

UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
CAMPUS DE SÃO JOÃO DA BOA VISTA

LUCAS MARIM DA SILVA

Redes neurais artificiais para compensação de distorção não linear em sistemas de comunicações ópticas coerentes

São João da Boa Vista

2019

Lucas Marim da Silva

Redes neurais artificiais para compensação de distorção não linear em sistemas de comunicações ópticas coerentes

Trabalho de Graduação apresentado ao Conselho de Curso de Graduação em Engenharia de Telecomunicações do Campus de São João da Boa Vista, Universidade Estadual Paulista, como parte dos requisitos para obtenção do diploma de Graduação em Engenharia de Telecomunicações .

Orientador: Prof^o Dr. Ivan Aritz Aldaya Garde

São João da Boa Vista

2019

Silva, Lucas Marim da

Redes neurais artificiais para compensação de distorção não linear em sistemas de comunicações ópticas coerentes / Lucas Marim da Silva -- São João da Boa Vista, 2019.
40 p. : il. color.

Trabalho de Conclusão de Curso – Câmpus Experimental de São João da Boa Vista – Universidade Estadual Paulista “Júlio de Mesquita Filho”.

Orientador: Prof. Dr. Ivan Aritz Aldaya Garde

Bibliografia

1. Comunicações ópticas 2. Efeito de Kerr 3. Redes neurais (Computação) 4. Telecomunicações

CDD 23. ed. – 621.382

Ficha catalográfica elaborada pela [Biblioteca-BJB](#)

Bibliotecário responsável: João Pedro Alves Cardoso – CRB-8/9717

UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
CÂMPUS EXPERIMENTAL DE SÃO JOÃO DA BOA VISTA
GRADUAÇÃO EM ENGENHARIA ELETRÔNICA E DE TELECOMUNICAÇÕES

TRABALHO DE CONCLUSÃO DE CURSO

**REDES NEURAS ARTIFICIAIS PARA COMPENSAÇÃO DE DISTORÇÃO NÃO
LINEAR EM SISTEMAS DE COMUNICAÇÕES ÓPTICAS COERENTES**

Aluno: Lucas Marim da Silva

Orientador: Prof. Dr. Ivan Aritz Aldaya Garde

Banca Examinadora:

- Ivan Aritz Aldaya Garde (Orientador)
- Leonardo de Melo João (Examinador)
- Marcelo Luís Francisco Abbade (Examinador)

A ata da defesa com as respectivas assinaturas dos membros encontra-se no prontuário do aluno (Expediente nº 19/2019)

São João da Boa Vista, 05 de dezembro de 2019

*"Se, porém, algum de vós necessita de sabedoria,
peça-a a Deus, que a todos dá liberalmente,
e nada lhes impropria; e ser-lhe-á concedida."
(Tiago 1:5)*

AGRADECIMENTOS

À Deus, primeiramente, por ter me concedido o dom da vida e me dado forças para superar cada obstáculo, guiando-me em meio a passos incertos, protegendo-me dia após dia, dando-me forças nos momentos de dificuldade e me amando incondicionalmente.

À meus pais, Silvia e Joel, que forneceram o meu sustento durante toda minha vida, me motivando fielmente durante a graduação, acreditando na minha capacidade e me apoiando durante as horas de cansaço.

À meus familiares, em especial meus irmãos Ana Carolina, Guilherme e Heloysa, que sempre foram grandes referências e fonte de inspiração, representando um porto seguro no qual sempre pude me firmar em durante momentos difíceis.

À meu avô, José Roberto, que foi um grande referencial em minha vida e sempre me incentivou a escolher os melhores caminhos, por mais difíceis que eles fossem, e que se orgulharia em me ver concluindo mais esta etapa em minha vida.

À todos os professores, que puderam compartilhar uma pequena parte de seus conhecimentos acadêmicos, me inspirando e incentivando a me tornar um profissional cada vez melhor.

Ao Prof. Dr. Ivan Aritz Aldaya Garde, que aceitou o desafio de me acompanhar nesse trabalho de conclusão de curso, provendo-me todo o apoio, suporte e confiança necessários para a elaboração deste trabalho.

À todos que direta ou indiretamente fizeram parte da minha formação e me apoiaram durante minha graduação.

RESUMO

Neste trabalho estudamos os impactos de impedimentos não lineares no desempenho de um sistema de comunicação óptica digital coerente e utilizamos redes neurais artificiais (ANNs) para a compensação deste tipo de adversidade. A compensação foi realizada por meio da classificação de símbolos utilizando aprendizagem de máquina, possibilitando a comparação entre os valores de taxa de erro de bit (BER) obtidos mediante classificação por ANNs e detecção por máxima verossimilhança. Os resultados revelam que, utilizando uma ANN com apenas uma camada escondida de 32 neurônios treinada com 12000 símbolos em 125 iterações de *backpropagation*, o valor ótimo de BER é reduzido em aproximadamente 23,55% se comparados à detecção convencional. Para esta topologia, o custo computacional para classificar cada símbolo é de 1152 operações de ponto flutuante.

PALAVRAS CHAVE: COMUNICAÇÕES ÓPTICAS; EFEITOS NÃO LINEARES; REDES NEURAIS ARTIFICIAIS.

ABSTRACT

In this work we study the impact of nonlinear impediments on the performance of a coherent digital optical communication system and we use artificial neural network (ANN) algorithms to compensate these adversities. Compensation was performed by classifying symbols using machine learning, making possible the comparison between the bit error rate (BER) obtained by ANNs classification and maximum likelihood detection. Results reveal that by utilizing an ANN with a single hidden layer of 32 neurons properly trained considering a training set of 12000 symbols and 125 backpropagation iterations, the optimal BER is reduced by approximately 23.55% compared to conventional detection. For this topology, the computational cost to classify each symbol is 1152 floating-point operations.

KEYWORDS: OPTICAL COMMUNICATIONS, NON-LINEAR EFFECTS, ARTIFICIAL NEURAL NETWORKS.

LISTA DE ILUSTRAÇÕES

Figura 1	Sistema LR-PON coerente simulado. S/P: Conversão serial-paralela (<i>serial-to-parallel conversion</i>). DAC: Conversor digital analógico (<i>digital-to-analog converter</i>). LD: Laser de diodo (<i>laser diode</i>). DP-MZM: Modulador Mach-Zenhder duplo em paralelo (<i>dual parallel-Mach-Zenhder modulator</i>). EDFA: Amplificador de fibra dopada a érbio (<i>erbium-doped fiber amplifier</i>). VOA: Atenuador óptico variável (<i>variable optical attenuator</i>). SSMF: Fibra monomodo padrão (<i>standard single mode fiber</i>). Att: Atenuador (<i>attenuator</i>). LPF: Filtro passa baixas (<i>low-pass filter</i>). ADC: Conversor analógico digital (<i>analog-to-digital converter</i>). DSP: Processamento digital de sinal (<i>digital signal processing</i>). . .	15
Figura 2	Constelações recebidas (16-QAM) em sistema óptico passivo <i>single channel</i> de 100 km de comprimento para potências transmitidas de (a) 3 mW, (b) 8 mW e (c) 13 mW. Valores de amplitude em unidades arbitrárias.	19
Figura 3	Neurônio artificial.	21
Figura 4	Função logística (sigmoide) considerando diferentes valores de β	22
Figura 5	Exemplo de uma rede neural artificial <i>feedforward</i> com n entradas, m saídas e duas camadas escondidas.	23
Figura 6	Função de custo de classificação utilizada no processo de treinamento de algoritmos de regressão logística	27
Figura 7	Constelações recebidas (16-QAM) em sistema óptico passivo <i>single channel</i> de 100 km de comprimento para potências transmitidas de (a) 2 mW, (b) 7 mW e (c) 12 mW. Regiões de decisão considerando detecção por máxima verossimilhança para as constelações recebidas considerando LOP de (d) 2 mW, (e) 7 mW e (f) 12 mW, e regiões de decisão considerando detecção por meio de algoritmo de redes neurais artificiais para LOP de (g) 2 mW, (h) 7 mW e (i) 12 mW. Valores de amplitude em unidades arbitrárias.	32
Figura 8	(a) Constelação recebida e corretamente classificada considerando LOP de 12 mW, classificação obtida por meio de (b) detecção por máxima verossimilhança e (c) redes neurais artificiais.	33
Figura 9	Taxa de erro de bit em função da potência transmitida considerando detecção por máxima verossimilhança e ANNs com diferentes quantidades de neurônios na camada escondida, todas submetidas a um treinamento com 16000 símbolos e 200 iterações.	34
Figura 10	Taxa de erro de bit em função do número de símbolos de treinamento considerando rede neural artificial com 32 neurônios na camada escondida e treinamento realizado em 200 iterações.	35

