa um custo de uma única vez. A otimização, que foi realizada a partir de um algoritmo heurístico, utiliza a ANN como forma de avaliar os MZMs candidatos. O algoritmo heurístico a ser utilizado aqui é a evolução diferencial (*Differential Evolution*, DE). Com esta configuração foi possível

obter MZMs de alto desempenho, e também nos permitiu analisar o compromisso entre os diferentes parâmetros de desempenho do MZM. Moduladores com BW_{EO} maior do que 40 GHz foram obtidos pela otimização por DE, enquanto estes apresentaram valores de V_{π} inferiores a 6,5 V e valores de IL inferiores a 4 dB, como apresentaremos na Seção 5.2.

1.5 Organização do Documento

O restante do documento é organizado da seguinte maneira: no Capítulo 2 discutimos a teoria e modelagem relacionadas ao MZM, no Capítulo 3 discutimos os algoritmos de otimização que serão utilizados no trabalho, no Capítulo 4 discutiremos a metodologia desenvolvida, no Capítulo 5 discutiremos os resultados obtidos e, por fim, no Capítulo 6 discutiremos nossas conclusões.

2 Fundamentos de Moduladores de Mach-Zehnder

No presente capítulo apresentamos os MZMs integrados na plataforma de silício. Na Seção 2.1 apresentamos o MZI, dispositivo passivo que serve como estrutura base para implementação de MZMs. Na Seção 2.2 discutimos sobre os MZMs, assim como suas configurações de alimentação e figuras de mérito. A Seção 2.3 aborda o método de deslocamento de fase geralmente utilizado em MZMs na plataforma de SOI. Por fim, na Seção 2.4, iremos apresentar como é realizada a alimentação da onda de RF no modulador.

2.1 Interferômetros de Mach-Zehnder

Para a interferometria ser realizada em um MZI, primeiramente um feixe de luz é dividido entre dois guias de ondas. Estes guias de onda são denominados braços do MZI. As ondas propagadas através destes são recombinadas para gerar uma onda de saída. O resultado da combinação (interferência) dos feixes tem seu comportamento dependente da diferença de fase das ondas na saída dos dois braços, assim como da intensidade das mesmas. A Figura 5 ilustra o esquemático de um MZI. A intensidade da onda de entrada, considerando propagação monomodo, é dada por I_{in} , e possui um campo elétrico de amplitude E_{in} , dado que $E_{in} = \sqrt{I_{in}}$. O braço superior e inferior terão, considerando um divisor de potência de 3 dB ideal (divide a potência uniformemente e sem perdas), uma amplitude de $E_{in1} = E_{in2} = E_{in}/\sqrt{2}$ como entrada (CHROSTOWSKI; HOCHBERG, 2015). Na saída dos guias de onda, temos os campos com amplitudes E_{out1} e E_{out2} , que podem ser dadas por:

$$E_{outi} = \frac{E_{in}}{\sqrt{2}} \exp\left(-j\beta_i L_i - \frac{\alpha_i}{2} L_i\right), \quad (3)$$

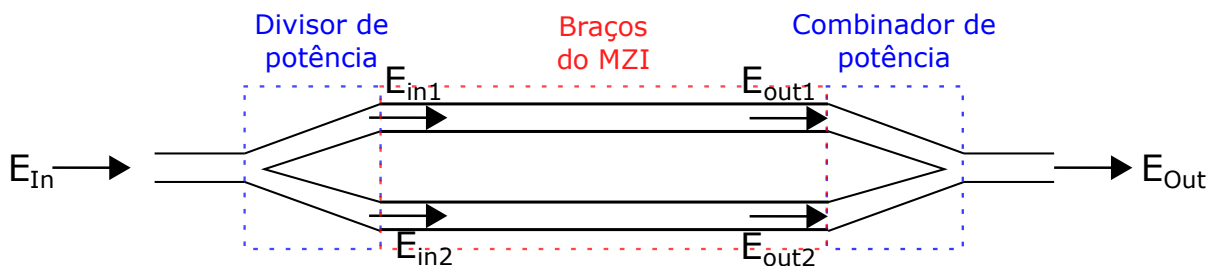


Figura 5 – Esquemático de um interferômetro de Mach-Zehnder.

em que $i = 1$ se refere ao braço superior e $i = 2$ ao braço inferior, $\beta_i = 2\pi n_i/\lambda$ é a constante de fase, α_i é o coeficiente de atenuação, L_i o comprimento e n_i o índice de refração, considerando o i -ésimo braço do MZI. Definimos a atenuação como $\alpha/2$ pois a Equação 3 se trata do campo elétrico, e não da intensidade. Na saída do MZI um combinador de potência é utilizado. Assumimos também que o combinador é ideal, apresentando uma combinação perfeita e sem perdas ópticas. A amplitude complexa do campo obtido na saída deste combinador, assumindo que os campos interferentes apresentam o mesmo estado de polarização, é dada por:

$$E_{out} = \frac{E_{in}}{2} \left[\exp(-j\beta_1 L_1 - \frac{\alpha_1}{2} L_1) + \exp(-j\beta_2 L_2 - \frac{\alpha_2}{2} L_2) \right]. \quad (4)$$

Esta equação mostra como um MZI afeta tanto a fase quanto a amplitude da luz que se propaga no dispositivo (AGRAWAL, 2014). Para obtermos a intensidade, basta elevarmos a magnitude do campo complexo de saída do MZI ao quadrado, o que resulta em:

$$I_{out} = \frac{E_{in}^2}{4} \left| \exp(-j\beta_1 L_1 - \frac{\alpha_1}{2} L_1) + \exp(-j\beta_2 L_2 - \frac{\alpha_2}{2} L_2) \right|^2, \quad (5)$$

A partir de manipulações trigonométricas e, desconsiderando as perdas para simplificar o resultado, podemos re-escrever a equação que descreve a intensidade de saída como:

$$I_{out} = \frac{I_{in}}{2} [1 + \cos(\beta_1 L_1 - \beta_2 L_2)]. \quad (6)$$

Da Equação 6, a função de transferência pode ser simplesmente descrita por:

$$T_m = \frac{I_{out}}{I_{in}} = \frac{1}{2} [1 + \cos(\beta_1 L_1 - \beta_2 L_2)] = \frac{1}{2} [1 + \cos(\phi_1 - \phi_2)], \quad (7)$$

em que $\phi_1(\phi_2)$ é a fase acumulada pelo braço 1(2). A Equação 7 nos mostra que a função de transferência do MZI depende do comprimento elétrico dos braços que, por sua vez, dependem do comprimento físico, do comprimento de onda e do índice de refração.

A Figura 6 ilustra a saída do MZI em relação à entrada, tanto para amplitude quanto para intensidade, desconsiderando perdas e em escala linear. Ao considerarmos apenas o desfasamento entre os braços como dado na Equação 7, em que $\Delta\phi = |\phi_1 - \phi_2|$, podemos analisar o efeito de interferometria de forma mais simplificada. Percebemos que quando não há desfasamento ou quando este é múltiplo de 2π , as ondas provenientes de

ambos os braços se combinam de forma construtiva, gerando uma intensidade igual a da onda de entrada do MZI, denominados pontos de pico. Porém, quando o desfasamento é múltiplo ímpar de π , há uma interferência destrutiva, tal que as ondas se anulam, denomina-se estes como pontos de nulo. Tem-se ainda as defasagens que são múltiplas ímpares de $\pi/2$, indicando que a intensidade de saída do MZI será equivalente à metade da intensidade máxima possível, denominados pontos de quadratura. Podemos perceber também como a amplitude normalizada do campo elétrico complexo de saída contém informação sobre a fase, nos permitindo explorar valores tanto positivos quanto negativos. A Figura 6 ilustra também a defasagem do campo elétrico complexo de saída. Como podemos observar, há uma defasagem de π radianos no campo elétrico complexo de saída, quando a diferença de fase entre os braços é múltiplo ímpar de π .

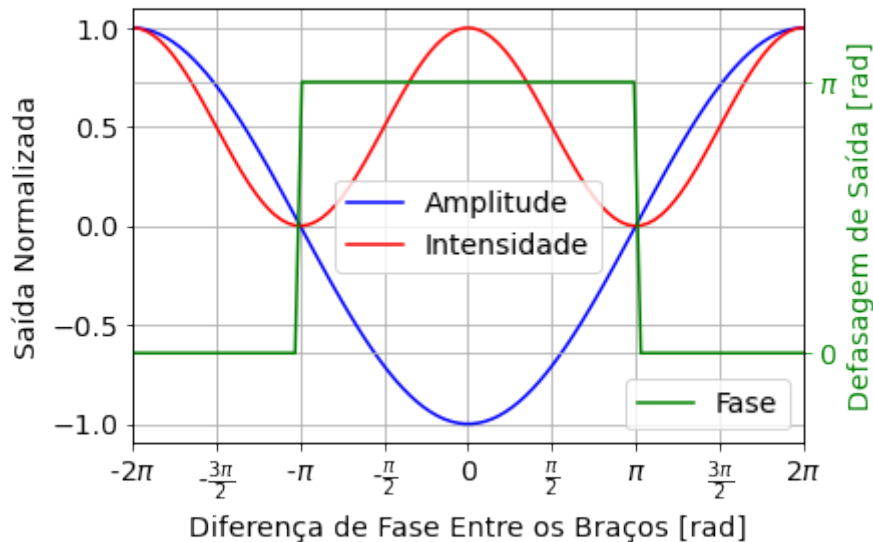


Figura 6 – Saída normalizada em relação à entrada, tanto para a intensidade quanto para a amplitude complexa do campo elétrico, e defasagem do campo elétrico complexo de saída em termos da diferença de fase.

A Figura 7(a) ilustra a função de transferência, em decibéis, de um MZI desbalanceado, ou seja, com diferentes comprimentos físicos entre os braços (GARDES *et al.*, 2008). Um dos braços considerados possui um comprimento de $500 \mu\text{m}$, enquanto o outro possui um comprimento de $600 \mu\text{m}$. Consideramos os braços do MZI como sendo guias de onda do tipo *strip* propagando o modo fundamental TE. Os valores do índice efetivo em relação aos comprimentos de onda foram calculados utilizando o *software* Lumerical Mode.

No caso em que os comprimentos físicos sejam iguais, esta configuração é denominada MZI balanceado. Por outro lado, quando os comprimentos ópticos são iguais, podemos

classificar o MZI como sendo simétrico, enquanto o termo assimétrico se refere àqueles MZIs com comprimentos ópticos distintos entre os braços. Tal assimetria pode ser gerada por um desbalanço nos comprimentos, mas também por um índice de refração diferente para cada braço. Quando adicionamos um certo desbalanceamento no MZI, haverá um grau de diferença na fase óptica dependente do comprimento de onda, o que ocasionará uma função de transferência aproximadamente periódica no domínio da frequência. Entretanto, nos cálculos da função de transferência mostrada na Equação 7, consideramos o índice efetivo do modo, tal qual é dependente do comprimento de onda em um meio dispersivo (tal como o silício). Esta dependência mostra que a função não é estritamente periódica quando consideramos a dispersão cromática do modo propagante. Para uma faixa de comprimento de ondas suficientemente estreito, porém, podemos considerar o comportamento como periódico. Neste caso, o período na função de transferência é denominado faixa espectral livre (*Free Spectral Range*, FSR) (AGRAWAL, 2014), e é dado por:

$$FSR[m] = \frac{\lambda^2}{n_g \Delta L_{ps}}, \quad (8)$$

em termos do comprimento de onda, ou podemos escrever como:

$$FSR[Hz] = \frac{c}{n_g \Delta L_{ps}}, \quad (9)$$

em termos da frequência, em que n_g é o índice de grupo. Na prática utilizamos os índices efetivos, ou seja, o índice de refração que é percebido pelo modo de propagação. A variável n_g é o índice de grupo, e esta é dada a partir do índice de refração efetivo n_{eff} por:

$$n_g = n_{eff}(\lambda) - \lambda \frac{dn_{eff}}{d\lambda}. \quad (10)$$

Podemos perceber pela Equação 8 que o FSR não é periódico, pois este depende do quadrado do comprimento de onda e do índice de grupo. O índice de grupo, por sua vez, também é dependente do comprimento de onda, entretanto este é constante para meios não dispersivos. No caso do FSR em relação à frequência, como dado pela Equação 9, temos um padrão periódico em meios não dispersivos.

A Figura 7(b) ilustra o FSR, calculado a partir da Equação 8, para o caso do MZI considerado na Figura 7(a). O índice de grupo foi obtido também utilizando o *software* *Lumerical Mode*. Pela Figura 7(b), fica clara a dependência do FSR em relação

ao comprimento de onda. Na próxima seção iremos discutir os MZMs, tais quais são dispositivos ativos e apresentam seu funcionamento baseado no MZI.

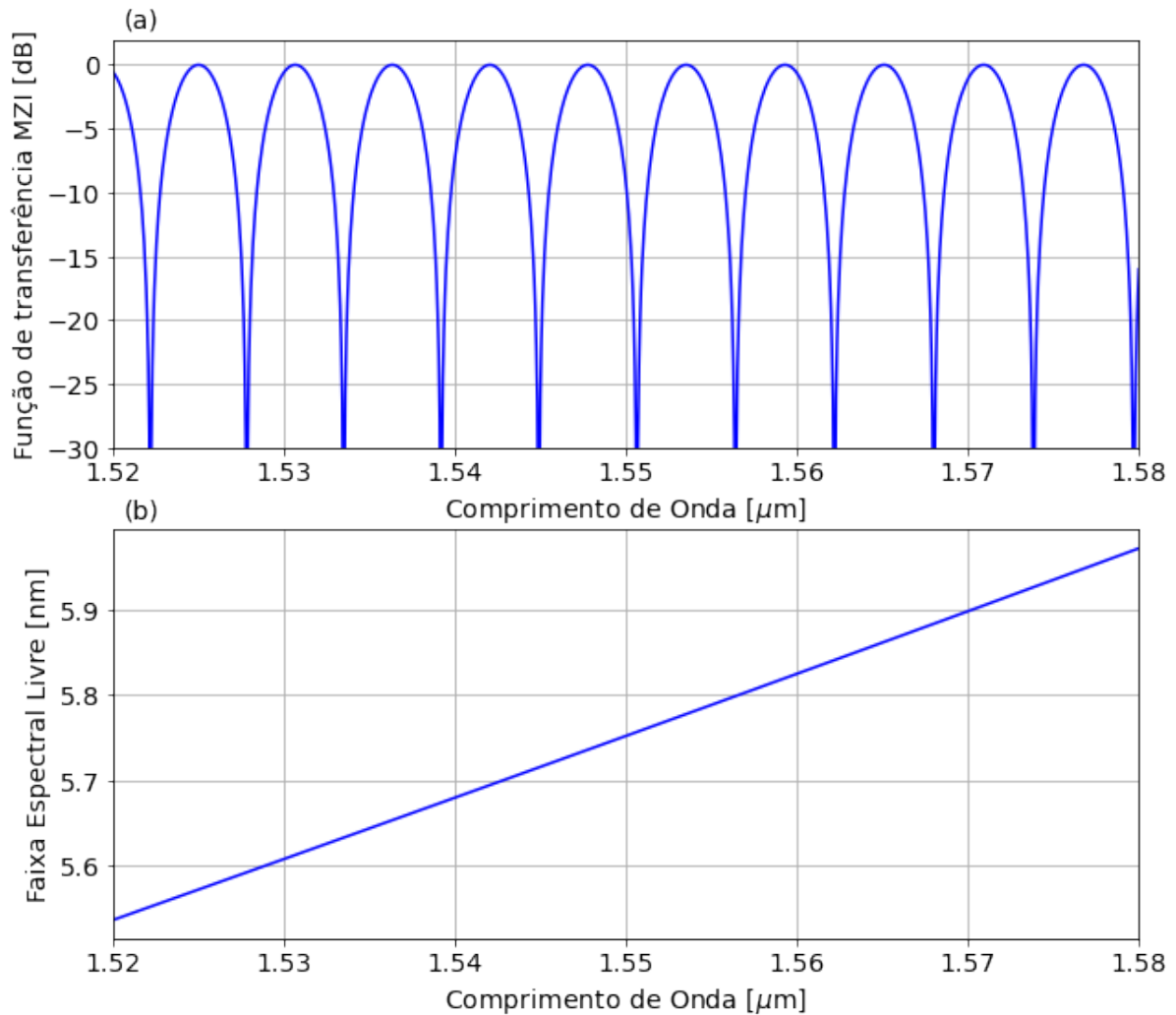


Figura 7 – (a) Função transferência de um MZI em relação ao comprimento de onda. Os braços possuem um desbalanço de $100 \mu\text{m}$ em relação ao comprimento físico destes. (b) FSR, em relação ao comprimento de onda, do MZI considerado.

2.2 Moduladores de Mach-Zehnder

Como mencionado na Seção 1.2, um MZM é baseado no MZI e opera a partir do efeito de interferometria. Entretanto, para podermos manipular a amplitude ou a fase do sinal de saída, é necessário variar a diferença da fase acumulada entre os braços. Na Seção 1.2 apresentamos os efeitos que podem ser explorados para esta variação, e concluímos que o PDE é o efeito mais eficiente na plataforma de SOI. A variação de cargas do PDE é realizada através de uma junção PN, como pode ser visualizado na

Subfigura 3(c). Maiores detalhes sobre a implementação destes deslocadores de fase serão dados na Seção 2.3. Na presente seção iremos tratar do funcionamento de MZMs de maneira geral.

Variando o índice de refração de um ou de ambos os braços do MZI, por meio do PDE, podemos controlar o comprimento óptico e, conseqüentemente, a relação de fase entre as duas ondas interferentes. Este controle do índice de refração de cada braço nos permite sintonizar as características de amplitude e de fase da onda de saída. Considerando a estrutura de MZM representada pela Figura 3(c), podemos definir a fase acumulada em um braço modulado em função da V_π (AGRAWAL, 2014):

$$\phi_j = \pi \frac{v_j(t)}{V_\pi}, \quad (11)$$

em que $v_j(t)$ é a tensão aplicada no j -ésimo braço. Tomando a Equação 4 e desconsiderando novamente as perdas de propagação, por simplicidade do resultado, e após algumas manipulações trigonométricas, podemos calcular a transmitância de campo complexo, $T_m^{(f)}$, como:

$$T_m^{(f)} = \frac{E_{out}}{E_{in}} = \cos \left[\frac{1}{2}(\phi_1 - \phi_2) \right] \exp \left(j \frac{\phi_1 + \phi_2}{2} \right). \quad (12)$$

O primeiro termo da Equação 12 corresponde à modulação de amplitude, enquanto o segundo termo corresponde à modulação de fase. Podemos observar que é possível operar o MZM como um modulador puramente de amplitude. Podemos realizar esta configuração implementando uma tensão $v_2(t) = -v_1(t)$, de forma que o segundo termo da Equação 12 é anulado e se tem o controle apenas da amplitude. Por outro lado, se considerarmos o caso em que $v_2(t) = -v_1(t) + V_{bias}$, como no modo de operação *Series Push-Pull* (SPP) (HENI *et al.*, 2015), a partir da Equação 11 temos que a soma entre as fases acumuladas será de:

$$\phi_1 + \phi_2 = \pi \frac{V_{bias}}{V_\pi}, \quad (13)$$

tal qual é dada por uma constante, nos permitindo controlar apenas o primeiro termo da Equação 12. Considerando este caso, a função de transferência terá a seguinte forma:

$$T_m(v_1(t)) = \cos^2 \left(\frac{\pi}{2V_\pi} [2v_1(t) - V_{bias}] \right). \quad (14)$$

O termo de tensão constante V_{bias} tem a função de ajustar o ponto de operação do modulador. Enquanto que a tensão aplicada $v_1(t)$ poderá realizar uma excursão em torno do ponto de operação definido por V_{bias} .

De forma análoga, podemos aplicar a relação de tensões dada por: $v_2(t) = v_1(t)$. Quando a tensão aplicada em ambos braços é equivalente, podemos provar que uma modulação puramente de fase ocorre. Neste caso, temos que a Equação 12 se resume a:

$$T_m^{(f)}(v_1(t)) = \exp\left(j\pi \frac{v_1(t)}{V_\pi}\right). \quad (15)$$

Até então, na presente seção, definimos como a tensão aplicada em cada braço do MZM pode modificar tanto a amplitude quanto a fase do sinal. Esta característica é útil para sistemas coerentes devido a possibilidade de modulação bipolar de amplitude, como no formato de modulação PAM (MELIKYAN *et al.*, 2020). Com uma estrutura como a demonstrada na Figura 4, podemos então combinar dois sinais modulados em PAM para gerar formatos mais complexos como QAM. Na seguinte subseção iremos citar algumas configurações de como a tensão pode ser aplicada no MZM.

2.2.1 Configurações de Alimentação

Existem variadas configurações de alimentação de tensão nos MZMs. Estas diferem entre si na forma como cada braço é alimentado. A presente subseção apresentará, em uma discussão breve, duas das configurações mais amplamente utilizadas para alimentação dos deslocadores de fase nos braços de um MZM. Estas configurações são denominadas *dual drive* e *series push-pull*. Em contrapartida, há também uma configuração de operação menos eficiente quando comparada às configurações citadas acima. Esta última é denominada *single drive*, e é caracterizada pela alimentação de apenas um dos braços do MZM.

A Figura 8 ilustra, de forma simplificada, as três configurações de alimentação de MZMs mencionadas. A Figura 8(a) mostra a configuração *single drive* de alimentação. Nesta configuração, apenas um dos braços é modulado, enquanto o outro apresenta um comportamento estático. A Figura 8(b) representa a configuração *dual drive* de alimentação. Como podemos notar, em *dual drive* temos a vantagem de que ambos os braços são alimentados. Comumente também utiliza-se a mesma tensão de polarização V_{bias} em ambos os braços.

Iremos apresentar uma breve análise a respeito da relação de consumo de potência entre a configuração *single drive*, P_{sd} e *dual drive*, P_{dd} . Apesar de necessitarmos de dois *drivers* na configuração *dual drive*, a potência dinâmica consumida ainda é aproximadamente metade do que aquela consumida pelo *single drive* para a mesma excursão de fase. Para causar um deslocamento de π na fase do sinal, no caso *single drive*, é necessária uma tensão de V_π injetada no braço alimentado. Neste caso, temos que $P_{sd} \propto V_\pi^2$. Já para o caso do *dual drive*, necessitamos de $V_{\pi/2}$ em um braço e $-V_{\pi/2}$ no outro braço para obtermos este mesmo deslocamento de fase. Desconsiderando a não linearidade da resposta do modulador, temos que a tensão necessária em cada braço pra realizar um deslocamento de $\pi/2$ ($-\pi/2$) na fase do sinal é equivalente a metade da tensão necessária para realizar um deslocamento de π ($-\pi$). Considerando os dois braços ,no caso do *dual drive*, temos que $P_{dd} \propto 2 \cdot (V_{\pi/2})^2$, o que indica uma diminuição pela metade na potência consumida em relação ao caso do *single drive*.

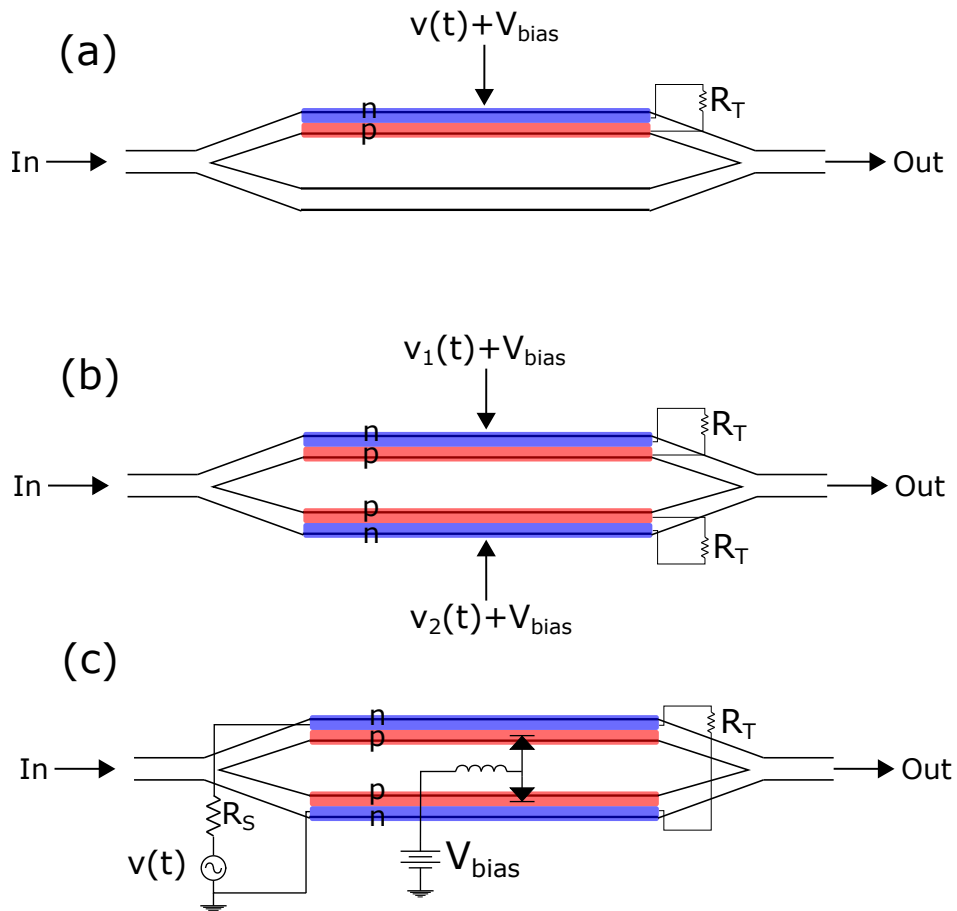


Figura 8 – Configurações de alimentação. (a) *Single drive* (b) *dual drive* e (c) *series push-pull*.

A Figura 8(c) ilustra um MZM utilizando configuração SPP. Podemos observar que a tensão de polarização V_{bias} é aplicada entre os diodos de cada braço. O indutor localizado entre a fonte DC e os diodos serve para filtrar as componentes AC da tensão de polarização. A configuração de SPP permite dobrar a resistência da junção enquanto reduz sua capacitância pela metade, o que pode aumentar a banda ao reduzir as perdas de RF (PATEL *et al.*, 2014). Uma vantagem de alta relevância pertencente à configuração de SPP é a remoção parcial do gorjeio (CHEN; DONG; CHEN, 2012). Em uma situação ideal, em que o deslocamento de fase é totalmente idêntico entre ambos os braços, o gorjeio seria totalmente eliminado. Entretanto, a não linearidade da resposta do MZM em SOI, em adição a variações induzidas pelo processo de fabricação, impedem a supressão total do gorjeio. Já em relação a potência dissipada na configuração em SPP, temos que esta é equivalente ao caso de *single drive*, já que a mesma tensão é aplicada em um único *driver* de forma a obter uma mudança de π na fase do sinal.

Na seguinte subseção iremos apresentar algumas formas de avaliar o desempenho do modulador.

2.2.2 Figuras de Mérito de um MZM

Nesta subseção serão apresentadas as figuras de mérito (*Figures Of Merit*, FOMs) para avaliação do desempenho do modulador. Especificamente, nos focaremos em três principais métricas para a realização do trabalho, sendo estas: BW_{EO} , IL e V_{π} . Formalizando a definição de IL , definimos esta métrica como sendo a razão entre a potência de entrada e a potência máxima de saída, e geralmente é representada em decibéis:

$$IL \text{ [dB]} = -10 \log_{10} \left(\frac{P_{out}}{P_{in}} \right) \quad (16)$$

Iremos também mencionar algumas outras figuras de mérito que são geralmente utilizadas na literatura para caracterização de um MZM. Um exemplo de indicador de desempenho que podemos citar é a razão de extinção (*Extinction Ratio*, ER). Há dois tipos de ER, um deles é o estático, em que o modulador não é alimentado eletricamente por um sinal variante no tempo. Por outro lado temos o ER dinâmico, em que consideramos a modulação do sinal para aferição. O ER estático tem a finalidade de quantificar a disparidade entre as perdas acumuladas de um braço do MZM em relação ao outro, ou

também a desproporção causada pelo divisor de potência. Podemos escrever o ER em função da potência máxima e da potência mínima obtida na saída do interferômetro:

$$ER_{static}[\text{dB}] = 10 \log_{10} \left(\frac{\max(P_{out})}{\min(P_{out})} \right). \quad (17)$$

O ER dinâmico, ou de modulação, mede a razão de extinção do sinal modulado. O ER dinâmico é dado por:

$$ER_{dyn}[\text{dB}] = 10 \log_{10} \left(\frac{P_H}{P_L} \right), \quad (18)$$

em que P_H e P_L representam, respectivamente, a potência máxima e a potência mínima do sinal modulado. Devido à não linearidade da resposta do modulador, em adição à não linearidade da resposta do deslocador de fase em silício, principalmente nos extremos da função de transferência, costuma-se limitar a faixa de excursão de modulação (SEPEHRIAN *et al.*, 2018). Desta forma, reduzimos distorções não lineares no sinal ao sacrifício de um menor valor de P_H . Em outras palavras, temos que na prática $P_H < \max(P_{out})$.

Apesar do ER ser uma medida muito utilizada na prática, por ser dada a partir de uma razão, a mesma pode resultar em uma má representação de desempenho se o ponto de operação estiver afastado do ponto de quadratura. Caso operamos o MZM próximo ao ponto de nulo, P_L irá ser reduzido de forma a ocasionar um ER_{dyn} alto, e o contrário ocorre ao operarmos o modulador próximo do ponto de pico. É possível definirmos uma medida que considera a diferença entre estes dois níveis de potência, de forma que o ponto de operação não afete a medida (PATEL *et al.*, 2014). Esta métrica é a amplitude de modulação óptica (*Optical Modulation Amplitude*, OMA), e é dada por:

$$OMA [\text{mW}] = P_H - P_L. \quad (19)$$

De forma análoga à Equação 16, podemos definir as perdas de potência do sinal modulado em relação à potência de entrada. As perdas de modulação α_{mod} são definidas pela razão entre a potência de entrada e a potência do símbolo modulado com maior potência:

$$\alpha_{mod}[\text{dB}] = -10 \log_{10} \left(\frac{P_H}{P_{in}} \right). \quad (20)$$

Podemos perceber que para obtermos um α_{mod} menor possível, além da minimização das perdas de propagação, é necessário operarmos em uma excursão próxima de V_π . Ao

evitarmos a porção não linear da curva de transferência, temos como consequência um aumento no α_{mod} (SEPEHRIAN *et al.*, 2019).

Há também formas de métricas compostas, que combinam duas ou mais métricas em apenas uma grandeza escalar. Por exemplo, para estimarmos a eficiência do modulador, é comum utilizarmos além de V_π , a métrica $V_\pi L_{ps}$. Esta medida considera a combinação de tensão e comprimento do deslocador de fase necessário para causar uma mudança de π na fase do sinal propagante. Entretanto, como o deslocamento de fase no silício possui um grau significativo de não linearidade, e também a variação de índice de refração deste satura com tensões de valores muito altos (ROSENBERG *et al.*, 2012), nem sempre há precisão ao estimar o V_π ou o L_{ps} tendo disponível o $V_\pi L_{ps}$. Outra medida interessante de eficiência, que contabiliza além do comprimento, as perdas, é dada pelo produto $V_\pi \cdot IL \cdot L_{ps}$ (TU *et al.*, 2011; GOYKHMANN *et al.*, 2013). Para tal métrica, é interessante sua minimização, de forma a reduzir as perdas, o *footprint* e a tensão aplicada na junção.

Podemos também estimar a energia gasta para representação de cada *bit* na modulação, uma métrica importante para circuitos integrados no âmbito comercial. Em (PATEL *et al.*, 2015) foi demonstrado um método para estimar tal métrica. Primeiramente, assumiu-se que a potência é dissipada apenas pela terminação e que há casamento de impedância entre fonte, linha de transmissão e terminação. Como na configuração de alimentação de SSP não há queda de tensão V_{bias} na resistência de terminação, podemos negligenciar o consumo de potência estática. A potência dinâmica, relativa a modulação, pode ser estimada considerando a razão entre o quadrado do valor eficaz da tensão do sinal modulante e a resistência de terminação, $P_{dyn} = V_{rms}^2/R_T$. A energia consumida por bit modulado pode ser dada então pela razão entre a potência dinâmica consumida e a taxa de bits.

Podemos perceber que estas figuras de mérito compostas não consideram a BW_{EO} do dispositivo. O trabalho realizado em (SEPEHRIAN *et al.*, 2018) propõe métrica baseada na penalidade de potência de transmissão (*Transmission Power Penalty*, TPP) dinâmica. Esta medida considera a redução na potência transmitida devido às perdas ópticas, $P_{\alpha L}$, ao ER finito, P_{ER} , e à banda finita, P_{ISI} , que está relacionada com a interferência inter-simbólica (*Intersymbol Interference*, ISI). O TPP é dado por:

$$TPP = P_{\alpha L} + P_{ER} + P_{ISI}. \quad (21)$$

Em (SEPEHRIAN *et al.*, 2018) estimou-se uma figura de mérito, considerando tais penalidades, sob a consideração que o deslocamento de fase é linear e que o sistema utiliza a modulação PAM-4. Esta FOM é dada por:

$$FOM_{PAM-4} = V_{\pi} \alpha L_{ps} \frac{M - 1}{1 - M \cdot \operatorname{erfc} \left(\frac{\pi}{2\sqrt{\ln(2)}} \frac{BW_{EO}}{BR} \right)}, \quad (22)$$

em que M indica o M -ésimo nível de modulação, BR é a taxa de baud e erfc é a função de erro complementar. Para o argumento z , a função de erro complementar é dada pela seguinte equação:

$$\operatorname{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^{\infty} e^{-t^2} dt. \quad (23)$$

Esta FOM apresentada em (SEPEHRIAN *et al.*, 2018) foi desenvolvida para o caso específico de modulação PAM-4. É conveniente a formulação de uma FOM generalizada, tal qual não levando em consideração as características da modulação, mas sim as características físicas da resposta do modulador. No presente trabalho, designamos FOMs para serem otimizados, e estes combinam IL, V_{π} e BW_{EO} . Mais detalhes sobre estas formas serão apresentados na Subseção 4.2.3. Na próxima seção iremos discutir o mecanismo de deslocamento de fase utilizado nos MZMs.

2.3 Deslocadores de Fase

Deslocadores de fase são essenciais na operação de MZMs, pois sem estes sua resposta seria estática, como no caso de MZIs. Como mencionado na Seção 1.2, utilizamos uma junção PN para induzir o PDE. Na Subseção 2.3.1 apresentaremos as diferentes estruturas de diodos deslocadores de fase encontradas na literatura. Na Subseção 2.3.2 apresentaremos o efeito de dispersão de plasma. Já na Subseção 2.3.3 mencionaremos as configurações de polarização de tensão em MZMs. Por fim, na Subseção 2.3.4 discutiremos a resposta óptica DC e de pequenos sinais do relacionado ao deslocador de fase.

2.3.1 Estruturas de Diodos

O deslocador de fase é implementado por meio de uma junção PN, também denominada diodo, inserida no núcleo do guia de onda. Dado este fato, é válido discutirmos

as diversas arquiteturas de junção PN que já foram estudadas na literatura, de forma a justificar nossa escolha para o presente trabalho. Utilizando um modelo simplificado de simulação modal e elétrica, os autores em (PEREZ-GALACHO *et al.*, 2015) estudaram diversas estruturas de diodos que compõem deslocadores de fase em silício. Na Figura 9 mostramos algumas destas estruturas. Cada arquitetura de organização das regiões PN estudadas apresentam vantagens e desvantagens específicas.

A Figura 9(a) ilustra uma estrutura de deslocador de fase baseado em um junção interdigital, em que I representa uma região intrínseca de silício. Este tipo de estrutura pode ser utilizada para maximizar a sobreposição do modo com a região de depleção, o que amplia a eficiência (diminuindo o V_π). A banda obtida por (PEREZ-GALACHO *et al.*, 2015) para tal estrutura é de apenas 10 GHz, porém um $V_\pi L_{ps}$ de somente 0,6 V·cm foi estimado. Apesar de uma alta eficiência de modulação, a BW_{EO} é inapropriada para sistemas que requerem altas taxas de transmissão.

A Figura 9(b) ilustra uma estrutura de junção PN lateral com regiões mais dopadas nas extremidades que servem como interface com o contato elétrico. Este tipo de estrutura favorece uma BW_{EO} alta, ao custo de uma eficiência menor de modulação. Em (PEREZ-GALACHO *et al.*, 2015) obteve-se uma BW_{EO} superior a 50 GHz, desconsiderando os efeitos interativos com a linha de transmissão, e um produto $V_\pi L_{ps}$ de 1,5 V·cm. Já na Figura 9(c) temos o esquemático de uma estrutura parecida, com a diferença que regiões com níveis de dopagens intermediárias foram inseridas entre as regiões P/N e P⁺⁺/N⁺⁺.