Figura 11 Taxa de erro de bit em função do número de iterações na etapa de treinamento considerando rede neural artificial com 32 neurônios e 12000 símbolos de *training set*. Processo realizado considerando blocos de treinamento (símbolos) diferentes, porém de mesmo tamanho 36

LISTA DE ABREVIATURAS E SIGLAS

- AI - *Artificial intelligence*
- ANN - *Artificial neural network*
- ASE - *Amplified spontaneous emission*
- BER - *Bit error rate*
- DBP - *Digital back propagation*
- DCF - *Dispersion-compensating fiber*
- EDFA - *Erbium doped fiber amplifiers*
- FWM - *Four-wave mixing*
- GVD - *Group-velocity dispersion*
- IVSTF - *Inverse Volterra Series Transfer Function*
- LOP - *Launch optical power*
- LR-PON - *Long reach passive optical network*
- OOK - *On-off keying*
- PMD - *Polarization mode dispersion*
- PON - *Passive optical network*
- QAM - *Quadrature amplitude modulation*
- SBS - *Stimulated Brillouin scattering*
- SER - *Symbol error rate*
- SNR - *Signal-to-noise ratio*
- SPM - *Self-phase modulation*
- SRS - *Stimulated Raman scattering*
- SSMF - *Standard single mode fiber*
- XPM - *Cross-phase modulation*
- WDM - *Wavelength-division Multiplex*

LISTA DE SÍMBOLOS

λ - comprimento de onda

n_1 - índice de refração do núcleo da fibra

n_2 - índice de refração da casca da fibra

\bar{n}_2 - coeficiente de índice não linear

Φ_{NL} - fase não linear

u - potencial de ativação

g - função de ativação

ω - peso sináptico

Ω - matriz de pesos sinápticos

J - função de custo médio

$Cost$ - função de custo

η - taxa de aprendizagem

δ_i - gradiente local

SUMÁRIO

1	INTRODUÇÃO	12
1.1	Evolução das comunicações ópticas	12
1.2	Descrição do problema e contribuição do trabalho	13
1.3	Organização do documento	13
2	SISTEMAS DE COMUNICAÇÕES ÓPTICAS	15
2.1	Sistemas de comunicações ópticas digitais coerentes	15
2.2	Impedimentos em sistemas de comunicações ópticas	16
2.2.1	Impedimentos lineares	16
2.2.2	Impedimentos não lineares	18
2.2.2.1	Modulação de fase não linear devido a efeito Kerr	18
2.2.2.1.1	<i>Automodulação de fase não linear</i>	19
3	REDES NEURAIS ARTIFICIAIS	21
3.1	Neurônios artificiais	21
3.2	Redes neurais artificiais	22
3.3	Treinamento de redes neurais artificiais <i>feedforward</i>	25
4	SETUP DE SIMULAÇÃO	30
4.1	Especificações técnicas do sistema de comunicação óptica considerado	30
4.2	Especificações do algoritmo de redes neurais artificiais implementado	30
5	RESULTADOS	31
5.1	Regiões de decisão considerando classificação por REDES NEURAIS ARTIFICIAIS	31
5.2	Otimização dos parâmetros do algoritmo de inteligência artificial	32
5.3	Análise de complexidade	37
6	CONSIDERAÇÕES FINAIS	38
	Referências	39

1 INTRODUÇÃO

1.1 EVOLUÇÃO DAS COMUNICAÇÕES ÓPTICAS

O desenvolvimento de tecnologias capazes de transmitir informação em fibras ópticas é algo relativamente recente. Os estudos preliminares que possibilitaram a implementação desta tecnologia tiveram início na década de 1960 com a invenção e demonstração do laser, que mais tarde viria a ser utilizado em grande parte dos sistemas de comunicações ópticas modernos (Agrawal 2012).

A evolução de estudos relacionados a materiais e processos utilizados na fabricação de dispositivos e fibras ópticas possibilitou o desenvolvimento dos primeiros sistemas de comunicação por fibra óptica, que surgiram por volta de 1975 e tornaram-se comerciais em 1980 (Sanferrare 1987). Embora os sistemas de primeira geração operassem com taxas de 45 Mb/s e spans de até 10 km, a capacidade de enlaces ópticos foi aumentando significativamente com o passar dos anos (Kogelnik 2000), sendo impulsionada por fatores como o desenvolvimento de lasers capazes de operar em comprimentos de onda nos quais a fibra possui menores perdas ($1,3 \mu m$ na segunda geração e $1,5 \mu m$ nas gerações posteriores) e a utilização de fibras monomodo, responsáveis por uma diminuição significativa na dispersão do canal (Agrawal 2012).

Um dos saltos tecnológicos mais significativos em sistemas de comunicações ópticas ocorreu no início da quarta geração com a utilização de multiplexação por divisão de comprimento de onda (WDM – *Wavelength-Division Multiplexing*), técnica que possibilitou a implementação de sistemas com taxas superiores a 1 Tb/s. O WDM tornou-se viável após o desenvolvimento dos primeiros amplificadores de fibra dopada de Érbio (EDFA – *Erbium Doped Fiber Amplifier*), no final da década de 1980, que possibilitaram a amplificação simultânea de múltiplos canais no domínio óptico, operando na faixa de 1530 – 1570 nm e dispensando a necessidade de conversão optoeletrônica e amplificação de múltiplos canais no domínio elétrico em nós intermediários da rede (4).

Embora a capacidade de sistemas de comunicações ópticas tenha crescido ao longo dos anos, a utilização de múltiplos canais causou o aumento da potência total transmitida em enlaces ópticos. Desta forma, o estudo de efeitos não lineares sobre a capacidade de sistemas de ondas luminosas despertou grande interesse nos últimos anos (Mittra e Stark 2001). Tais estudos baseiam-se no conceito de capacidade de canal, proposto por Shannon no contexto da teoria da informação (Shannon 1948), e demonstram que as não linearidades reduzem a eficiência espectral do sistema quando altas potências são transmitidas pela fibra (Essiambre et al. 2010; Mecozzi e Essiambre 2012).

Tem-se que, independentemente do comprimento do enlace, existe uma SNR ótima que define, dados os parâmetros do sistema, o ponto ao qual o canal passa a ser limitado por efeitos não lineares (Mecozzi e Essiambre 2012). Desta forma, a compensação desses efeitos pode resultar no deslocamento da potência de transmissão ótima, resultando no aumento das taxas de bit e/ou do comprimento máximo do enlace.

Estudos recentes vêm sendo realizados com o propósito de compensar não linearidades em sistemas de comunicações ópticas. As primeiras técnicas utilizadas foram *digital back propagation* (DBP) (Rafique et al. 2011) e *Inverse Volterra Series Transfer Function* (IVSTF) (Liu et al. 2011), porém

o alto custo computacional dessas técnicas impossibilita a aplicação em tempo real. Em contrapartida, a compensação pode ser feita utilizando algoritmos de inteligência artificial (IA – *Artificial intelligence*), que reduzem significativamente o número de operações requeridas (Mata et al. 2018; Jarajreh et al. 2014; 6). A mitigação de não linearidades pode ser realizada mediante diversos tipos de algoritmos, que por sua vez podem apresentar desempenho e complexidade diferentes. O presente trabalho busca apresentar resultados obtidos aplicando algoritmos de redes neurais artificiais (ANN – *Artificial neural network*) de classificação para redução de taxas de erro de bit em sistemas de comunicações ópticas digitais coerentes.

1.2 DESCRIÇÃO DO PROBLEMA E CONTRIBUIÇÃO DO TRABALHO

Nos últimos anos, diversos autores vêm desenvolvendo estudos buscando compensar diferentes tipos de adversidades por meio da utilização de algoritmos baseados em inteligência artificial. Atualmente, existem trabalhos publicados na literatura que buscam compensar efeitos como ruído de fase (Torres et al. 2016), dispersão por modo de polarização (PMD – *Polarization mode dispersion*) (Wu et al. 2009) e não linearidades (Jarajreh et al. 2014; 6).