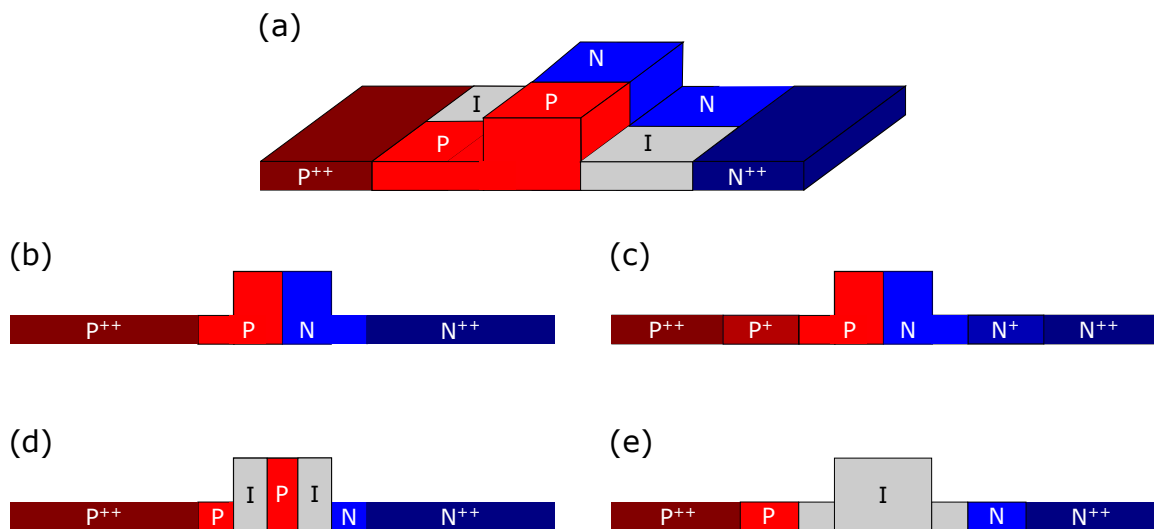


Figura 9 – Diferentes estruturas de diodos. (a) Estrutura de junção interdigital. (b) Junção PN lateral convencional e (c) junção PN lateral com regiões de dopagens intermediárias. (d) Estrutura PIPIN. (e) Estrutura PIN.

Estas regiões intermediárias são capazes de diminuir a resistência em série, sem interferir nas perdas ópticas, com o intuito de melhorar a banda (SEPEHRIAN *et al.*, 2019). Este é o tipo de estrutura a ser utilizada no presente trabalho, pois apresenta uma larga BW_{EO} e com um V_{π} relativamente baixo.

A Figura 9(d) ilustra uma estrutura PIPIN, a qual possui o propósito de diminuir as perdas ópticas no meio. A maior desvantagem de tais estruturas é a redução no BW_{EO} e um aumento no V_{π} . Como podemos ver pela figura, uma região P é inserida no guia intrínseco. A banda estimada é de 20 GHz enquanto o $V_{\pi}L_{ps} = 1,37$ V·cm (PEREZ-GALACHO *et al.*, 2015). As perdas na simulação foram estimadas como apenas 2,2 dB.

A Figura 9(e) ilustra uma estrutura PIN. Neste caso em particular, o funcionamento se dá a partir da injeção de portadores livres. A operação por injeção de portadores possibilita uma melhor eficiência ao custo de uma banda reduzida. Com um $V_{\pi}L_{ps}$ de apenas 0,029 V·cm, esta estrutura é ultra-eficiente, porém a BW_{EO} estimada é de apenas 43 MHz (PEREZ-GALACHO *et al.*, 2015). Tal arquitetura pode ser utilizada em chaveadores ou outros dispositivos em que operação em alta velocidade não seja crítica. A Tabela 2.3.1 sumariza os resultados obtidos em (PEREZ-GALACHO *et al.*, 2015), incluindo o caso de junção PN lateral se houvessem seções com dopagens intermediárias entre o núcleo e o contato. Na próxima subseção discutiremos como a tensão aplicada em um diodo pode alterar o índice de refração e a atenuação do meio.

Tabela 2 – Resultados de simulação obtidos para diferentes estruturas de diodos. Tais resultados incluem as métricas BW_{EO} , IL e V_{π} obtidos em (PEREZ-GALACHO *et al.*, 2015)

Métrica/Estrutura	Interdigital	Lateral	Lateral com Seções Intermediárias	PIPIN	PIN
BW_{EO} [GHz]	10	50	>50	20	0,043
IL [dB]	3,9	4,7	≈4,7	2,2	1,57
$V_{\pi}L_{PS}$ [V · cm]	0,6	1,5	≈1,5	1,37	0,023

2.3.2 Efeito de Dispersão de Plasma

Como discutido na Seção 1.2, o silício possui seu índice de refração pouco variante em relação à temperatura e aos efeitos Pockels e Kerr. O efeito de PDE se aproveita do fenômeno de que o índice de refração em um meio se varia de forma relacionada com a variação da densidade de portadores livres. Na prática, controla-se esta variação

externamente, a partir da aplicação de uma corrente elétrica. Tal variação do número de portadores, sendo realizada tanto por injeção quanto por extração dos mesmos, resulta em uma alteração no valor do índice de refração. A partir da observação do fenômeno em dados experimentais, Soref e Bennett (SOREF; BENNETT, 1987) deduziram um modelo semi-empírico para descrever a variação do índice de refração e da absorção em relação à variação de portadores. Tal modelo foi atualizado em 2011 com dados e métodos mais recentes (NEDELJKOVIC; SOREF; MASHANOVICH, 2011). Para um comprimento de onda de 1550 nm, as equações que descrevem a variação de índice de refração e de atenuação, respectivamente, são dadas por:

$$\Delta n = -5,4 \times 10^{-22} \Delta N^{1,011} - 1,53 \times 10^{-18} \Delta P^{0,838}, \quad (24)$$

e:

$$\Delta \alpha = 8,88 \times 10^{-21} \Delta N^{1,167} + 5,84 \times 10^{-20} \Delta P^{1,109}, \quad (25)$$

em que os termos ΔN e ΔP representam a densidade em excesso de elétrons e de lacunas em cm^{-3} , respectivamente. Com injeção de elétrons aumentamos ΔN e reduzimos ΔP , e o contrário acontece com extração de portadores. É válido notar que estas variações de índice e de atenuação são também dependentes do comprimento de onda, porém, por simplicidade, nos limitaremos ao caso de $\lambda = 1550$ nm.

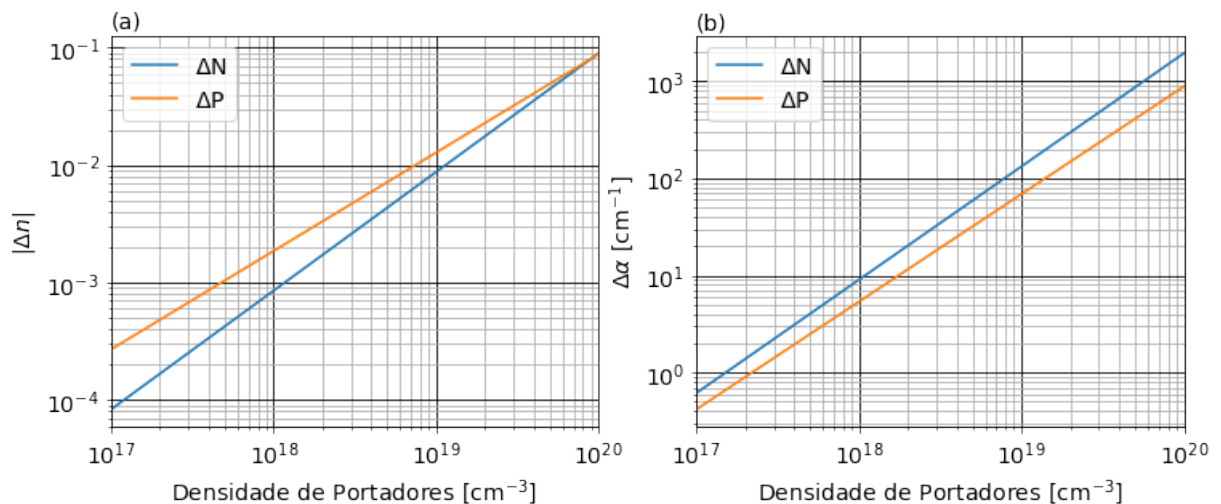


Figura 10 – Variação de (a) índice de refração e (b) atenuação em relação às variações de elétrons e de lacunas.

A Figura 10 ilustra como a variação dos portadores modificam tanto o índice de refração quanto o coeficiente de atenuação. A Figura 10(a) mostra o comportamento da Equação 24, indicando como a variação de índice é afetado em relação às variações de densidade de elétrons e lacunas. A Figura 10(b) mostra o comportamento da Equação 25, isto é, a variação da atenuação em relação às variações dos portadores. Podemos perceber que a variação do índice de refração devido à variação da densidade de lacunas é mais significativa do que devido à variação de elétrons. Também fica evidente que a variação de densidade de elétrons afeta de forma mais significativa a mudança de atenuação. Observando estes fenômenos, podemos concluir que as variações de densidade de lacunas são preferíveis devido à maior modulação de índice e menor atenuação causada em relação às variações de densidade de elétrons. Outro ponto a considerarmos no projeto de deslocadores de fase é se a tensão a ser aplicada tem uma polarização direta ou reversa em relação à polarização do diodo. Na próxima subseção iremos discutir estas formas de operação.

2.3.3 Polarização Direta e Reversa da Junção PN

Como a própria nomenclatura indica, polarização direta se refere a injeção de portadores, ou seja, quando se introduz portadores livres no diodo. Por outro lado, a polarização reversa ocorre ao extrair portadores livres da estrutura (REED *et al.*, 2010). A Figura 11 ilustra o circuito equivalente, considerando pequenos sinais, para ambos os modos de operação. Na Figura 11(a) temos o modelo equivalente de um diodo polarizado diretamente, em que g_d é a condutância, c_j é a capacitância da junção, c_d é a capacitância

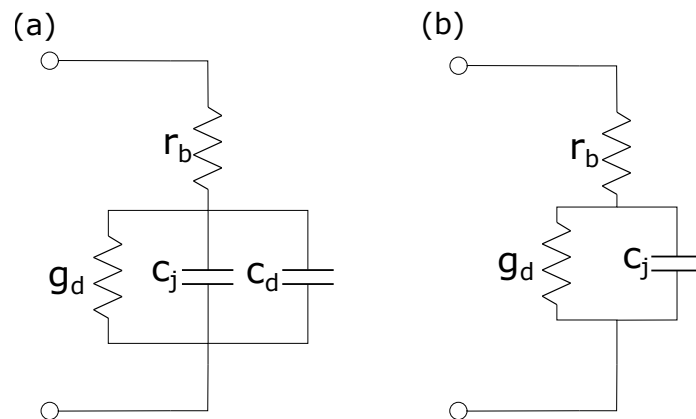


Figura 11 – Circuito elétrico equivalente, considerando pequenos sinais, de um (a) diodo polarizado diretamente e (b) diodo polarizado reversamente.

de difusão e r_b contabiliza as resistências de forma geral, desde a margem da região de depleção até o contato. Na Figura 11(b), por outro lado, temos a ilustração do circuito elétrico equivalente de um diodo operando por polarização reversa.

A partir da Figura 11 podemos perceber que o termo c_d se aplica somente ao caso de polarização direta, pois somente neste caso haverá difusão. O resultado de uma capacitância menor é uma BW_{EO} maior, já que a banda é inversamente proporcional à capacitância da junção, como iremos verificar a seguir. Outra diferença que iremos discutir também é na largura da região de depleção. Enquanto uma tensão reversa tende a ampliar a área da região de depleção, uma tensão direta irá ocasionar uma redução da mesma. Apesar de apresentar uma BW_{EO} maior, a operação por depleção apresenta a desvantagem de requerir uma excursão maior de tensão para modulação. Utilizando polarização direta, há uma maior eficiência de modulação ao custo de uma BW_{EO} significativamente reduzida, a tal ponto que é impraticável sua aplicação em sistemas de alta transmissão de dados. Em alguns trabalhos recentes, como (SOBU *et al.*, 2021), um equalizador passivo RC foi utilizado para controlar os parâmetros RC da junção, permitindo uma BW_{EO} superior a 40 GHz ser alcançada, porém com o custo de um V_π de aproximadamente 20 V.

Além da polarização de operação, há diversas formas de estruturar a junção PN dentro do guia de ondas. A Figura 12 representa o guia de onda utilizado no braço do MZM. Este tipo de guia de onda, denominado *rib*, possui uma seção de silício de menor espessura, denominado *slab*, formando uma espécie de base em um guia de ondas retangular (GOELL, 1973). Utiliza-se deste tipo de estruturas para implementação do deslocador de fase, de forma que as extremidades do guia são utilizadas para conectar o contato elétrico. Para

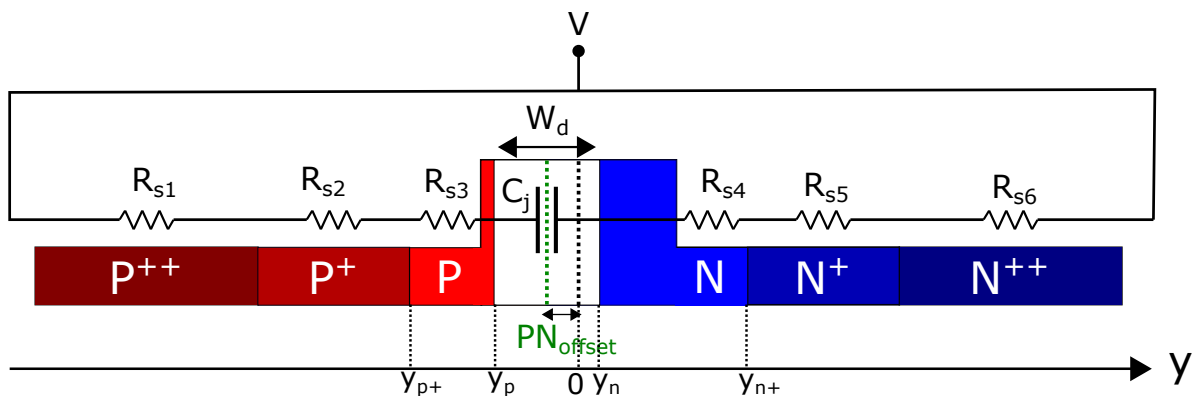


Figura 12 – Seção transversal de um guia do tipo *rib* com três níveis de dopagem. No centro do guia óptico há a formação de uma junção PN. A figura também mostra uma simplificação de um circuito elétrico equivalente à junção.

desenvolver um deslocador de fase, implementamos uma junção PN lateral no centro do guia de ondas, ocasionando a sobreposição do modo propagante com a junção.

Para o cálculo da largura de depleção da junção PN, podemos assumir uma distribuição linear nas densidades de portadores minoritários entre a região de depleção e o meio dopado (CHROSTOWSKI; HOCHBERG, 2015). Esta aproximação se aplica nos casos em que o comprimento de difusão é muito maior do que a largura de depleção. A largura de depleção, W_d , é dada pela equação:

$$W_d = \sqrt{\frac{2\epsilon_0\epsilon_r(N_A + N_D)(V_j - V)}{qN_A N_D}}, \quad (26)$$

em que q é a carga elementar, ϵ_0 é a permissividade do vácuo, ϵ_r é a permissividade relativa, V_j é a barreira de potencial da junção PN e V é a tensão externa aplicada. Os termos N_A e N_D são as densidades de impurezas, respectivamente, de aceitadores (lacunas) e doadores (elétrons), que são introduzidos no silício. O V_j pode ser calculado por:

$$V_j = \frac{K_B T}{q} \ln \left(\frac{N_A N_D}{n_{int}^2} \right), \quad (27)$$

sendo que K_B é a constante de Boltzmann, T é a temperatura e n_{int} é a concentração intrínseca de portadores. Na próxima subseção iremos mostrar como a resposta óptica depende dos conceitos discutidos na presente subseção.

2.3.4 Resposta Óptica

O guia óptico *rib* ilustrado pela Figura 12 possui 6 níveis de dopagem. No centro do guia, na interface entre as dopagens do tipo P e do tipo N , há a formação da junção PN. Como discutido na seção anterior, a injeção ou extração de portadores livres na junção PN ocasionará uma mudança tanto do índice efetivo do modo quanto da atenuação. A mudança de índice, ao aplicar uma tensão V , é dada por:

$$\Delta n_{eff}(V) = n_{eff}(V) - n_{eff}(0), \quad (28)$$

em que $n_{eff}(V)$ é o índice efetivo dada uma tensão aplicada V , e $n_{eff}(0)$ é o índice efetivo quando não há tensão aplicada. Com a mudança de índice efetivo Δn_{eff} , podemos calcular a mudança de fase causada na onda propagante no guia, que é dada por:

$$\Delta\phi(V) = \frac{2\pi L_{ps}\Delta n_{eff}(V)}{\lambda}. \quad (29)$$

O índice efetivo para uma dada tensão pode ser calculado levando em consideração a mudança de índice de refração do núcleo e a integral de sobreposição com o modo (CHROSTOWSKI; HOCHBERG, 2015), dado por:

$$n_{eff}(V) = n_{eff,int} + \frac{\int E^*(y)\Delta n(y,V)E(y)dy}{\int E^*(y)E(y)dy} \cdot \frac{dn_{eff}}{dn_c}, \quad (30)$$

em que y é a direção perpendicular à junção PN, indicado na Figura 12, $n_{eff,int}$ é o índice efetivo intrínseco do meio (sem dopagem), $\Delta n(y,V)$ é a variação do índice do meio de propagação, dada uma tensão aplicada, e $\frac{dn_{eff}}{dn_c}$ é a variação do índice efetivo em relação à variação do índice no núcleo, e é aproximadamente 1. A região de depleção forma uma área aproximadamente intrínseca de silício, a qual possui um índice de refração diferente daquela das regiões dopadas. Controlando a largura da região de depleção, controla-se a densidade de portadores na área modal efetiva. De forma parecida, a atenuação na junção dada uma certa tensão pode ser calculada por:

$$\alpha_j(V) = \frac{\int E^*(y)\Delta\alpha(y,V)E(y)dy}{\int E^*(y)E(y)dy}, \quad (31)$$

em que $\Delta\alpha(y,V)$ é a variação de perdas no núcleo dada uma tensão V .

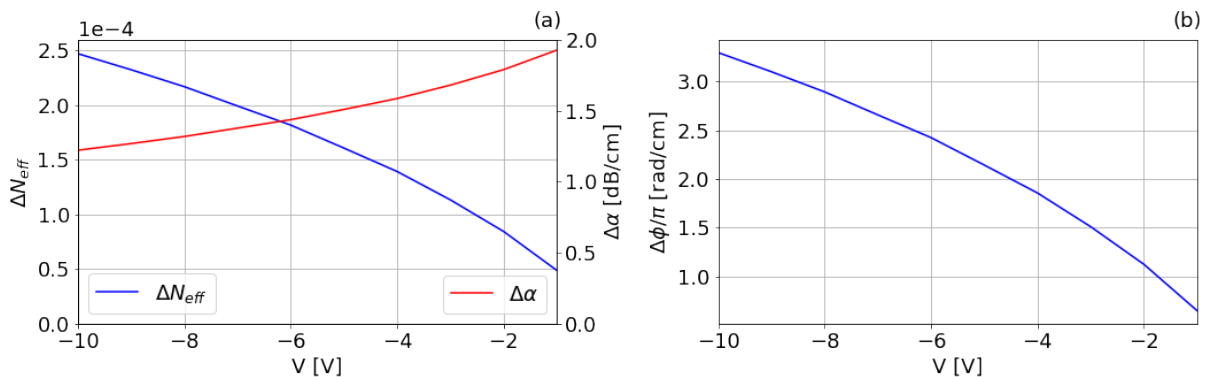


Figura 13 – (a) Variação do índice efetivo e das perdas ao aplicarmos uma tensão reversamente polarizada e (b) variação da fase do sinal para os mesmos valores de tensão.

Utilizando as Equações 28-31, as variações de índice efetivo, perdas e diferença de fase foram calculadas para uma faixa de -10 V até -1 V aplicados na junção. A Figura 13(a) mostra as mudanças de índice efetivo em relação à tensão reversa aplicada. Podemos observar que este se mantém na ordem de $1 \cdot 10^{-4}$ e é mais intenso para tensões reversas maiores. Já para o caso das perdas podemos ver uma redução da mesma ao aumentarmos a tensão reversa. A Figura 13(b) mostra a variação de fase de acordo com a tensão reversa aplicada. Como esperado, dada a Equação 29, a variação de fase é proporcional à variação do índice efetivo, e estas apresentam um comportamento não linear. Podemos observar também que o V_π neste caso é de aproximadamente -1,7 V para 1 cm de comprimento, ou seja, um $V_\pi L_{ps}$ de 1,7 V·cm de tensão reversa.

A V_π define a tensão necessária para um braço do MZM realizar uma mudança de π na fase da onda propagante. Esta métrica está relacionada ao consumo de potência do MZM, fator importante ao considerar um sistema fotônico integrado comercial, dadas as limitações de potência nestes circuitos. Além do mais, devemos observar que *drivers* possuem uma limitação de tensão possível de ser aplicada, o que faz com que a diminuição de V_π seja altamente desejável. Para atingir um defasamento de π entre as ondas de cada braço do MZM, outro fator relevante é o comprimento do deslocador de fase L_{ps} . Como discutido na Subseção 2.2.2, podemos definir a métrica $V_\pi L_{ps}$, tal qual leva em consideração tanto a eficiência de modulação como o comprimento do deslocador de fase. Esta métrica pode ser estimada pelo modelo de pequeno sinal e dada pela equação (ZHOU *et al.*, 2016):

$$V_\pi L_{ps} \cdot \frac{2\pi}{\lambda} \frac{dn_{eff}}{dV} = \pi. \quad (32)$$

Resposta a Pequenos Sinais

É possível calcularmos a resistência e a capacitância da junção PN, de forma a prever o efeito da mesma na BW_{EO} (CHROSTOWSKI; HOCHBERG, 2015). A Figura 12 mostra um modelo simplificado de circuito elétricos representando as resistências de folhas e de capacitância do guia de ondas *rib*. A resistência da junção PN pode ser calculada por meio de:

$$R_j = \sum_i R_{si} W_i, \quad (33)$$

onde R_{si} e W_i são, respectivamente, a resistência da secção e a largura da camada de dopagem do i -ésimo bloco que compõe a junção PN. Por outro lado, a capacitância da junção PN é dada como:

$$c_j = C_{dep} + C_{fr} = t_{RIB} \sqrt{\frac{q\epsilon_0\epsilon_s}{2(1/N_D + 1/N_A)(V_j - V)}} + \frac{\epsilon_0\epsilon_{SiO_2}}{\pi} \ln\left(\frac{2\pi t_{RIB}}{W_d}\right), \quad (34)$$

em que ϵ_s é a permissividade elétrica relativa do silício, t_{RIB} é a espessura do *rib*, C_{fr} é a capacitância devido aos campos marginais e ϵ_{SiO_2} é a permissividade elétrica relativa da sílica (ZHOU *et al.*, 2016). A frequência de corte de 3 dB da junção PN, que indica a BW_{EO} , pode ser então aproximada pela constante RC:

$$BW_{EO} = \frac{1}{2\pi R_j c_j}. \quad (35)$$

Como discutido na Seção 2.3.3, com o aumento da tensão reversa aplicada há uma diminuição na resistência e na capacitância da junção PN, conseqüentemente causando um aumento na largura da região de depleção e na BW_{EO} (CHROSTOWSKI; HOCHBERG, 2015). Por outro lado, para maior eficiência de modulação, um MZM possui um comprimento relativamente longo, de forma a aumentar a distância de interação do modo óptico com a junção. No caso de uma junção mais longa, a capacitância é ampliada, o que limita a BW_{EO} , indicando um compromisso entre esta métrica e a eficiência de modulação. A frequência de corte da junção RC pode ser superior a 100 GHz, porém na prática há outros fatores limitantes que impedem a operação da mesma em tal largura de banda, pois a resposta eletro-óptica do MZM considera também a interação entre os eletrodos e o modulador.

Na próxima seção iremos discutir como o sinal de RF é inserido no modulador, como são modeladas as linhas de transmissão que servem de eletrodos e como suas características afetam a resposta do MZM.

2.4 Eletrodos de Ondas Viajantes

Na presente seção iremos discutir como o sinal modulante, este no domínio de RF, é aplicado no MZM. As linhas de transmissão que conduzem tal sinal são, geralmente, implementadas em uma camada superior aos guias de ondas ópticos do MZM. O projeto das linhas de transmissão deve levar em consideração que componentes de altas frequências, com banda de até 50 GHz, devem ser conduzidas por estas. Em geral, podemos classificar

os eletrodos que alimentam um MZM como sendo de dois tipos, concentrados (*lumped*) e distribuídos.

Eletrodos concentrados são eletrodos de relativamente pequenas dimensões, que permitem uma baixa ocupação de área do MZM e baixas perdas de RF (ALFERNES, 1982). Entretanto, o eletrodo concentrado requer uma maior tensão aplicada, o que indica um aumento no consumo de potência. Além do mais, eletrodos concentrados não permitem a obtenção de uma impedância alta o suficiente para realizar casamento de impedância que, como discutiremos na Subseção 2.4.3, é um fator determinante para o aumento da BW_{EO} . Um MZM ultra-compacto de eletrodo concentrado foi otimizado em um trabalho recente (CONG *et al.*, 2021). No MZM desenvolvido em tal trabalho, valores de BW_{EO} menores que 30 GHz foram obtidos, além de um alto $V_{\pi}L_{ps}$ de 2,2 V·cm para apenas 150 μm de comprimento de deslocador de fase.

Por outro lado, eletrodos distribuídos, ou eletrodos de onda viajante (*Traveling Wave Electrode*, TWE), possuem um comprimento significativamente maior que o comprimento de onda de RF. Neste caso, é válida uma análise de parâmetros distribuídos, que é dada com maiores detalhes no Apêndice A. Para sistemas que requerem altas taxa de dados, TWEs são recomendados no projeto de linhas de transmissão. De maneira oposta a eletrodos concentrados, TWEs permitem o sinal de RF se propagar ao longo de todo o comprimento do deslocador de fase, ou seja, o efeito de modulação é distribuído ao longo do guia óptico do MZM (REED *et al.*, 2014). Este tipo de eletrodo apresenta uma maior BW_{EO} e uma maior eficiência de modulação ao custo de ocupar uma maior área. As configurações mais comumente utilizadas para TWEs são o guia de onda coplanar (*Coplanar Waveguide*, CPW) e o guia de onda de tira coplanar (*Coplanar Stripline*, CPS) (CHEN; CHOU, 1997), principalmente por estes serem compatíveis com o processo de fabricação CMOS. No presente trabalho iremos dar maior foco no tipo CPS de linha de transmissão, pois esta é a utilizada nas simulações do MZM.

2.4.1 Guia de Ondas de Tiras Coplanares sem Carga

Como mencionado, o modelo de linhas CPS permite a integração no processo de fabricação CMOS. Outra vantagem apresentada pelo CPS é que esta configuração permite a conexão de elementos paralelos à linha de transmissão (MOTTA, 2018). No

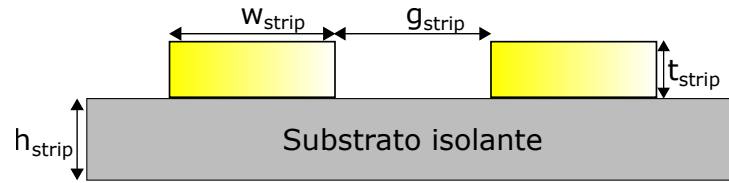


Figura 14 – Seção transversal de uma linha CPS simétrica com única camada de substrato.

caso discutido aqui, a junção PN do MZM é vista como uma conexão RC, como visto na Figura 11, e na prática esta é disposta em uma configuração paralela ao TWE. Além do mais, linhas CPS possuem uma geometria extremamente simples. Um CPS é formado por dois condutores paralelos sobre um substrato isolante, assim como representado pela Figura 14, em que se ilustra um CPS simétrico, ou seja, que ambos condutores possuem a mesma largura. Diferentemente de CPWs, CPSs apresentam apenas dois condutores, que podem ser utilizados como sinal e terra. CPWs, por outro lado, possuem três condutores, em que podemos utilizar uma configuração em que um condutor é destinado ao sinal e os outros dois são para aterramento deste. Esta configuração pode ocasionar o surgimento de modos parasitas caso a junção ou algum outro elemento se conecte a apenas em um dos terras, necessitando de configurações blindadas como em (TU *et al.*, 2014). Na configuração de CPS a blindagem não é necessária, devido à existência de apenas um condutor terra.

Iremos discutir na presente seção as linhas CPS descarregadas. Linhas descarregadas não consideram a junção PN (deslocador de fase óptico) na análise, que, como mencionado, é percebida como uma carga RC pela linha de transmissão. Existem dois métodos principais que analisam o comportamento de tais linhas de transmissão e estimam os valores de impedância característica, velocidade de fase da onda de RF e as perdas de propagação. Há um método de análise dinâmico que é capaz de ser aplicado para qualquer gama de frequências e modos (NGUYEN, 2001). Por outro lado, há um método de análise quasi-estático, baseado em mapeamento conforme, que apresenta precisão limitada para operação em DC e modos TEM. Entretanto, além de sua simplicidade de implementação, métodos quasi-estáticos já foram utilizados para modelar linhas de transmissão de maneira eficiente e suficientemente precisa (MALYUTIN, 2019; CHOO; CHO; CHOO, 2015).

Em (BAHL; BOZZI; GARG, 2013) um método de análise quasi-estático analítico foi utilizado. Este modelo fez a consideração de que o substrato possui uma espessura infinita. Na continuação apresentamos as principais expressões para o cálculo da impedância

característica de uma linha CPS. Considerando o caso de um CPS simétrico, assim como o ilustrado na Figura 14, definimos um parâmetro auxiliar Δ dado por:

$$\Delta = \frac{1,25t_{strip}}{\pi} \left[1 + \ln \left(\frac{4\pi W_{strip}}{t_{strip}} \right) \right], \quad (36)$$

em que t_{strip} é a espessura das tiras, enquanto W_{strip} é a largura das mesmas. O parâmetro Δ simplifica a definição de um espaçamento (*gap*) g_e efetivo e uma largura de strip efetiva W_e , dados por:

$$g_e = g_{strip} - \Delta, \quad (37)$$

$$W_e = W_{strip} + \Delta, \quad (38)$$

em que g_{strip} é o espaçamento entre as tiras. Podemos definir a razão de aspecto entre o espaçamento e a largura dos substratos como:

$$k^{(1)} = \frac{a}{b} = \frac{g_{strip}}{g_{strip} + 2W}, \quad (39)$$

enquanto a razão de aspecto efetiva é dada por:

$$k_e = \frac{g_e}{g_e + 2W_e}. \quad (40)$$

O complemento da razão de aspecto e da razão de aspecto efetiva podem ser dadas pelas seguintes equações:

$$k^{(1)'} = \sqrt{1 - k^{(1)2}}, \quad (41)$$

$$k_e' = \sqrt{1 - k_e^2}. \quad (42)$$

Considerando um valor de constante dielétrica relativa do meio sendo de $\epsilon_r \geq 9$ e que a proporção entre espessura e largura das tiras seja de $t_{strip}/W_{strip} < 0,1$, definimos a constante dielétrica efetiva como sendo:

$$\epsilon_r' = \epsilon_r + \frac{1,4(\epsilon_r - 1)t}{\left| \frac{K(k^{(1)})}{K(k^{(1)'})} \right| g + 1,4t}, \quad (43)$$

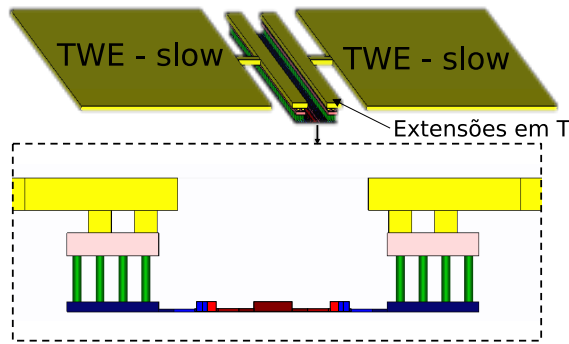


Figura 15 – Perspectiva isométrica e seção transversal da conexão do TWE com os guias de ondas do MZM.

sendo $K(x)$ a integral elíptica completa de primeiro tipo aplicada no argumento x . A impedância característica de um CPS pode ser calculada como:

$$Z_{0-cps} = \frac{120\pi K(k_e)}{\sqrt{\epsilon'_r} K(k'_e)}. \quad (44)$$

Tiras Coplanares de Onda Lenta

O índice de refração do CPS tem um valor demasiadamente baixo comparado ao da onda de luz no guia de onda óptico. Como iremos mencionar na Subseção 2.4.3, uma diferença (descasamento) entre a velocidade da onda óptica e a onda de RF pode limitar significativamente a BW_{EO} . Para aumentarmos o índice de refração no CPS, de forma a reduzirmos a velocidade da onda RF, utilizamos um tipo de eletrodos denominados eletrodos de onda lenta (*Slow-Wave Electrodes*). A Figura 15 mostra, em perspectiva isométrica e em corte transversal, o TWE-*slow* conectado com o MZM. Como podemos observar, tais eletrodos possuem extensões com um formato de T , e são amplamente aplicados em MZMs de alta velocidade (CHEN; DONG; CHEN, 2012; DING *et al.*, 2014; JAEGER; LEE, 1992).

2.4.2 Linha de Transmissão com Carga

As análises feitas até então consideraram uma linha de transmissão sem carga. No *design* de MZMs, a linha de transmissão terá uma carga distribuída ao longo do comprimento de toda linha, esta representando a junção PN. A inserção da junção PN

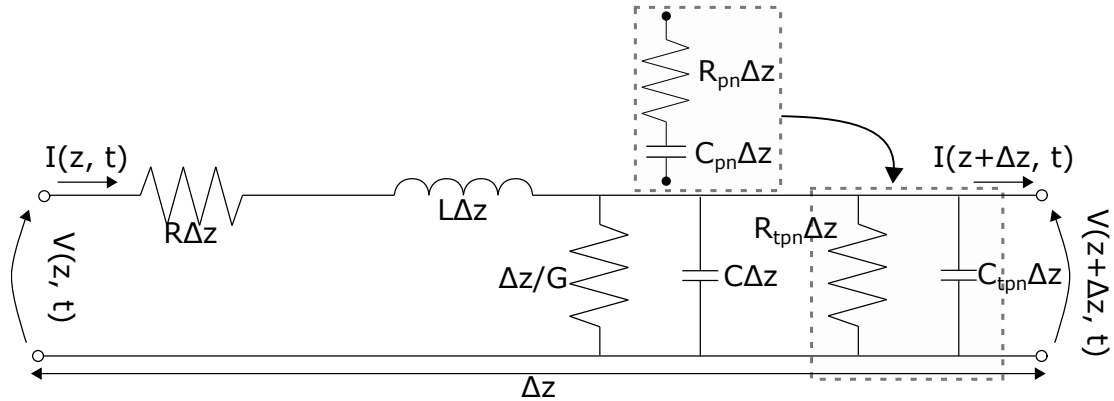


Figura 16 – Circuito equivalente RC paralelo da linha de transmissão com a junção PN como terminação. A janela superior mostra os elementos da junção PN antes de tomarmos o circuito RC paralelo equivalente. Figura adaptada de (POZAR, 2005).

é interpretada pela rede como uma capacitância, C_{pn} , e uma resistência, R_{pn} , ambos posicionados em série entre os condutores do CPS (DING *et al.*, 2012). Podemos observar na janela superior da Figura 16 que ambos os elementos da junção PN estão em série entre si. Para facilitar a análise, é conveniente realizar uma equivalência nos elementos em série R_{pn} e C_{pn} , de forma a obtermos um circuito equivalente RC paralelo (PATEL *et al.*, 2014). A Figura 16 mostra o modelo de circuito equivalente RC paralelo. Esta transformação é válida quando, para uma frequência angular ω , o fator Q do circuito em série e do circuito em paralelo apresentam os mesmos valores (YANG; ZHAO; CHEN, 2010) O fator Q em relação aos parâmetros do circuito RC em série e em paralelo é dado por:

$$Q = \frac{1}{\omega C_{pn} R_{pn}} = \omega C_{tpn} R_{tpn}, \quad (45)$$

em que ω é a frequência angular. Para redefinirmos os elementos como paralelos entre si e entre os condutores, temos as seguintes equivalências:

$$R_{tpn} = R_{pn}(1 + Q^2), \quad (46)$$

$$C_{tpn} = C_{pn} \frac{Q^2}{1 + Q^2}. \quad (47)$$

A impedância característica da linha carregada Z_{0l} e a constante de propagação γ_l , em relação aos elementos da linha de transmissão, podem ser dadas por:

$$Z_{0l} = \sqrt{\frac{R + j\omega L}{(G + R^{-1}) + j\omega(C + C_{tpn})}}, \quad (48)$$

$$\gamma_l = \alpha_l + j\beta_l = \sqrt{(R + j\omega L) + (|G + R_{tpn}^{-1}| + j\omega[C + C_{tpn}])}. \quad (49)$$

Nas Equações 48 e 49, o parâmetro R indica a resistência em série por unidade de comprimento, dado em Ω/m , enquanto o L indica a indutância em série, também por unidade de comprimento, dada em H/m . Os outros dois parâmetros são a capacitância C e a condutância G em paralelo, ambos por unidade de comprimento, dados em F/m e S/m , respectivamente. Maiores detalhes sobre os parâmetros distribuídos são discutidos no Apêndice A. Os parâmetros β_l e α_l são as constantes de fase e de atenuação de RF para a linha de transmissão com carga. Sob condições de baixas perdas, podemos definir as perdas na linha de transmissão RF como a soma das contribuições dos condutores e dos dielétricos da seguinte forma:

$$\alpha_l[\text{Np}/\text{m}] = \alpha_c + \alpha_d = \frac{R}{Z_{0l}} + (G + R_{tpn}^{-1}Z_{0l}), \quad (50)$$

em que os parâmetros α_c e α_d são as constantes de atenuação devido aos condutores e o dielétrico, respectivamente. A impedância da linha carregada, dada as condições de baixas perdas, pode ser simplificada para:

$$Z_{0l} = \sqrt{\frac{L}{C + C_{tpn}}}. \quad (51)$$

A constante de fase do guia carregado, β_l , e o índice de refração efetivo do mesmo, $n_{eff,l}$ podem ser dados por:

$$\beta_l = \omega \sqrt{L(C + C_{tpn})}, \quad (52)$$

$$n_{eff,l} = c \sqrt{L(C + C_{tpn})}. \quad (53)$$

Com a modelagem realizada até então, podemos estimar a BW_{EO} a partir dos parâmetros das linhas de transmissão. Na seguinte subseção iremos mostrar como esta relação pode ser representada.

2.4.3 Banda Eletro-Óptica Considerando os Parâmetros da Linha de Transmissão

O calculo da BW_{EO} do MZM previamente discutida, dada pela Equação 35, considera unicamente a junção PN do dispositivo. Na prática, como a arquitetura elétrica é necessária para aplicar a tensão na junção PN, o modelo apresenta certas limitações. Em (PATEL *et al.*, 2015) demonstrou-se que os efeitos elétricos que mais afetam a BW_{EO} são: o descasamento de impedância entre o terminal do *chip* e o eletrodo, o descasamento de velocidade de grupo entre as ondas RF e óptica e as perdas de micro-ondas. Para estimar a BW_{EO} , consideramos um sinal elétrico com amplitude V_g e operando a uma frequência $\omega_g = 2\pi f_g$, sendo que a tensão média aplicada entre os eletrodos do modulador é dada por (YU; BOGAERTS, 2012):

$$V_{avg}(\omega) = \frac{V_g (1 + p_1)(V_+ + p_2 V_-) \exp(j\beta_o L_{ps})}{2 [\exp(\gamma_l L_{ps}) + p_1 p_2 \exp(-\gamma_l L_{ps})]}, \quad (54)$$

em que:

$$V_{\pm} = \exp(\pm j\phi_{\pm}) \frac{\text{sen}(\phi_{\pm})}{\phi_{\pm}}, \quad (55)$$

$$\phi_{\pm} = -(j\gamma_l \pm \beta_o) \frac{L_{ps}}{2}, \quad (56)$$

$$p_1 = \frac{Z_{0l} - Z_s}{Z_{0l} + Z_s}, \quad (57)$$

$$p_2 = \frac{Z_T - Z_{0l}}{Z_T + Z_{0l}}, \quad (58)$$

$$\beta_o = \frac{\omega_g n_g}{c}. \quad (59)$$

Na Equação 54 definimos o V_{avg} como função de ω , isto ocorre pois o mesmo depende de γ_l , β_o e das impedâncias, que por sua vez são funções de ω . Nas Equações 55-59, Z_s é a impedância da fonte, Z_T é a impedância do terminal da linha de transmissão, ϕ representa a variação de fase resultante da propagação pelo braço do MZM, c é a velocidade da luz no vácuo, n_g é o índice de grupo óptico e β_o é a constante de fase óptica. Operando em regime de pequenos sinais, temos que a tensão V_{avg} é equivalente à média da tensão dos

sinais modulantes que são aplicados na junção PN. Temos que a resposta em frequência eletro-óptica do modulador é dada por:

$$m(\omega) = \left| \frac{(1 + j\omega C_j R_j) V_{avg}(\omega)}{((1 + j\omega C_j R_j) V_{avg}(\omega_0))} \right|, \quad (60)$$

em que $V_{avg}(\omega_0)$ é a tensão média de mais baixa frequência da fonte elétrica.

A Equação 60 consegue descrever a resposta eletro-óptica em frequência do MZM considerando unicamente a dinâmica de portadores na junção PN e a tensão aplicada. Porém, para observarmos o comportamento em frequência do MZM considerando as limitações do TWE, é conveniente descrevermos a resposta $m(\omega)$ explicitamente em função dos parâmetros da linha de RF. Apesar de desenvolvido para moduladores LiNbO_3 , a equação derivada em (GHIONE, 2009) possui um comportamento aproximadamente parecido ao observado em MZMs de silício (PATEL *et al.*, 2014), e é dada por:

$$m'(\omega) = \frac{R_T + R_S}{R_T} \left| \frac{Z_{in}}{Z_{in} + Z_S} \right| \left| \frac{(Z_T + Z_{0l})F(u_+) + (Z_T - Z_{0l})F(u_-)}{(Z_T + Z_{0l}) \exp(\gamma_l L_{ps}) + (Z_T - Z_{0l}) \exp(-\gamma_l L_{ps})} \right|, \quad (61)$$

em que R_T e R_S indicam a resistência de terminação e de fonte, Z_{in} é a impedância de entrada da linha de transmissão carregada, dada por:

$$Z_{in} = Z_{0l} \frac{Z_T + Z_{0l} \tanh(\gamma_l L_{ps})}{Z_{0l} + Z_T \tanh(\gamma_l L_{ps})}, \quad (62)$$

e o termo $F(u_{\pm})$ é dado por:

$$F(u_{\pm}) = \frac{1 - \exp(u_{\pm})}{u_{\pm}}, \quad (63)$$

e $u_{\pm}(\omega)$ é dado por:

$$u_{\pm}(\omega) = \pm \alpha_l L_{ps} + j \frac{\omega}{c} (\pm n_{RFg} - n_g) L_{ps}, \quad (64)$$

em que n_{RFg} é o índice de grupo de RF. O trabalho desenvolvido em (PATEL *et al.*, 2014), por meio da Equação 61, analisou os três aspectos principais na interação entre a onda óptica e a onda de RF que atuam na BW_{EO} . Realizaremos no presente trabalho uma breve revisão sobre estes aspectos, pois são de grande importância no *design* de um MZM.

Descasamento de Impedância

O descasamento de impedância está relacionado com o efeito de reflexão nas interfaces. A condição de casamento é de que a impedância de todos componentes sejam equivalentes, ou seja, $Z_T = Z_S = Z_{0l}$. O trabalho realizado em (YU; BOGAERTS, 2012) explora um efeito interessante que ocorre ao forçarmos um descasamento entre a impedância de carga e a da fonte. Ao manter $Z_S = 50 \Omega$ e $Z_T = 35 \Omega$, notou-se alguns picos (*ripples*) em frequências baixas, por volta de 12 GHz, no trabalho de (ALAM *et al.*, 2021). Estes *ripples* funcionam como uma pré-enfase, e são capazes de estender o BW_{EO} do MZM. Este aumento de banda vem com o custo de causar uma menor amplitude na resposta do MZM para frequências baixas. Tomando a Equação 61 e considerando um casamento de velocidade perfeito e desconsiderando as perdas (GHIONE, 2009), podemos organizar a Equação 61 da seguinte forma:

$$m(\omega) = \left| \frac{(1 - \Gamma_T \Gamma_S)[1 + \Gamma_T F(U_-)]}{1 - \Gamma_T \Gamma_S \exp(-2j\beta_l L)} \right|, \quad (65)$$

em que Γ_T e Γ_S são os coeficientes de reflexão na interface com a terminação e com a fonte, respectivamente. Temos que U_- é dado por:

$$U_- = -j \frac{2\omega}{c} n_g L_{ps}. \quad (66)$$

Entretanto, ao causar tal descasamento entre a impedância de entrada da linha e a impedância da fonte, ocorrerá uma redução da tensão V_S que é transferida para linha. Neste caso, podemos estimar a nova tensão V_{in} que será injetada na linha, considerando tal descasamento, dada pela seguinte relação:

$$V_{in} = V_S \frac{Z_{in}}{Z_S + Z_{in}}. \quad (67)$$

Descasamento de Velocidade

Considerando uma linha de transmissão com impedância perfeitamente casada e sem perdas, a resposta do MZM se resume a:

$$m(\omega) = \left| \text{sinc} \left(\frac{\pi f \Delta n_{RF-o} L_{ps}}{c} \right) \right|, \quad (68)$$

em que Δn_{RF-o} é a diferença entre o índice de refração da onda óptica e de RF (GHIONE, 2009). A BW_{EO} de 3 dB pode ser aproximada por:

$$BW_{EO}[\text{Hz}] = \frac{1,89c}{\pi|n_{RFg} - n_g|L_{ps}}. \quad (69)$$

Podemos perceber que a BW_{EO} do MZM tende a diminuir conforme aumentamos a disparidade entre o índice de grupo óptico e de RF. Como pode ser observado, o comprimento do deslocador de fase também é um parâmetro que diminui o BW_{EO} .

Perdas de RF

Para análise do impacto das perdas de RF na resposta eletro-óptica consideramos agora uma situação de casamento perfeito entre as impedâncias da linha, fonte e terminação, assim como um casamento perfeito entre as velocidades de grupo das ondas ópticas e elétricas. Dadas estas condições, podemos re-escrever a Equação 61 como:

$$m(\omega) = \exp\left(-\alpha_l(\omega)\frac{L_{ps}}{2}\right) \left| \frac{\sinh\left(\alpha_l(\omega)\frac{L_{ps}}{2}\right)}{\alpha_l(\omega)\frac{L_{ps}}{2}} \right|. \quad (70)$$

Podemos observar que a resposta do MZM decai conforme aumentamos as perdas de RF ou, da mesma forma, o comprimento do deslocador de fase.

3 Fundamentos de Inteligência Artificial

Na presente seção iremos discutir alguns conceitos e algoritmos de inteligência artificial (IA). A discussão sobre IA aqui inclui tanto a subárea de aprendizagem de máquina (*Machine Learning*, ML) quanto algoritmos heurísticos baseados em processos biológicos, mais especificamente, o DE. Iremos discutir também a área de aprendizagem profunda (*Deep Learning*, DL), que é um subconjunto de algoritmos de ML. Neste trabalho, exploram-se tanto modelos de DL quanto o DE para a otimização do MZM. Na Seção 3.1 definiremos ML de forma geral, introduzindo esta subárea da IA e também algumas de suas aplicações em sistemas de telecomunicações. Nas discussões, uma maior ênfase será direcionada aos modelos de DL supervisionados e de regressão, pois estes serão os utilizados no presente trabalho. Modelos supervisionados são aqueles que têm seu treinamento realizado a partir da observação de saídas conhecidas, enquanto regressão se refere àqueles que representam a saída a partir de valores contínuos. Na Seção 3.1 discutiremos também o funcionamento de ANNs, assim como ferramentas que podem ser utilizadas para aprimorá-las e métricas de desempenho geralmente usadas. Já na Seção 3.2 discutiremos o algoritmo heurístico utilizado otimização do *design* do MZM, isto é, o DE. Por fim, na Seção 3.3, apresentaremos a configuração de otimização que integra um modelo de ANN e o algoritmo heurístico DE.

3.1 Princípios de Aprendizagem de Máquinas e Aprendizagem Profunda

Na presente seção iremos discutir os fundamentos básicos de ML e de DL. Na Subseção 3.1.1 iremos apresentar a área de ML, assim como seus fundamentos básicos. Na Subseção 3.1.2 iremos discutir como se dá o funcionamento de um ANN, tanto no treinamento quanto na inferência de amostras. Na Subseção 3.1.3 iremos discutir técnicas de aprimoramento de desempenho de ANNs. Por fim, na Subseção 3.1.4 iremos discutir algumas métricas para avaliação de desempenho da ANN.

3.1.1 Aprendizagem de Máquinas

Com a capacidade de computadores atuais de processar enormes quantidades de dados, sobretudo de alta dimensionalidade, algoritmos de ML apresentam uma poderosa ferramenta para solução de problemas complexos. ML se refere a uma categoria de algoritmos que são desenvolvidos com o intuito de imitar a habilidade de aprendizado humano. Estes algoritmos aprendem sem serem programados explicitamente (SAMUEL, 1959). Além do mais, algoritmos de ML conseguem extrair informação e padrões de um conjunto complexo de dados, que seriam de difícil modelagem. Com o avanço tecnológico atual, diversos dispositivos eletrônicos que fazem parte do cotidiano utilizam massivamente de tais algoritmos. Como exemplos de aplicação prática, ML capacita que sites *online* direcionem anúncios específicos para um dado usuário, filtrar *spam* em *e-mails*, e também reconhecimento facial e de voz em celulares (SARKER, 2021).

Em geral, podemos classificar os algoritmos de ML de acordo com sua natureza de aprendizagem. Há dois tipos fundamentais neste quesito, sendo estes categorizados como supervisionados ou não supervisionados. Em algoritmos supervisionados, o aprendizado se dá a partir de entradas previamente rotuladas, ou seja, conhece-se a saída esperada para uma amostra a partir dos dados de treinamento (MOHRI; ROSTAMIZADEH; TALWALKAR, 2018). Em algoritmos não supervisionados, o conjunto de treinamento dispõe apenas dos valores de entrada, sendo o objetivo do algoritmo encontrar uma síntese de tais dados (HINTON; SEJNOWSKI, 1999).

Se tratando de algoritmos não supervisionados, uma das finalidades mais comuns neste tipo se referem a técnicas de clusterização (*clustering*). Clusterização tem o desígnio de agrupar dados de acordo com o seu grau de semelhança, formando então os conjuntos denominados *clusters*. Geralmente o grau de semelhança utilizado é baseado na distância entre os valores das amostras, utilizando características estatísticas como média e variância para geração dos agrupamentos.

Por outro lado, algoritmos supervisionados observam o conjunto de treinamento, conseguindo aprender uma função que mapeia um conjunto de entradas para um conjunto de saídas. Em relação aos algoritmos supervisionados, podemos ainda categorizá-los como sendo de regressão ou de classificação (SHALEV-SHWARTZ; BEN-DAVID, 2014). O que diferencia ambas categorias é se o tipo de saída dos dados possui natureza contínua ou

discreta. No caso de algoritmos de regressão, os dados de entrada são mapeados em uma saída contínua.

Já para o problema de classificação, os algoritmos mapeiam as entradas em uma saída discreta, ou seja, as saídas do algoritmo apresentam uma gama limitada de valores, estes representando diferentes classes. Neste trabalho, utilizaremos o algoritmo de ANN como forma supervisionada de regressão.

Como forma de exemplificar o funcionamento de um algoritmo supervisionado de regressão, tomemos o problema que consiste em prever o preço de um imóvel de acordo com o tamanho de sua área. Os dados de treinamento são compostos de um conjunto de imóveis, cujas áreas (entradas) e preços (saídas) são conhecidas. Durante a fase de treinamento, o algoritmo otimiza seus parâmetros de acordo com os dados de treinamento. O intuito desta fase é capacitar o modelo em prever preço de imóveis, inclusive de amostras que não estão incluídas no conjunto de treinamento. Já na fase de inferência, podemos obter uma estimativa do preço de um imóvel arbitrário ao introduzirmos a área do mesmo no algoritmo.

Aprendizagem profunda (*Deep Learning*, DL) é uma subárea de ML que representa modelos compostos de múltiplos níveis de abstração (LECUN; BENGIO; HINTON, 2015). Estes tipos de arquiteturas permitem extrair padrões de sistemas complexos e apresentam a premissa de reduzir a interação de humanos no desenvolvimento dos modelos. Algoritmos de DL são baseados em arquiteturas de redes neurais artificiais, e a profundidade em ANNs se remete à uma maior quantidade de camadas de processamento.

Modelos de DL são amplamente aplicados na área de telecomunicações, como em DSP (LIU *et al.*, 2020), em sistemas de comunicações sem fio (DAI *et al.*, 2020; LIAO; WEI; ZOU, 2020), em sistemas de comunicações ópticas (KUTZ, 2021; SUN *et al.*, 2020) e, mais recentemente, em fotônica (MA *et al.*, 2020). Assim, em sistemas fotônicos, modelos de DL nos permite inferir a influência de cada parâmetro de um dispositivo na manipulação da luz. Podemos citar alguns dispositivos fotônicos produzidos utilizando ANNs profundas, tais como nanoestruturas plasmônicas (MALKIEL *et al.*, 2018), meta-materiais (MA; CHENG; LIU, 2018) e nanopartículas que aproximam o espalhamento da luz (PEURIFOY *et al.*, 2018). Entretanto, até onde sabemos, não há trabalhos que otimizem, por meio de DL, o *design* de moduladores ópticos integrados, especialmente na plataforma de silício.

Como mencionado, modelos supervisionados de ANNs requerem um conjunto de treinamento para otimizar seus parâmetros. Devido à alta capacidade de ANNs profundas

de se adaptarem aos dados, ocasionalmente o modelo pode memorizar o conjunto de treinamento, porém apresentando ineficiência em prever a saída para novos dados de entrada. Ao realizarmos muitas iterações, ou ao utilizarmos arquiteturas excessivamente robustas, o algoritmo irá também se adaptar aos ruídos e irregularidades dos dados do conjunto de treinamento. Tal efeito, denominado *overfitting*, faz com que o algoritmo apresente um bom desempenho ao processar os dados de treinamento, porém seja incapaz de generalizar seus resultados para dados não vistos. De fato, é provado que uma ANN com um número suficiente de parâmetros pode aproximar, com alto grau de acurácia, qualquer função mensurável, incluindo ruídos (HORNİK; STINCHCOMBE; WHITE, 1989). Por outro lado, se o algoritmo demonstra uma baixa performance tanto no conjunto de treinamento quanto no conjunto de testes, mesmo depois de ter convergido durante o processo de treinamento, ele provavelmente sofre de *underfitting*. Em contraste com o caso de *overfitting*, tal efeito ocorre quando o modelo é simples demais para se adequar aos dados de treinamento.

Devido ao *overfitting*, ao utilizar apenas os próprios exemplos de treinamento na avaliação do modelo, temos resultados que podem superestimar seu desempenho. Considerando este fenômeno, é válido reservar uma parcela dos dados disponíveis para se avaliar a capacidade de generalização do algoritmo. Este conjunto de dados é denominado conjunto de testes (SHALEV-SHWARTZ; BEN-DAVID, 2014). Além do conjunto de testes, podemos definir o denominado conjunto de validação. Dos dados que restaram após definir-se o conjunto de treinamento, uma parcela destes será distribuída para o conjunto de validação e o restante para o conjunto de testes. Os dados de validação servem para avaliarmos o desempenho do algoritmo a cada iteração e para compararmos diferentes configurações do modelo, enquanto que os dados de teste servem para ser utilizados após o treinamento do modelo, de forma a estimar seu desempenho em dados futuros.

A Figura 17 ilustra a operação de três modelos de regressão em um conjunto de dados de uma dimensão. Os dados estão separados entre um subconjunto de treinamento e, para fins de ilustração, utilizaremos o subconjunto restante tanto para validação quanto para teste dos modelos. Dos modelos, um destes apresenta *overfitting*, o outro *underfitting* e o último apresenta um comportamento adequado. A Subfigura 17 (a) ilustra as curvas de mapeamento obtidas pelos diferentes modelos em relação ao conjunto de treinamento e de teste. A Subfigura 17 (b) ilustra o erro quadrático médio (*Mean Squared Error*, MSE) do modelo, em escala logarítmica, obtido tanto para o conjunto de treinamento quanto para

o conjunto de validação, em relação às iterações de treinamento. Esta curva é denominada curva de aprendizado, e será discutida com maiores detalhes na Subseção 3.1.4.

Como podemos observar na Subfigura 17 (a), no caso em que o modelo é linear, este apresenta o efeito de *underfitting* neste conjunto de dados. Não importa quantas iterações este realiza durante o treinamento, o erro em ambos conjuntos não irá ser reduzido a partir de um certo ponto. Já para o caso em que ocorre *overfitting*, como podemos observar na Subfigura 17 (a), o modelo se esforça para zerar o erro em todas amostras de treinamento. Como podemos observar na Subfigura 17 (b), este comportamento tende a aumentar o erro no conjunto de validação enquanto o diminui no conjunto de treinamento. Já para o caso adequado, observamos que a curva segue a distribuição dos dados. Isto resulta em um erro de validação próximo do erro de treinamento, e este converge para um valor menor em relação aqueles obtidos pelos outros modelos.

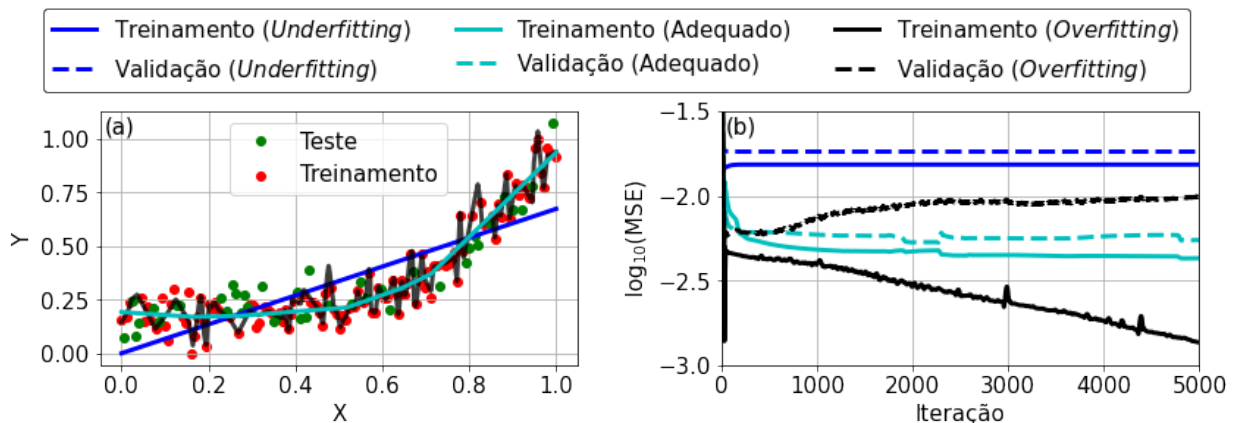


Figura 17 – Comparação entre três modelos de regressão, em que um deles apresenta *underfitting*, o outro apresenta *overfitting*, e o último demonstra um mapeamento adequado. (a) Conjunto de dados, de teste e de treinamento, e as curvas de mapeamento obtidas pelos modelos de regressão. (b) Erro do conjunto de treinamento e de validação em relação às iterações de treinamento.

Modelos de ANNs são compostos de uma variedade de parâmetros internos, tais quais determinam a saída de acordo com uma entrada recebida. Neste caso então, a fase de treinamento se resume em encontrar os valores destes parâmetros que melhor consigam descrever a distribuição de dados. Esta otimização dos parâmetros se dá de forma a minimizar as diferenças entre as saídas reais dos exemplos de treinamento e as saídas do

algoritmo. Para o caso de regressão, dado um conjunto de treinamento \mathbf{S} , o qual possui m exemplos de treinamento, podemos defini-lo como:

$$\mathbf{S} = (\mathbf{x}^{(1)}, \mathbf{y}^{(1)}), (\mathbf{x}^{(2)}, \mathbf{y}^{(2)}), \dots, (\mathbf{x}^{(m)}, \mathbf{y}^{(m)}), \quad \mathbf{x}^{(i)} \in \mathbb{R}^n, \mathbf{y}^{(i)} \in \mathbb{R}^k, \quad 1 \leq i \leq m, \quad (71)$$

em que cada exemplo de treinamento tem n atributos de entrada e k saídas.

Um conceito que é válido introduzir é o de redimensionamento de atributos, que consiste em manter os valores dos atributos entre os mesmos intervalos de valores e com uma distribuição estatística similar. Tal prática geralmente resulta em uma convergência mais rápida dos algoritmos (BURKOV, 2019). Para realizarmos tal redimensionamento, iremos utilizar a estandardização (ou *Z-score*). Desta forma, podemos redefinir os valores do j -ésimo atributo dos m exemplos de treinamento como sendo:

$$x'_j = \frac{x_j - \mu}{\sigma}, \quad j \in [1, n], \quad (72)$$

em que μ é a média das amostras considerando o j -ésimo atributo, assim como σ é o desvio padrão destas. Tal transformação assegura que as distribuições terão média nula e desvio padrão unitário.

A Figura 18 ilustra como a estandardização transforma uma distribuição. Podemos observar na Figura 18(a) uma distribuição normal com média 10 e desvio padrão de 4. Na Figura 18(b), após a aplicação da Equação 72, a distribuição possui agora uma média nula e um desvio padrão de 1. Apesar do exemplo da Figura 18 se tratar de uma distribuição normal, a estandardização pode ser aplicada à qualquer tipo de distribuição, desde que tenhamos informações sobre a média e a variância da mesma (MENDENHALL; SINCICH, 2006).

3.1.2 Funcionamento de Redes Neurais Artificiais

Superficialmente baseadas no funcionamento do cérebro, uma ANN consiste de diversos módulos (unidades neurais) conectados entre si em uma estrutura de rede (BASHEER; HAJMEER, 2000). Uma unidade neural recebe um conjunto de entradas para reproduzir uma saída. As entradas são combinadas linearmente com os parâmetros (pesos) do neurônio, e então aplica-se uma função de ativação nesta combinação. A partir de um conjunto de treinamento rotulado dado pela Equação 71, um neurônio pondera tais entradas a partir

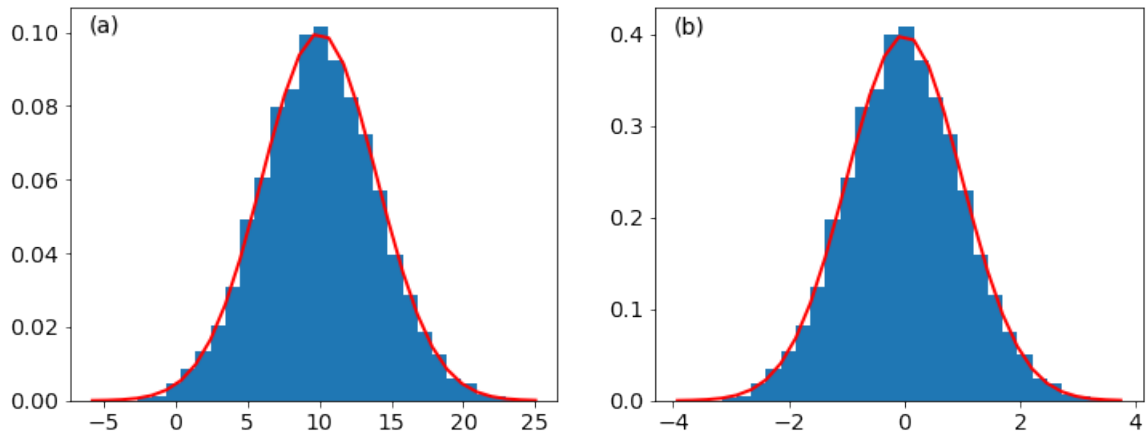


Figura 18 – (a) Uma distribuição gaussiana com média 10 e desvio padrão de 4 e (b) a mesma distribuição após a estandardização.

de um conjunto de pesos $\mathbf{w} \in \mathbb{R}^n$. Podemos definir um potencial de ativação u , o qual depende desta combinação, e é dado por:

$$u = u(\mathbf{x}) = \sum_{k=1}^n w_k \cdot x_k + b = \mathbf{w}^T \cdot \mathbf{x} + b, \quad (73)$$

em que o termo b representa o termo de *bias*, tal qual introduz *offset* na função de ativação. Neste caso, x_k indica a k -ésima entrada do unidade neural, e o superescrito T indica a matriz transposta.

Para o mapeamento não linear da ANN, é essencial aplicarmos uma função de ativação em u . Para tal tarefa utilizaremos uma função denominada função unidade linear retificada (*Rectified Linear Unit*, ReLu) (AGARAP, 2019). Ao utilizar a função ReLu, o

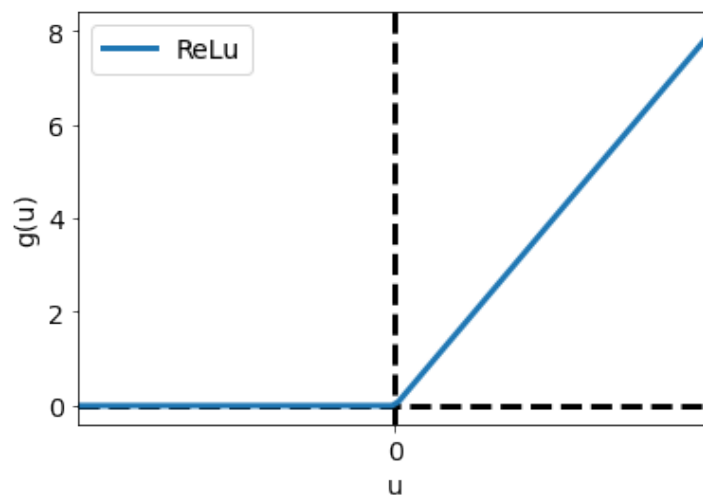


Figura 19 – Função de ativação unidade linear retificada (ReLU).

trabalho feito em (KRIZHEVSKY; SUTSKEVER; HINTON, 2017) mostrou que o modelo é capaz de convergir até 6 vezes mais rápido em comparação com a utilização de outras funções de ativação tradicionais, como a tangente hiperbólica. Atribui-se estes resultados ao fato de que o ReLu é uma função não saturante, o que pode aumentar a magnitude do gradiente na otimização. A função ReLu g , aplicada em u , definirá a saída do neurônio, e esta é dada por:

$$h(\mathbf{x}) = g(u) = \max(0, u), \quad (74)$$

tal como mostra a Figura 19.

Até então tratamos de uma configuração com apenas um único neurônio, porém uma ANN é a conexão de diversos neurônios que podem estar estruturados em camadas, com variadas quantidades de neurônios. Iremos discutir aqui o caso de uma rede totalmente conectada, sendo denominada também como perceptron multicamada (*Multilayer Perceptron*, MLP). Em um MLP, cada unidade neural recebe como entrada as saídas de todos neurônios da camada anterior, realiza o procedimento explicado nas Equações 73 e 74, e então transmite a informação para as camadas seguintes.

A Figura 20 ilustra uma ANN do tipo totalmente conectada, MLP. Os neurônios são representados pelos círculos em azul e os enlaces que ligam os neurônios pelas linhas em preto. As ANNs discutidas neste projeto se tratam das de arquitetura *feedforward*, tais quais não possuem retroalimentações (HAYKIN, 2007). Este tipo de redes são flexíveis e permitem um treinamento relativamente simples e intuitivo.

A rede ilustrada pela Figura 20 é composta de um número arbitrário de camadas. A primeira camada é conhecida como camada de entrada, e recebe as informações de uma dada amostra. Cada neurônio de entrada remete a um atributo x_j , $j = 1, 2, 3, \dots, n$. A última camada é denominada camada de saída, e é responsável por apresentar os resultados, além de também realizar parte do processamento (SILVA; SPATTI; FLAUZINO, 2010). As camadas intermediárias, entre a de entrada e a de saída, são denominadas camadas escondidas, e possuem a finalidade de extrair atributos dos exemplos de treinamento.

Em uma rede como descrita pela Figura 20, cada neurônio que compõe a camada escondida ou de saída, de acordo com as Equações 73 e 74, é alimentado pelas saídas dos neurônios da camada anterior. Este processo de inserção de uma amostra e obtenção de uma saída é denominado alimentação direta (*Feedforward Propagation*, FP).

Como mencionado, uma ANN faz uso do conjunto de treinamento para se adaptar aos exemplos. Primeiramente, define-se uma função custo, esta com a finalidade de avaliar o desempenho da rede a cada iteração. O objetivo do treinamento é minimizar a função custo a partir da otimização dos valores dos pesos da ANN. Utilizamos a definição de gradiente para obter a variação dos parâmetros que causará uma redução na função custo. Como a rede possui múltiplos nós, para o cálculo do gradiente é necessária a aplicação da regra da cadeia. A regra da cadeia é implementada via um algoritmo denominado propagação retroativa (BP, *Backward Propagation*) (SHALEV-SHWARTZ; BEN-DAVID, 2014). Após o BP encontrar os valores dos gradientes, utilizamos um algoritmo denominado gradiente descendente (GD, *Gradient Descent*) para atualizar os pesos. Tal algoritmo representa um método de otimização iterativo, que busca um mínimo local de uma função seguindo a direção de descida mais inclinada. Dado uma função custo J , o algoritmo GD

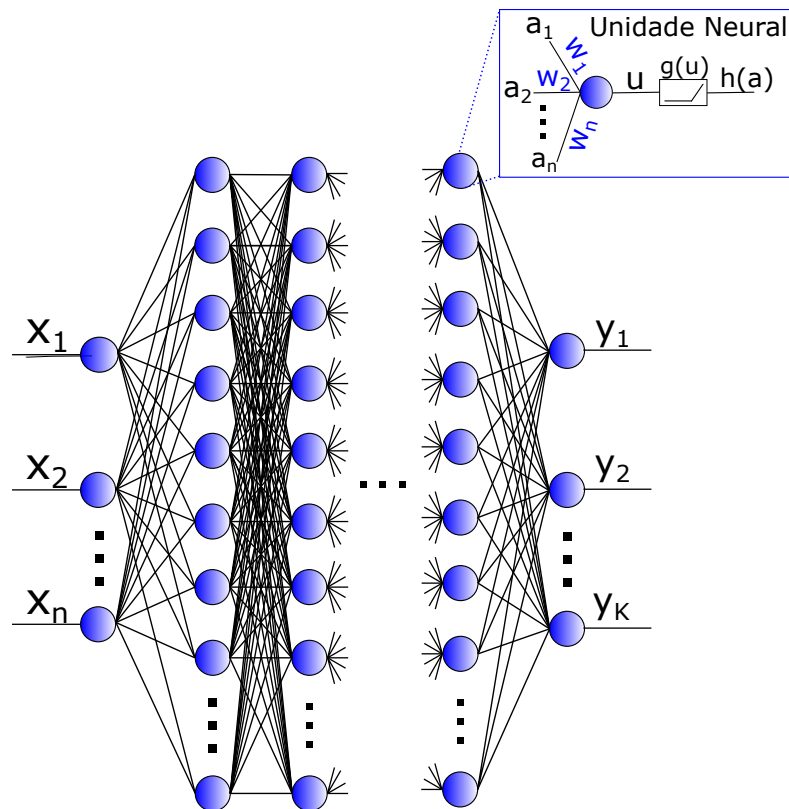


Figura 20 – Uma ANN totalmente conectada composta de L_c camadas com quantidades arbitrárias de unidades neurais. Tem-se n neurônios de entrada K neurônios de saída.

tem como objetivo realizar uma atualização dos pesos para minimizar a mesma, dada como:

$$\begin{aligned} \Delta b &= -\eta \cdot \frac{\partial J}{\partial b}, \\ \Delta w_k &= -\eta \cdot \frac{\partial J}{\partial w_k}, \text{ para } k \in \{1, \dots, n\} \end{aligned} \quad (75)$$

onde η representa a taxa de aprendizado, uma constante que define o tamanho do deslocamento (passo) rumo ao mínimo de J . Uma função custo amplamente utilizada para algoritmos de regressão é, compreensivelmente, o MSE, que é dado por:

$$J = J(\mathbf{w}, b) = \frac{1}{m} \sum_{i=1}^m (h(x^{(i)}) - y^{(i)})^2, \quad (76)$$

considerando os m exemplos de treinamento. Na equação 76, $h(x^{(i)})$ e $y^{(i)}$ representam, respectivamente, a saída da ANN e as saídas reais das amostras para o i -ésimo exemplo de treinamento. É válido mencionar que $h(x^{(i)})$ é função dos pesos sinápticos e dos termos de *bias* de cada camada, como observado nas Equações 73 e 74.