Estudos recentes apresentam a compensação de não linearidades utilizando ANNs por intermédio da equalização não linear do canal, assim como apresentado em (Jarajreh et al. 2014). Esta equalização pode ser implementada utilizando duas ANNs, cada uma responsável pela equalização de uma das componentes do sinal (fase e quadratura). Neste caso, a saída da ANN apresenta valores contínuos, tratando-se de um algoritmo de regressão baseado em aprendizado de máquina (*machine learning*).

Diferentemente das aplicações conhecidas na literatura, o presente trabalho busca implementar algoritmos de ANNs para a compensação de não linearidades por meio de classificação de símbolos. Estes algoritmos devem ser capazes de determinar, mediante treinamento, regiões de decisão ótimas, buscando reduzir as taxas de erro de bit (BER – *Bit error ratio*) com relação ao modelo de detecção convencional (detecção por máxima verossimilhança). Nesta configuração, a compensação das componentes em fase e quadratura do sinal é realizada simultaneamente.

1.3 ORGANIZAÇÃO DO DOCUMENTO

Neste trabalho são implementados algoritmos de ANNs para a compensação de efeitos não lineares em sistemas de comunicações ópticas digitais coerentes. O documento é dividido como se segue:

O **Capítulo 2** apresenta uma descrição teórica do funcionamento de sistemas de comunicações ópticas digitais coerentes e dos principais impedimentos físicos que atuam sobre este tipo de sistema.

O **Capítulo 3** introduz o funcionamento de redes neurais artificiais, detalhando sua operação como equalizador não linear, bem como o processo de treinamento baseado no método de *backpropagation*.

O **Capítulo 4** descreve o arranjo de simulação, apresentando as especificações técnicas do sistema de comunicação óptica considerado e detalhando informações topológicas das ANNs implementadas.

O **Capítulo 5** apresenta os resultados obtidos utilizando ANNs para classificação de símbolos, comparando o desempenho dos algoritmos com o do método convencional de detecção (máxima verossimilhança). Além disso, são estudados os procedimentos para otimização da ANN com a

finalidade de obter uma boa relação entre custo computacional e desempenho, mensurado por meio da BER.

Por fim, o **Capítulo 6** destaca as principais conclusões obtidas e descreve possíveis trabalhos que poderão ser realizados futuramente.

2 SISTEMAS DE COMUNICAÇÕES ÓPTICAS

2.1 SISTEMAS DE COMUNICAÇÕES ÓPTICAS DIGITAIS COERENTES

Embora grande parte dos sistemas de comunicações ópticas atuais sejam baseados em detecção direta empregando modulação OOK (On-Off keying), sistemas ópticos digitais coerentes passaram a se popularizar a partir dos anos 2010, principalmente por possibilitarem a utilização de modulações com maior eficiência espectral (Agrawal 2012).

Sistemas de comunicações ópticas digitais coerentes modernos operam com demodulação intradina, que permite separar as componentes em fase e quadratura do sinal transmitido após a fotodetecção. Inicialmente, estes sistemas foram desenvolvidos para enlaces de longa distância, mas passaram a ser progressivamente utilizados em enlaces menores, como redes ópticas passivas (*Passive optical network*, PON) (Lavery et al. 2010). Como exemplo, a Figura 1 apresenta um diagrama de blocos do sistema implementado no presente trabalho, que constitui uma rede óptica passiva coerente de longa distância (*Long reach passive optical network*, LR-PON).

É possível notar no diagrama de blocos (Figura 1) que a informação a ser transmitida é direcionada a um conversor serial-paralelo, responsável por ordenar sequencialmente os bits. O conversor é então conectado a um mapeador QAM, cuja finalidade é mapear a constelação desejada, seguido por dois conversores digital-analógico, que geram os sinais modulantes das componentes em fase e quadratura. A modulação óptica se dá por meio de um modulador Mach-Zenhder duplo (*Mach-Zenhder modulator*,

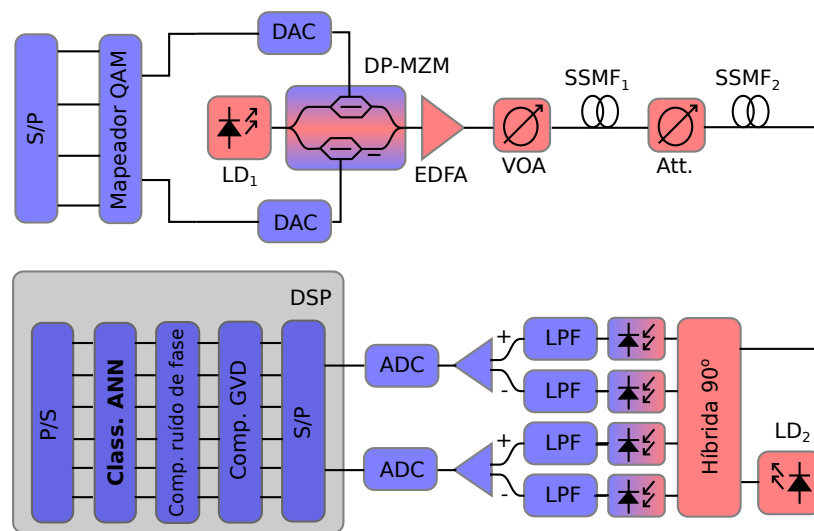


Figura 1 – Sistema LR-PON coerente simulado. S/P: Conversão serial-paralela (*serial-to-parallel conversion*). DAC: Conversor digital analógico (*digital-to-analog converter*). LD: Laser de diodo (*laser diode*). DP-MZM: Modulador Mach-Zenhder duplo em paralelo (*dual parallel-Mach-Zenhder modulator*). EDFA: Amplificador de fibra dopada a érbio (*erbium-doped fiber amplifier*). VOA: Atenuador óptico variável (*variable optical attenuator*). SSMF: Fibra monomodo padrão (*standard single mode fiber*). Att: Atenuador (*attenuator*). LPF: Filtro passa baixas (*low-pass filter*). ADC: Conversor analógico digital (*analog-to-digital converter*). DSP: Processamento digital de sinal (*digital signal processing*).

MZM), conectado à um *laser* de diodo (LD_1). Ainda no transmissor, utilizam-se um amplificador EDFA e um atenuador óptico variável (*Variable optical attenuator*, VOA) para ajustar a potência do sinal transmitido.

Na Figura 1, o canal óptico é representado por dois trechos de fibra monomodo padrão (*Standard single mode fiber*, SSMF), separados por um atenuador. O primeiro trecho representa um enlace de longa distância conectado a um divisor óptico passivo (*Splitter*), que divide o sinal para N usuários. Um atenuador (*Attenuator*, Att.) é utilizado para representar as perdas por divisão do sinal, ao passo que o segundo trecho de fibra representa o enlace que conecta o usuário ao divisor de sinal.

Por fim, ao se analisar o receptor do sistema, nota-se que o sinal óptico é inicialmente combinado com o sinal gerado pelo *laser* de diodo local (LD_2), que opera próximo à frequência nominal da portadora. Este procedimento é realizado por meio de uma rede híbrida de 90° , de modo que nas saídas o oscilador local é defasado em 0° , 90° , 180° e 270° . Posteriormente, as saídas passam por pares de fotodetectores balanceados limitados em banda, sendo esta limitação representada por filtros passa baixas (*Low-pass filter*, LPF) com frequência de corte igual a 75% da banda do sinal recebido. Após a fotodetecção, os sinais são diferencialmente amplificados, possibilitando a separação das componentes em fase e quadratura do sinal.

Posteriormente à digitalização do sinal mediante conversores analógico digitais (*Analog-to-digital converter*, ADC), as componentes em fase e quadratura são submetidas ao bloco de processamento digital de sinal (*Digital signal processing*, DSP). Este bloco é dividido em subsistemas responsáveis pela compensação de adversidades, sendo o subsistema de compensação de não linearidades o foco de implementação e discussão do presente trabalho.

2.2 IMPEDIMENTOS EM SISTEMAS DE COMUNICAÇÕES ÓPTICAS

De modo geral, sistemas de comunicações são suscetíveis a adversidades que degradam a qualidade do sinal transmitido, aumentando a probabilidade de ocorrência de erros na detecção da informação no receptor. Em particular, canais ópticos são afetados por diversos tipos de impedimentos, que por sua vez estão relacionados a efeitos lineares e não lineares (Agrawal 2012).

As subseções a seguir apresentarão uma breve descrição dos principais impedimentos lineares e não lineares que afetam um sistema de comunicação óptica implementado com fibra SSMF.

2.2.1 Impedimentos lineares

Os principais efeitos lineares que afetam sistemas de comunicações ópticas estão relacionados a atenuação, dispersão e ruído (Agrawal 2012). Considerando um enlace implementado com fibras monomodo, podemos observar as seguintes adversidades:

- **Dispersão cromática:** também conhecida como dispersão de velocidade de grupo (GVD - *Group-velocity dispersion*), este efeito está relacionado com a velocidade de grupo associada ao modo de propagação fundamental da fibra, que varia em função da frequência. Desta forma, as diferentes componentes espectrais do sinal viajam com velocidades diferentes, ocasionando interferência intersimbólica (Agrawal 2012). Este efeito pode ser compensado por meio da

utilização de fibras compensadoras de dispersão (DCF - *Dispersion-compensating fiber*) ou por meio de processamento digital de sinais. O sistema implementado no presente trabalho faz uso de compensação de GVD por meio de DSP.

- **Dispersão do modo de polarização:** A PMD é um efeito que está relacionado aos pequenos desvios da perfeita simetria cilíndrica, que levam à birrefringência da fibra. Considerando birrefringência constante, a elipsidade do núcleo quebra a degenerescência das duas polarizações do modo fundamental da fibra, ocasionando velocidades de grupo distintas para cada polarização, e conseqüentemente ocasionando interferência intersimbólica. Devido à natureza estocástica da PMD, modelos analíticos que descrevem este efeito são bastante complexos (Agrawal 2012). Técnicas de DSP conseguem compensar a contribuição linear deste efeito.
- **Atenuação:** A atenuação em fibras ópticas está relacionada com os processos de absorção e espalhamento de luz no material. O mecanismo fundamental de perda em fibras é conhecido como espalhamento Rayleigh, e ocorre devido à existência de flutuações microscópicas de densidade em uma escala menor que o comprimento de onda do sinal óptico (λ) (Agrawal 2012). Este fenômeno é um efeito elástico, uma vez que os fótons mantêm sua frequência original após o processo de espalhamento (Agrawal 2000). A atenuação também pode ocorrer em outros dispositivos passivos da rede de distribuição, normalmente devido à perdas de inserção e derivação. A compensação de atenuação se dá por meio da utilização de amplificadores, que podem atuar de forma localizada ou distribuída.
- **Ruído de fase:** O ruído de fase é particularmente crítico em sistemas coerentes (Torres et al. 2016), e está relacionado com a ocorrência de emissão espontânea nos lasers do transmissor e receptor, que causam variações de fase no campo gerado. O ruído de fase é um processo estocástico não ergódico, que gera uma rotação da constelação, e está diretamente relacionado com a largura de linha dos *lasers*. Deste modo, quanto menor é este parâmetro, maior o tempo de coerência da fase (5). O sistema implementado no presente trabalho faz uso de compensação de ruído de fase por meio de DSP.
- **Ruído aditivo:** Em sistemas de comunicações ópticas, o ruído aditivo pode ocorrer no domínio óptico e elétrico. No domínio óptico, a principal fonte de ruído é denominada *amplified spontaneous emission* (ASE) e está relacionado com a ocorrência de emissões espontâneas nos amplificadores, gerando fótons com amplitude e fase aleatórias que podem ser amplificados conjuntamente ao sinal (4). O ruído ASE pode ser desconsiderado em sistemas ópticos passivos. No domínio elétrico, o ruído ocorre durante a fotodetecção, podendo ser classificado como térmico, causado pelo movimento aleatório dos elétrons, e *shot*, relacionado à natureza discreta da corrente elétrica (elétrons são quantizados). A densidade espectral de potência do ruído *shot* é diretamente proporcional à corrente gerada no fotodetector (Agrawal 2012). Sendo assim, esse tipo de ruído é significativo em sistemas ópticos digitais coerentes, uma vez que o receptor mistura o sinal à um oscilador local com um nível de potência consideravelmente alto.

2.2.2 Impedimentos não lineares

Em sistemas de comunicações ópticas, as não linearidades dominantes são geradas pelas fibras, impactando diretamente na capacidade destes sistemas. Estes efeitos estão diretamente relacionados com a polarização do meio dielétrico (sílica) devido a propagação de campos eletromagnéticos intensos (Agrawal 2012). Os principais impedimentos não lineares são:

- **Espalhamento estimulado de luz:** diferentemente do espalhamento Rayleigh (explicado na subseção anterior), espalhamentos estimulados de luz são fenômenos inelásticos. No caso de Stokes, a frequência da luz espalhada é menor que a do sinal que a originou. Neste tipo de efeito, o espalhamento de um fóton origina outro fóton e um fônon, ambos com energia menor à do fóton que os originou (Agrawal 2000). Em fibras ópticas existem dois tipos de espalhamento estimulado de luz, denominados espalhamento estimulado de Brillouin (SBS - *stimulated Brillouin scattering*) e espalhamento estimulado de Raman (SRS - *stimulated Raman scattering*), e ocorrem somente quando a potência do sinal é superior a seus respectivos limiares. O limiar de Brillouin é consideravelmente menor que o de Raman, porém, normalmente os efeitos de ambos são desprezíveis, uma vez que sistemas de comunicações ópticas são projetados para operarem abaixo desses limiares (Agrawal 2012; Agrawal 2000).
- **Efeito Kerr:** este efeito está relacionado com a modulação de fase não linear em sistemas de comunicações ópticas (Agrawal 2000), e é o principal impedimento não linear a ser compensado neste trabalho. Desta forma, este fenômeno será descrito com maiores detalhes na subseção 2.2.2.1.

2.2.2.1 Modulação de fase não linear devido a efeito Kerr

Os sistemas de comunicações ópticas digitais coerentes sofrem modulação de fase não linear devido ao efeito Kerr (em particular o efeito Kerr óptico) (Agrawal 2012). Esta modulação está relacionada ao comportamento não linear do índice de refração do material, que varia de acordo com a intensidade do sinal óptico aplicado (Agrawal 2000). Para descrever esta variação, podem-se modelar os índices de refração do núcleo (n'_1) e da casca (n'_2) da fibra de acordo com a seguinte equação:

$$n'_j = n_j + I\bar{n}_2, \quad j = 1, 2 \quad (2.1)$$

sendo I é a intensidade óptica localizada e \bar{n}_2 o coeficiente de índice não linear, que é da ordem de $2,6 \times 10^{-20} \text{ m}^2/W$ para sílica fundida e pode variar de acordo com a dopagem do núcleo da fibra. Embora $\bar{n}_2 \ll n_1, n_2'$, a modulação não linear de fase torna-se significativa para enlaces ópticos longos e/ou altos valores de potência de transmissão (Agrawal 2000).

O efeito Kerr causa três fenômenos distintos: automodulação de fase, modulação de fase cruzada e mistura de quatro ondas. De modo geral, modulação de fase cruzada e mistura de quatro ondas são efeitos que ocorrem exclusivamente quando se faz uso de múltiplos canais, como é o caso de sistemas

WDM, enquanto que a automodulação de fase ocorre em qualquer tipo de sistema de comunicação óptica.

Neste trabalho, o sistema óptico implementado utiliza um único canal e é afetado, portanto, somente pela automodulação de fase, que será apresentada com maiores detalhes na subseção a seguir.

2.2.2.1.1 Automodulação de fase não linear

Devido à existência de \bar{n}_2 , a constante de fase do sinal óptico sofre alterações que geram uma fase não linear. Considerando um enlace com baixa dispersão e comprimento L , a fase não linear gerada pode ser descrita de acordo com a equação abaixo:

$$\Phi_{NL} = \int_0^L k_0 \left(\frac{P(z)}{A_{eff}} \right) \bar{n}_2 dz = \int_0^L \left(\frac{2\pi}{\lambda} \frac{\bar{n}_2}{A_{eff}} \right) P dz = \int_0^L \gamma P(z) dz = \gamma P_{in} L_{eff} \quad (2.2)$$

sendo P é a potência óptica, A_{eff} a área modal efetiva, γ é o parâmetro não linear, P_{in} a potência óptica instantânea na entrada da fibra e L_{eff} o comprimento efetivo da fibra. Devido à variação de P_{in} com relação ao tempo, Φ_{NL} também apresenta dependência temporal. Desta forma, a alteração da fase não linear em função de P_{in} gera uma modulação autoinduzida, denominada automodulação de fase (Agrawal 2000).