De modo a considerar todos neurônios da camada de saída da rede, podemos re-definir a função custo J como:

$$J = \frac{1}{m} \sum_{i=1}^m \sum_{k=1}^K \left[\left(h_k(x^{(i)}) - y_k^{(i)} \right)^2 \right]. \quad (77)$$

Por meio da BP, reajustam-se os pesos dos neurônios, camada por camada em direção contra propagante (SCHMIDHUBER, 2015). O gradiente da função custo ∇J , de acordo com a regra da cadeia, dada uma rede com L_c camadas, pode ser definido como:

$$\nabla J = \frac{\partial J}{\partial \mathbf{w}'^{(L_c)}} = \frac{\partial J}{\partial a_i^{(L_c)}} \cdot \frac{\partial a_i^{(L_c)}}{\partial u_i} \cdot \frac{\partial u_i}{\partial \mathbf{w}'^{(L_c)}}. \quad (78)$$

Na Equação 78, a matriz $\mathbf{w}'^{(L_c)}$ de pesos sinápticos é uma concatenação do vetor de *bias* $b^{(L_c)}$ e da matriz de pesos $\mathbf{w}^{(L_c)}$. Os termos $a_i^{(L_c)}$ e u_i representam, respectivamente, a saída do neurônio i e a função potencial de ativação, relacionados a camada de saída, ou seja, $a_i^{(L_c)} = h_i(x)$. As derivadas das Equações 73, 74 e 77 podem ser dadas por:

$$\frac{\partial J}{\partial a_i^{(L_c)}} = -2(y_i - a_i^{(L_c)}), \quad (79)$$

$$\frac{\partial a_i^{(L_c)}}{\partial u_i} = g'(u_j), \quad (80)$$

$$\frac{\partial u_i}{\partial \mathbf{w}'^{(L_c)}} = a_i^{(L_c-1)}. \quad (81)$$

A derivada da função ReLU $g(u)$ em relação a u é simplesmente a derivada do maior valor entre 0 e u , podendo ser então 0 ou 1, respectivamente. Apesar da função ReLU $g(u)$ não ser diferenciável em $u = 0$, na prática a definimos como sendo nula, ou seja:

$$g'(u) = \begin{cases} 1, & \text{se } u > 0, \\ 0, & \text{se } u \leq 0, \end{cases} \quad (82)$$

o que representa uma função degrau unitário, ou função de Heaviside. A partir das derivadas obtidas pelas equações 79, 80 e 81, é possível substituímos seus resultados na Equação 78, reescrevendo-a como:

$$\nabla J = \frac{\partial J}{\partial \mathbf{w}'^{(L_c)}} = -2(y_i - a_i^{(L_c)}) \cdot g'(u_j) \cdot a_i^{(L_c-1)} = -\delta_j \cdot a_i^{(L_c-1)}, \quad (83)$$

em que δ_j é denominado gradiente local, dado por:

$$\delta_j = -2(y_i - a_i^{(L_c)}) \cdot g'(u_j). \quad (84)$$

O próximo passo é atualizarmos os valores dos pesos, tal atualização, considerando GD comum, é dada por:

$$\Delta \mathbf{w}'^{(L_c)} = -\eta \frac{\partial J}{\partial \mathbf{w}'^{(L_c)}} = \eta \cdot \delta_j \cdot a_i^{(L_c-1)}, \quad (85)$$

que descreve simplesmente a trajetória dos valores dos pesos em direção ao mínimo da função de custo.

Utilizando os pesos atualizados da última camada, as camadas anteriores realizarão o mesmo processo (ALPAYDIN, 2010). Da mesma forma descrita pela Equação 83, sendo os termos \mathbf{w}' , a_i e u_i referentes à camada a ser processada, pode-se provar que o ajuste de pesos sinápticos para a l -ésima camada escondida é dada por:

$$\Delta \mathbf{w}'^{(l)} = \eta \cdot \delta_j^{(l)} \cdot a_i^{(l-1)}, \quad (86)$$

sendo $a_{(i)}^{(l-1)}$ as saídas da camada anterior à l -ésima camada. o gradiente local $\delta_j^{(i)}$ é dado por:

$$\delta_j^{(l)} = g'(u_j) \sum_{k \in \text{layer}} \delta_k^{(l+1)} \mathbf{w}'^{(l+1)}. \quad (87)$$

O procedimento é realizado até que se alcance a primeira camada. É denominado época (*epoch*) uma iteração completa de FP e em seguida BP. Uma ANN realiza um determinado número de épocas antes de convergir para os valores ótimos das matrizes de pesos $\mathbf{w}'^{(l)}$ para as l camadas. É válido ressaltar que os pesos devem ser inicializados de forma aleatória, de forma a evitar simetrias entre os estados iniciais dos neurônios da rede (ALPAYDIN, 2010).

3.1.3 Aprimoramento de Redes Neurais Artificiais

Na presente subseção, iremos discutir técnicas utilizadas para melhorar o treinamento e a capacidade de generalização de ANNs. Primeiramente, iremos discutir sobre como é implementado o método Kaiming He de inicialização de pesos em ANNs. Discutiremos também sobre otimizadores como o gradiente descendente estocástico (*Stochastic Gradient Descent*, SGD) e o otimizador *AdamW*, utilizado em nosso trabalho. Então iremos falar sobre ferramentas que melhoram o desempenho de ANNs profundas, entre eles o *batch normalization* (BN), *drop out* (DO) e *residual connections* (RES). Estas permitem um aprimoramento tanto na velocidade de treinamento quanto na capacidade de generalização da ANN (KIM, 2021).

Método uniforme Kaiming He de inicialização de pesos

Ao inicializarmos os pesos usando uma distribuição normal com desvio padrão fixo, uma ANN profunda pode apresentar problemas de convergência (SIMONYAN; ZISSERMAN, 2014). Neste caso, como a variância dos pesos iniciais não é controlada, a ANN pode sofrer dos problemas de dissipação de gradiente ou de explosão de gradiente. Estes ocorrem quando os gradientes tomam valores demasiadamente baixos ou altos, respectivamente. O método uniforme de Kaiming He de inicialização de pesos consiste em contornar este problema para camadas que utilizem a função de ativação ReLu. O

trabalho realizado em (HE *et al.*, 2015) provou que, para a l -ésima camada de uma ANN, a condição para evitar valores extremos de gradientes é dada por:

$$\frac{1}{2}s_l \text{Var} [\mathbf{w}'^{(l)}] = 1. \quad (88)$$

Na Equação 88, o termo s_l representa a quantidade de unidades neurais da l -ésima camada. Esta equação implica que a inicialização dos pesos da l -ésima camada seguirá uma distribuição gaussiana:

$$\mathbf{w}'^{(l)} \sim \mathcal{N} \left(0, \frac{2}{s_l} \right). \quad (89)$$

Métodos Avançados de Gradiente Descendente

Nesta subsubseção introduzimos uma variação do GD denominada SGD. Nesta configuração, o conjunto de dados de treinamento é dividido em conjuntos menores (*mini-batches*) e a otimização é aplicada em cada conjunto separadamente para cada época (KLEINER; WOLFOWITZ, 1952; RUDER, 2016). Desta forma, além de economizar espaço na memória do computador, é comprovado que a convergência é alcançada mais rapidamente e a otimização pode evitar mínimos locais. Entretanto, quanto menor o tamanho dos *mini-batches*, mais ruidosa são as atualizações dos pesos. É possível observar também que, devido ao menor paralelismo dos cálculos no FP e BP, o custo computacional é aumentado.

Um outro algoritmo de otimização baseada em gradiente de primeira ordem que se destaca é o *Adam* (KINGMA; BA, 2014). Tal algoritmo, além de apresentar uma alta eficiência computacional, é capaz de adaptar a taxa de aprendizado para cada parâmetro de acordo com o primeiro e o segundo momento dos gradientes. Em outras palavras, o *Adam* ajusta o tamanho do passo de aprendizado de acordo com a magnitude e a variação do gradiente, para um dado parâmetro em uma dada iteração. Considerando o otimizador *Adam*, a atualização de peso é dada, considerando o t -ésimo época, por:

$$\mathbf{w}'_{t+1} = \mathbf{w}'_t - \eta \frac{\hat{\mathbf{m}}_t}{\sqrt{\hat{\mathbf{v}}_t + \epsilon_{ADAM}}}, \quad (90)$$

em que ϵ_{ADAM} é um parâmetro de correção para impedir divisão por zero e $\hat{\mathbf{m}}_t$ e $\hat{\mathbf{v}}_t$ são o primeiro e o segundo momento de gradiente corrigidos, respectivamente. A correção se dá

para evitar que os momentos sejam inicializados como nulos (KINGMA; BA, 2014), e esta é dada por:

$$\hat{\mathbf{m}}_t = \frac{\mathbf{m}_t}{1 - \beta_t^{(1)}}, \quad (91)$$

$$\hat{\mathbf{v}}_t = \frac{\mathbf{v}_t}{1 - \beta_t^{(2)}}, \quad (92)$$

em que \mathbf{m}_t e \mathbf{v}_t são o primeiro e o segundo momento de gradiente, enquanto $\beta_t^{(1)}$ e $\beta_t^{(2)}$ são os coeficientes de primeiro e de segundo momento. Os termos \mathbf{m}_t e \mathbf{v}_t são também denominados como médias móveis, em que \mathbf{m}_t é a média do gradiente e \mathbf{v}_t é a variância descentralizada do mesmo. Os coeficientes $\beta_t^{(1)}$ e $\beta_t^{(2)}$ controlam a velocidade de decaimento das médias móveis. Os momentos de primeiro e de segundo grau do gradiente são atualizados, a cada época, da seguinte forma:

$$\mathbf{m}_t = \beta^{(1)} \cdot \mathbf{m}_{t-1} + (1 - \beta^{(1)}) \cdot \nabla J_t, \quad (93)$$

$$\mathbf{v}_t = \beta^{(2)} \cdot \mathbf{v}_{t-1} + (1 - \beta^{(2)}) \cdot \nabla J_t^2, \quad (94)$$

em que ∇J_t^2 é a multiplicação matricial elemento por elemento de ∇J_t com ela mesma, ($\nabla J_t \odot \nabla J_t$).

Baseado no algoritmo *Adam*, o algoritmo *AdamW* implementa uma funcionalidade adicional no mesmo (LOSHCHILOV; HUTTER, 2017). Esta funcionalidade está relacionada com o decaimento de pesos, uma técnica utilizada para regularização da ANN a fim de reduzir o *overfitting*. Para maior entendimento do *AdamW*, iremos discutir a implementação de uma técnica de decaimento de peso denominada regularização L2 (NG, 2004). Ao adicionarmos a somatória dos quadrados dos pesos das unidades neurais na função de custo da ANN, iremos forçar a otimização a reduzir a magnitude dos mesmos. Tomando a Equação 77, podemos atualizá-la como:

$$J_{\text{reg}} = J + L2 = \frac{1}{m} \sum_{i=1}^m \sum_{k=1}^K \left[\left(h_k(x^{(i)}) - y_k^{(i)} \right)^2 \right] + \frac{\lambda_{\text{reg}}}{m} \sum_{l=1}^{L_c-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_{l+1}} (w_{ji}^{(l)})^2, \quad (95)$$

em que $L2$ é o termo de custo de regularização e λ_{reg} é o parâmetro de regularização, tal qual controla a intensidade da regularização. Podemos perceber que o somatório envolve

todos os pesos da ANN. O gradiente de J_{reg} então irá apresentar o termo adicional, provindo da regularização, tal que:

$$\nabla J_{\text{reg}} = \nabla J + \nabla L2, \quad (96)$$

desta forma, a atualização de pesos, dada pela Equação 85, toma a seguinte forma:

$$\Delta \mathbf{w}'_{\text{reg}}{}^{(L_c)} = -\eta \nabla J_{\text{reg}} = -\eta \left(\nabla J + \lambda_{\text{reg}} \mathbf{w}' \right) \quad (97)$$

Ao aplicarmos o gradiente da função custo regularizada, dada pela equação 96, diretamente na otimização por *Adam*, podemos observar que as médias móveis dadas pelas Equações 93 e 94 irão levar em consideração também o termo de regularização. É possível provar que esta limitação reduz o desempenho da ANN (LOSHCHILOV; HUTTER, 2017). Uma forma equivalente de decaimento de pesos, porém de prática implementação para o otimizador *Adam*, consiste em desacoplar tal decaimento do cálculo do gradiente. Desta forma, de acordo com a Equação 97, podemos aplicar o decaimento de peso somente na atualização dos pesos realizada pelo *Adam* (Equação 90). Para o caso do *AdamW* então temos que a atualização de pesos é dada por:

$$\mathbf{w}'_{t+1} = \mathbf{w}'_t - \eta \cdot \frac{\beta^{(1)} \cdot \mathbf{m}_{t-1} + (1 - \beta^{(1)}) \cdot (\nabla J_t + \lambda_{\text{reg}} \mathbf{w}'_t)}{\sqrt{\hat{\mathbf{v}}_t + \epsilon_{\text{ADAM}}}}. \quad (98)$$

Batch Normalization

Uma dificuldade em treinar redes neurais profundas se dá no fato de que, como os pesos sinápticos estão variando a cada época, a distribuição de entradas para cada camada também é variado. Este impedimento é denominado deslocamento covariável interno (*internal covariate shift*) (SHIMODAIRA, 2000). Durante o treinamento, a ANN pode desperdiçar recursos tentando aprender as novas distribuições de entradas nas camadas. O método de BN foi desenvolvido para contornar este impedimento (IOFFE; SZEGEDY, 2015). Além disto, observou-se que o BN implementa uma certa regularização na ANN, e também este reduz a importância dos valores iniciais dos pesos sinápticos. Outra consequência do BN é que este suaviza a função custo na otimização, o que aumenta a velocidade de convergência (SANTURKAR *et al.*, 2018).

Assim como aplicamos a estandardização Z -score nos dados de entrada da ANN, o BN propõe uma normalização das entradas de cada camada escondida. Tal normalização é implementada considerando as propriedades estatísticas de um *mini-batch*. Para cada neurônio da l -ésima camada de uma ANN, cada entrada do *mini-batch* a ser processado tem seu valor confinado entre uma pequena faixa de valores. Após tal normalização, a distribuição das entradas deve possuir média nula e variância unitária. Considerando um *mini-batch* com m_b entradas, temos que a média para o j -ésimo neurônio da camada l pode ser calculada como:

$$\mu_{\text{BN}j}^{(l)} = \frac{1}{m_b} \sum_{i=1}^{m_b} u_{ji}^{(l)}, \quad (99)$$

em que u é o potencial de ativação (Equação 73) da camada $l - 1$. Da mesma forma, a variância do j -ésimo neurônio da l -ésima camada pode ser calculada como:

$$\left(\sigma_{\text{BN}j}^{(l)}\right)^2 = \frac{1}{m_b} \sum_{i=1}^{m_b} \left(u_{ji}^{(l)} - \mu_{\text{BN}j}^{(l)}\right)^2. \quad (100)$$

Após o cálculo das propriedades estatísticas, temos que as amostras providas do *mini-batch* em análise tomará o valor de:

$$\hat{u}_{ji}^{(l)} = \frac{u_{ji}^{(l)} - \mu_{\text{BN}j}^{(l)}}{\sqrt{\sigma_{\text{BN}j}^{(l)} + \epsilon_{\text{BN}}}}, \quad (101)$$

em que ϵ_{BN} é um escalar de baixo valor, utilizado para evitar divisão por zero. Após a normalização das entradas para a distribuição obter média zero e variância unidade, o BN ainda apresenta a capacidade de moldar a distribuição da forma que a ANN melhor apresentar desempenho de treinamento. Desta forma, para cada neurônio da l -ésima camada, introduzimos dois parâmetros treináveis. A ANN então irá aprender a distribuição que minimiza o erro de treinamento, e esta é dada por:

$$\hat{u}_{ji}^{(l)} = \gamma_{\text{BN}j}^{(l)} \cdot \hat{u}_{ji}^{(l)} + \beta_{\text{BN}j}^{(l)}, \quad (102)$$

em que $\gamma_{\text{BN}j}^{(l)}$ e $\beta_{\text{BN}j}^{(l)}$ são os parâmetros treináveis que são responsáveis por escalonar e deslocar, respectivamente, a distribuição das entradas. Os parâmetros treináveis representam grafos adicionais na otimização por GD.

Dropout

O DO é uma técnica de regularização amplamente utilizada em ANNs profundas (SRIVASTAVA *et al.*, 2014). A cada época do treinamento da ANN, algumas interconexões entre unidades neurais são desconsideradas de forma aleatória. Esta técnica impede os neurônios de se adaptarem excessivamente ao conjunto de treinamento. Definimos uma probabilidade ρ_{DO} em que os neurônios serão cortados da ANN. Considerando a ativação da i -ésima unidade neural da l -ésima camada, $a_i^{(l)}$, a multiplicamos pelo termo de probabilidade de DO $r_i^{(l)} \sim \mathcal{B}(\rho_{\text{DO}})$, em que \mathcal{B} é uma distribuição de Bernoulli. Considerando o vetor de saídas da camada $l - 1$, $\mathbf{a}^{(l-1)}$, aplicando DO temos que:

$$\tilde{\mathbf{a}}^{(l-1)} = \mathbf{r}^{(l-1)} \cdot \mathbf{a}^{(l-1)}, \quad (103)$$

em que $\mathbf{r}^{(l-1)}$ é o vetor de termos de DO da camada anterior, e cada elemento indica uma amostragem da distribuição de Bernoulli para uma unidade neural. Na camada sob análise l , temos que o potencial de ativação será dado então por:

$$\tilde{u}^{(l)} = \mathbf{w}^{(l)T} \cdot \tilde{\mathbf{a}}^{(l-1)} + b^{(l-1)}, \quad (104)$$

de forma que a ativação da j -ésima unidade neural da camada l então será dada por:

$$\tilde{\mathbf{a}}^{(l)} = g(\tilde{u}^{(l)}). \quad (105)$$

Cabe mencionar que as equações descritas até aqui descrevem o comportamento da ANN com DO na fase de treinamento. Na fase de inferência, o DO deve ser desativado e os pesos aprendidos durante o treinamento mantidos.

Conexões Residuais

A ferramenta RES é de simples implementação, porém apresenta uma grande melhoria no treinamento de ANNs profundas. Introduzido inicialmente na ANN denominada ResNet, RES pode melhorar o aprendizado de ANNs com até mais de 1000 camadas (HE *et al.*, 2016). A convergência no treinamento de ANNs profundas tende a ser de difícil obtenção devido a problemas como dissipação de gradiente, entretanto este pode ser parcialmente resolvido a partir da normalização das entradas das camadas. Outro problema que foi

observado é o de degradação, em que o erro da ANN se satura, tanto o de treinamento quanto o de validação. Esta degradação de ambos os erros não indica um *overfitting*, mas sim uma dificuldade enigmática no aprendizado em ANNs profundas.

A proposta de caminhos residuais se dá na conexão de uma camada com a outra a partir de um atalho paralelo. Estas conexões pulam uma ou mais camadas, e podem realizar um mapeamento identidade, ou seja, transportam seu valor para uma camada mais adiante. No caso de mapeamento identidade, o RES não adiciona parâmetros adicionais na ANN, e o GD pode ser aplicado normalmente. Dado o vetor de mapeamento de atributos realizado pela l -ésima camada, $\mathbf{a}^{(l)}$, podemos somar, elemento por elemento, uma entrada pertencente à alguma camada anterior $\mathbf{a}^{(k)}$, em que $k < l$. Desta forma, geramos um novo mapeamento, dado por:

$$\mathcal{H}^{(l)} = \mathbf{a}^{(l)} + \mathbf{a}^{(k)}, \quad \text{para } [1 \leq k < l]. \quad (106)$$

A soma da Equação 106 é realizada elemento por elemento nos vetores. Caso algum dos vetores apresente um tamanho diferente, podemos simplesmente preencher este com elementos nulos de forma a obter um vetor com a mesma dimensão.

Como a ANN profunda é um mapeador que combina diversas camadas não lineares, a função identidade pode ser de difícil solução. Gerando uma conexão residual contendo em si esta função identidade, uma ANN pode facilmente encontrá-la quando for necessário. Outra vantagem desta técnica é que a propagação dos gradientes podem fluir diretamente através das camadas, o que diminui problemas como dissipação de gradiente.

3.1.4 Métricas Do Mapeamento da ANN

Na presente subseção iremos discutir algumas métricas de desempenho em relação ao modelo de ANN. Além da métrica de perda MSE, como dada pela Equação 77, também utilizamos o valor absoluto médio (*Mean Absolute Error*, MAE) para visualização da diferença entre o valor simulado e a saída da ANN. O MAE é dado por:

$$MAE = \frac{1}{m} \sum_{i=1}^m |h(x^{(i)}) - y^{(i)}|, \quad (107)$$

em que $y^{(i)}$ é o valor resultante do modelo de simulação e $h(x^{(i)})$ é o valor previsto pela rede, para o i -ésimo exemplo de treinamento.

Calculamos também a correlação de Pearson (FREEDMAN; PISANI; PURVES, 2007) entre os dados do conjunto de testes simulados e os previstos pela ANN. Tal métrica quantifica o grau de correlação entre duas variáveis, e apresenta valores entre -100% , correlação negativa perfeita, e 100% , correlação positiva perfeita. A correlação de Pearson é dada pela seguinte equação:

$$r_p = \frac{\sum_{i=1}^m \left(h(x^{(i)}) - \overline{\mathbf{h}(\mathbf{x})} \right) \left(y^{(i)} - \bar{y} \right)}{\sqrt{\sum_{i=1}^m \left(h(x^{(i)}) - \overline{\mathbf{h}(\mathbf{x})} \right)^2 \sum_{i=1}^m \left(y^{(i)} - \bar{y} \right)^2}}, \quad (108)$$

em que $\overline{\mathbf{h}(\mathbf{x})}$ é a média das saídas da ANN e \bar{y} é a média dos valores simulados, em relação ao conjunto considerado.

Analizamos também a curva de aprendizado da ANN. A curva de aprendizado é simplesmente a representação gráfica do valor de erro, tanto do conjunto de dados de treinamento como de validação, obtido em termos de época. É importante analisarmos as curvas de treinamento para identificarmos se o algoritmo está aprendendo, a partir dos erros de treinamento. Além do mais, estas curvas servem para nos conceder uma noção da capacidade de generalização da rede. A partir da evolução dos erros de validação em relação aos de treinamento, podemos prever se *overfitting* está ocorrendo. Como os valores de erros começam relativamente altos nas primeiras épocas e convergem para valores de algumas ordens inferiores, as figuras obtidas apresentarão uma escala logarítmica em relação ao erro.

3.2 Evolução Diferencial

O algoritmo de DE é um algoritmo heurístico iterativo baseado no processo evolutivo das espécies biológicas. O DE, e outros algoritmos evolutivos, apresentam algumas vantagens. Estas incluem o fato de que, como a otimização não depende do cálculo do GD, a função a ser otimizada não tem o requerimento de ser diferenciável. Outras vantagens incluem a possibilidade de solucionar problemas de alta dimensionalidade, com valores discretos ou contínuos, e ainda o DE pode ser aplicado em problemas de múltiplos objetivos.

Originalmente apresentado em (STORN; PRICE, 1995), o DE opera seguindo as etapas mostradas na Figura 21. O DE tem seu funcionamento semelhante ao do GA, este discutido na Subseção 1.1 do Apêndice B, entretanto, o DE apresenta alguns aprimora-

mentos nas etapas de cruzamento e de mutação. Em um trabalho comparativo, o DE se mostrou superior ao GA em 16 testes de otimizações multi-objetivas (GEORGIOUDAKIS; PLEVRIS, 2020), apresentando uma maior homogeneidade ao redor da solução ótima em relação ao GA e à otimização por enxame de partículas (PSO, *Particle Swarm Optimization*) (ISLAM *et al.*, 2019). É válido mencionar que o DE já foi utilizado para otimização de design de dispositivos fotônicos, como divisores de potência (XIE *et al.*, 2020), pentes de frequências ópticas baseada em MZMs (PENDIUK; NEVES; POHL, 2020) e focalização da luz em escala menores ao comprimento de onda (BOR; TURDUEV; KURT, 2016). Um vasto estudo sobre implementação de DE para problemas de eletromagnetismo é realizado em (QING, 2010).

Discutiremos agora a operação do DE, de acordo com a Figura 21. No primeiro bloco, uma população de indivíduos é gerada, estes apresentando parâmetros de valores aleatórios seguindo uma distribuição normal entre limites pré-definidos. Para quantificar o desempenho de cada indivíduo, definimos uma função *fitness*. Tal função pode ser uma simples operação entre os parâmetros do indivíduo ou o resultado de um mapeamento mais complexo como, por exemplo, um modelo de simulação eletromagnética aplicado nos parâmetros. No bloco de avaliação, calcula-se a função *fitness* para cada indivíduo da população. O bloco seguinte ao de avaliação consiste de três fases, estas sendo mutação, cruzamento e seleção. Na mutação, para cada i -ésimo candidato $\mathbf{S}_t^{(i)}$ da população na iteração t , geramos um vetor doador $\mathbf{v}_t^{(i)}$, sendo este a combinação de três candidatos aleatórios amostrados da população. Considerando os três candidatos aleatórios $\mathbf{S}_t^{(a)}$, $\mathbf{S}_t^{(b)}$ e $\mathbf{S}_t^{(c)}$, dado que $a \neq b \neq c \neq i$, temos que $\mathbf{v}_t^{(i)}$ é dado por:

$$\mathbf{v}_t^{(i)} = \mathbf{S}_t^{(a)} + M \cdot (\mathbf{S}_t^{(b)} - \mathbf{S}_t^{(c)}), \quad (109)$$

em que M é a taxa de mutação. A taxa M apresenta valores de até 1 e serve para escalar a diferença de vetores para a geração do vetor doador.

Na recombinação, temos a geração do vetor tentativa $\mathbf{u}_t^{(i)}$ ao mesclar parâmetros entre o vetor doador e o vetor em observação. Já na fase de seleção, o algoritmo dará preferência em selecionar o melhor vetor entre o vetor em observação e o vetor tentativa. Há vários métodos de recombinação e seleção (DAS; SUGANTHAN, 2011), entretanto aqui iremos citar apenas os utilizados no presente trabalho, que é a recombinação binomial e a seleção egoísta (*greedy selection*). No caso da recombinação binomial, a recombinação

é realizada parâmetro a parâmetro, com a probabilidade de haver troca de parâmetro definida pela taxa de recombinação. Na seleção egoísta, o vetor em observação é substituído pelo vetor tentativa caso o último apresente melhores resultados. Devemos citar também que há alguns métodos de penalização geralmente utilizados para se manter os valores dos parâmetros dentro de certos limites (PRICE; STORN; LAMPINEN, 2006). Um dos mais simples de se implementar é o método denominado *bounce-back*, este método redefine o valor do parâmetro que excede o limite inferior ou superior previamente definido. Na Subseção 4.2.2 daremos maiores detalhes sobre a implementação do DE no presente trabalho.

3.3 Otimização Integrando Algoritmos Heurísticos e Redes Neurais Artificiais

Métodos simples de otimização, como varredura de parâmetros e o método Simplex (HUSSAIN; QAYYUM; HUSSAIN, 2019), muitas vezes não são capazes de resolver problemas complexos. Algoritmos heurísticos apresentam uma abordagem de desempenho satisfatório ao resolver problemas não lineares de múltiplas variáveis (PEARL, 1984). A essência destes algoritmos se dá no descobrimento de soluções sub-ótimas enquanto complexidade computacional do processo é reduzida ao mínimo possível. Apesar destes mostrarem um tempo de execução relativamente baixo, a avaliação da FOM ainda pode ser

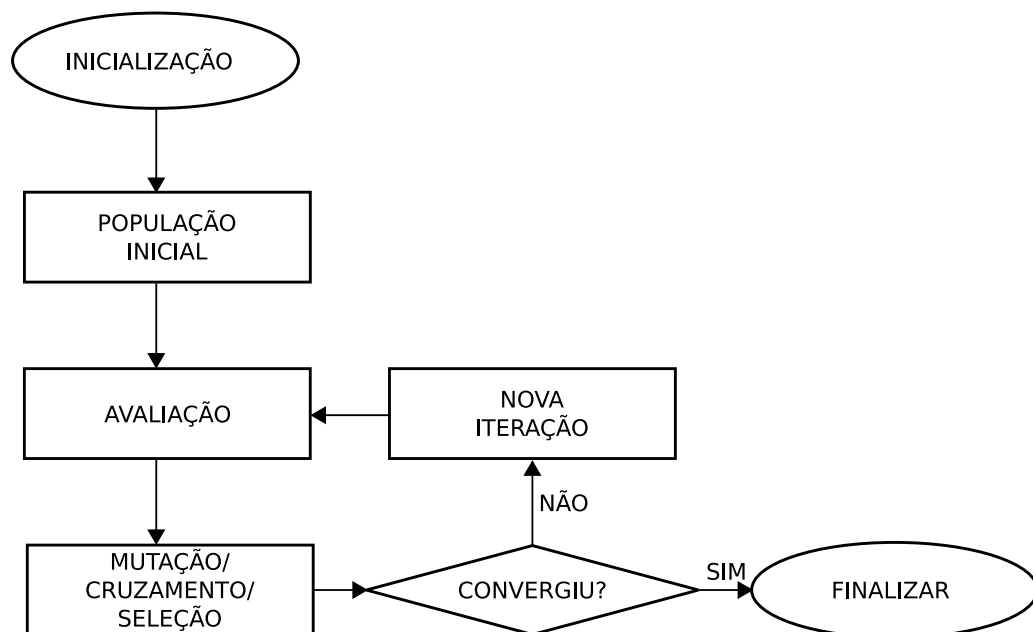


Figura 21 – Diagrama de blocos do algoritmo de DE.

um processo que exibe um alto custo computacional. Em especial, no caso de otimização de *design* de dispositivos que requerem modelos computacionais complexos, como simulações eletromagnéticas, a avaliação das amostras pode ocasionar uma demora significativa no processo. No presente trabalho propomos um modelo conjunto de algoritmos de otimização com uma ANN, esta utilizada para calcular a FOM (VILLARRUBIA *et al.*, 2018).

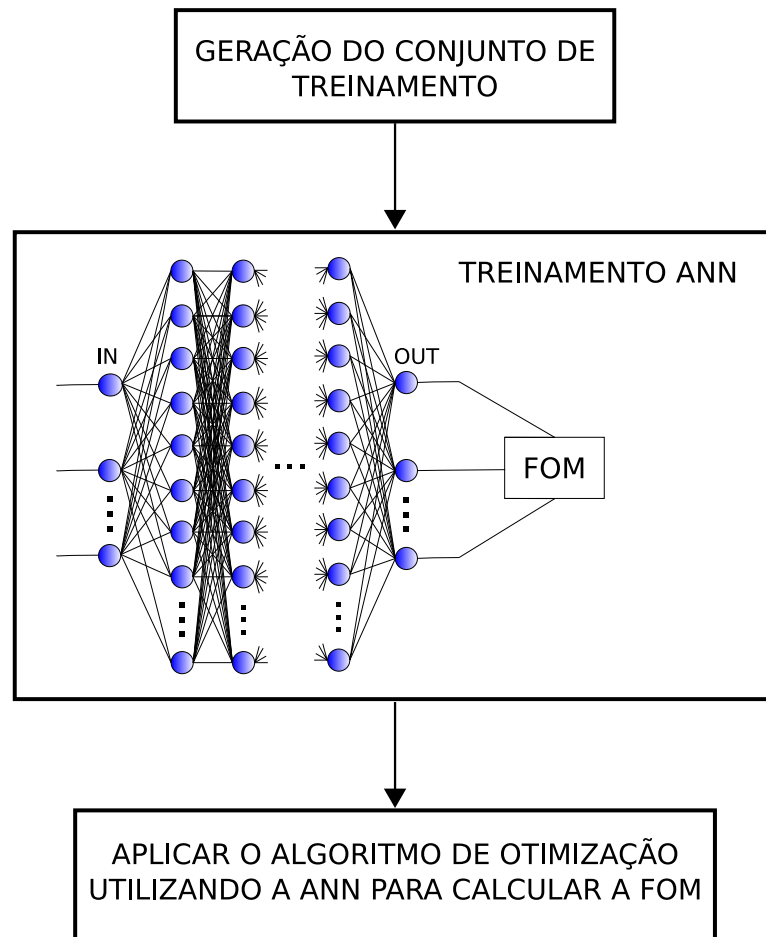


Figura 22 – Diagrama de blocos do nosso esquema de otimização conjunta de ANNs com algoritmos de otimização.

A Figura 22 ilustra o diagrama de blocos da otimização de um dispositivo utilizando modelo baseado em ANN. O primeiro bloco se refere à geração do conjunto de dados que será utilizado para o treinamento da ANN. O treinamento deve mapear as variáveis de otimização para um escalar, este representando a FOM. Após o treinamento da ANN, o algoritmo de otimização pode ser aplicado de forma a maximizar uma determinada FOM, com seu valor estimado pela ANN.

Este método de otimização inclui as fases adicionais de obtenção do conjunto de dados e do treinamento da ANN. Entretanto, este representa um custo computacional

de uma única vez. Após a ANN estar funcional, diversas configurações de FOMs e de algoritmos de otimização podem ser utilizados com a vantagem de apresentarem um custo computacional significativamente reduzido. É válido ressaltar também que, neste esquema, adiciona-se a etapa da configuração dos hiper-parâmetros da ANN, como o tipo, quantidade e tamanho de camadas neurais, funções de ativação e outros recursos de DL.

4 Método de Otimização

No presente capítulo, iremos discutir o método de otimização utilizado no presente trabalho. Primeiramente, a partir do modelo de simulação, iremos gerar um conjunto de dados de MZMs com configurações aleatórias. Utilizando estes dados, iremos treinar uma ANN de forma a emular o modelo de simulação. Por fim, utilizando o algoritmo heurístico DE em conjunto com a ANN, iremos otimizar os parâmetros de *design* do MZM. A Seção 4.1 descreve o método para a simulação do MZM e para a geração dos dados. Para o cálculo da BW_{EO} , da V_{π} e da IL do MZM utilizamos o modelo proposto em (MOTTA *et al.*, 2017) e descrito no Capítulo 2. A partir deste modelo, geramos os dados de treinamento para a ANN utilizando a linguagem de programação Python. Na Seção 4.2 descrevemos como a otimização foi realizada empregando ANN e o algoritmo DE.

4.1 Geração de Dados

Para a geração do conjunto de dados de MZMs que servirão tanto para o treinamento, quanto para a determinação da acurácia de predição da ANN, aplicamos os modelos semi-empíricos descritos no Capítulo 2 e desenvolvidos em (MOTTA *et al.*, 2017). A Tabela 3 indica as variáveis de otimização consideradas e seus respectivos limites para geração do conjunto de dados. Geramos 10.000 dados, com os parâmetros dados por uma distribuição uniforme entre os limites definidos pela tabela. Já a Figura 23 ilustra a arquitetura completa do MZM, assim como os parâmetros a serem variados. A parte superior ilustra os TWEs com extensões em T , assim como o parâmetro L_{ps} , indicando o comprimento do deslocador de fase. Abaixo desta temos a seção transversal do dispositivo, indicando ambos os braços e onde a V_{bias} é aplicado. Já na parte inferior temos a seção transversal de um braço, com a largura das quatro regiões dopadas, do guia de ondas e o deslocamento da junção PN em relação ao centro do guia. É válido ressaltar que ambos os braços são idênticos, então ambos apresentam os mesmos parâmetros variáveis.

A Figura 24 ilustra o diagrama de blocos para geração dos dados a partir do modelo de simulação proposto em (MOTTA *et al.*, 2017). Utilizando o *software* de simulação eletromagnética CST *Microwave Studio*, obteve-se a matriz de parâmetros de transmissão para o CPS sem carga. Para contabilizar a carga na linha, foi implementado o modelo

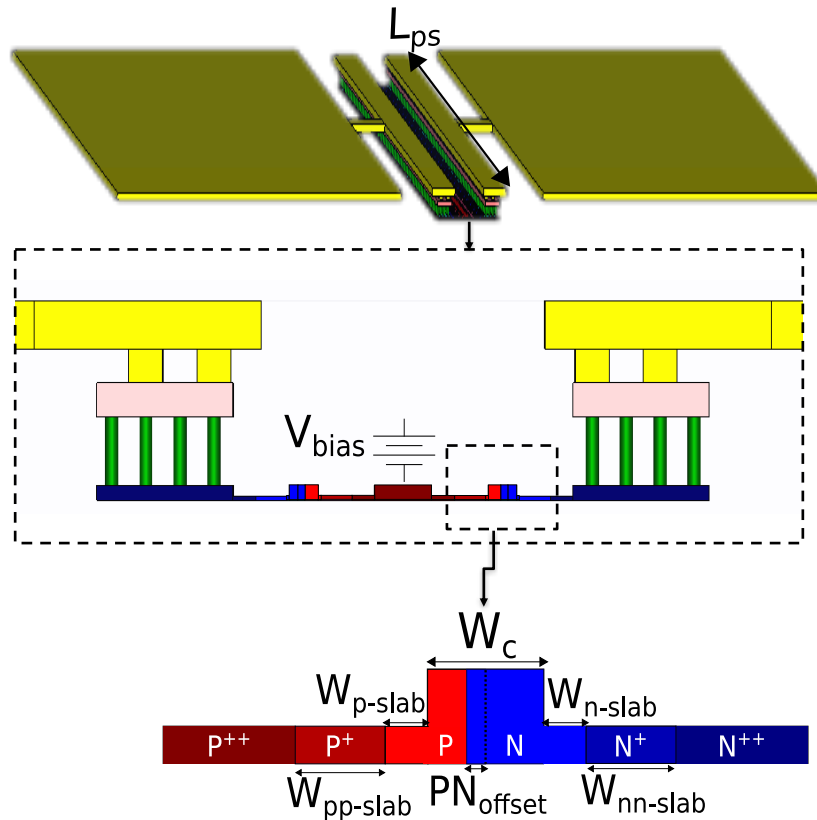


Figura 23 – Ilustração do MZM a ser implementado. A parte superior representa uma seção do modulador e dos guias de transmissão, enquanto a parte do meio representa a seção transversal do modulador completo e abaixo temos a seção transversal de um dos braços do modulador. Figura adaptada de (MOTTA *et al.*, 2017).