Um fato importante a se destacar é que caso a constelação transmitida possua símbolos com diferentes valores de energia, Φ_{NL} não será uniforme, uma vez que símbolos com energia mais elevada possuem maiores valores de amplitude e sofrem maiores rotações. Além disso, a distorção da constelação gerada pela automodulação de fase tende a ser mais significativa para potências de transmissão (LOP - *Launch optical power*) mais elevadas.

A Figura 2 mostra as constelações recuperadas por meio de detecção intradina em um sistema de comunicação óptica digital coerente com baixa dispersão considerando diferentes valores de LOP. Nota-se que o efeito Kerr é pouco significativo quando a potência transmitida é de 3 mW, e que neste caso o sistema é limitado por ruído. Considerando uma potência de transmissão de 8 mW, a

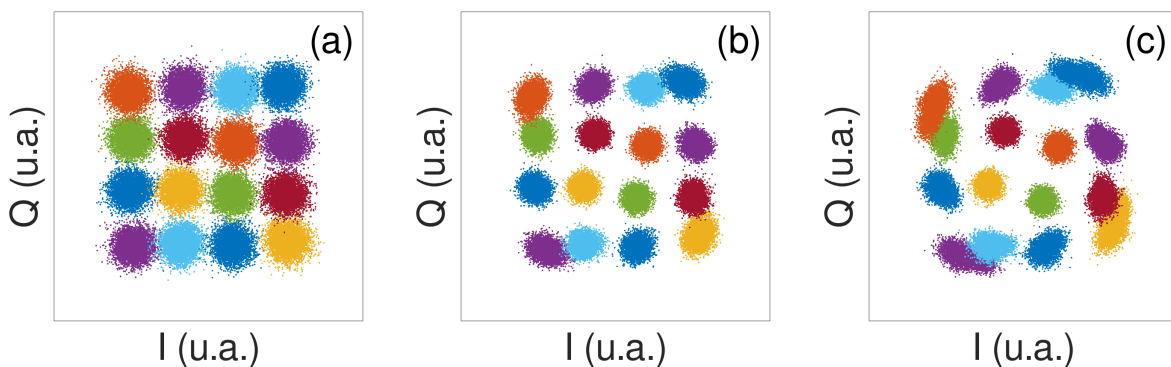


Figura 2 – Constelações recebidas (16-QAM) em sistema óptico passivo *single channel* de 100 km de comprimento para potências transmitidas de (a) 3 mW, (b) 8 mW e (c) 13 mW. Valores de amplitude em unidades arbitrárias.

automodulação de fase passa a ser perceptível, e tende a distorcer significativamente símbolos com maiores amplitudes. Para 13 mW de transmissão, a constelação é bastante distorcida, ocasionando em um aumento significativo nas taxas de erro de símbolo (SER - *Symbol error rate*). Para este caso, faz-se necessária a utilização de algoritmos de detecção com regiões de decisão otimizadas, uma vez que os símbolos são dispostos de maneira não uniforme.

3 REDES NEURAIS ARTIFICIAIS

Neste trabalho o algoritmo desenvolvido para a compensação de efeitos não lineares em sistemas de comunicações ópticas digitais coerentes é baseado em ANNs.

ANNs são algoritmos computacionais que se inspiram no funcionamento do sistema nervoso de seres vivos, sendo capazes de adquirir e manter o conhecimento baseado em informações prévias. Neste tipo de modelo computacional, definem-se unidades de processamento (neurônios artificiais) que são interconectados por meio de sinapses artificiais representadas por vetores/matrizes de pesos (Silva e Spatti 2010). Na seção a seguir, será apresentada uma breve descrição do modelo de funcionamento de neurônios artificiais.

3.1 NEURÔNIOS ARTIFICIAIS

De forma geral, neurônios artificiais são elementos computacionais não lineares que tipicamente fornecem saídas contínuas. Estas unidades processadoras recebem parâmetros de entrada, representados por meio do conjunto $\{x_1, x_2, x_3, \dots, x_n\}$, e os ponderam por meio das conexões de entrada, definidas como conjunto de pesos sinápticos $\{w_1, w_2, w_3, \dots, w_n\}$. Após multiplicar cada uma das entradas $\{x_i\}$ pelo seu respectivo peso $\{w_i\}$, estes sinais são agregados e produzem a soma ponderada das entradas, denotada por potencial de ativação (u). De u , também subtrai-se um valor θ constante, definido como limiar de ativação, que especifica o patamar que pode gerar um disparo em relação à saída do neurônio (Silva e Spatti 2010). A Figura 3 apresenta a estrutura básica de um neurônio artificial.

Partindo do procedimento previamente descrito, pode-se descrever o potencial de ativação de acordo com a seguinte expressão:

$$u = \sum_{i=1}^n w_i \cdot x_i - \theta \quad (3.1)$$

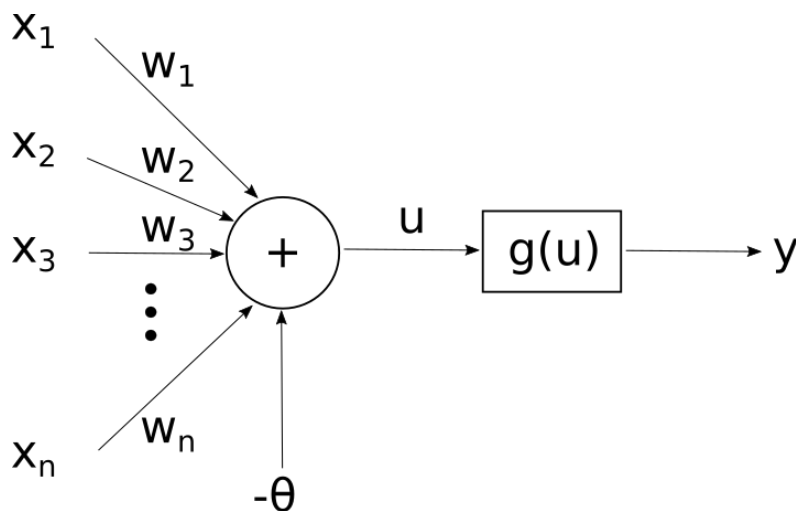


Figura 3 – Neurônio artificial.

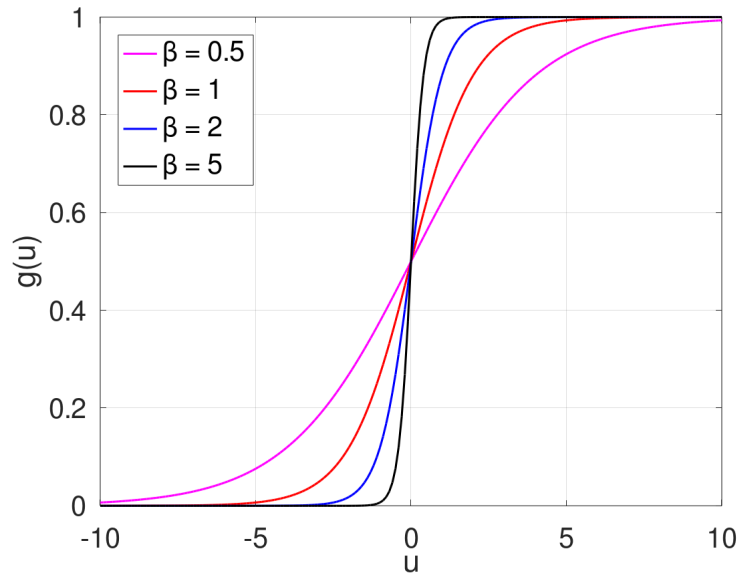


Figura 4 – Função logística (sigmoide) considerando diferentes valores de β

Por fim, o neurônio artificial opera sobre u aplicando uma função de ativação cuja finalidade é realizar uma mistura não linear das entradas e limitar a saída do neurônio dentro de um intervalo de valores razoáveis para o problema. Desta forma, produz-se uma saída y , tal que $y = g(u)$, que pode ou não ser direcionada a outros neurônios.

Algoritmos de classificação baseados em ANN normalmente utilizam uma função $g(u)$ (função de ativação) limitada entre 0 e 1, de modo que 0 representa hipótese falsa, e 1 verdadeira (Silva e Spatti 2010; Alpaydin 2009). Neste trabalho, a função de ativação utilizada é a função logística (sigmóide), cuja descrição matemática é expressa por:

$$g(u) = \frac{1}{1 + e^{-\beta \cdot u}}, \quad (3.2)$$

na qual β é uma constante real relacionada à inclinação da função e seu ponto de inflexão.

É possível observar na Figura 4 a forma da função logística considerando diferentes valores de β . Nota-se que ao adotar valores altos para β , a inclinação no ponto de inflexão $(0, g(0))$ tende a ser cada vez maior, de modo que, conforme $\beta \rightarrow \infty$, $g(u)$ tende à função degrau (Alpaydin 2009). Especificamente neste trabalho, assumiremos o caso trivial, em que $\beta = 1$.

3.2 REDES NEURAIS ARTIFICIAIS

De modo geral, redes neurais artificiais são constituídas por modelos computacionais que se inspiram no funcionamento de um cérebro. Sua arquitetura define a forma como os neurônios são arranjados e interconectados, formando uma estrutura robusta composta de estruturas computacionais (neurônios artificiais) interconectadas. Uma mesma arquitetura pode possuir diferentes tipos de topologia, que apresentam variações nas quantidades de neurônios, função de ativação, número de

camadas, etc (Alpaydin 2009).

ANNs são algoritmos baseados em aprendizagem supervisionada, ou seja, exigem um processo de treinamento com a finalidade de ajustar os pesos sinápticos, e conseqüentemente os valores de limiar de ativação de cada neurônio (Silva e Spatti 2010). O processo de treinamento é realizado considerando uma seqüência de dados (*training set*) conhecida tanto na entrada quanto na saída da rede, utilizada para otimizar o algoritmo visando a redução de erros na saída (transmissor e receptor). De acordo com (Silva e Spatti 2010) uma ANN é composta por três tipos de camadas, denominadas:

- **Camada de entrada:** Recebe as informações de entrada do sistema, normalmente de forma normalizada. A quantidade de neurônios de entrada depende da quantidade de parâmetros a serem considerados no problema;
- **Camadas intermediárias (escondidas ou ocultas):** São camadas de neurônios que possuem a finalidade de extrair características do processo considerado. São responsáveis pela maior parte do processamento interno do algoritmo.
- **Camada de saída:** Além de realizar parte do processamento interno, a camada de saída também é responsável pela apresentação dos resultados finais. Constitui a última camada de neurônios, e é indispensável para qualquer arquitetura/topologia de ANN.

As ANNs implementadas no presente trabalho seguem a arquitetura *feedforward*, sendo constituídas de neurônios artificiais interconectados em um único sentido (positivo). Estes algoritmos também são conhecidos como redes *Perceptron* e suportam somente fluxo unidirecional de informação, sendo classificados de duas formas: redes de camada simples (única), compostas somente por camadas de entrada e saída, ou redes de múltiplas camadas, que possuem uma ou mais camadas escondidas (Haykin 2007). A Figura 5 apresenta uma ANN *feedforward* com duas camadas escondidas considerando n entradas e m saídas.

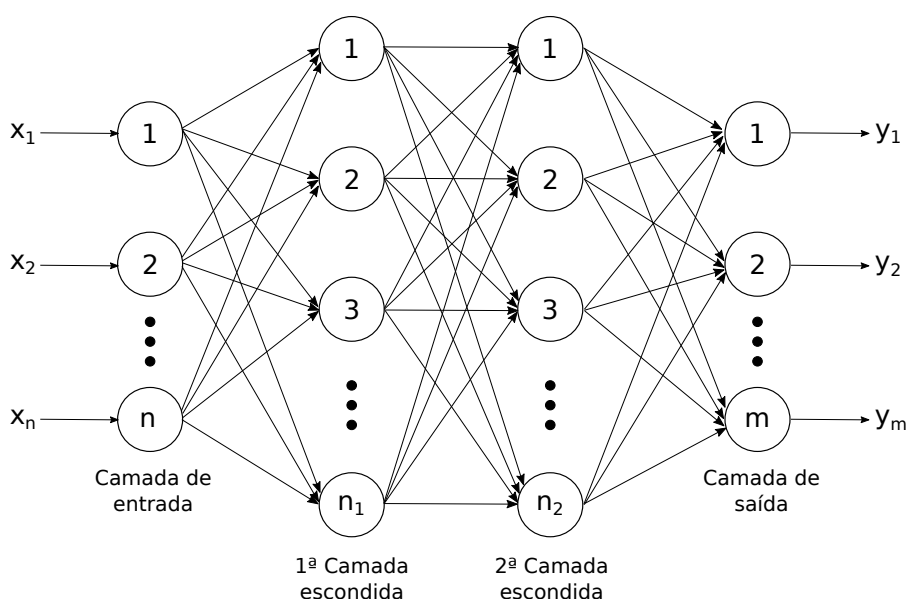


Figura 5 – Exemplo de uma rede neural artificial *feedforward* com n entradas, m saídas e duas camadas escondidas.

Para descrever o funcionamento de uma rede neural artificial *feedforward*, podemos assumir uma rede com a topologia apresentada na Figura 5. Nesta arquitetura, cada neurônio artificial que compõe as camadas escondida ou de saída é alimentado pelas saídas dos neurônios da camada anterior, seguindo o procedimento descrito na seção 3.1. Como exemplo, a equação abaixo apresenta as saídas de cada neurônio considerando uma camada escondida de comprimento j .

$$\begin{aligned}
 a_1 &= g(\omega_{01}x_0 + \omega_{11}x_1 + \omega_{21}x_2 + \dots + \omega_{n1}x_n) \\
 a_2 &= g(\omega_{02}x_0 + \omega_{12}x_1 + \omega_{22}x_2 + \dots + \omega_{n2}x_n) \\
 a_3 &= g(\omega_{03}x_0 + \omega_{13}x_1 + \omega_{23}x_2 + \dots + \omega_{n3}x_n) \\
 &\vdots \\
 a_j &= g(\omega_{0j}x_0 + \omega_{1j}x_1 + \omega_{2j}x_2 + \dots + \omega_{nj}x_n),
 \end{aligned} \tag{3.3}$$

sendo $x_0 = 1$ e $g(u)$ a função logística representada pela equação 3.2 considerando $\beta = 1$ (caso trivial).

As entradas e saídas de cada uma das camadas que compõe uma ANN podem ser representadas e processadas na forma vetorial com a finalidade de facilitar a implementação computacional deste tipo de algoritmo (Alpaydin 2009; Haykin 2007). Para isso, é preciso inicialmente descrever as entradas e saídas de cada camada como vetores colunas, de modo que o vetor saída seja resultante de uma operação entre o vetor de entrada e uma matriz composta por pesos sinápticos. Considerando uma camada com n entradas e m neurônios artificiais, é possível descrever os vetores de entrada (X), saída (Y) e a matriz de pesos (Ω) de acordo com:

$$X = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{3.4}$$

$$\Omega = \begin{bmatrix} \omega_{01} & \omega_{02} & \omega_{03} & \dots & \omega_{0m} \\ \omega_{11} & \omega_{12} & \omega_{13} & \dots & \omega_{1m} \\ \omega_{21} & \omega_{22} & \omega_{23} & \dots & \omega_{2m} \\ \vdots & \vdots & \vdots & & \vdots \\ \omega_{n1} & \omega_{n2} & \omega_{n3} & \dots & \omega_{nm} \end{bmatrix} \tag{3.5}$$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_m \end{bmatrix} \tag{3.6}$$

Na equação 3.5, cada coluna da matriz Ω representa os pesos sinápticos de um neurônio, e que cada linha corresponde aos pesos de cada neurônio atribuídos a uma entrada específica. Sendo assim, a expressão que relaciona a saída da camada com x e Ω é descrita por:

$$Y = g(U) = g(\Omega^T X) = \frac{1}{1 + e^{-\Omega^T X}}, \quad (3.7)$$

em que U é denominado vetor potencial de ativação, podendo ser descrito de acordo com a equação 3.8.

$$U = \Omega^T X. \quad (3.8)$$

Considerando o fluxo natural do algoritmo (unidirecional), as operações das camadas posteriores repetem o mesmo procedimento previamente descrito, assumindo que a entrada da i -ésima camada seja tal que $X_i = Y_{i-1}$. A propagação da informação pela rede partindo da camada de entrada até a saída é denominada *forward propagation* (Alpaydin 2009).

A próxima seção descreverá com maiores detalhes o processo de treinamento de redes neurais artificiais baseado no método de *backpropagation*.