Tabela 3 – Parâmetros variáveis do MZM no modelo de simulação, com seus respectivos alcances de valores.

Parâmetro	Mínimo	Máximo
PN_{offset}	-225 nm	225 nm
W_c	450 nm	500 nm
W_{p-slab}	50 nm	500 nm
W_{n-slab}	50 nm	500 nm
$W_{pp-slab}$	0,6 μm	1 μm
$W_{nn-slab}$	0,6 μm	1 μm
V_{bias}	-10 V	-2,5 V
L_{ps}	0,5 mm	4 mm

descrito na Subseção 2.4.2, utilizando a linguagem de programação Python. O índice de grupo óptico do guia de ondas foi calculado utilizando o *software* Lumerical *Mode* e teve seu valor incluído no cálculo do modelo de banda, também via Python. Para a estimativa da V_π e da IL, é necessário analisar a interação da luz com o material e sua variação de densidade de cargas. Para isto, em (MOTTA *et al.*, 2017) implementou-se o modelo de um dos braços

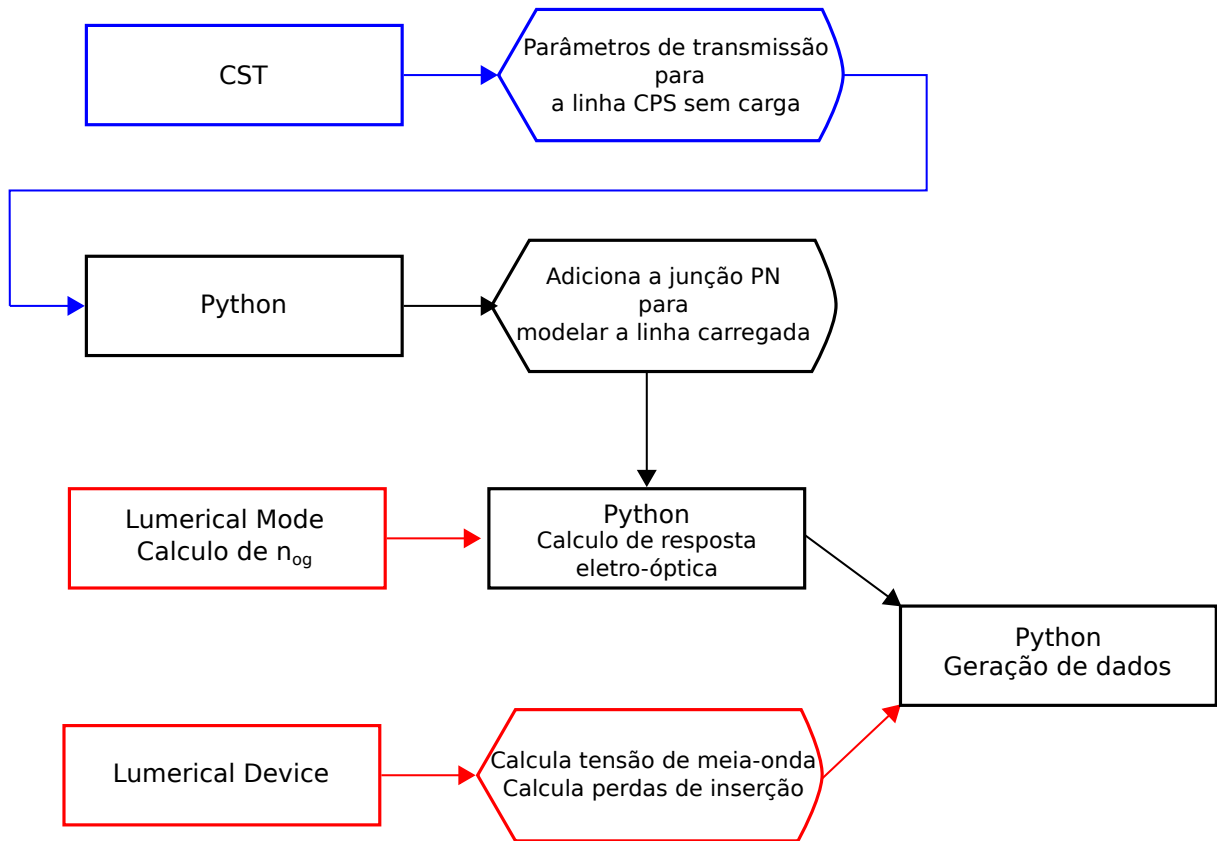


Figura 24 – Diagrama de blocos da geração de dados a partir do modelo de simulação. Os blocos com contornos em preto indicam a utilização da linguagem de programação de alto nível, enquanto que os blocos com contornos vermelhos e azuis indicam, respectivamente, simulações ópticas e elétricas.

do MZM, de acordo com o desenvolvimento descrito na Subseção 2.3.4, no *software* de simulação Lumerical *Device*. Como ambos os braços são iguais, a simulação de apenas um dos deslocadores de fase é o suficiente para extração dos parâmetros do modulador. É válido ressaltar que, apesar da escolha dos *softwares* de simulação electromagnética para a seção elétrica, CST, e para seção óptica, Lumerical, outros *softwares* parecidos poderiam ter sido utilizados para obtenção de resultados equivalentes.

4.1.1 Simulação da Seção Elétrica

Como discutido na Subseção 2.4.1, utilizamos uma estrutura de guias de onda CPS *Slow*. Utilizando CST, em (MOTTA *et al.*, 2017), a estrutura ilustrada na Figura 14 foi implementada, e 62 pontos no intervalo de frequência entre 1 GHz e 50 GHz foram considerados. A espessura do guia elétrico foi escolhida de forma a este se restringir nos

padrões definidos pela *foundry* IMEC, sendo esta de $0,61 \mu\text{m}$. Como moduladores de MZM em silício apresentam um comprimento na ordem de milímetros, geralmente entre 1 e 4 mm, a modelagem completa para uma análise eletromagnética se torna extremamente custosa computacionalmente. Portanto, segmentou-se a estrutura total em blocos de poucos micrômetros. A estrutura completa foi avaliada concatenando os parâmetros de transmissão dos blocos (Seção 1 do Apêndice A).

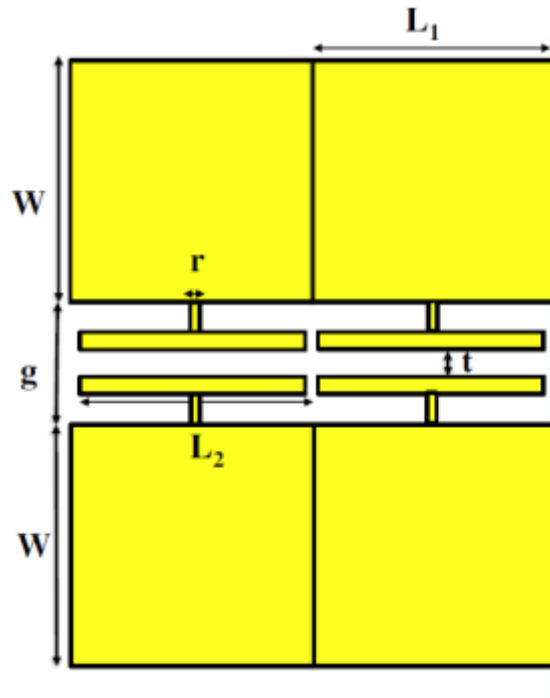


Figura 25 – *Design* de um CPS com extensões em formato de T (MOTTA *et al.*, 2017).

A Figura 25 mostra a visão superior de dois blocos de simulação do TWE-*slow*. Os parâmetros de comprimento escolhidos para cada seção foram $L_1 = 50 \mu\text{m}$ e $L_2 = 47 \mu\text{m}$. A espessura do eletrodo é de $0,7 \mu\text{m}$, e o mesmo é composto de alumínio. De forma a minimizar as perdas, maximizar o índice efetivo elétrico e obter um valor de impedância característica de aproximadamente 50Ω , os valores dos parâmetros do CPS foram otimizados. Os valores ótimos encontrados foram de $W = 50 \mu\text{m}$, $g = 25 \mu\text{m}$, $r = 2 \mu\text{m}$ e $t = 5 \mu\text{m}$ (MOTTA *et al.*, 2017).

Até então, considerou-se uma linha de transmissão sem carga. A resistência e a capacitância relacionadas à junção PN foram incluídas utilizando a linguagem de programação de alto nível Python. Considerando os valores das dopagens, de acordo com as restrições impostas pela *foundry* IMEC, as concentrações de elétrons e de lacunas dentro do guia de onda foram escolhidas como sendo, respectivamente, $\Delta N = 3 \cdot 10^{17} \text{ cm}^{-3}$ e

$\Delta P = 5 \cdot 10^{17} \text{ cm}^{-3}$. No caso específico em que analisamos aqui, as resistências de folha foram disponibilizadas pela *founndry*, sendo estas iguais a 4930 Ω/m , 2108 Ω/m , 1486 Ω/m , 942 Ω/m , 50 Ω/m e 160 Ω/m , para as regiões P, N, P+, N+, P++ e N++ respectivamente. Dados os valores de largura de cada seção, podemos calcular a resistência da junção PN utilizando a Equação 33. Para uma tensão aplicada na junção, podemos estimar a largura da região de depleção pela Equação 26 e então a capacitância da junção utilizando a Equação 34.

Considerando os parâmetros de transmissão da linha de transmissão CPS sem carga, podemos calcular a constante de propagação complexa, a impedância característica, o índice efetivo e as perdas da linha pelas Equações 135-138. A partir destes, calculamos os parâmetros distribuídos inerentes à linha de transmissão, dados pelas Equações 139-142. Como estamos interessados no comportamento da linha carregada, devemos implementar o modelo do circuito equivalente RC paralelo, em que os elementos são atualizados conforme as Equações 46 e 47. Finalmente, podemos obter a impedância característica, a constante de propagação, as perdas e o índice efetivo para a linha carregada pelas Equações 48, 49, 50 e 53.

Para maximizar a BW_{EO} , assim como discutido na Subsubseção 2.4.3, consideramos uma impedância de terminação como sendo de $Z_T = 35 \Omega$, a da fonte como $Z_s = 50 \Omega$ e, como mencionado, a impedância da é de aproximadamente $Z_{0l} = 50 \Omega$ até uma frequência elétrica de 50 GHz.

4.1.2 Simulação da Seção Óptica

Simultaneamente às simulações elétricas descritas na Subseção 4.1.1, o índice de grupo óptico para o guia de onda *rib* utilizado no MZM foi calculado utilizando o *software* Lumerical *Mode*, como ilustrado na Figura 24. Obtidos tais resultados, podemos estimar a banda do modulador pela Equação 60. Por outro lado, para obtenção da V_π e da IL, utilizou-se um modelo implementado no *software* de simulação eletromagnética Lumerical *Device* (MOTTA *et al.*, 2017). O modelo 1D apresentado na Subseção 2.3.4 nos dá uma noção de como é a resposta óptica do MZM. Entretanto, o mesmo apresenta um alto grau de imprecisão, pois este realiza diversas aproximações e negligencia fenômenos como a distribuição de portadores na vertical e a sobreposição do modo propagante com as regiões

de dopagem intermediária (CHROSTOWSKI; HOCHBERG, 2015). O *software* Lumerical *Device* nos permite realizar as simulações de elementos finitos de modos (*Finite Element Eigenmode*, FEEM) (JIN, 2002) enquanto contabiliza as variações de densidades de cargas no material dopado, levando em consideração a diferença de potencial elétrico aplicado.

A parte inferior da Figura 23 representa o modelo implementado no *Device*, com seus parâmetros variáveis indicados na mesma. As regiões com dopagens intermediárias, $P+$ e $N+$, são utilizadas como interface entre as regiões com dopagens fracas e fortes. Estas são responsáveis por reduzir as atenuações ópticas causadas pelas regiões de maiores dopagens, tais quais servem como interface para o contato elétrico (ZHOU *et al.*, 2020). Primeiramente, devemos definir os valores de concentração dos portadores livres para cada região do modulador. Como mencionado na Subseção 4.1.1, os valores de ΔN e de ΔP são de, respectivamente, $3 \cdot 10^{17} \text{ cm}^{-3}$ e $5 \cdot 10^{17} \text{ cm}^{-3}$. Já para as outras regiões, definimos como $\Delta N^+ = \Delta P^+ = 4 \cdot 10^{18} \text{ cm}^{-3}$ e de $\Delta N^{++} = \Delta P^{++} = 1 \cdot 10^{20} \text{ cm}^{-3}$ para as regiões de nível de dopagem intermediária e forte, respectivamente.

Inicialmente, com o simulador de carga do *Device*, obtivemos a distribuição de elétrons e de lacunas para cada ponto de malha da geometria representada na Figura 23 (MOTTA *et al.*, 2017). As equações de Soref (Equações 24 e 25) foram aplicadas, para $\lambda_0 = 1550 \text{ nm}$, de forma a estimar a mudança de índice de refração e de atenuação para as concentrações dadas. Podemos obter a variação de absorção em relação à atenuação utilizando a Equação 2. Com os valores de variação de índice e de absorção, podemos estimar a matriz de índice de refração complexo, dada pela Equação 1, para cada célula do espaço de simulação. O simulador considera a tensão aplicada, realizando o cálculo de *drift* dos portadores e da região de depleção formada. O FEEM, por sua vez, realiza a sobreposição do modo de propagação com as regiões dopadas e de depleção, nos permitindo obter o índice efetivo complexo do modo em função da tensão aplicada.

Após a obtenção do índice de refração efetivo complexo, nos foi permitido estimar a V_π e a IL. O valor da V_π é estimado a partir da Equação 32. A derivada é estimada ao calcular a variação obtida em n_{eff} a partir de duas simulações aplicando uma pequena perturbação em V_{bias} . Já em relação às perdas, tomamos a componente imaginária, de forma que $\kappa = \Im(\overline{n_{eff}})$ e $\alpha = 4\pi\kappa/\lambda$, e utilizamos a Equação 50 para obtermos a atenuação por unidade de comprimento. Como o comprimento do deslocador de fase é conhecido, é possível a obtenção do valor da IL.

4.2 Otimização do Modulador

Na presente seção iremos discutir o método adotado para a otimização do MZM. Primeiro discutiremos sobre como a ANN será implementada, e então discutiremos sobre o algoritmo de otimização .

4.2.1 Mapeamento por Rede Neural

Após a geração de dados de 10.000 moduladores com valores de parâmetros aleatórios, podemos treinar o modelo de ANN. Esta deve considerar como entrada as dimensões das diferentes regiões, assim como a tensão de polarização, e na saída esta apresenta a V_{π} , a BW_{EO} e a IL. Das 10.000 configurações, utilizamos 90% como dados de treinamento e os outros 10% foram utilizados tanto como validação do modelo quanto como teste. Aplicamos a estandardização, como indicada pela Equação 72, para cada atributo de entrada e de saída individualmente. Como é comum na otimização de ANNs, dividimos o conjunto de treinamento em diversos blocos denominados de *mini-batches* (KESKAR *et al.*, 2017). Cada *mini-batch* possui 500 exemplos de treinamento. Este método, como discutido na Subseção 3.1.3, além de economizar memória de acesso aleatório (*Random Access Memory*, RAM), ainda é capaz de encontrar a solução em um número menor de iterações. A Figura 26(a) mostra uma ANN de quantidade arbitrária de camadas e neurônios escondidos aplicada para modelagem do MZM. Tal modelo, como mencionado, recebe 8 parâmetros de *design* e nos retorna os 3 parâmetros de métrica do MZM. A Figura 26(b) indica o diagrama de blocos do modelo utilizado na otimização adotada no presente trabalho. A escolha desta configuração é discutida na Subseção 5.1.1. É válido ressaltar que os modelos e o treinamento e os aprimoramentos dos mesmos foram implementados a partir da biblioteca *PyTorch* (PASZKE *et al.*, 2019).

Rede Neural Artificial Base

No presente trabalho, definimos uma ANN base que consiste de uma ANN que não faz utilização de métodos de aprimoramento de DL, mais especificamente o BN, o DO e o RES. A ANN base servirá como referência para a otimização das configurações

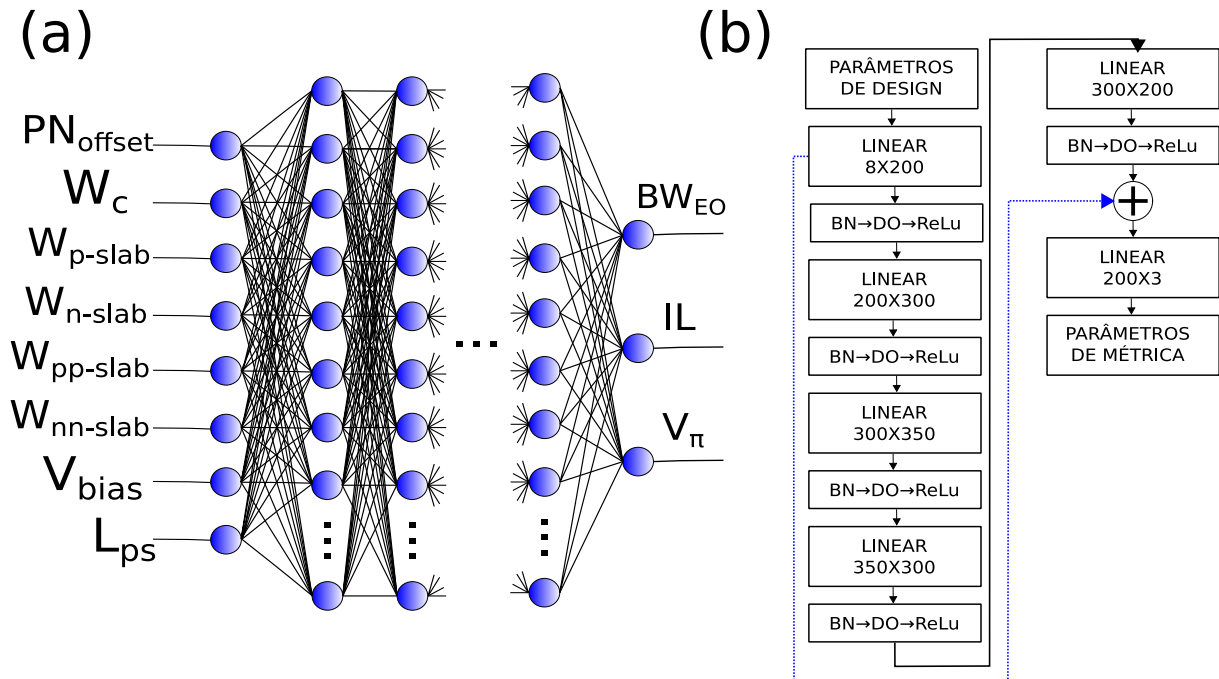


Figura 26 – Ilustração da ANN utilizada no trabalho. As entradas são os parâmetros de *design* e a saída são os parâmetros de métricas referentes ao MZM. (a) ANN arbitrária aplicada ao nosso problema e (b) ANN utilizada na otimização aqui presente.

de outros modelos. O modelo proposto para referência possui 4 camadas escondidas totalmente conectadas (*Fully Connected*, FC). No final de cada camada da rede, com exceção da camada de saída, aplicamos a ativação não linear ReLu. Os pesos sinápticos são inicializados via método de inicialização uniforme de Kaiming He. Os números de neurônio por camada escondida são de 200, 300, 300 e 200.

Para a avaliação da ANN e otimização dos pesos sinápticos, a função de custo utilizada é a MSE. O método de otimização selecionado para minimizar a função de custo durante o treinamento é a estimativa de momento adaptativo com decaimento de peso, *AdamW*. O decaimento de peso é de 0,05 para todas as ANNs implementadas. Devemos nos atentar que uma taxa de aprendizado muito reduzida pode causar uma demora para a convergência do modelo e também uma maior sensibilidade a mínimos locais, enquanto uma taxa muito alta pode causar dificuldades na convergência (GOODFELLOW; BENGIO; COURVILLE, 2016). Considerando isto, implementamos um decaimento de taxa de aprendizado. Este decaimento irá determinar a taxa de aprendizado máxima que um parâmetro pode tomar durante a otimização via *AdamW*. Um decaimento de taxa estabelece uma maior taxa de aprendizado durante as primeiras épocas e o mesmo tem

seu valor reduzido conforme a evolução das iterações. O decaimento de taxa ajuda na velocidade de otimização e também na capacidade de generalização da rede (YOU *et al.*, 2019).

Em relação ao otimizador *AdamW*, escolhemos os parâmetros de acordo com os padrões utilizados na literatura, assim como em (KINGMA; BA, 2014), em que os coeficientes de primeiro e de segundo momento são dados respectivamente por $\beta^{(1)} = 0,9$ e $\beta^{(2)} = 0,99$. O fator de correção, tal qual serve para evitar divisões por zero, é dado por $\epsilon_{ADAM} = 1 \cdot 10^{-8}$. A taxa de aprendizado inicial é fixada a 0,5, e o decaimento é dado de maneira que há uma redução de 80% no valor da taxa de aprendizado a cada 50 iterações. Ao total, o modelo foi treinado por 400 iterações.

Redes Neurais Aprimoradas

Além da ANN base, iremos implementar ANNs com mecanismos de DL que servem para melhorar a convergência e a generalização das mesmas. O DO implementado nos modelos apresenta uma probabilidade de *drop out* de $\rho_{DO} = 0,5$. A respeito do BN, quando utilizado, é aplicado entre todas as camadas escondidas da ANN. Para o caso do RES, implementamos uma conexão residual entre a camada de entrada e a penúltima camada escondida da ANN. O BN é aplicado antes da ativação ReLu, como recomendado em (IOFFE; SZEGEDY, 2015) e também aplicado antes do DO, como recomendado em (LI *et al.*, 2019).

A Tabela 4 mostra as diferentes ANNs utilizadas no nosso trabalho. A quantidade de camadas escondidas utilizadas varia de 3 a 6, e cada camada possui entre 200, 300 ou 350 unidades neurais. Além de comparar as ANNs bases, comparamos também as ANNs utilizando diferentes ferramentas de aprimoramento. No caso da ANN com 4 camadas escondidas, aplicamos as ferramentas BN, DO e RES isoladamente, todas as combinações

Tabela 4 – Número de camadas escondidas por configuração de ANN, número de neurônios pertencentes a cada camada e configurações utilizadas.

Número de camadas escondidas	Número de neurônios por camada escondida	Configurações utilizadas
3	200-300-200	ANN base e completa
4	200-300-300-200	ANN base, DO, BN, RES, DO+BN, DO+RES, BN+RES e completa
5	200-300-350-300-200	ANN base e completa
6	200-300-350-350-300-200	ANN base e completa

possíveis entre duas destas e as três juntas, que denominamos ANN completa. No caso das configurações de 3, 5 e 6 camadas escondidas, aplicamos apenas as configurações de ANN base e de ANN completa. A ANN completa de 5 camadas é mostrada na Figura 26.

4.2.2 Otimização por Evolução Diferencial

Com o modelo de ANN treinado, podemos estimar as métricas de um MZM de maneira precisa e mais rápida que o simulador eletromagnético. Esta redução na complexidade computacional nos permite a aplicação de algoritmos heurísticos que requerem diversas avaliações de FOMs em cada iteração. Para tal função escolhemos a configuração de ANN que apresentou o melhor desempenho.

A população inicial do DE foi composta por 600 moduladores gerados aleatoriamente, com seus parâmetros limitados entre os valores indicados na Tabela 3. A otimização por DE foi realizada durante 50 gerações e, para cada indivíduo da população, geramos um vetor tentativa. O coeficiente de mutação foi estabelecido como 0,5. Definimos também uma probabilidade de 0,7 para o cruzamento gene a gene entre o indivíduo a ser analisado e o vetor doador. Para assegurar que as condições definidas na Tabela 3 sejam verificadas, aplicamos o método de penalização *bounce-back*. Neste caso, retificamos o valor do parâmetro caso este ultrapasse o limite máximo ou mínimo definido pela Tabela 3. Caso o valor seja menor que o limite inferior, este é atualizado para ser igual ao mínimo, assim como este tomará o valor máximo caso ultrapasse o limite superior.

4.2.3 Métricas de Otimização

Na presente subseção discutiremos algumas métricas de MZMs consideradas na otimização por DE. A Tabela 5 ilustra as FOMs utilizadas na otimização pelo algoritmo DE em conjunto com a ANN. Devido ao compromisso entre as métricas de desempenho do MZM, a otimização de uma destas irá limitar as outras. Ao otimizar apenas duas, enquanto mantemos a outra dentro de um determinado limite, podemos observar qual é a configuração ótima encontrada para tais métricas de desempenho. As variáveis BW_{thr} , IL_{thr} e V_{thr} se referem aos limites impostos nas métricas de BW_{EO} , IL e V_{π} respectivamente. Na FOM definida, quando um dos parâmetros é limitado, podemos perceber que há duas

Tabela 5 – FOMs utilizados durante a otimização. A tabela inclui os FOMs que foram utilizados ao limitar alguma das métricas, de forma que seu valor não seja ultrapassado

Métrica a limitar	FOM
BW_{EO}	$(V_{\pi}^2 IL)^{-1}$, se $BW_{EO} > BW_{thr}$ e $BW_{EO} \cdot 10^{-6}$, caso contrário.
IL	$BW_{EO}^2 V_{\pi}^{-2}$, se $IL < IL_{thr}$ e $IL^{-1} \cdot 10^{-6}$, caso contrário.
V_{π}	$BW_{EO}^2 IL^{-1}$, se $V_{\pi} < V_{thr}$ e $V_{\pi}^{-1} \cdot 10^{-6}$, caso contrário.
Sem limitações	$\frac{BW_{EO}^2}{V_{\pi}^{1,8}}$

condições. Quando a condição de limiar é obedecida, o algoritmo otimiza as outras duas variáveis. Quando tal condição não é obedecida, o algoritmo otimiza a variável a ser limitada e a multiplica por um número de valor diminuto (10^{-6}), de forma a criar regiões de FOM evitáveis.

Utilizamos também uma FOM que considera todas as métricas de desempenho do MZM. Verificamos que ao dar um peso levemente maior para o parâmetro BW_{EO} , o algoritmo DE foi capaz de encontrar soluções mais balanceadas. Mantivemos a IL excluída desta FOM pois, maximizando o BW_{EO} , estamos minimizando o IL de forma indireta. Isto ocorre pois o comprimento do modulador é um dos parâmetros com maior influência em ambos. As perdas são dependentes da distância de propagação, e como a região do silício dopado apresenta perdas ópticas significativas, o comprimento do dispositivo tem um impacto relevante no IL. Já para o caso do BW_{EO} , um maior comprimento do MZM ocasiona maiores perdas de RF nas linhas de transmissão do TWE, e também o impacto do descasamento de velocidade entre as ondas ópticas e de RF é mais acentuado. Estas limitações são demonstradas pelas Equações 70 e 69.

5 Resultados

No presente capítulo iremos discutir os resultados obtidos no trabalho. Na Seção 5.1 iremos apresentar os resultados obtidos da modelagem do MZM utilizando ANN. Na Seção 5.2 iremos apresentar os resultados da otimização de *design* implementando DE em conjunto com a ANN.

5.1 Resultados da Rede Neural Artificial

Na presente seção iremos discutir os resultados obtidos pela ANN, tanto durante seu aprendizado quanto no seu desempenho final.

5.1.1 Curvas de Aprendizado

A Figura 27 mostra os resultados obtidos pelas configurações de ANN de 4 camadas escondidas, tais quais são listadas na terceira coluna da Tabela 4. A Figura 27(a) se refere ao MSE em escala logarítmica para conjunto de treinamento, enquanto a Figura 27(b) se refere ao conjunto de testes. Podemos observar que o comportamento do erro para o conjunto de treinamento é diferente em relação aquele do testes. Enquanto que a utilização do RES, tanto individualmente quanto combinado com DO ou BN, aumenta o erro de treinamento em relação à ANN base, este foi o mecanismo que mais melhorou a capacidade de generalização da ANN. Podemos concluir que o RES foi o que apresentou o menor erro no conjunto de testes. O BN, por si só, apresentou piores resultados que o DO. Já na combinação de DO com BN, ocorre um evento intrigante, em que um resultado pior do que apenas com BN ou apenas com DO é encontrado após as 400 iterações. Uma hipótese é que, devido às variações aleatórias das distribuições de saída de cada camada escondida causada pelo DO, o BN não consegue encontrar seus parâmetros ótimos (LI *et al.*, 2019). Entretanto, a ANN completa mostra que o RES é capaz de contornar esta desarmonia entre BN e DO. Podemos perceber também que, em todos os casos que utilizamos RES, o erro de validação requiriu de um número de menor de iterações para convergir para um valor. Outro fato interessante a respeito do RES é que, nos casos em que este é considerado, o erro de validação não torna a aumentar. Como apresentado na Figura 17, uma das

características do *overfitting* é justamente um aumento do erro de validação enquanto o de treinamento é diminuído. Apesar do RES não ser considerado uma forma de regularização, este reduziu o *overfitting* no aprendizado deste conjunto de dados em questão.

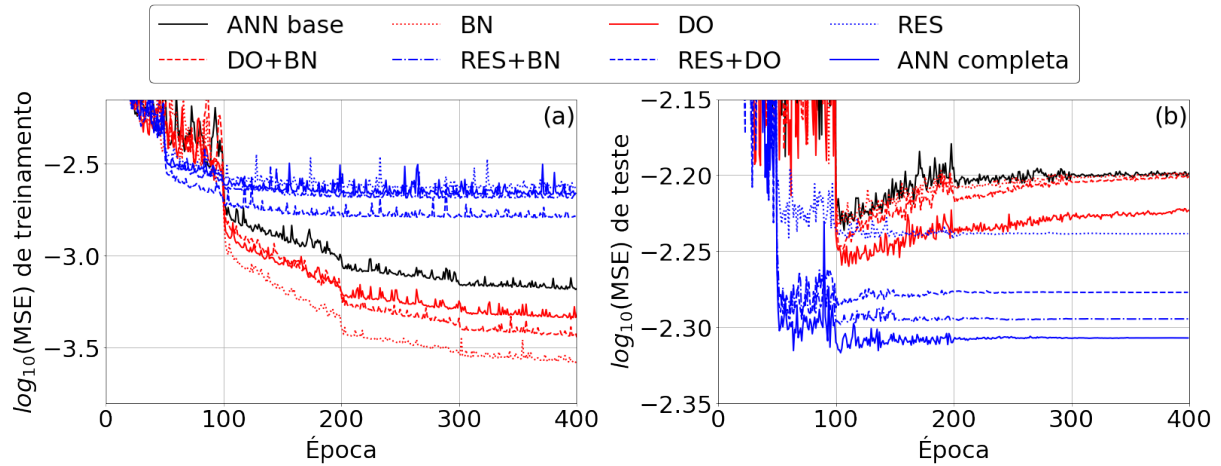


Figura 27 – Curva de aprendizado para o caso da ANN que demonstra 4 camadas escondidas. A figura compara a otimização da ANN ao utilizar BN, DO, RES, todas combinações de 2 destes e o modelo completo. As curvas de aprendizado são em relação ao (a) conjunto de treinamento e (b) conjunto de testes.

A Figura 28 ilustra as curvas de aprendizado tanto para o caso do (a) conjunto de treinamento, quanto para o caso do (b) conjunto de testes, para ANNs de 3 a 6 camadas escondidas. A figura também compara as curvas de aprendizado para ANNs base e completa. Assim como no caso da Figura 28, a utilização de RES aumentou o erro no conjunto de treinamento mas reduziu o erro do conjunto de testes. A ANN completa de 5 camadas escondidas foi a que mais reduziu o erro de teste, entretanto a de 4 camadas foi a que mais reduziu o erro de treinamento, tanto no caso da ANN base como da completa. Já para o caso das ANNs com 3 e 6 camadas, um pior resultado pode ser observado. Nossa hipótese é de que, com 3 camadas, a ANN não possui um número suficiente de parâmetros para alcançar o resultado ótimo, apresentando *underfitting*. Por outro lado, com 6 camadas escondidas, a quantidade de parâmetros da ANN pode ser alta demais para o problema, o que pode significar um leve *overfitting*. Podemos perceber também o efeito do RES em reduzir o *overfitting*, para todas as configurações de ANN. Escolhemos a ANN completa de 5 camadas escondidas para a otimização conjunta com o DE devido ao seu menor erro no conjunto de testes.

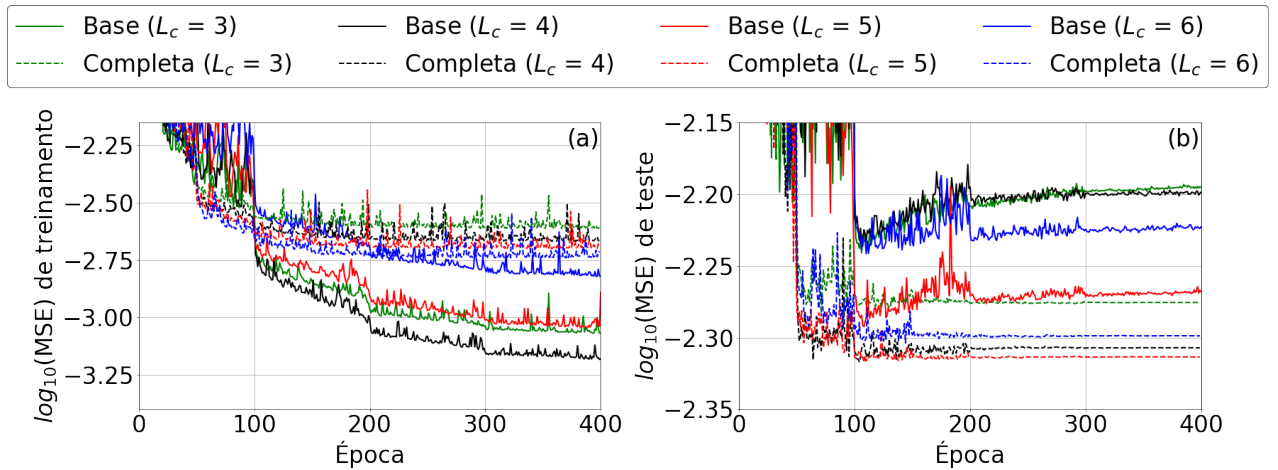


Figura 28 – Curvas de aprendizado comparando as ANNs de 3 a 6 camadas escondidas, tanto para o caso da ANN base quanto para a ANN completa. Curvas de aprendizado em relação ao (a) conjunto de treinamento e (b) conjunto de testes.

Finalmente, a Figura 29 mostra o decaimento da taxa de aprendizado em relação às iterações. Nas Figuras 27 e 28 podemos observar o decaimento do erro mais abrupto a cada redução da taxa de aprendizado, principalmente na iteração 50 e na 100. O η inicial é de $5 \cdot 10^{-3}$ e, no final das 400 épocas, este converge para $6,4 \cdot 10^{-4}$.

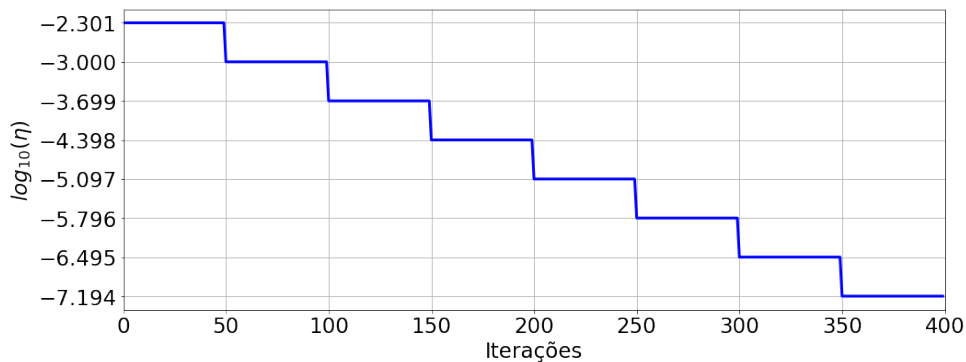


Figura 29 – Decaimento da taxa de aprendizado em relação às épocas do treinamento da ANN.

5.1.2 Desempenho da Rede Neural Artificial

Na presente subseção iremos analisar o desempenho da ANN completa de 5 camadas, que foi a que apresentou melhores resultados, como observado na subseção anterior. Para a avaliação de desempenho da ANN nós utilizamos o conjunto de testes, previamente

reservado para este fim. A Figura 30 mostra os resultados de predição obtidos pela rede em relação aos resultados obtidos pelo modelo de simulação. Primeiramente, nas Subfiguras 30(a), (b) e (c), selecionamos 50 amostras aleatórias do conjunto de testes e comparamos a saída da rede com os valores simulados para os parâmetros BW_{EO} , IL e V_{π} , respectivamente. O valor simulado é representado por um marcador em círculo azul, enquanto o previsto pela ANN é representado por um marcador em círculo laranja. A partir destas figuras podemos observar uma alta precisão na inferência da ANN considerando os dados de teste, apesar de ser possível perceber uma maior dispersão entre os valores de V_{π} .

Como outra forma de analisar o desempenho de predição da rede no conjunto de dados de teste, as Subfiguras 30(d), (e) e (f) mostram a relação entre os valores previstos e os de simulação para as métricas BW_{EO} , IL e V_{π} , respectivamente. Cada ponto indicado nos gráficos representa uma amostra de modulador, tendo seu valor de métrica simulado e previsto representado em ambos os eixos. A linha tracejada representa a relação ideal entre os valores simulados e os previstos, caso tal em que não há erros. Podemos perceber que para o caso da Figura 30(f) há uma maior dispersão de pontos para valores

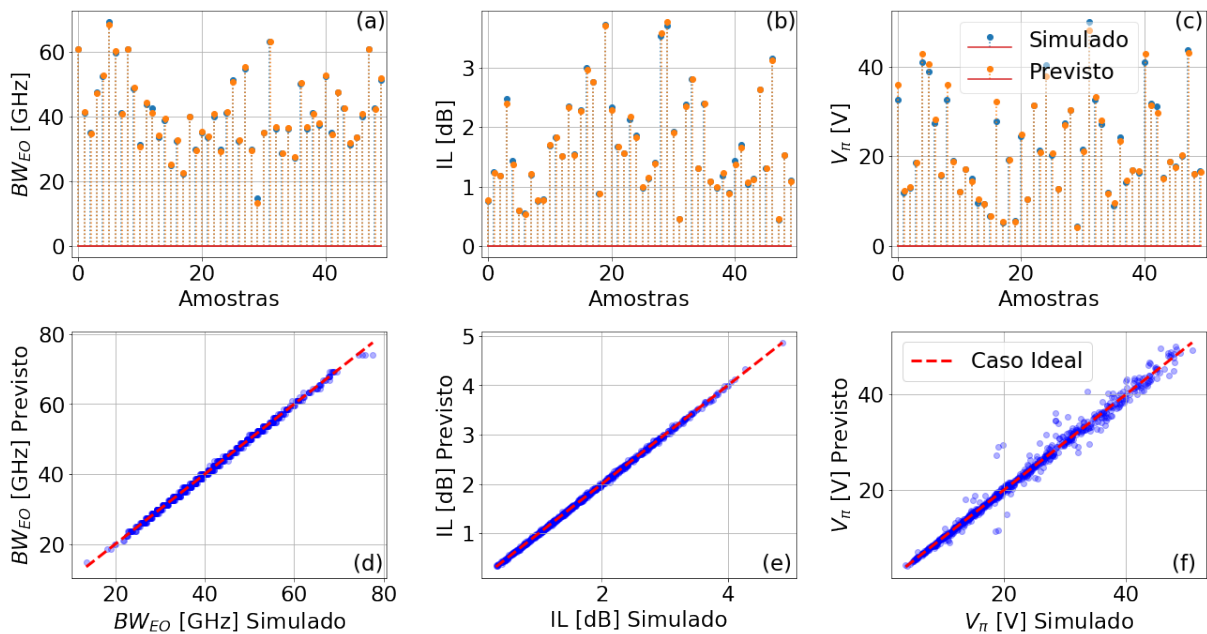


Figura 30 – Sobreposição de um subconjunto dos valores de teste simulados e previstos pelo modelo baseado em ANN para (a) BW_{EO} , (b) IL e (c) V_{π} . Relação entre os valores simulados e previstos, considerando todo o conjunto de testes para (d) BW_{EO} , (e) IL e (f) V_{π} . A linha tracejada indica os valores ideais.

Tabela 6 – Valores de MSE, MAE e r_p para as métricas BW_{EO} , IL e V_π e para as amostras em que $V_\pi < 10$ V. Não calculamos o r_p para o caso em que $V_\pi < 10$ V pois a correlação de Pearson para conjunto de dados com diferentes amostras pode não ser muito precisa.

	BW_{EO}	IL	V_π	$V_\pi < 10$ V
MAE	0,37	0,016	0,63	0,19
MSE	0,23	$5,00 \cdot 10^{-4}$	1,56	0,13
r_p [%]	99,91	99,93	99,36	–

elevados de V_π . Porém, esta região corresponde a valores de V_π que não são de interesse e, conseqüentemente, não afeta significativamente o processo de otimização. Os valores de correlação de Pearson calculados segundo a Equação 108 são de 99,91%, 99,93% e 99,36% para o BW_{EO} , o IL e a V_π , respectivamente.

A Tabela 6 mostra os valores de MSE e MAE, considerando os parâmetros BW_{EO} , IL e V_π e também para as amostras em que $V_\pi < 10$ V. Os valores de MAE obtidos para os parâmetros BW_{EO} , IL e V_π são dados por 0,37, 0,0165 e 0,63, respectivamente. Cabe mencionar que, apesar da variável V_π apresentar um MAE maior, a maior parte dos erros é ocasionada por valores de V_π elevados. Se considerarmos apenas as amostras que possuem um $V_\pi < 10$ V, que é a faixa de interesse na otimização, o valor de MAE decai para apenas 0,19, ou seja, menos do que um terço do MAE considerando todas amostras. Nossa hipótese é de que isto ocorre pois algumas combinações aleatórias de parâmetros na nossa geração de dados resulta em alguns valores de V_π exorbitantes, de até 50 V, o que gera *outliers* no nosso conjunto.

É válido mencionar que, o modelo de simulação demora cerca de 3 minutos para realizar a inferência de uma única amostra, considerando um CPU Intel Xeon E5-2650. Enquanto isso, a ANN completa de 5 camadas demora cerca de $17,2 \mu s$ para realizar uma inferência na mesma amostra, considerando uma CPU similar. Podemos concluir então que a ANN apresenta um tempo de inferência 10 milhões de vezes inferior em relação aquele apresentado pelo modelo de simulação.

5.2 Resultados de Otimização Utilizando Evolução Diferencial

A Figura 31 mostra os resultados da otimização por DE ao utilizar a FOM geral, descrita na última linha da Tabela 5. Executamos o algoritmo de DE 100 vezes, com populações inicializadas aleatoriamente para cada uma das execuções. Podemos perceber

que a otimização por DE encontrou valores ligeiramente menores de V_π e de BW_{EO} em relação ao conjunto de treinamento, enquanto valores intermediários de IL foram alcançados. A otimização resultou em MZMs que apresentaram valores de BW_{EO} entre 37,5 GHz e 39 GHz, IL entre 3,8 dB e 3,9 dB e V_π entre 5,9 V e 6,2 V. Podemos perceber que, mesmo sem otimizar o IL diretamente, este ainda apresentou valores intermediários e pouca variação entre as diferentes execuções do DE.

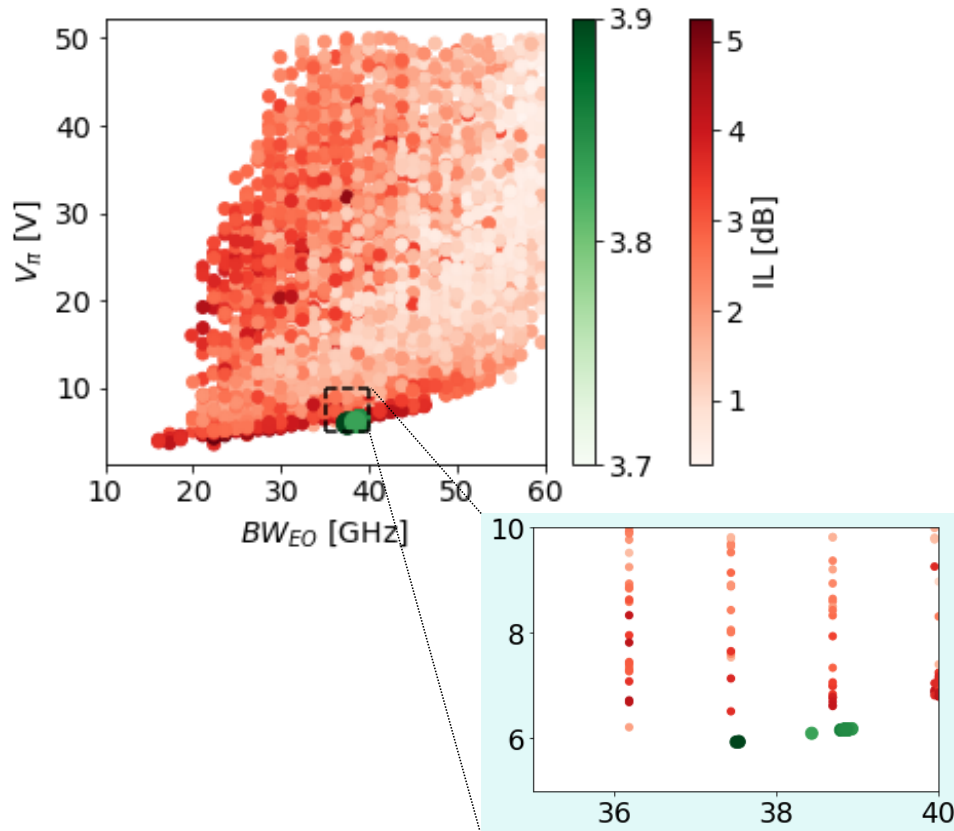


Figura 31 – Métricas de desempenho do MZM do conjunto de treinamento da ANN e as métricas obtidas por 100 execuções do algoritmo DE. Os eixos são dados pela BW_{EO} e pelo V_π , e a intensidade de cor representa os valores de IL.

A Figura 32 ilustra a curva de convergência do DE para as 100 inicializações aleatórias, de acordo com a Figura 31, levando em conta o melhor FOM encontrado na geração. Os valores convergiram para uma média de FOM de aproximadamente 56,86 e um desvio padrão ao redor da média de aproximadamente 0,17. Estes resultados nos mostram a eficiência do DE ao otimizar os parâmetros do MZM em conjunto com a ANN, e também a baixa variância dos resultados para diferentes condições iniciais. Uma comparação com outros algoritmos heurísticos pode encontrada na Seção 3 do Apêndice B. Por outro lado,

estas características do DE já foram analisadas, para outros problemas de otimização, em (GEORGIODAKIS; PLEVRIS, 2020).

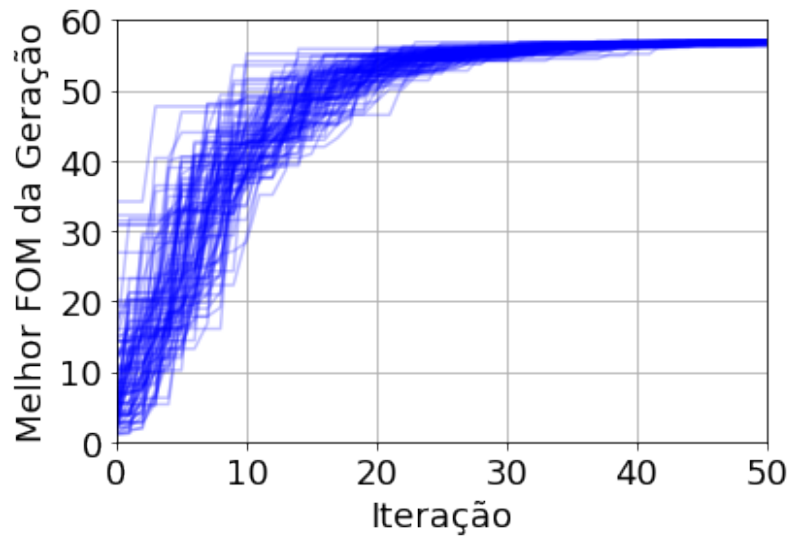


Figura 32 – FOM do melhor indivíduo da geração considerando 100 inicializações aleatórias.

A Figura 33 ilustra os valores das métricas de diferentes configurações otimizadas de MZMs. Executamos o DE 100 vezes para cada valor de limiar diferente, e para cada parâmetro de desempenho limitado. Os centroides de cada conjunto de otimizações são representados por marcadores em estrela. As três subfiguras da Figura 33 demonstram a mesma relação entre as métricas do MZM, porém desde pontos de vistas diferentes. É válido ressaltar que os valores das métricas limitadas encontrados pela otimização sempre vão ser aproximadamente equivalentes ao valor limiar.

A Subfigura 33(a) mostra a otimização de acordo com a FOM dada pela primeira linha da Tabela 5, em que os valores escolhidos de limiar para o BW_{EO} são de 30 GHz, 35 GHz, 40 GHz e 45 GHz. Podemos perceber que ao limitar o valor de BW_{EO} para apenas 30 GHz, a otimização encontrou valores de V_{π} de menos que 5 V. Entretanto, ao aumentarmos o limite inferior de BW_{EO} , o V_{π} tende a aumentar, o que indica uma menor eficiência de modulação ao custo de uma resposta mais rápida de modulação. É evidente também que o IL tende a diminuir ao aumentarmos a BW_{EO} . Como discutido na Subseção 4.2.3, a maximização do BW_{EO} tem grande interação com a minimização do IL. No caso do IL, na Subfigura 33(b), otimizamos a FOM dada pela segunda linha da Tabela 5, utilizando valores de 3 dB, 3,5 dB, 4 dB, 4,5 dB e 5 dB como limites superiores de IL. Já para na Subfigura 33(c) otimizamos a FOM dada pela terceira linha da Tabela 5,

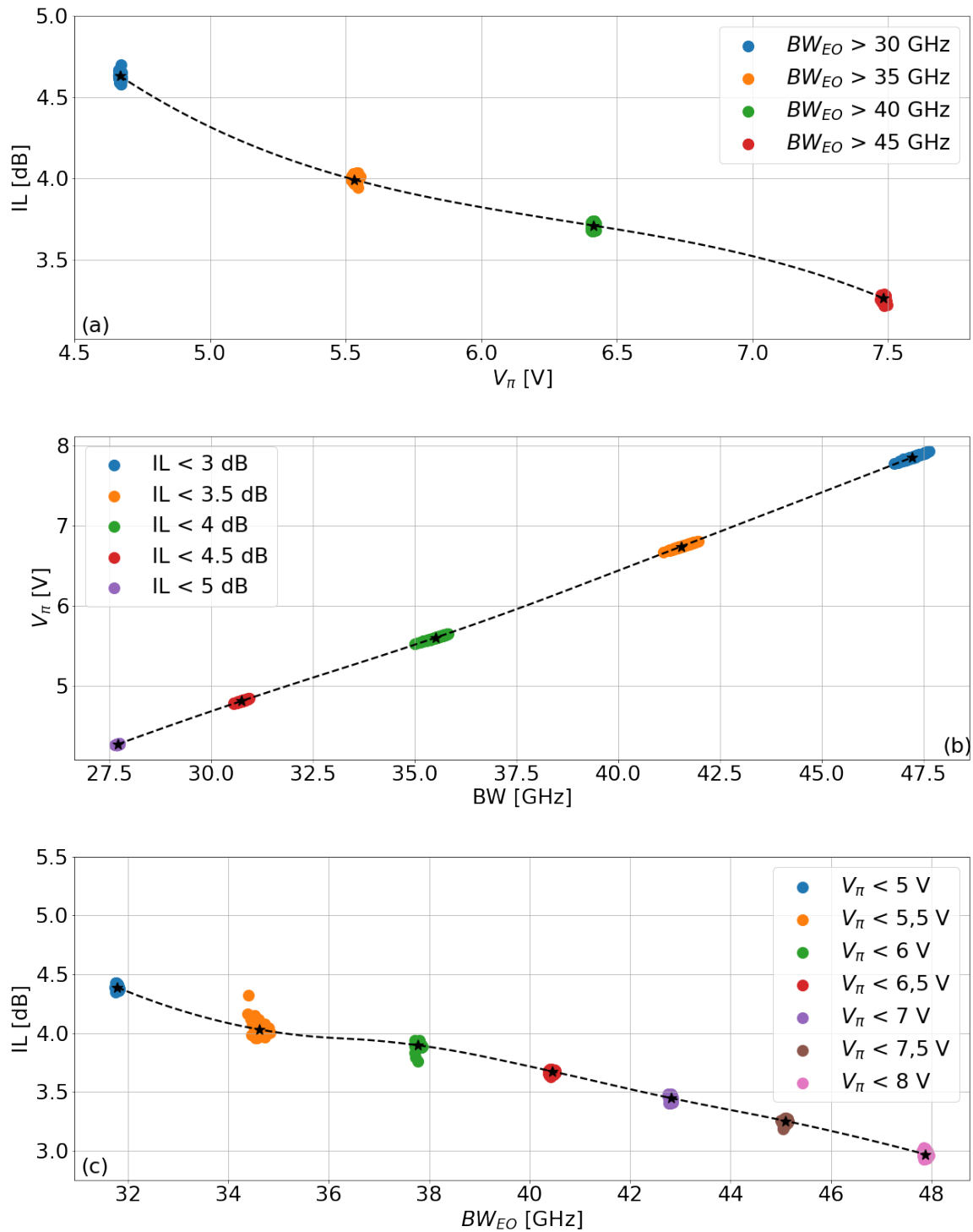


Figura 33 – Otimização por DE de apenas duas métricas de desempenho do MZM, enquanto o outro é limitado para determinado valor de limiar. (a) Métricas obtidas ao limitar o BW_{EO} , (b) métricas obtidas ao limitar o IL e (c) métricas obtidas ao limitar o V_{π} .

com os seguintes valores de limite superior para V_π : 5 V, 5,5 V, 6 V, 6,5 V, 7 V, 7,5 V e 8 V. Nesta Subfiguras fica evidente a relação do aumento do V_π à medida que o BW_{EO} também é aumentado. Evidentemente, isto também resulta em uma redução do IL, pois o comprimento do modulador tende a ser menor.

5.3 Validação dos Moduladores Otimizados com o Modelo de Simulação

Os MZMs obtidos pela otimização são diferentes daqueles observados pelo modelo de ANN durante o treinamento. Por conseguinte, é interessante validarmos as configurações otimizadas utilizando o modelo de simulação eletromagnética. Realizamos tal análise utilizando as 100 configurações de MZMs otimizadas por DE, ilustrados na Figura 31, e mais 300 configurações de MZMs otimizadas por outros algoritmos heurísticos, ilustrados na Figura 42. Mais detalhes sobre a obtenção destes últimos são dados na Seção 3 do Apêndice B.

A Figura 34 ilustra a mesma análise de acurácia realizada no conjunto de treinamento, como é dada na Figura 30. Nas Subfiguras 34 (a), (b) e (c), selecionamos 50 configurações otimizadas de MZMs, e foram comparados as previsões da ANN e os valores obtidos pelo modelo de simulação para as métricas BW_{EO} , IL e V_π , respectivamente. Como podemos observar, as amostras estão bem próximas para qualquer um dos três parâmetros de métrica. Já as Subfiguras 34 (d), (e) e (f) mostram a relação entre as métricas de desempenho obtidas pela ANN e pelo modelo de simulação considerando os 400 MZMs otimizados. Como podemos observar, os resultados estão bem próximos do caso ideal, tal qual representado pela linha tracejada. Entretanto, podemos reparar que a ANN ligeiramente subestima os valores das três métricas.

A Tabela 7 mostra as métricas obtidas ao compararmos os valores previstos com os simulados. Como podemos visualizar, os valores de MAE são menores que 0,5 para todos os casos, e os valores de MSE apresentados são adequadamente baixos. O coeficiente de correlação de Pearson é maior do que 98% para as três métricas. A tabela também mostra a média das métricas obtidas tanto pelo modelo de simulação quanto pela ANN. Podemos comprovar por estes valores que, como mencionamos, a ANN tende a subestimar os valores das métricas, apresentando erros percentuais de no máximo 2,5% em relação às suas médias.

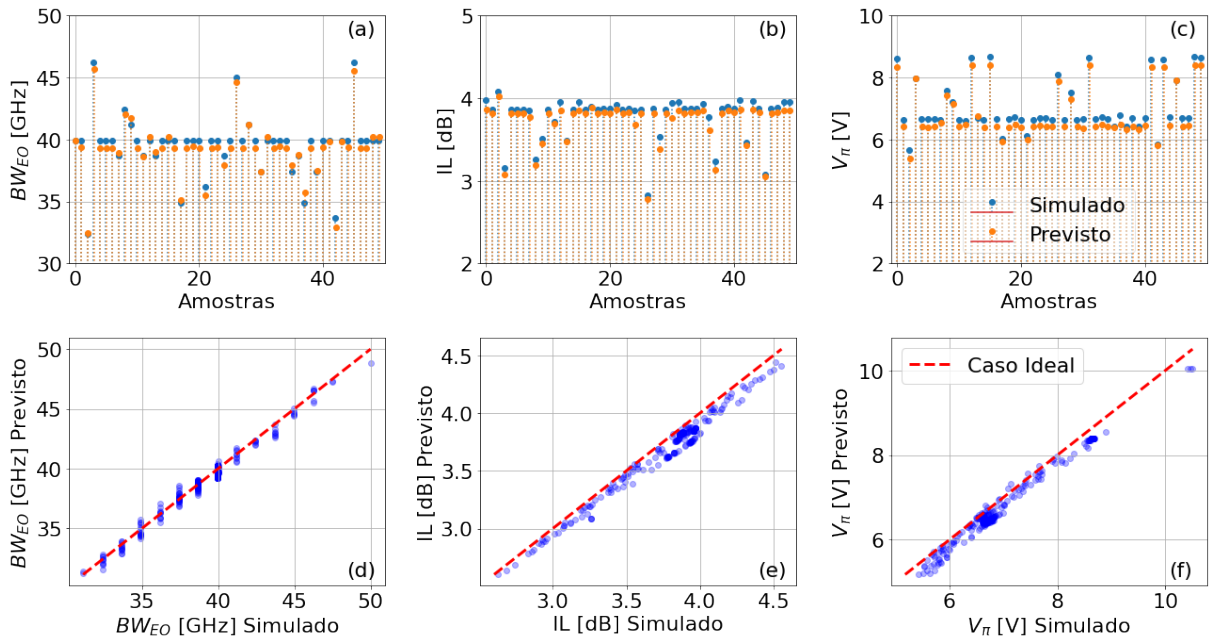


Figura 34 – Sobreposição de um subconjunto dos valores de métricas relativas aos MZMs otimizados pelos algoritmos heurísticos, previstos pela ANN e pelo modelo de simulação eletromagnética, para (a) BW_{EO} , (b) IL e (c) V_{π} . Relação entre os valores simulados e previstos, considerando os 400 MZM otimizados, para (d) BW_{EO} , (e) IL e (f) V_{π} . A linha tracejada indica a relação ideal.

Tabela 7 – Avaliação das métricas obtidas pelos 400 MZMs otimizados pelos algoritmos heurísticos. A tabela inclui o MAE, MSE e r_p ao relacionar as métricas obtidas pelos diferentes métodos. A tabela também inclui a média das métricas para cada método, assim como o erro percentual entre estas.

Métrica	MAE	MSE	r_p	Média de Simulação	Média de Predição	Erro Percentual das Médias
BW_{EO}	0,451	0,255	98,57%	39,32 GHz	39,03 GHz	-0,74%
IL	0,075	0,007	98,81%	3,81 dB	3,74 dB	-1,87%
V_{π}	0,219	0,055	99,33%	6,96 V	6,79 V	-2,5 %

6 Conclusões e Trabalho Futuro

A partir de modelos analíticos e empíricos encontrados na literatura, foi utilizado um modelo de simulação para estimar a BW_{EO} , a IL e a V_π de um MZM (MOTTA *et al.*, 2017). Este modelo inclui a simulação de linhas de transmissão pelo *software* CTS e, incluindo a junção PN como a carga em tal linha, estimamos a BW_{EO} utilizando uma linguagem de programação de alto nível. As interações entre a onda óptica e a densidade de portadores no MZM foram simuladas com o *software* Lumerical *Device*, o qual nos permitiu extrair o V_π e o IL do dispositivo. Variando aleatoriamente os parâmetros geométricos da junção PN, e também a V_{bias} e o L_{ps} , obtivemos 10.000 dados de MZMs, que foram utilizados para treinar e avaliar uma ANN.

Desenvolvemos uma ANN capaz de modelar as relações entre as características do MZM e as métricas de desempenho BW_{EO} , IL e V_π . Apesar do custo computacional para gerar os dados de treinamento, após ser treinado, o modelo pode substituir as simulações eletromagnéticas de forma muito mais eficiente em relação ao tempo de inferência. Esta economia de tempo nos permite executar algoritmos de otimização, como o DE, com uma vasta população e por diversas gerações, em poucos minutos. Até onde sabemos, esta é a primeira vez que uma ANN está sendo implementada para modelar um MZM em silício.

Implementamos ferramentas de DL na ANN, como o BN, o DO e o RES, de forma a melhorar a convergência e a generalização da mesma. Otimizamos também a quantidade de camadas escondidas do modelo. Utilizando o DE em conjunto com a ANN, obtivemos diversos MZMs de alto desempenho, apresentando BW_{EO} de quase 40 GHz, V_π em torno de 6 V e IL de menos de 4 dB. Podemos também adaptar a otimização para diferentes aplicações ao modificarmos a FOM. A partir disto, analisamos o compromisso entre os parâmetros de desempenho do MZM na otimização.

Trabalhos futuros poderão incluir uma ANN de mapeamento inverso para a otimização do MZM. Neste caso, os parâmetros de desempenho são tomados como entrada e a ANN retorna os parâmetros de *design* responsáveis por este. Encontramos problemas na convergência de tal ANN, que podem ser contornados em trabalhos futuros. Outra sugestão para trabalhos futuros consiste numa análise da complexidade computacional da ANN e do método de otimização conjunto da ANN com o algoritmo heurístico.

Referências

- AGARAP, A. F. *Deep Learning using Rectified Linear Units (ReLU)*. 2019. Disponível em: <https://arxiv.org/abs/1803.08375>. Citado na página 83.
- AGRAWAL, G. *Sistemas de comunicação por fibra óptica*. [S.l.]: Elsevier Brasil, 2014. v. 4. Citado 7 vezes nas páginas 27, 30, 35, 36, 47, 49 e 51.
- ALAM, M. S. *et al.* Net 220 Gbps/ λ IM/DD transmsion in O-band and C-band with silicon photonic traveling-wave MZM. *Journal of Lightwave Technology*, OSA, v. 39, n. 13, p. 4270–4278, Jul 2021. Disponível em: <http://www.osapublishing.org/jlt/abstract.cfm?URI=jlt-39-13-4270>. Citado na página 75.
- ALFERNES, R. Waveguide electrooptic modulators. *IEEE Transactions on Microwave Theory and Techniques*, v. 30, n. 8, p. 1121–1137, 1982. Citado na página 67.
- ALPAYDIN, E. *Introduction to Machine Learning*. 2nd. ed. [S.l.]: The MIT Press, 2010. ISBN 026201243X, 9780262012430. Citado 2 vezes nas páginas 87 e 88.
- BAEHR-JONES, T. *et al.* Ultralow drive voltage silicon traveling-wave modulator. *Optics Express*, OSA, v. 20, n. 11, p. 12014–12020, May 2012. Disponível em: <http://www.opticsexpress.org/abstract.cfm?URI=oe-20-11-12014>. Citado na página 43.
- BAHL, I. *et al.* *Microstrip Lines and Slotlines, Third Edition*. 3rd. ed. [S.l.]: Artech House, 2013. Citado na página 68.
- BASHEER, I.; HAJMEER, M. Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods*, v. 43, n. 1, p. 3–31, 2000. ISSN 0167-7012. Neural Computing in Microbiology. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0167701200002013>. Citado na página 82.
- BENNER, A. Optical interconnect opportunities in supercomputers and high end computing. In: *Optical Fiber Communication Conference*. Optical Society of America, 2012. p. OTu2B.4. Disponível em: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2012-OTu2B.4>. Citado na página 26.
- BOR, E. *et al.* Differential evolution algorithm based photonic structure design: Numerical and experimental verification of subwavelength $\lambda/5$ focusing of light. *Scientific Reports*, v. 6, p. 30871, 08 2016. Citado na página 96.
- BURKOV, A. *The Hundred-Page Machine Learning Book*. Andriy Burkov, 2019. ISBN 9781999579517. Disponível em: <https://books.google.com.br/books?id=0jbxwQEACAAJ>. Citado na página 82.
- CAMPOPIANO, C.; GLAZER, B. A coherent digital amplitude and phase modulation scheme. *IRE Transactions on Communications Systems*, v. 10, n. 1, p. 90–95, 1962. Citado na página 28.
- CELLER, G.; CRISTOLOVEANU, S. Frontiers of silicon-on-insulator. *Journal of Applied Physics*, v. 93, p. 4955–4978, 05 2003. Citado na página 30.

CHAHAR, V. *et al.* A review on genetic algorithm: Past, present, and future. *Multimedia Tools and Applications*, v. 80, 02 2021. Citado na página 142.

CHANG, H.-H. *et al.* Integrated hybrid silicon triplexer. *Optics Express*, OSA, v. 18, n. 23, p. 23891–23899, Nov 2010. Disponível em: <http://www.opticsexpress.org/abstract.cfm?URI=oe-18-23-23891>. Citado na página 31.

CHEN, E.; CHOU, S. Characteristics of coplanar transmission lines on multilayer substrates: modeling and experiments. *IEEE Transactions on Microwave Theory and Techniques*, v. 45, n. 6, p. 939–945, 1997. Citado na página 67.

CHEN, H.-W. *et al.* Forty Gb/s hybrid silicon Mach-Zehnder modulator with low chirp. *Optics Express*, OSA, v. 19, n. 2, p. 1455–1460, Jan 2011. Disponível em: <http://www.opticsexpress.org/abstract.cfm?URI=oe-19-2-1455>. Citado na página 43.

CHEN, L. *et al.* Chirp and dispersion tolerance of a single-drive push-pull silicon modulator at 28 Gb/s. *IEEE Photonics Technology Letters*, v. 24, n. 11, p. 936–938, 2012. Citado 2 vezes nas páginas 54 e 70.

CHMIELAK, B. *et al.* Pockels effect based fully integrated, strained silicon electro-optic modulator. *Optics Express*, OSA, v. 19, n. 18, p. 17212–17219, Aug 2011. Disponível em: <http://www.opticsexpress.org/abstract.cfm?URI=oe-19-18-17212>. Citado na página 36.

CHOO, J. *et al.* Simple quasi-static analysis for coplanar stripline within multilayer dielectrics. *IEICE Electronics Express*, v. 12, n. 17, p. 20150361–20150361, 2015. Citado na página 68.

CHROSTOWSKI, L.; HOCHBERG, M. *Silicon Photonics Design: From Devices to Systems*. [S.l.]: Cambridge University Press, 2015. Citado 10 vezes nas páginas 30, 31, 39, 40, 46, 63, 64, 65, 66 e 105.

CISCO. *Cisco Visual Networking Index: Forecast and Trends, 2018–2023 White Paper*. 2022. <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html>. Acessado em: 22/04/2022. Citado na página 25.

COCORULLO, G.; RENDINA, I. Thermo-optical modulation at 1.5 μm in silicon etalon. *Electronics Letters*, v. 28, p. 83 – 85, 02 1992. Citado na página 35.

COLDREN, L. A. *et al.* Diode lasers and photonic integrated circuits. In: _____. *Diode Lasers and Photonic Integrated Circuits*. John Wiley & Sons, Ltd, 2012. cap. 8, p. 451–507. ISBN 9781118148167. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118148167.ch8>. Citado na página 29.

CONG, G. *et al.* Ultra-compact non-travelling-wave silicon carrier-depletion mach-zehnder modulators towards high channel density integration. *IEEE Journal of Selected Topics in Quantum Electronics*, v. 27, n. 3, p. 1–11, 2021. Citado na página 67.

DAI, L. *et al.* Deep learning for wireless communications: An emerging interdisciplinary paradigm. *IEEE Communications Magazine*, PP, 05 2020. Citado na página 79.

DAS, S.; SUGANTHAN, P. N. Differential evolution: A survey of the state-of-the-art. *IEEE Transactions on Evolutionary Computation*, v. 15, n. 1, p. 4–31, 2011. Citado na página 96.

DATA, I. W. *Internet World Data*. 2022. [⟨https://www.internetworldstats.com/sa/br.htm⟩](https://www.internetworldstats.com/sa/br.htm). Acessado em: 10/05/2022. Citado na página 25.

DATAPROT. *IoT Statistics*. 2022. [⟨https://dataprot.net/statistics/iot-statistics/⟩](https://dataprot.net/statistics/iot-statistics/). Acessado em: 10/05/2022. Citado na página 25.

DENIEL, L. *et al.* DAC-less PAM-4 generation in the O-band using a silicon Mach-Zehnder modulator. *Optics Express*, OSA, v. 27, n. 7, p. 9740–9748, Apr 2019. Disponível em: [⟨http://www.opticsexpress.org/abstract.cfm?URI=oe-27-7-9740⟩](http://www.opticsexpress.org/abstract.cfm?URI=oe-27-7-9740). Citado na página 43.

DERR, F. Optical QPSK transmission system with novel digital receiver concept. *Electronics Letters*, v. 27, p. 2177–2179, nov 1991. Citado na página 29.

DILAO, R. From charles darwin to evolutionary genetic algorithms. In: _____. [S.l.: s.n.], 2009. p. 1–11. Citado na página 142.

DING, J. *et al.* Low-voltage, high-extinction-ratio, Mach-Zehnder silicon optical modulator for CMOS-compatible integration. *Optics Express*, OSA, v. 20, n. 3, p. 3209–3218, Jan 2012. Disponível em: [⟨http://www.osapublishing.org/oe/abstract.cfm?URI=oe-20-3-3209⟩](http://www.osapublishing.org/oe/abstract.cfm?URI=oe-20-3-3209). Citado na página 71.

DING, R. *et al.* High-speed silicon modulator with slow-wave electrodes and fully independent differential drive. *Journal of Lightwave Technology*, v. 32, n. 12, p. 2240–2247, 2014. Citado na página 70.

DOYLEND, J. K. *et al.* Silicon photonic resonator-enhanced defect-mediated photodiode for sub-bandgap detection. *Optics Express*, OSA, v. 18, n. 14, p. 14671–14678, Jul 2010. Disponível em: [⟨http://www.opticsexpress.org/abstract.cfm?URI=oe-18-14-14671⟩](http://www.opticsexpress.org/abstract.cfm?URI=oe-18-14-14671). Citado na página 31.

FENG, S. *et al.* Silicon photonics: from a microresonator perspective. *Laser & Photonics Reviews*, v. 6, n. 2, p. 145–177, 2012. Disponível em: [⟨https://onlinelibrary.wiley.com/doi/abs/10.1002/lpor.201100020⟩](https://onlinelibrary.wiley.com/doi/abs/10.1002/lpor.201100020). Citado 2 vezes nas páginas 37 e 39.

FIGUEIREDO, R. C. *et al.* Adjustable chirp to enhance chromatic dispersion tolerance for PAM4 transmission in intra data center communications. In: *Proceedings of the Brazilian Technology Symposium*. [S.l.: s.n.], 2017. v. 1, p. 1–4. Citado na página 35.

FOGEL, D. B. *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*. [S.l.]: IEEE Press, 1995. ISBN 0780310381. Citado na página 143.

FREEDMAN, D. *et al.* Statistics (international student edition). *Pisani, R. Purves, 4th edn. WW Norton & Company, New York*, 2007. Citado na página 95.

GARDES, F. Y. *et al.* Optical modulators in silicon photonic circuits. In: _____. *Silicon Photonics*. John Wiley & Sons, Ltd, 2008. cap. 4, p. 95–145. ISBN 9780470994535. Disponível em: [⟨https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470994535.ch4⟩](https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470994535.ch4). Citado na página 48.

GEORGIODAKIS, M.; PLEVRIS, V. A comparative study of differential evolution variants in constrained structural optimization. *Frontiers in Built Environment*, v. 6, 07 2020. Citado 2 vezes nas páginas 96 e 117.