3.3 TREINAMENTO DE REDES NEURAIS ARTIFICIAIS *FEEDFORWARD*

De modo geral, ANNs são algoritmos de aprendizagem supervisionada, portanto, exigem uma etapa de treinamento responsável pela otimização dos pesos sinápticos atribuídos a cada neurônio artificial que compõe a rede (Silva e Spatti 2010). Durante esta etapa, faz-se necessária a utilização de um bloco de dados com valores de saída conhecidos (*training set*), que são utilizados para a otimização dos pesos por meio da minimização de uma função de custo média (J), que deve assumir valores cada vez menores conforme as saídas do algoritmo se aproximam dos valores esperados. Esta função pode variar de acordo com o tipo de algoritmo, e deve levar em conta o tipo de saída apresentada pela rede neural, que pode ser contínua (regressão) ou discreta (classificação) (Alpaydin 2009).

O treinamento de ANNs *Perceptrons* multicamadas faz uso do algoritmo de *backpropagation*, normalmente aplicado sucessivamente em duas etapas específicas (Haykin 2007; Alpaydin 2009). A primeira delas é denominada *forward propagation* (apresentada na seção 3.2), e consiste na inserção das amostras do bloco de teste na rede e na propagação da informação até a saída. Durante esta etapa, obtém-se as respostas da ANN mantendo-se inalterados os pesos e limiares.

Posteriormente, as saídas são comparadas aos valores esperados, uma vez que o *training set* é conhecido pelo receptor. Baseado nesta comparação, será calculada a função de custo médio que será utilizada para ajustar os pesos de todos os neurônios artificiais que compõem a rede. Este ajuste é realizado durante a segunda etapa, denominada *backward propagation* (Silva e Spatti 2010).

A minimização da função de custo médio em algoritmos de ANN durante o *backward propagation* pode ser implementada utilizando o algoritmo de Gradiente Descendente de grupo (*Batch gradient descent*), que consiste na localização do mínimo global de uma função (neste caso a função de custo médio) por meio do cálculo da raiz de seu gradiente ($\nabla J = 0$). Este algoritmo é muitas vezes chamado

na literatura como regra Delta (Widrow e Hoff 1960), e representa um método de otimização iterativo que minimiza ∇J por meio do ajuste dos pesos das entradas (Ω), sendo também aplicado em outros algoritmos de inteligência artificial. A relação iterativa entre ∇J e $\Delta\Omega$ é expressa por:

$$\Delta\Omega = -\eta \cdot \nabla J, \quad (3.9)$$

sendo η um valor constante denominado taxa de aprendizagem capaz de indicar o quão rápido o processo de otimização estará se deslocando rumo ao ponto mínimo de J (Widrow e Hoff 1960).

Para que a otimização seja convergente, é necessário que J seja convexa, uma vez que essa condição garante a existência de um único ponto mínimo, denominado mínimo global. Caso essa condição não seja satisfeita, o gradiente descendente pode convergir para um ponto de mínimo local diferente do ponto ótimo, resultando em uma configuração sub-ótima (Widrow e Hoff 1960). Da literatura, descreve-se J de acordo com a equação 3.10.

$$J(\Omega) = \frac{1}{m} \sum_{j=1}^m Cost, \quad (3.10)$$

em que $Cost$ é o custo individual de cada uma das m (ou simplesmente função de custo) amostras que compõem o *training set* e Ω é a matriz de pesos sinápticos da camada de saída. A função de custo, por sua vez, pode variar de acordo com a ANN implementada, assim como já mencionado.

Dentre as possíveis funções de custo ($Cost$), uma opção que é normalmente utilizada em ANNs é o erro quadrático. Em particular, as redes neurais implementadas no presente trabalho são algoritmos de classificação, e não utilizam esta função de custo, mas sim uma função semelhante à aplicada em algoritmos de regressão logística, denominada entropia cruzada e expressa de acordo com a seguinte equação:

$$Cost = -(d)\log(y) - (1 - d)\log(1 - y). \quad (3.11)$$

Por se tratar de um algoritmo de classificação, os valores esperados na saída de um algoritmo de regressão logística estão restritos à 0 e 1, de modo que 0 representa falso e 1 verdadeiro. Analisando a equação 3.11, caso a saída esperada seja 0 ($d = 0$), então $Cost \rightarrow 0$ se $y \rightarrow 0$ e $Cost \rightarrow \infty$ se $y \rightarrow 1$. De forma análoga, caso a saída esperada seja 1 ($d = 1$), então $Cost \rightarrow 0$ se $y \rightarrow 1$ e $Cost \rightarrow \infty$ se $y \rightarrow 0$ (Alpaydin 2009). Este comportamento pode ser observado na Figura 6.

Embora a equação 3.11 seja adequada para problemas de classificação, é necessário lembrar que ANNs são compostas por múltiplos neurônios artificiais, de modo que a função de custo deve levar em conta todos os neurônios da camada de saída (Alpaydin 2009). Assim, a função de custo adequada para este caso é:

$$Cost = - \sum_{i=1}^n [(d_i(k))\log(y_i(k)) - (1 - d_i(k))\log(1 - y_i(k))]. \quad (3.12)$$

Partindo das equações 3.10 e 3.12, é possível definir o custo médio (J) da seguinte forma:

$$J(\Omega) = -\frac{1}{m} \sum_{j=1}^m \sum_{i=1}^n [(d_i(k)) \log(y_i(k)) - (1 - d_i(k)) \log(1 - y_i(k))]. \quad (3.13)$$

Definidas as expressões para as funções de custo ($Cost$) e custo médio (J) ao final da *forward propagation*, o ajuste dos pesos sinápticos pode ser realizado camada a camada durante o *backward propagation*, que se inicia na camada de saída e segue na direção contra propagante até a primeira camada escondida da ANN. O ajuste tem o objetivo de minimizar o erro obtido entre as saídas produzidas ($y(k)$) e desejadas ($d(k)$) (Silva e Spatti 2010). Aplicando-se então a definição de gradiente conjuntamente à regra da cadeia, é possível definir ∇J na camada de saída de acordo com a equação 3.14.

$$\nabla J = \frac{\partial J}{\partial \Omega^{out}} = \frac{\partial J}{\partial Y_i} \cdot \frac{\partial Y_i}{\partial U_i} \cdot \frac{\partial U_i}{\partial \Omega^{out}}, \quad (3.14)$$

em que Ω^{out} é a matriz de pesos sinápticos, Y_i o vetor de saída e U_i o vetor potencial de ativação, todos relacionados à camada de saída.

Partindo das equações 3.1, 3.7 e 3.13 e assumindo que a entrada da camada de saída é igual a saída da última camada escondida ($X_i = Y_{i-1}$), é possível obter as seguintes relações:

$$\frac{\partial J}{\partial Y_i} = -(D_i - Y_i) \quad (3.15)$$

$$\frac{\partial Y_i}{\partial U_i} = g'(U_i) \quad (3.16)$$

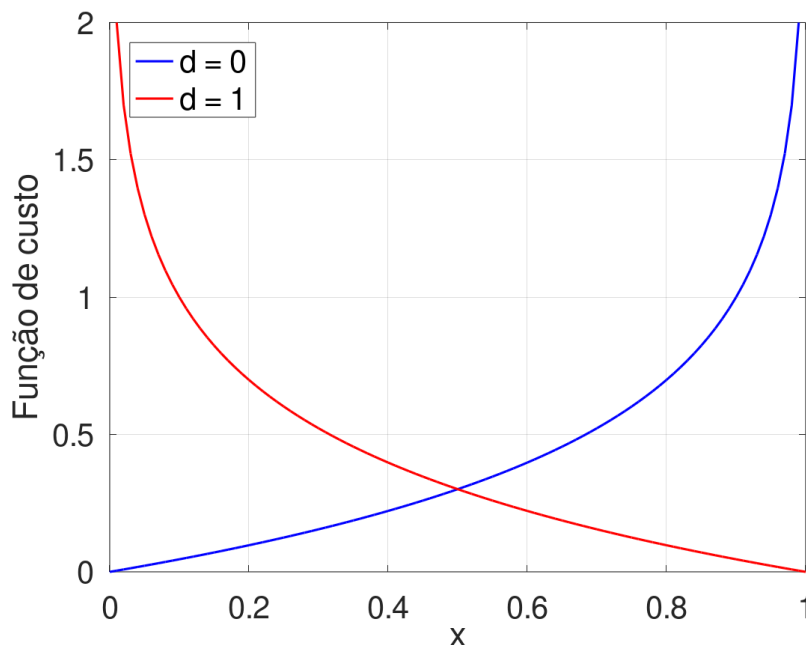


Figura 6 – Função de custo de classificação utilizada no processo de treinamento de algoritmos de regressão logística

$$\frac{\partial U_i}{\partial \Omega^{out}} = Y_{(i-1)}, \quad (3.17)$$

em que D_i é o vetor de saídas esperadas.