- GHIONE, G. *Semiconductor Devices for High-Speed Optoelectronics*. 1st. ed. USA: Cambridge University Press, 2009. ISBN 0521763444. Citado 3 vezes nas páginas 74, 75 e 76.
- GOELL, J. E. Rib waveguide for integrated optical circuits. *Applied Optics*, OSA, v. 12, n. 12, p. 2797–2798, Dec 1973. Disponível em: <http://www.osapublishing.org/ao/abstract.cfm?URI=ao-12-12-2797>. Citado na página 62.
- GOODFELLOW, I. *et al. Deep Learning*. MIT Press, 2016. (Adaptive computation and Machine learning). ISBN 9780262035613. Disponível em: <https://books.google.co.in/books?id=Np9SDQAAQBAJ>. Citado na página 107.
- GOYKHMANN, I. *et al.* Optimization of efficiency-loss figure of merit in carrier-depletion silicon Mach-Zehnder optical modulator. *Optics Express*, OSA, v. 21, n. 17, p. 19518–19529, Aug 2013. Disponível em: <http://www.osapublishing.org/oe/abstract.cfm?URI=oe-21-17-19518>. Citado na página 56.
- GRIFFIN, R. *et al.* Integrated DQPSK transmitter for dispersion-tolerant and dispersion-managed DWDM transmission. In: *OFC 2003 Optical Fiber Communications Conference, 2003*. [S.l.: s.n.], 2003. p. 770–771 vol.2. Citado na página 42.
- GUAN, H. *et al.* Compact and low loss 90° optical hybrid on a silicon-on-insulator platform. *Optics Express*, OSA, v. 25, n. 23, p. 28957–28968, Nov 2017. Disponível em: <http://www.osapublishing.org/oe/abstract.cfm?URI=oe-25-23-28957>. Citado na página 33.
- GUTIERREZ, A. M. *et al.* Analytical model for calculating the nonlinear distortion in silicon-based electro-optic Mach-Zehnder Modulators. *Journal of Lightwave Technology*, v. 31, n. 23, p. 3603–3613, 2013. Citado na página 42.
- HAKANSSON, A. *et al.* High-efficiency defect-based photonic-crystal tapers designed by a genetic algorithm. *Journal of Lightwave Technology*, v. 23, n. 11, p. 3881–3888, 2005. Citado na página 142.
- HAYKIN, S. *Redes Neurais: Princípios e Prática*. Artmed, 2007. ISBN 9788577800865. Disponível em: <https://books.google.com.br/books?id=bhMwDwAAQBAJ>. Citado na página 84.
- HE, K. *et al.* Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *Computing Research Repository*, abs/1502.01852, 2015. Disponível em: <http://arxiv.org/abs/1502.01852>. Citado na página 89.
- HE, K. *et al.* Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [s.n.], 2016. p. 770–778. Disponível em: <https://arxiv.org/abs/1512.03385?context=cs>. Citado na página 93.
- HENI, W. *et al.* High speed plasmonic modulator array enabling dense optical interconnect solutions. *Optics Express*, v. 23, p. 29746–29757, 11 2015. Citado na página 51.
- HINAKURA, Y. *et al.* 64 Gbps Si photonic crystal slow light modulator by electro-optic phase matching. *Optics Express*, OSA, v. 27, n. 10, p. 14321–14327, May 2019. Disponível em: <http://www.opticsexpress.org/abstract.cfm?URI=oe-27-10-14321>. Citado na página 43.

- HINTON, G.; SEJNOWSKI, T. J. *Unsupervised Learning: Foundations of Neural Computation*. The MIT Press, 1999. ISBN 9780262288033. Disponível em: <https://doi.org/10.7551/mitpress/7011.001.0001>. Citado na página 78.
- HOLLAND, J. H. *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI: University of Michigan Press, 1975. Second edition, 1992. Citado na página 142.
- HORNIK, K. *et al.* Multilayer feedforward networks are universal approximators. *Neural Networks*, v. 2, p. 359–366, 1989. Citado na página 80.
- HUNSPERGER, R. G. *Integrated Optics: Theory and Technology*. 6th. ed. [S.l.]: Springer Publishing Company, Incorporated, 2009. ISBN 0387897747. Citado na página 31.
- HUSSAIN, M. *et al.* Effect of seven steps approach on simplex method to optimize the mathematical manipulation. *International Journal of Recent Technology and Engineering*, v. 7, p. 34–43, 01 2019. Citado na página 97.
- IOFFE, S.; SZEGEDY, C. *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. arXiv, 2015. Disponível em: <https://arxiv.org/abs/1502.03167>. Citado 2 vezes nas páginas 91 e 108.
- IOT-ANALYTICS. *Number of Connected IoT Devices*. 2022. <https://iot-analytics.com/number-connected-iot-devices/>. Acessado em: 10/05/2022. Citado na página 25.
- ISLAM, M. R. *et al.* A comparison of performance of GA, PSO and differential evolution algorithms for dynamic phase reconfiguration technology of a smart grid. In: *2019 IEEE Congress on Evolutionary Computation (CEC)*. [S.l.: s.n.], 2019. p. 858–865. Citado na página 96.
- IVANOVICH, D. *et al.* Polarization division multiplexing for optical data communications. In: *SPIE - Society of Photo-Optical Instrumentation Engineers*. [S.l.: s.n.], 2018. p. 50. Citado na página 32.
- JACQUES, M. *et al.* Modulator material impact on chirp, DSP, and performance in coherent digital links: comparison of the lithium niobate, indium phosphide, and silicon platforms. *Optics Express*, OSA, v. 26, n. 17, p. 22471–22490, Aug 2018. Disponível em: <http://www.opticsexpress.org/abstract.cfm?URI=oe-26-17-22471>. Citado na página 42.
- JACQUES, M. *et al.* 240 Gbit/s silicon photonic Mach-Zehnder modulator enabled by two 2.3-Vpp drivers. *Journal of Lightwave Technology*, v. 38, n. 11, p. 2877–2885, 2020. Citado na página 43.
- JACQUES, M. *et al.* Net 212.5 Gbit/s transmission in O-band with a SiP MZM , one driver and linear equalization. In: *2020 Optical Fiber Communications Conference and Exhibition (OFC)*. [S.l.: s.n.], 2020. p. 1–3. Citado na página 43.
- JAEGER, N.; LEE, Z. Slow-wave electrode for use in compound semiconductor electrooptic modulators. *IEEE Journal of Quantum Electronics*, v. 28, n. 8, p. 1778–1784, 1992. Citado na página 70.
- JAFARI, O. *et al.* High-efficiency silicon photonic modulator using coupled Bragg grating resonators. *Journal of Lightwave Technology*, v. 37, n. 9, p. 2065–2075, 2019. Citado na página 43.

- JAFARI, O. *et al.* Efficiency-speed tradeoff in slow-light silicon photonic modulators. *IEEE Journal of Selected Topics in Quantum Electronics*, v. 27, n. 3, p. 1–11, 2021. Citado na página 43.
- JAIN, N. *et al.* A review of particle swarm optimization. *Journal of The Institution of Engineers (India): Series B*, v. 99, 03 2018. Citado na página 144.
- JIN, J. *The Finite Element Method in Electromagnetics*. Wiley, 2002. ISBN 9780471438182. Disponível em: <https://books.google.com.br/books?id=N6yVQgAACAAJ>. Citado na página 105.
- KAZOVSKY, L. G. Optical heterodyning versus optical homodyning: A comparison. *Journal of Optical Communications*, v. 6, n. 1, p. 18–24, 1985. Disponível em: <https://doi.org/10.1515/JOC.1985.6.1.18>. Citado na página 29.
- KAZOVSKY, L. G. Coherent optical fiber communications. *IEEE Journal of Quantum Electronics*, QE-25, June 1989 1989. Citado na página 28.
- KENNEDY, J.; EBERHART, R. Particle swarm optimization. In: *Proceedings of ICNN'95 - International Conference on Neural Networks*. [S.l.: s.n.], 1995. v. 4, p. 1942–1948 vol.4. Citado na página 144.
- KERRINCKX, E. *et al.* Photonic crystal fiber design by means of a genetic algorithm. *Optics Express*, OSA, v. 12, n. 9, p. 1990–1995, May 2004. Disponível em: <http://opg.optica.org/oe/abstract.cfm?URI=oe-12-9-1990>. Citado na página 142.
- KESKAR, N. S. *et al.* *On Large-Batch Training for Deep Learning: Generalization Gap and Sharp Minima*. 2017. Disponível em: <https://arxiv.org/abs/1609.04836?context=cs>. Citado na página 106.
- KIEFER, J.; WOLFOWITZ, J. Stochastic Estimation of the Maximum of a Regression Function. *The Annals of Mathematical Statistics*, Institute of Mathematical Statistics, v. 23, n. 3, p. 462 – 466, 1952. Disponível em: <https://doi.org/10.1214/aoms/1177729392>. Citado na página 89.
- KIKUCHI, K. Fundamentals of coherent optical fiber communications. *Journal of Lightwave Technology*, v. 34, n. 1, p. 157–179, 2016. Citado 2 vezes nas páginas 28 e 29.
- KIM, T. *Generalizing MLPs With Dropouts, Batch Normalization, and Skip Connections*. arXiv, 2021. Disponível em: <https://arxiv.org/abs/2108.08186>. Citado na página 88.
- KINGMA, D. P.; BA, J. *Adam: A Method for Stochastic Optimization*. 2014. Disponível em: <https://arxiv.org/abs/1412.6980>. Citado 3 vezes nas páginas 89, 90 e 108.
- KIRKPATRICK, S. *et al.* Optimization by simulated annealing. *Science*, v. 220, n. 4598, p. 671–680, 1983. Disponível em: <https://www.science.org/doi/abs/10.1126/science.220.4598.671>. Citado na página 146.
- KRIZHEVSKY, A. *et al.* Imagenet classification with deep convolutional neural networks. *Communications of Association for Computing Machinery*, Association for Computing Machinery, New York, NY, USA, v. 60, n. 6, p. 84–90, maio 2017. ISSN 0001-0782. Disponível em: <https://doi.org/10.1145/3065386>. Citado na página 84.

- KUTZ, J. N. Deep learning for control of nonlinear optical systems. In: JALALI, B.; KITAYAMA, K. ichi (Ed.). *AI and Optical Data Sciences II*. Society of Photo-Optical Instrumentation Engineers, 2021. v. 11703, p. 63 – 70. Disponível em: <https://doi.org/10.1117/12.2576998>. Citado na página 79.
- LECUN, Y. *et al.* Deep learning. *Nature*, v. 521, p. 436–44, 05 2015. Citado na página 79.
- LI, G. Recent advances in coherent optical communication. *Advanced Optics Photonics*, OSA, v. 1, n. 2, p. 279–307, Apr 2009. Disponível em: <http://aop.osa.org/abstract.cfm?URI=aop-1-2-279>. Citado 2 vezes nas páginas 28 e 29.
- LI, X. *et al.* Understanding the disharmony between dropout and batch normalization by variance shift. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2019. p. 2677–2685. Citado 2 vezes nas páginas 108 e 111.
- Li, X. *et al.* Highly efficient silicon Michelson interferometer modulators. *IEEE Photonics Technology Letters*, v. 25, n. 5, p. 407–409, 2013. Citado na página 40.
- LIAO, F. *et al.* Deep learning methods in communication systems: A review. *Journal of Physics: Conference Series*, IOP Publishing, v. 1617, p. 012024, Aug 2020. Disponível em: <https://doi.org/10.1088/1742-6596/1617/1/012024>. Citado na página 79.
- LIU, A. *et al.* Recent development in a high-speed silicon optical modulator based on reverse-biased PN diode in a silicon waveguide. *Semiconductor Science and Technology*, v. 23, p. 064001, 05 2008. Citado na página 43.
- LIU, Y. *et al.* A brief review on deep learning in application of communication signal processing. In: *2020 IEEE 5th International Conference on Signal and Image Processing (ICSIP)*. [S.l.: s.n.], 2020. p. 51–54. Citado na página 79.
- LOSHCHILOV, I.; HUTTER, F. *Decoupled Weight Decay Regularization*. arXiv, 2017. Disponível em: <https://arxiv.org/abs/1711.05101>. Citado 2 vezes nas páginas 90 e 91.
- MA, W. *et al.* Deep-learning-enabled on-demand design of chiral metamaterials. *American Chemical Society Nano*, v. 12, n. 6, p. 6326–6334, 2018. PMID: 29856595. Disponível em: <https://doi.org/10.1021/acsnano.8b03569>. Citado na página 79.
- MA, W. *et al.* Deep learning for the design of photonic structures. *Nature Photonics*, p. 1–14, 2020. Citado na página 79.
- MAK, J. C. C. *et al.* Binary particle swarm optimized 2×2 power splitters in a standard foundry silicon photonic platform. *Optics Letters*, OSA, v. 41, n. 16, p. 3868–3871, Aug 2016. Disponível em: <http://opg.optica.org/ol/abstract.cfm?URI=ol-41-16-3868>. Citado na página 144.
- MALKIEL, I. *et al.* Plasmonic nanostructure design and characterization via deep learning. *Light: Science & Applications*, v. 7, 12 2018. Citado na página 79.
- MALYUTIN, G. Quasi-static analysis of a coplanar strip line with shielding planes used as a distributed sensor. In: *2019 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON)*. [S.l.: s.n.], 2019. p. 0125–0129. Citado na página 68.

MARQUES-HUESO, J. *et al.* Genetic algorithm designed silicon integrated photonic lens operating at 1550 nm. *Applied Physics Letters*, v. 97, p. 071115 – 071115, 09 2010. Citado na página 142.

MASTRONARDI, L. *et al.* High-speed Si/GeSi hetero-structure electro absorption modulator. *Optics Express*, OSA, v. 26, n. 6, p. 6663–6673, Mar 2018. Disponível em: <http://www.opticsexpress.org/abstract.cfm?URI=oe-26-6-6663>. Citado na página 38.

MELIKYAN, A. *et al.* Differential drive i/q modulator based on silicon photonic electro-absorption modulators. *Journal of Lightwave Technology*, v. 38, n. 11, p. 2872–2876, 2020. Citado na página 52.

MENDENHALL, W.; SINCICH, T. *Statistics for Engineering and the Sciences (5th Edition)*. USA: Prentice-Hall, Inc., 2006. ISBN 0131877062. Citado na página 82.

MILLER, D. A. B. Device requirements for optical interconnects to silicon chips. *Proceedings of the IEEE*, v. 97, n. 7, p. 1166–1185, 2009. Citado 2 vezes nas páginas 26 e 38.

MOHRI, M. *et al.* *Foundations of Machine Learning*. 2. ed. Cambridge, MA: MIT Press, 2018. (Adaptive Computation and Machine Learning). ISBN 978-0-262-03940-6. Citado na página 78.

MOSHAEV, V. *et al.* Optimizations of Si PIN diode phase-shifter for controlling MZM quadrature bias point using SoI rib waveguide technology. *Optics & Laser Technology*, v. 138, p. 106844, 2021. ISSN 0030-3992. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0030399220314778>. Citado na página 43.

MOTTA, D. A. *et al.* Design of a 40 GHz bandwidth slow-wave silicon modulator. In: *2017 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC)*. [S.l.: s.n.], 2017. p. 1–5. Citado 8 vezes nas páginas 9, 100, 101, 102, 103, 104, 105 e 121.

MOTTA, D. de A. *Projeto e caracterização de modulador Mach-Zehnder em fotônica de silício*. Tese (Dissertação) — Unicamp, 2018. Citado 2 vezes nas páginas 67 e 139.

NEDELJKOVIC, M. *et al.* Free-carrier electrorefraction and electroabsorption modulation predictions for silicon over the 1–14- μm infrared wavelength range. *IEEE Photonics Journal*, v. 3, p. 1171–1180, 12 2011. Citado na página 60.

NG, A. Y. Feature selection, L_1 vs. L_2 regularization, and rotational invariance. In: *Proceedings of the Twenty-First International Conference on Machine Learning*. New York, NY, USA: Association for Computing Machinery, 2004. (ICML '04), p. 78. ISBN 1581138385. Disponível em: <https://doi.org/10.1145/1015330.1015435>. Citado na página 90.

NGUYEN, C. Analysis methods for RF, microwave, and millimeter-wave planar transmission line structures. IOP Publishing, v. 12, n. 10, p. 1745–1745, sep 2001. Disponível em: <https://doi.org/10.1088/0957-0233/12/10/705>. Citado na página 68.

NORIMATSU, S. *et al.* Psk optical homodyne detection using external cavity laser diodes in Costas loop. *IEEE Photonics Technology Letters*, v. 2, n. 5, p. 374–376, 1990. Citado na página 29.

- OIF. *Optical Internetworking Forum*. 2021. Acessado em: 13/01/2022. Disponível em: <https://www.oiforum.com/>. Citado na página 33.
- OLIVARES, I. *et al.* Enhancing Pockels effect in strained silicon waveguides. *Optics Express*, OSA, v. 27, n. 19, p. 26882–26892, Sep 2019. Disponível em: <http://www.opticsexpress.org/abstract.cfm?URI=oe-27-19-26882>). Citado na página 36.
- PANTOUVAKI, M. *et al.* Active components for 50 Gb/s NRZ-OOK optical interconnects in a silicon photonics platform. *Journal of Lightwave Technology*, v. 35, n. 4, p. 631–638, 2017. Citado na página 43.
- PARK, J. W. *et al.* High-modulation efficiency silicon Mach-Zehnder optical modulator based on carrier depletion in a PN diode. *Optics Express*, OSA, v. 17, n. 18, p. 15520–15524, Aug 2009. Disponível em: <http://www.opticsexpress.org/abstract.cfm?URI=oe-17-18-15520>). Citado na página 43.
- PASZKE, A. *et al.* Pytorch: An imperative style, high-performance deep learning library. In: *Advances in Neural Information Processing Systems 32*. Curran Associates, Inc., 2019. p. 8024–8035. Disponível em: <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>). Citado na página 106.
- PATEL, D. *et al.* Design, analysis, and transmission system performance of a 41 GHz silicon photonic modulator. *Optics Express*, OSA, v. 23, n. 11, p. 14263–14287, Jun 2015. Disponível em: <http://www.opticsexpress.org/abstract.cfm?URI=oe-23-11-14263>). Citado 3 vezes nas páginas 43, 56 e 73.
- PATEL, D. *et al.* High-speed compact silicon photonic Michelson interferometric modulator. *Optics Express*, OSA, v. 22, n. 22, p. 26788–26802, Nov 2014. Disponível em: <http://www.opticsexpress.org/abstract.cfm?URI=oe-22-22-26788>). Citado 5 vezes nas páginas 40, 54, 55, 71 e 74.
- PEARL, J. *Heuristics: Intelligent Search Strategies for Computer Problem Solving*. USA: Addison-Wesley Longman Publishing Co., Inc., 1984. ISBN 0201055945. Citado na página 97.
- PENDIUK, G. F. I. *et al.* Use of a differential evolution algorithm for determining input driving signals in optical frequency combs. *OSA Continuum*, OSA, v. 3, n. 8, p. 2232–2242, Aug 2020. Disponível em: <http://opg.optica.org/osac/abstract.cfm?URI=osac-3-8-2232>). Citado na página 96.
- PEREZ-GALACHO, D. *et al.* Comparison among silicon modulators based on free-carrier plasma dispersion effect. *International Conference on Transparent Optical Networks*, v. 2015, 08 2015. Citado 3 vezes nas páginas 11, 58 e 59.
- PEURIFOY, J. *et al.* Nanophotonic particle simulation and inverse design using artificial neural networks. *Science Advances*, v. 4, n. 6, p. eaar4206, 2018. Disponível em: <https://www.science.org/doi/abs/10.1126/sciadv.aar4206>). Citado na página 79.
- POZAR, D. M. *Microwave engineering; 3rd ed.* Hoboken, NJ: Wiley, 2005. Disponível em: <https://cds.cern.ch/record/882338>). Citado 4 vezes nas páginas 8, 10, 71 e 137.

PRICE, K. *et al.* *Differential Evolution: A Practical Approach to Global Optimization*. Springer Berlin Heidelberg, 2006. (Natural Computing Series). ISBN 9783540313069. Disponível em: <https://books.google.com.br/books?id=hakXI-dEhTkC>. Citado na página 97.

QING, A. *Differential evolution in electromagnetics / Anyong Qing and Ching Kwang Lee*. Berlin ;: Springer-Verlag, 2010. (Evolutionary, learning, and optimization ; v. 4). ISBN 9783642128691. Citado na página 96.

RAHIM, A. *et al.* Taking silicon photonics modulators to a higher performance level: state-of-the-art and a review of new technologies. *Advanced Photonics*, SPIE, v. 3, n. 2, p. 1 – 23, 2021. Disponível em: <https://doi.org/10.1117/1.AP.3.2.024003>. Citado na página 35.

REED, G. *et al.* Silicon optical modulators [review]. *Nature Photonics*, v. 4, p. 518–526, 07 2010. Citado 3 vezes nas páginas 35, 36 e 61.

REED, G. T. *et al.* Silicon photonics: the early years. In: KUBBY, J. A.; JABBOUR, G. E. (Ed.). *Optoelectronic Integration on Silicon II*. Society of Photo-Optical Instrumentation Engineers, 2005. v. 5730, p. 1 – 18. Disponível em: <https://doi.org/10.1117/12.596921>. Citado na página 30.

REED, G. T. *et al.* Recent breakthroughs in carrier depletion based silicon optical modulators. *Nanophotonics*, v. 3, n. 4-5, p. 229–245, 2014. Disponível em: <https://doi.org/10.1515/nanoph-2013-0016>. Citado na página 67.

ROSENBERG, J. C. *et al.* A 25 Gbps silicon microring modulator based on an interleaved junction. *Optics Express*, OSA, v. 20, n. 24, p. 26411–26423, Nov 2012. Disponível em: <http://www.osapublishing.org/oe/abstract.cfm?URI=oe-20-24-26411>. Citado na página 56.

RUDER, S. *An overview of gradient descent optimization algorithms*. arXiv, 2016. Disponível em: <https://arxiv.org/abs/1609.04747>. Citado na página 89.

SAMANI, A. *et al.* Silicon photonic Mach–Zehnder modulator architectures for on chip PAM-4 signal generation. *Journal of Lightwave Technology*, v. 37, n. 13, p. 2989–2999, 2019. Citado na página 43.

SAMUEL, A. L. Some studies in Machine learning using the game of checkers. *IBM Journal on Research and Development*, v. 3, p. 210–229, 1959. Citado na página 78.

SANTURKAR, S. *et al.* *How Does Batch Normalization Help Optimization?* arXiv, 2018. Disponível em: <https://arxiv.org/abs/1805.11604>. Citado na página 91.

SARKER, I. Machine learning: Algorithms, real-world applications and research directions. *Springer Nature Computer Science*, v. 2, 05 2021. Citado na página 78.

SCHMIDHUBER, J. Deep learning in neural networks: An overview. *Neural Networks*, Elsevier BV, v. 61, p. 85–117, Jan 2015. ISSN 0893-6080. Disponível em: <http://dx.doi.org/10.1016/j.neunet.2014.09.003>. Citado na página 86.

SEPEHRIAN, H. *et al.* Silicon photonic IQ modulators for 400 Gb/s and beyond. *Journal of Lightwave Technology*, OSA, v. 37, n. 13, p. 3078–3086, Jul 2019. Disponível em: <http://jlt.osa.org/abstract.cfm?URI=jlt-37-13-3078>. Citado 3 vezes nas páginas 43, 56 e 59.

SEPEHRIAN, H. *et al.* Assessing performance of silicon photonic modulators for pulse amplitude modulation. *IEEE Journal of Selected Topics in Quantum Electronics*, v. 24, n. 6, p. 1–10, 2018. Citado 3 vezes nas páginas 55, 56 e 57.

SHALEV-SHWARTZ, S.; BEN-DAVID, S. *Understanding Machine learning: from theory to algorithms*. Cambridge University Press, 2014. ISBN 9781107057135 1107057132. Disponível em: http://www.worldcat.org/search?qt=worldcat_org_all&q=9781107057135. Citado 3 vezes nas páginas 78, 80 e 85.

SHIMODAIRA, H. Improving predictive inference under covariate shift by weighting the log-likelihood function. *Journal of Statistical Planning and Inference*, v. 90, p. 227–244, 10 2000. Citado na página 91.

SHIRATORI, R. *et al.* Particle swarm optimization of silicon photonic crystal waveguide transition. *Optics Letters*, OSA, v. 46, n. 8, p. 1904–1907, Apr 2021. Disponível em: <http://opg.optica.org/ol/abstract.cfm?URI=ol-46-8-1904>. Citado na página 144.

SHOKOOH-SAREMI, M.; MAGNUSSON, R. Particle swarm optimization and its application to the design of diffraction grating filters. *Optics Letters*, OSA, v. 32, n. 8, p. 894–896, Apr 2007. Disponível em: <http://opg.optica.org/ol/abstract.cfm?URI=ol-32-8-894>. Citado na página 144.

SIEW, S. Y. *et al.* Review of silicon photonics technology and platform development. *Journal of Lightwave Technology*, v. 39, n. 13, p. 4374–4389, 2021. Citado 2 vezes nas páginas 29 e 30.

SIEW, S. Y. *et al.* Review of silicon photonics technology and platform development. *Journal of Lightwave Technology*, OSA, v. 39, n. 13, p. 4374–4389, Jul 2021. Disponível em: <http://opg.optica.org/jlt/abstract.cfm?URI=jlt-39-13-4374>. Citado na página 29.

SILVA, I. N. d. *et al.* *Redes neurais artificiais para engenharia e ciências aplicadas*. [S.l.]: Artliber Editora, 2010. Citado na página 84.

SIMONYAN, K.; ZISSERMAN, A. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. arXiv, 2014. Disponível em: <https://arxiv.org/abs/1409.1556>. Citado na página 88.

SOBU, Y. *et al.* High-speed optical digital-to-analog converter operation of compact two-segment all-silicon Mach–Zehnder modulator. *Journal of Lightwave Technology*, v. 39, n. 4, p. 1148–1154, 2021. Citado na página 62.

SOREF, R.; BENNETT, B. Electrooptical effects in silicon. *IEEE Journal of Quantum Electronics*, v. 23, n. 1, p. 123–129, 1987. Citado 3 vezes nas páginas 36, 37 e 60.

SRIVASTAVA, N. *et al.* Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, JMLR.org, v. 15, n. 1, p. 1929–1958, jan 2014. ISSN 1532-4435. Citado na página 93.

- STAMPOULIDIS, L. *et al.* Cost-effective broadband GaAs IQ modulator array for long-reach OFDM-PONs. In: *39th European Conference and Exhibition on Optical Communication (ECOC 2013)*. [S.l.: s.n.], 2013. p. 1–3. Citado na página 42.
- STATISTA. *Statista: Number of Internet Users in Brazil*. 2022. <<https://www.statista.com/statistics/255208/number-of-internet-users-in-brazil/>>. Acessado em: 10/05/2022. Citado na página 25.
- STORN, R.; PRICE, K. Differential evolution: A simple and efficient adaptive scheme for global optimization over continuous spaces. *Journal of Global Optimization*, v. 23, 01 1995. Citado na página 95.
- STRESHINSKY, M. *et al.* Low power 50 Gb/s silicon traveling wave Mach-Zehnder modulator near 1300 nm. *Optics Express*, OSA, v. 21, n. 25, p. 30350–30357, Dec 2013. Disponível em: <<http://www.opticsexpress.org/abstract.cfm?URI=oe-21-25-30350>>. Citado na página 43.
- STUCCHI, M. *et al.* On-chip optical interconnects versus electrical interconnects for high-performance applications. *Microelectronic Engineering*, v. 112, p. 84 – 91, 2013. ISSN 0167-9317. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167931713002979>>. Citado na página 26.
- SUN, C. *et al.* Deep reinforcement learning for optical systems: A case study of mode-locked lasers. *Machine Learning: Science and Technology*, v. 1, 09 2020. Citado na página 79.
- SZU, H.; HARTLEY, R. Fast simulated annealing. *Physics Letters A*, v. 122, n. 3, p. 157–162, 1987. ISSN 0375-9601. Disponível em: <<https://www.sciencedirect.com/science/article/pii/0375960187907961>>. Citado na página 147.
- TAKENAKA, M.; TAKAGI, S. Strain engineering of plasma dispersion effect for sige optical modulators. *IEEE Journal of Quantum Electronics*, v. 48, n. 1, p. 8–16, 2012. Citado na página 36.
- TAMURA, Y. *et al.* Lowest-ever 0.1419-dB/km loss optical fiber. In: *Optical Fiber Communication Conference Postdeadline Papers*. Optica Publishing Group, 2017. p. Th5D.1. Disponível em: <<http://opg.optica.org/abstract.cfm?URI=OFC-2017-Th5D.1>>. Citado na página 43.
- THOMSON, D. *et al.* Roadmap on silicon photonics. *Journal of Optics*, IOP Publishing, v. 18, n. 7, p. 073003, jun 2016. Disponível em: <<https://doi.org/10.1088/2040-8978/18/7/073003>>. Citado 2 vezes nas páginas 30 e 31.
- THRASKIAS, C. A. *et al.* Survey of photonic and plasmonic interconnect technologies for intra-datacenter and high-performance computing communications. *IEEE Communications Surveys Tutorials*, v. 20, n. 4, p. 2758–2783, 2018. Citado na página 25.
- TSALLIS, C.; STARIOLO, D. A. Generalized simulated annealing. *Physica A: Statistical Mechanics and its Applications*, Elsevier BV, v. 233, n. 1–2, p. 395–406, Nov 1996. ISSN 0378-4371. Disponível em: <[http://dx.doi.org/10.1016/S0378-4371\(96\)00271-3](http://dx.doi.org/10.1016/S0378-4371(96)00271-3)>. Citado 2 vezes nas páginas 146 e 147.

- TU, X. *et al.* Silicon optical modulator with shield coplanar waveguide electrodes. *Optics Express*, OSA, v. 22, n. 19, p. 23724–23731, Sep 2014. Disponível em: <http://www.osapublishing.org/oe/abstract.cfm?URI=oe-22-19-23724>. Citado na página 68.
- TU, X. *et al.* Fabrication of low loss and high speed silicon optical modulator using doping compensation method. *Optics Express*, OSA, v. 19, n. 19, p. 18029–18035, Sep 2011. Disponível em: <http://www.osapublishing.org/oe/abstract.cfm?URI=oe-19-19-18029>. Citado na página 56.
- VARDHAN, H. *et al.* Wireless data center with millimeter wave network. In: *2010 IEEE Global Telecommunications Conference GLOBECOM 2010*. [S.l.: s.n.], 2010. p. 1–6. Citado na página 26.
- VILLARRUBIA, G. *et al.* Artificial neural networks used in optimization problems. *Neurocomputing*, v. 272, p. 10–16, 2018. ISSN 0925-2312. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0925231217311116>. Citado na página 98.
- WANG, C. *et al.* Nanophotonic lithium niobate electro-optic modulators. *Optics Express*, Optica Publishing Group, v. 26, n. 2, p. 1547–1555, Jan 2018. Disponível em: <http://opg.optica.org/oe/abstract.cfm?URI=oe-26-2-1547>. Citado na página 42.
- WATTS, M. R. *et al.* Low-voltage, compact, depletion-mode, silicon Mach–Zehnder modulator. *IEEE Journal of Selected Topics in Quantum Electronics*, v. 16, n. 1, p. 159–164, 2010. Citado na página 43.
- WENG, Q. *et al.* The propagation characteristics of rectangular waveguides filled with inhomogeneous double-negative dielectrics using a semianalytical method. *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, v. 34, n. 4, p. e2860, 2021. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jnm.2860>. Citado 2 vezes nas páginas 32 e 42.
- WILLIAMS, D. Traveling waves and power waves: Building a solid foundation for microwave circuit theory. *IEEE Microwave Magazine*, v. 14, n. 7, p. 38–45, 2013. Citado na página 139.
- XIANG, Y. *et al.* Generalized simulated annealing for global optimization: The gensa package. *The R Journal Volume 5(1):13-29, June 2013*, v. 5, 06 2013. Citado na página 147.
- XIE, Y. *et al.* Design of an arbitrary ratio optical power splitter based on a discrete differential multiobjective evolutionary algorithm. *Applied Optics*, OSA, v. 59, n. 6, p. 1780–1785, Feb 2020. Disponível em: <http://opg.optica.org/ao/abstract.cfm?URI=ao-59-6-1780>. Citado na página 96.
- XU, H. *et al.* High-speed silicon modulator with band equalization. *Optics Letters*, OSA, v. 39, n. 16, p. 4839–4842, Aug 2014. Disponível em: <http://ol.osa.org/abstract.cfm?URI=ol-39-16-4839>. Citado na página 43.
- XUAN, Z. *et al.* Silicon microring modulator for 40 Gb/s NRZ-OOK metro networks in O-band. *Optics Express*, OSA, v. 22, n. 23, p. 28284–28291, Nov 2014. Disponível em: <http://www.opticsexpress.org/abstract.cfm?URI=oe-22-23-28284>. Citado na página 39.

YANG, T. liang *et al.* Feedback analysis of transcutaneous energy transmission with a variable load parameter. *ETRI Journal*, v. 32, 2010. Citado na página 71.

YOU, K. *et al.* *How Does Learning Rate Decay Help Modern Neural Networks?* 2019. Disponível em: [⟨https://arxiv.org/abs/1908.01878⟩](https://arxiv.org/abs/1908.01878). Citado na página 108.

YU, H.; BOGAERTS, W. An equivalent circuit model of the traveling wave electrode for carrier-depletion-based silicon optical modulators. *Journal of Lightwave Technology*, v. 30, n. 11, p. 1602–1609, 2012. Citado 2 vezes nas páginas 73 e 75.

ZHANG, F. *et al.* High baud rate transmission with silicon photonic modulators. *IEEE Journal of Selected Topics in Quantum Electronics*, v. 27, n. 3, p. 1–9, 2021. Citado na página 43.

ZHANG, Y. *et al.* 200 Gbit/s optical PAM4 modulation based on silicon microring modulator. In: *2020 European Conference on Optical Communications (ECOC)*. [S.l.: s.n.], 2020. p. 1–4. Citado na página 39.

ZHENG, L. *et al.* Intrinsic chirp analysis of a single dual-drive silicon PAM-4 optical modulator. *Optics Express*, OSA, v. 26, n. 24, p. 32014–32025, Nov 2018. Disponível em: [⟨http://www.opticsexpress.org/abstract.cfm?URI=oe-26-24-32014⟩](http://www.opticsexpress.org/abstract.cfm?URI=oe-26-24-32014). Citado na página 43.

ZHOU, J. *et al.* Silicon photonics carrier depletion modulators capable of 85 GBaud 16QAM and 64 GBaud 64QAM. In: *Optical Fiber Communication Conference (OFC) 2019*. Optical Society of America, 2019. p. Tu2H.2. Disponível em: [⟨http://www.osapublishing.org/abstract.cfm?URI=OFC-2019-Tu2H.2⟩](http://www.osapublishing.org/abstract.cfm?URI=OFC-2019-Tu2H.2). Citado na página 43.

ZHOU, J. *et al.* High baud rate all-silicon photonics carrier depletion modulators. *Journal of Lightwave Technology*, v. 38, n. 2, p. 272–281, 2020. Citado 2 vezes nas páginas 43 e 105.

ZHOU, X. *et al.* Datacenter optics: requirements, technologies, and trends (invited paper). *Chinese Optics Letters*, OSA, v. 15, n. 5, p. 120008, May 2017. Disponível em: [⟨http://col.osa.org/abstract.cfm?URI=col-15-5-120008⟩](http://col.osa.org/abstract.cfm?URI=col-15-5-120008). Citado na página 25.

ZHOU, Y. *et al.* Modeling and optimization of a single-drive push-pull silicon Mach-Zehnder modulator. *Photonics Research*, OSA, v. 4, n. 4, p. 153–161, Aug 2016. Disponível em: [⟨http://www.osapublishing.org/prj/abstract.cfm?URI=prj-4-4-153⟩](http://www.osapublishing.org/prj/abstract.cfm?URI=prj-4-4-153). Citado 3 vezes nas páginas 43, 65 e 66.

Apêndices

Apêndice A – Análise de Parâmetros Distribuídos

Como no desenvolvimento do MZM utilizamos eletrodos viajantes baseados em linhas de transmissão, discutidos na Subseção 2.4, é válido discutirmos como se dá a análise de parâmetros distribuídos. Esta seção é baseada na análise realizada em (POZAR, 2005). A Figura A mostra uma seção infinitesimal de uma linha de transmissão. Os parâmetros indicados no esquemático são parâmetros distribuídos. Os parâmetros R e G representam perdas de propagação no meio, R representando a resistência devido a condutância finita dos condutores e G é devido às perdas dielétricas que ocorrem no material interface entre ambos condutores. Por outro lado, a indutância L é a autoindutância causada pelo par de condutores enquanto a capacitância ocorre devido a distância próxima entre os condutores. Dados os parâmetros e a tensão e corrente de entrada para este segmento, podemos aplicar a lei de Kirchoff como:

$$v(z, t) - R\Delta z i(z, t) - L\Delta z \frac{\partial i(z, t)}{\partial t} - v(z + \Delta z, t) = 0, \quad (110)$$

$$i(z, t) - G\Delta z v(z + \Delta z, t) - C\Delta z \frac{\partial v(z + \Delta z, t)}{\partial t} - i(z + \Delta z, t) = 0. \quad (111)$$

Dividindo as Equações 110 e 111 por Δz , tal que $\Delta z \rightarrow 0$, obtemos as seguintes equações diferenciais denominadas equações do telégrafo, tais quais são dadas por:

$$\frac{\partial v(z, t)}{\partial z} = -Ri(z, t) - L \frac{\partial i(z, t)}{\partial t}, \quad (112)$$

$$\frac{\partial i(z, t)}{\partial z} = -Gv(z, t) - C \frac{\partial v(z, t)}{\partial t}. \quad (113)$$

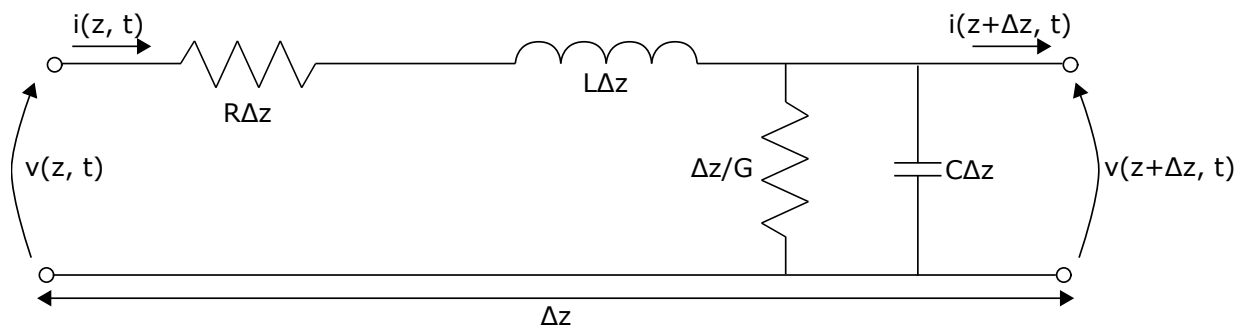


Figura 35 – Seção infinitesimal de uma linha de transmissão, com a tensão e a corrente propagante em função do tempo e os parâmetros distribuídos ilustrados. Figura adaptada de (POZAR, 2005).

No domínio da frequência e utilizando fasores com comportamento cossenoidal, simplificamos as Equações 112 e 113 para as seguintes formas:

$$\frac{dV(z)}{dz} = -(R + j\omega L)I(z), \quad (114)$$

$$\frac{dI(z)}{dz} = -(G + j\omega C)V(z), \quad (115)$$

e resolvendo simultaneamente ambas equações, temos que:

$$\frac{d^2V(z)}{dz^2} - \gamma^2V(z) = 0, \quad (116)$$

$$\frac{d^2I(z)}{dz^2} - \gamma^2I(z) = 0, \quad (117)$$

$$(118)$$

dado que:

$$\gamma = \alpha + j\beta = \sqrt{(R + j\omega L)(G + j\omega C)}. \quad (119)$$

As soluções para as Equações 116 e 117 podem ser dadas por:

$$V(z) = V_0^+ \exp(-\gamma z) + V_0^- \exp(\gamma z), \quad (120)$$

$$I(z) = I_0^+ \exp(-\gamma z) + I_0^- \exp(\gamma z). \quad (121)$$

A impedância característica indica a razão entre a tensão se movendo na direção positiva e a corrente se movendo na direção negativa, e é dada por:

$$Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}}. \quad (122)$$

É importante observar que a velocidade de fase da onda de RF no meio é dada por:

$$v_{p-wg} = \frac{\omega}{\beta} = \lambda_{RF}f, \quad (123)$$

em que λ_{RF} é o comprimento de onda de RF. Podemos definir o coeficiente de reflexão a partir da impedância da linha e as ondas de tensão em ambas direções.

$$\Gamma(z) = \frac{V^-}{V^+} = \frac{Z(z) - Z_0}{Z(z) + Z_0}, \quad (124)$$

em que $Z(z)$ é a impedância no ponto z da linha.

A.1 Parâmetros S e de Transmissão

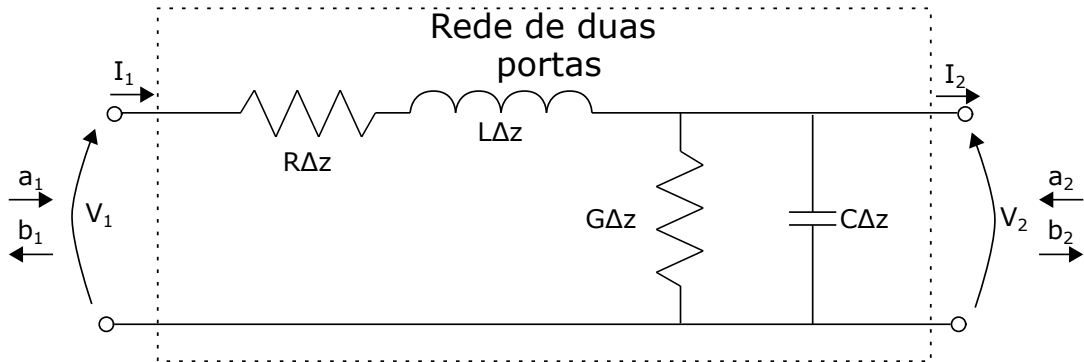


Figura 36 – Linha de transmissão representada como uma rede de duas portas.

Podemos analisar um segmento de uma linha de transmissão como uma rede de duas portas. A matriz de espalhamento nos permite relacionar as ondas que entram e as que saem, tanto como transmissão quanto como reflexões. Primeiramente, podemos definir os parâmetros a_1 , b_1 , a_2 e b_2 em função das tensões e correntes de entrada (WILLIAMS, 2013; MOTTA, 2018):

$$a_1 = \frac{V_1 + Z_{ref} I_1}{2\sqrt{Z_{ref}}}, \quad (125)$$

$$b_1 = \frac{V_1 - Z_{ref} I_1}{2\sqrt{Z_{ref}}}, \quad (126)$$

$$a_2 = \frac{V_2 - Z_{ref} I_2}{2\sqrt{Z_{ref}}}, \quad (127)$$

$$b_2 = \frac{V_2 + Z_{ref} I_2}{2\sqrt{Z_{ref}}}, \quad (128)$$

em que Z_{ref} é uma impedância de referência, geralmente definida como $Z_{ref} = 50 \Omega$. Podemos então definir a relação entre estes termos a partir da matriz de espalhamento:

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \quad (129)$$

em que S_{21} e S_{12} representam, respectivamente, o ganho de transmissão direto e o inverso, enquanto S_{11} e S_{22} representam o coeficiente de reflexão na porta 1 e na porta 2, respecti-

vamente. Para representarmos a concatenação de duas ou mais redes, utiliza-se a matriz de parâmetros transmissão (matriz ABCD). Podemos definir a matriz de transmissão como:

$$\begin{bmatrix} V_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V_1 \\ -I_2 \end{bmatrix}. \quad (130)$$

Podemos definir os valores dos termos da matriz de transmissão, assim como relacionarmos com a matriz de espalhamento, a partir das seguintes equações:

$$A = \cosh(\gamma l) = \frac{(Z_{ref1}^* + S_{11}Z_{ref1})(1 - S_{22}) + S_{12}S_{21}Z_{ref1}}{2S_{21}\sqrt{R_{ref1}R_{ref2}}}, \quad (131)$$

$$B = Z_0 \sinh(\gamma l) = \frac{(Z_{ref1}^* + S_{11}Z_{ref1})(Z_{ref2}^* + S_{22}Z_{ref2}) - S_{12}S_{21}Z_{ref1}Z_{ref2}}{2S_{21}\sqrt{R_{ref1}R_{ref2}}}, \quad (132)$$

$$C = \frac{1}{Z_0} \sinh(\gamma l) = \frac{(1 - S_{11})(1 - S_{22}) - S_{12}S_{21}}{2S_{21}\sqrt{R_{ref1}R_{ref2}}}, \quad (133)$$

$$D = \cosh(\gamma l) = \frac{(1 - S_{11})(Z_{ref2}^* + S_{22}Z_{ref2}) + S_{12}S_{21}Z_{ref2}}{2S_{21}\sqrt{R_{ref1}R_{ref2}}}, \quad (134)$$

em que Z_{refi} e R_{refi} são, respectivamente, a impedância e a resistência de referência da i -ésima porta, geralmente com valores de 50Ω e l é a distância de propagação considerada. A partir da Equação 131 podemos definir a constante de propagação como:

$$\gamma = \frac{\cosh^{-1}(A)}{l} = \frac{1}{l} \ln \left(\frac{A + D}{2} + \sqrt{BC} \right), \quad (135)$$

e a impedância característica como:

$$Z_0 = \sqrt{\frac{B}{C}}. \quad (136)$$

Podemos definir o índice efetivo no guia RF como:

$$n_{eff-mw} = \frac{\Im(\gamma)c}{\omega}, \quad (137)$$

sendo que $\Im\gamma$ se refere à componente imaginária de γ . As perdas em Np/cm podem ser dadas por:

$$\alpha[\text{Np/m}] = |\Re(\gamma) \cdot 10^{-3} \cdot 20 \cdot \log_{10}(e)|, \quad (138)$$

em que $\Re(\gamma)$ indica a componente real de γ . Obtido os valores de constante de propagação complexa, impedância característica, índice efetivo e as perdas, podemos calcular os parâmetros da linha da transmissão dados por:

$$R = \Re(\gamma \cdot Z_0), \quad (139)$$

$$L = \frac{\Im(\gamma \cdot Z_0)}{\omega}, \quad (140)$$

$$G = \Re\left(\frac{\gamma}{Z_0}\right), \quad (141)$$

$$C = \frac{1}{\omega} \Im\left(\frac{\gamma}{Z_0}\right). \quad (142)$$

Apêndice B – Outros Algoritmos Heurísticos

No presente Anexo iremos discutir outros algoritmos heurísticos que foram testados para a otimização de *design* do MZM.

B.1 Funcionamento dos Algoritmos Heurísticos

Na presente seção iremos discutir os princípios dos algoritmos GA, PSO e DA.

B.1.1 Algoritmo Genético

Baseado na teoria da evolução, GA é um algoritmo heurístico que busca a melhor solução para dado problema (HOLLAND, 1975). De fato, seu funcionamento foi baseado na teoria de Darwin sobre a seleção natural (DILAO, 2009). Como sua otimização não depende do cálculo de GD, a função a ser otimizada pelo GA não necessita ser diferenciável. GA é comum no solucionamento de problemas com alta dimensionalidade, e ainda pode ser aplicado em problemas multi-objetivos de valores discretos ou contínuos (CHAHAR; KATOCH; CHAUHAN, 2021). Algumas aplicações em fotônica incluem o *design* de uma fibra de cristal fotônico (KERRINCKX *et al.*, 2004), um taper de cristal fotônico (HAKANSSON *et al.*, 2005) e uma lente fotônica integrada (MARQUES-HUESO *et al.*, 2010).

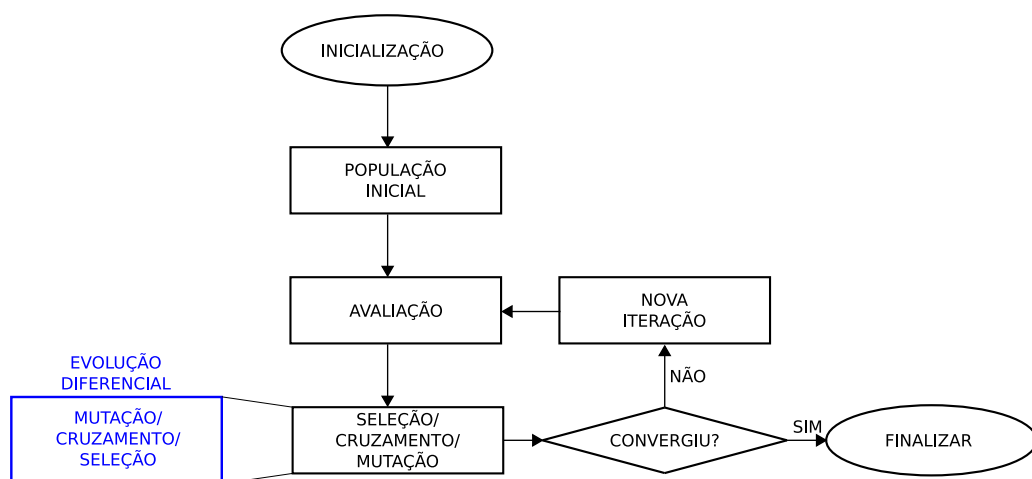


Figura 37 – Diagrama de fluxo do algoritmo genético. A caixa em azul se refere ao algoritmo de evolução diferencial.

A Figura 37 ilustra o diagrama de fluxo do algoritmo genético. Primeiramente, é definida uma população inicial. A população é formada de indivíduos (cromossomos) que possuem atributos (genes) aleatórios. Uma função de avaliação (*fitness*) deve ser definida como forma de mensurar o desempenho de cada indivíduo (FOGEL, 1995). Escolhe-se os indivíduos mais aptos, ou seja, os que apresentam maior *fitness*. A partir deste subconjunto selecionado, recombina-se estes cromossomos mais aptos para formar novos cromossomos (filhos) para a próxima geração, esta fase é denominada como cruzamento. Para aumentar a variedade da população e evitar mínimos locais, aplica-se uma mutação aleatória em alguns parâmetros dos cromossomos filhos. Os novos filhos substituem os cromossomos menos aptos da população da geração anterior. Este procedimento é realizado de forma iterativa até uma condição de convergência for obedecida.

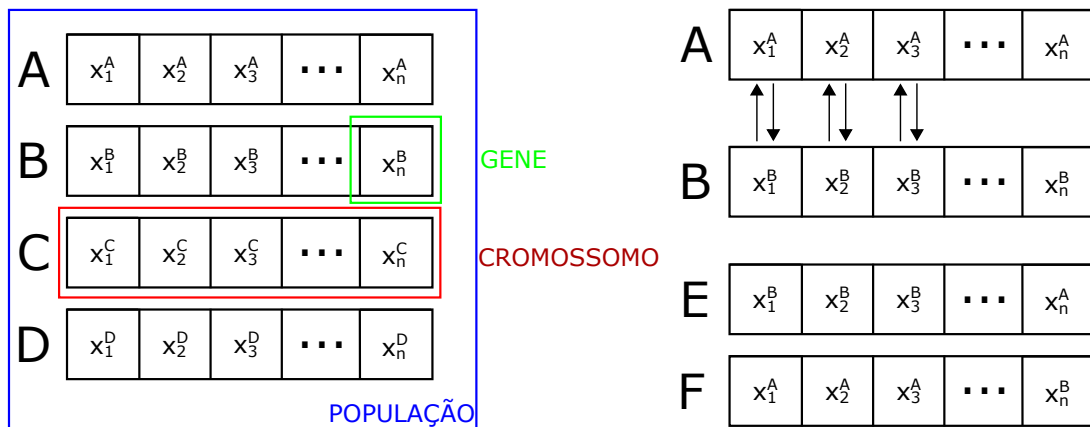


Figura 38 – Esquemático do cruzamento entre dois cromossomos.

A Figura 38 ilustra uma população inicial com quatro cromossomos, cada um com n genes. Supondo que após o cálculo de avaliação temos que os cromossomos **A** e **B** são os mais aptos, estes serão selecionados para o cruzamento. No cruzamento, alguns genes são trocados, como demonstrado no lado direito da figura. O cruzamento aqui considera a troca dos três primeiros genes entre ambos cromossomos. Após o cruzamento, há a geração de dois novos cromossomos, indicados pelas letras **E** e **F**.

A função de avaliação deve ser definida de forma a considerar o objetivo da otimização. Esta deve sempre representar uma métrica de cada cromossomo como um escalar. Já sobre o cruzamento, temos que um dos mais utilizados na literatura e que apresenta um bom funcionamento é o cruzamento de um ponto. Um gene aleatório de dois pais é selecionado aleatoriamente, tal qual será o ponto de cruzamento. Então os genes à direita ou à esquerda deste gene serão trocados entre os cromossomos pais para formar

um novo indivíduo. Dada a escolha dos pais mais aptos, definimos a taxa de cruzamento como sendo a probabilidade de ocorrer cruzamento entre dois pais. No caso da mutação temos também a taxa de mutação, que deve ser mantida abaixo da taxa de cruzamento. Uma forma comum de mutação para problemas contínuos consiste em somar um número aleatório, seguindo uma distribuição uniforme ou normal, dentro de uma faixa de valor para um ou mais genes. Caso o valor de um gene ultrapasse os valores permitidos, limitamos seu valor para o mantermos dentro da faixa desejada.

B.1.2 Otimização por Enxame de Partículas

Assim como o GA, o algoritmo de otimização por enxame de partículas (PSO, *Particle Swarm Optimization*) é baseado em processos biológicos. Entretanto, PSO se baseia no comportamento que uma revoada de pássaros realiza ao se deslocar (*flocking*) pelos céus (KENNEDY; EBERHART, 1995). Proposto em 1995 por Kennedy e Eberhart, PSO propõe a otimização de uma solução, de forma iterativa, levando em consideração uma função de avaliação. Requerendo poucos parâmetros para o programador definir, PSO apresenta uma implementação mais simples que algoritmos evolucionários, porém este ainda é capaz de solucionar problemas não lineares com eficiência (JAIN; NANGIA; JAIN, 2018). Outra vantagem em comum com algoritmos evolucionários é que o PSO não requer que a função de avaliação seja diferenciável. Na área de fotônica, PSO já se mostrou útil para otimização de estruturas de múltiplos parâmetros, como um filtro baseado em grades de Bragg (SHOKOOH-SAREMI; MAGNUSSON, 2007), *tapers* (SHIRATORI *et al.*, 2021) e divisores de potência (MAK *et al.*, 2016) baseados em cristais fotônicos.

A Figura 39 ilustra o diagrama de blocos do algoritmo PSO. A inicialização do algoritmo se dá de forma estocástica, em que cada candidato é denominado como *partícula*. As partículas se deslocam com uma certa velocidade, ou seja, na iteração $t + 1$, a i -ésima partícula se desloca da seguinte forma:

$$\mathbf{S}_{t+1}^{(i)} = \mathbf{S}_t^{(i)} + \mathbf{V}_{t+1}^{(i)}, \quad (143)$$

em que $\mathbf{S}_t^{(i)}$ é a posição da i -ésima partícula na iteração t e $\mathbf{V}_{t+1}^{(i)}$ é a velocidade atualizada. A velocidade de cada partícula é inicializada como um número aleatório que segue uma distribuição uniforme entre -1 e 1. No PSO, a velocidade e a direção de deslocamento da

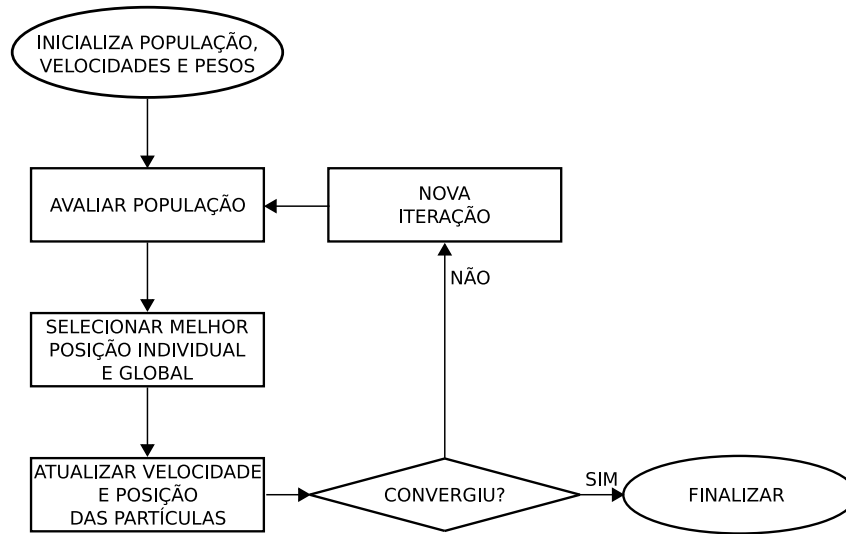


Figura 39 – Diagrama de fluxo do algoritmo de otimização por enxames de partículas.

i -ésima partícula depende de três fatores: (*i*) seu momento acumulado, dada a iteração anterior, (*ii*) de sua posição que apresentou o melhor resultado, e da (*iii*) direção da partícula que possui a melhor avaliação entre o enxame. Em outras palavras, cada partícula tenta encontrar a solução baseando-se em sua intuição (de forma cognitiva), mas também baseiam-se em um fator social. A atualização da velocidade da i -ésima partícula, para a iteração $t + 1$, é dada então por:

$$\mathbf{V}_{t+1}^{(i)} = W\mathbf{V}_t^{(i)} + c_1r_1(\mathbf{S}_b^{(i)} - \mathbf{S}_t^{(i)}) + c_2r_2(\mathbf{S}_b^{(k)} - \mathbf{S}_t^{(i)}). \quad (144)$$

Na Equação 144, o parâmetro W define o peso do momento da partícula i , e c_1 e c_2 são as constantes cognitivas e sociais, respectivamente. As constantes c_1 e c_2 ponderam a direção da otimização entre a melhor alcançada pelo indivíduo e a melhor alcançada pelo enxame em geral. Por outro lado, temos dois pesos, r_1 e r_2 , que são dados seguindo uma distribuição uniforme entre 0 e 1. A posição com a melhor avaliação da partícula i é dada pelo vetor $\mathbf{S}_b^{(i)}$ e a posição da partícula que apresenta o melhor resultado entre o enxame é dada por $\mathbf{S}_b^{(k)}$. De fato, as subtrações vetoriais da Equação 144 influenciam na direção da partícula, de forma a levar esta seguir a melhor posição obtida pela mesma ($\mathbf{S}_b^{(i)} - \mathbf{S}_t^{(i)}$) e também considerando todo o enxame ($\mathbf{S}_b^{(k)} - \mathbf{S}_t^{(i)}$). O processo é realizado iteração por iteração, até que a melhor posição possível seja encontrada. No trabalho realizado aqui, definimos $w = 0,5$, $c_1 = 1$ e $c_2 = 2$, de forma a priorizar a direção que aponta para melhor posição encontrada no enxame.

B.1.3 Dual Annealing

O algoritmo de recozimento duplo (DA, *Dual Annealing*) ou algoritmo de recozimento generalizado (GSA, *Generalized Simulated Annealing*) é um algoritmo heurístico de otimização baseado em busca aleatória (TSALLIS; STARIOLO, 1996). Sua operação se dá combinando o algoritmo de recozimento clássico (CSA, *Classic Simulated Annealing*), que realiza uma busca com maior foco global, com o algoritmo de recozimento rápido (FSA, *Fast Simulated Annealing*), que realiza uma busca com maior foco local.

O CSA é um algoritmo baseado no processo de recozimento de metais (KIRKPATRICK; GELATT; VECCHI, 1983). Este tratamento possibilita a alteração de propriedades físicas dos metais a partir da termodinâmica. O recozimento concede uma maior capacidade de deformação plástica e dureza nos metais, deixando-o manipulável para metalurgia. Se inicia o processo elevando a temperatura do metal acima da temperatura de recristalização e abaixo da temperatura de fusão. Esta temperatura é mantida por uma certa quantia de tempo e lentamente reduzida, causando o desenvolvimento de novos grãos no metal e o deixando mais maleável a processos de manufatura.

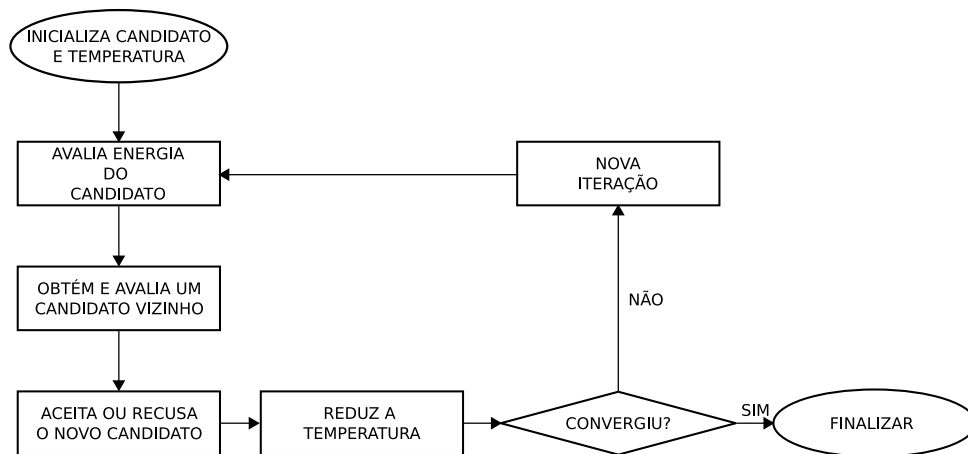


Figura 40 – Diagrama de fluxo do algoritmo de DA.

A Figura 40 mostra o diagrama de blocos para o algoritmo de DA. O CSA tem, na sua inicialização, apenas um único candidato $\mathbf{S}(t = 0)$, este apresentando parâmetros aleatórios. Definimos também um valor inicial, relativamente alto, de temperatura. A função objetiva nesse cenário é denominada energia de estado do candidato. A cada iteração t , uma solução parecida (vizinha) $\mathbf{S}'(t)$ é gerada. A energia do candidato atual é comparado com a de seu vizinho. Se a energia do vizinho é menor, o algoritmo aceita este

como solução. Entretanto, se este apresenta uma energia maior, há uma probabilidade de aceitação, esta sendo proporcional ao valor da temperatura. O valor da probabilidade desejado no fim da iteração é de um, e para isto a temperatura diminui logarithmicamente com as iterações (SZU; HARTLEY, 1987).

A distribuição de perturbação nos parâmetros do candidato atual, para geração do vizinho, é denominada distribuição de visitação (DV). No caso do CSA, assume-se que a DV é uma distribuição gaussiana. Por outro lado, no caso do FSA, a DV segue uma distribuição de Cauchy Lorentz, a qual realiza uma busca local com pouca probabilidade de explorar pontos distantes. No DA, uma DV que combina as distribuições do caso do CSA e do FSA é apresentada (XIANG *et al.*, 2013). A DV é dada então por:

$$g(\Delta\mathbf{S}(t)) \propto \frac{[T(t)]^{\left(-\frac{D}{3-q_v}\right)}}{\left[1 + (q_v - 1) \frac{(\Delta\mathbf{S}(t))^2}{[T(t)]^{\left(\frac{2}{3-q_v}\right)}}\right]^{\left(\frac{1}{q_v-1} + \frac{D-1}{2}\right)}}, \quad (145)$$

em que $T(t)$ é a temperatura na iteração t , D é a quantidade de dimensões e q_v é o parâmetro de visitação, tal qual controla o formato da DV, este mantido entre 1 e 3. Na Equação 145, o termo $\Delta\mathbf{S}$ é denominado salto de teste da variável \mathbf{S} , o qual definirá o candidato vizinho, este gerado pela DV. Como mencionado, o vizinho será aceito se sua energia for menor que do candidato atual. No caso que esta é maior, a aceitação do vizinho como solução dependerá da probabilidade de aceitação, dada por:

$$p_{q_a} = \min \{1, [1 - (1 - q_a)\beta\Delta E]^{\left(\frac{1}{1-q_a}\right)}\}, \quad (146)$$

em que β é o parâmetro de Lagrange, dado por $\beta \equiv 1/kT$ e k é uma constante positiva (TSALLIS; STARIOLO, 1996). O parâmetro de aceitação q_a controla a chance de aceitação do vizinho como solução, e ΔE é a diferença de energia entre o candidato atual e o vizinho. A temperatura tem seu valor reduzido seguindo a equação:

$$T(t+1) = T(t) \frac{2^{(q_v-1)} - 1}{(1+t)^{(q_v-1)} - 1}. \quad (147)$$

Podemos perceber pela Equação 147 que, quando $q_v = 1$ temos o CSA como equivalente, e quando $q_v = 2$ temos o FSA. Para $q_v > 2$, é possível provar que a otimização é mais rápida que ambos CSA e FSA, pois os saltos são maiores e a redução de T é mais rápida (XIANG *et al.*, 2013). O algoritmo é executado até uma iteração definida ou até atingir uma condição de convergência.

B.2 Metodologia dos Algoritmos Heurísticos

Assim como no caso do DE, realizamos a otimização conjunta dos outros algoritmos heurísticos, assim como ilustrado pela Figura 22. Na presente seção iremos discutir os parâmetros e configurações utilizados pelos algoritmos heurísticos de otimização.

Em relação ao GA, utilizamos uma população inicial de 10.000 MZMs com parâmetros aleatórios. O quantidade de pais selecionados para cruzamento é de 100 por iteração, e o algoritmo foi executado durante 50 iterações. A taxa de cruzamento foi de 0,7, e a taxa de mutação é de 1/8, ou seja, em média um dos parâmetros de cada MZM descendente será modificado. No caso do PSO utilizamos uma populaçã de 2.000 MZM candidatos, e o algoritmo foi executado durante 50 iterações. O momento de inércia foi configurado para ser de 0,5. O coeficiente social e cognitivo são de, respectivamente, 1 e 2. Já para o DA, definimos a temperatura inicial como 5200. O parâmetro de visitação e de aceitação são de 2,62 e -5, respectivamente. O DA também foi executado durante 50 iterações.

B.3 Resultados dos Algoritmos Heurísticos

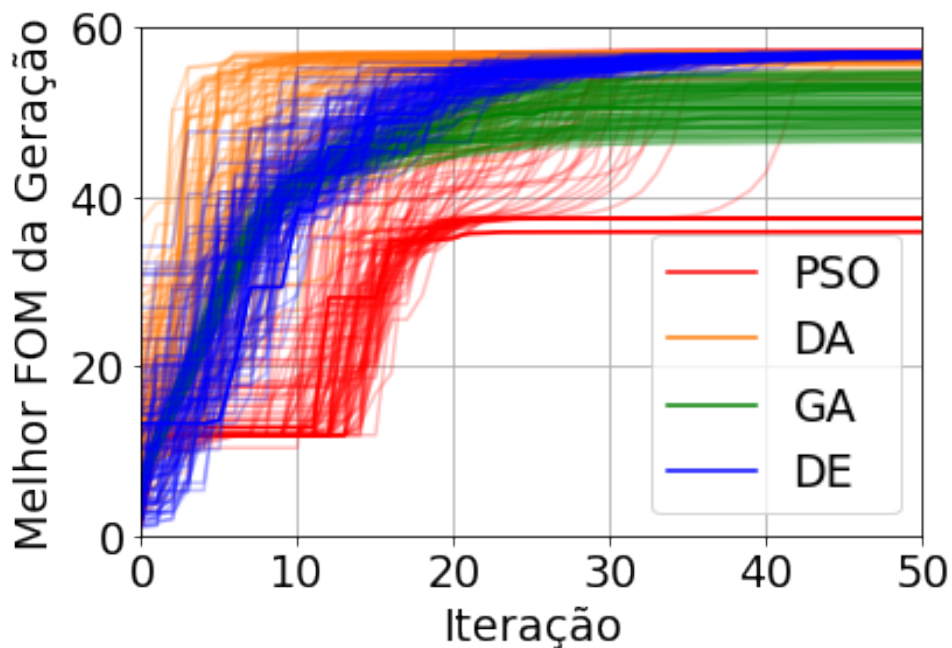


Figura 41 – Curva de convergência da otimização pelos algoritmos heurísticos, considerando 100 inicializações aleatórias para cada um destes.

Tabela 8 – Média, desvio padrão e valor máximo para a FOM obtida pelos algoritmos heurísticos, considerando 100 inicializações aleatórias. A tabela também mostra os valores mínimos e máximos obtidos para cada um dos parâmetros de métrica do MZM.

Algoritmos	DE	GA	PSO	DA
Média	56,86	51,06	48,4	56,5
Desvio padrão	0,17	2,31	9,8	0,53
Valor máximo	57,09	54,86	57,15	57,11
BW_{EO} [GHz] min/max	37,5/38,9	30/44,8	30,9/40,1	32,2/39,3
IL [dB] min/max	3,8/3,9	2,9/4,4	3,8/4,65	3,8/4,3
V_{π} [V] min/max	5,9/6,2	5,1/7,6	4,8/8,1	5,1/6,4

A Figura 41 mostra as curvas de convergência para os algoritmos DE, GA, DA e PSO na otimização da FOM da terceira linha da Tabela 5, considerando 100 inicializações aleatórias. Podemos perceber que, em média, o DE apresenta uma FOM superior aos outros algoritmos, assim como uma variância menor.

A Tabela 8 mostra os valores de média, desvio padrão e valor máximo de FOM obtida pelos algoritmos heurísticos para as 100 execuções. Podemos perceber que o DE apresenta a maior média e o menor desvio padrão entre estes. O PSO, por outro lado, apresenta a menor média e o maior desvio padrão, entretanto, este resultado mostra uma vantagem. Devido à alta variância, uma grande diversidade de MZMs podem ser obtidos a partir de uma única FOM. Além do mais, o maior FOM máximo foi obtido a partir do PSO. Vale destacar também que o DA apresentou bons resultados, com valores próximos ao obtido pelo DE e com uma convergência relativamente rápida, como pode ser visto na Figura 41.

A Subfigura 42(a) mostra as métricas de desempenho dos MZMs com *design* otimizado pelos algoritmos heurístico, em relação aos dados de treinamento da ANN. As Subfiguras 42(b), (c) e (d) mostram os dados obtidos pelo GA, PSO e DA, respectivamente. Podemos perceber que o GA e o DA encontram diversas soluções espalhadas pelo espaço do FOM, com maior efeito sobre o GA. Enquanto isso, o PSO encontrou basicamente 7 regiões de máximos locais, entretanto estas estão relativamente distantes uma da outra. A Tabela 8 lista os valores mínimos e máximos para os parâmetros de desempenho do MZM, encontrados pelos algoritmos de otimização. No corpo principal do texto utilizamos o DE para realizarmos a análise da otimização do MZM em conjunto com a ANN. Esta escolha é justificada pela baixa variância entre o desempenho dos MZMs encontrados pelo algoritmo, e ainda que estes apresentam um alto valor de FOM.

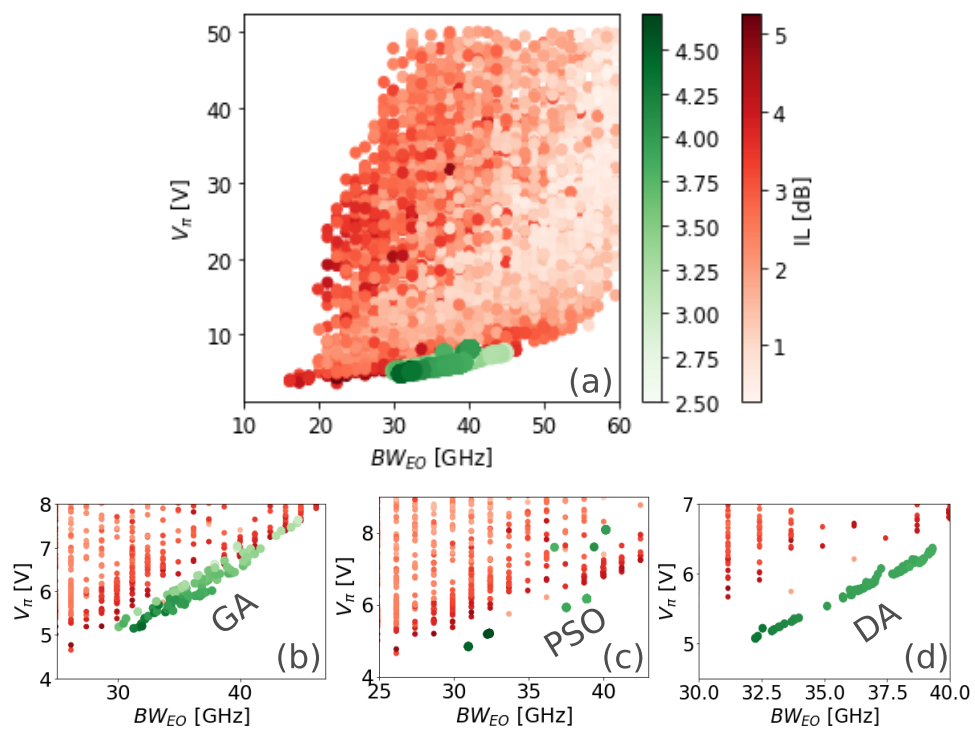


Figura 42 – Métricas de desempenho dos MZMs obtidos pela otimização via algoritmos heurísticos para 100 inicializações aleatórias. (a) Conjunto de dados do treinamento da ANN e os MZMs otimizados pelo GA, PSO e DA. (b) Resultados da otimização por GA, por (b) PSO e por (d) DA.