Assim, ao substituir as equações 3.15, 3.16 e 3.16 na expressão 3.14, obtém-se ∇J da seguinte forma:

$$\nabla J = \frac{\partial J}{\partial \Omega^{out}} = -(D_i - Y_i) \cdot g'(U_j) \cdot Y_{(i-1)} = -\delta_j \cdot Y_{(i-1)}, \quad (3.18)$$

sendo δ_j , definido como gradiente local, calculado da seguinte forma:

$$\delta_j = (D_i - Y_i) \cdot g'(U_j). \quad (3.19)$$

Partindo do desenvolvimento matemático descrito, é possível obter o ajuste da matriz de pesos sinápticos por meio do algoritmo de *Gradient descent*, apresentado na equação 3.9. Assim, podemos definir $\Delta \Omega^{out}$ de acordo com a seguinte equação:

$$\Delta \Omega^{out} = -\eta \frac{\partial J}{\partial \Omega^{out}} = \eta \cdot \delta_j \cdot Y_{(i-1)}. \quad (3.20)$$

Considerando a equação 3.20, é possível otimizar sucessivamente os pesos sinápticos da camada de saída do algoritmo por meio da comparação entre as saídas da ANN e os valores esperados, uma vez que o bloco de treinamento é previamente conhecido. É necessário então otimizar os pesos das camadas anteriores por meio da retropropagação do erro, que é realizada mediante a ponderação do mesmo pelo pesos ajustados. Desta forma, a resposta desejada para neurônios das camadas anteriores é determinada pelos neurônios da camada imediatamente à frente, a qual foi devidamente otimizada no passo anterior (Haykin 2007; Silva e Spatti 2010).

O processo de otimização dos pesos sinápticos das camadas escondidas é semelhante ao descrito para a camada de saída, levando em conta a mesma expressão para ∇J apresentada na equação 3.14, porém utilizando matriz Ω e vetores y_i e u_i relacionados à i -ésima camada a ser otimizada. Por intermédio das definições anteriores, têm-se para estes casos as seguintes relações:

$$\frac{\partial J}{\partial Y_i} = \sum_{k=1}^m \frac{\partial J}{\partial U_{(i+1)}} \cdot \frac{\partial U_{(i+1)}}{\partial Y_i} = \sum_{k=1}^m \frac{\partial J}{\partial U_{(i+1)}} \cdot \frac{\partial (\sum_{k=1}^m \Omega_{(i+1)} \cdot Y_i)}{\partial Y_i} = \sum_{k=1}^m \frac{\partial J}{\partial U_{(i+1)}} \cdot \Omega_{(i+1)} \quad (3.21)$$

$$\frac{\partial Y_i}{\partial U_i} = g'(U_i) \quad (3.22)$$

$$\frac{\partial u_i}{\partial \Omega} = Y_{(i-1)}, \quad (3.23)$$

em que m representa a quantidade de neurônios da camada à frente, cujo índice é representado como $(i + 1)$. Analogamente ao que foi previamente descrito, é possível definir genericamente o ajuste de

pesos sinápticos para a i -ésima camada escondida de acordo com a equação 3.24.

$$\Delta\Omega = \eta \cdot \delta_j^{(i)} \cdot Y_{(i-1)}, \quad (3.24)$$

sendo

$$\delta_j^{(i)} = \left(\sum_{k=1}^m \delta_k^{(i+1)} \cdot \Omega_{i+1} \right) \cdot g'(U_j). \quad (3.25)$$

O procedimento descrito é implementado no *backward* até que se otimize a primeira camada escondida. Cada iteração do processo de *backpropagation*, composto pelas etapas de propagação e contrapropagação, é responsável pelo ajuste gradual dos pesos de modo que, conforme as iterações são realizadas, as matrizes Ω referêntes a cada camada tendem a se aproximar de seus valores ótimos (Alpaydin 2009).

Neste trabalho, utilizou-se o software OCTAVE para implementar computacionalmente as ANNs. Durante o processo de treinamento, fez-se uso da função *fmincg*, responsável por aplicar o *Gradient descent* e otimizar os pesos relacionados a cada neurônio que compõe a rede.

4 SETUP DE SIMULAÇÃO

4.1 ESPECIFICAÇÕES TÉCNICAS DO SISTEMA DE COMUNICAÇÃO ÓPTICA CONSIDERADO

Para este trabalho, considerou-se um sistema de comunicação óptica digital coerente mostrado no esquemático da Figura 1. Sendo assim, o sistema representa uma rede de distribuição LR-PON com mecanismos de compensação de ruído de fase e dispersão cromática por meio de DSP. As principais especificações técnicas do sistema considerado encontram-se descritas abaixo:

- **Comprimento:** Enlace *single span* de 100 km;
- **Taxa de transmissão:** 56 Gbps por polarização (inicialmente operando com uma única polarização);
- **Largura de linha dos lasers:** 100 kHz;
- **Formato de modulação:** 16 QAM mapeado utilizando codificação de Gray;
- **Número de usuários:** 64, representados por meio de uma perda de ≈ 18 dB referentes à divisão do sinal;
- **Potência de transmissão (*Launch optical power, LOP*):** De 2 a 12 mW (passo de 1 mW);

O sistema foi implementado computacionalmente no software *VPI Transmission Maker* e simulado considerando um total de 81500 símbolos (326000 bits) para cada valor de potência especificado. Os símbolos recebidos foram classificados por meio de detecção por máxima verossimilhança e algoritmos de ANNs, ambos implementados utilizando o software *OCTAVE*.

4.2 ESPECIFICAÇÕES DO ALGORITMO DE REDES NEURAIAS ARTIFICIAIS IMPLEMENTADO

Para classificar os símbolos, implementaram-se ANNs *feedforward*, cujo funcionamento foi previamente apresentado na Seção 3.2. As topologias utilizadas são compostas por 3 camadas: uma de entrada, uma escondida e uma de saída.

Especificamente para este trabalho, os algoritmos foram implementados considerando-se uma camada de entrada composta por 2 neurônios, sendo cada um deles responsável por receber uma das componentes do sinal (fase e quadratura). Em contrapartida, a camada de saída foi implementada com 16 neurônios, sendo cada um deles responsável por representar um símbolo dentre os possíveis, de acordo com a modulação considerada (16 QAM). O número de neurônios da camada escondida foi um dos parâmetros otimizados no trabalho, e será discutido com maiores detalhes na Seção 5.2.

5 RESULTADOS

5.1 REGIÕES DE DECISÃO CONSIDERANDO CLASSIFICAÇÃO POR REDES NEURAIS ARTIFICIAIS

Após definir as especificações técnicas do sistema de comunicação óptica digital coerente e implementar o algoritmo de redes neurais artificiais para classificação de símbolos, foi possível analisar o desempenho do sistema considerando diferentes topologias de ANNs. Inicialmente, buscou-se analisar o funcionamento do algoritmo para posteriormente otimizar seus principais parâmetros: o número de neurônios na camada escondida, o comprimento do *training set* e o número de iterações do *backpropagation* durante a etapa de treinamento.

Para avaliar o funcionamento do sistema, inicialmente propôs-se uma ANN com 64 neurônios na camada escondida, *training set* composto por 16000 símbolos e 200 iterações. Em um primeiro momento, buscou-se analisar as regiões de decisão obtidas por meio de ANN, comparando-as com regiões de classificação por máxima verossimilhança (Lathi 1998). A Figura 7 apresenta as constelações e regiões de decisão obtidas considerando potências de transmissão de 2, 7 e 12 mW.

É possível notar que, para um LOP de 2 mW, o sistema está limitado por ruído, uma vez que a modulação de fase não linear devido ao efeito Kerr é imperceptível. Especificamente para este caso, as regiões de decisão obtidas utilizando ANN e detecção por máxima verossimilhança são consideravelmente próximas, e as fronteiras (limiares de decisão) normalmente passam pelo ponto médio entre símbolos adjacentes.

Considerando uma potência de transmissão de 7 mW, é perceptível que a melhoria na SNR do sinal é alcançada em detrimento de uma modulação de fase não linear mais significativa. Para este caso, embora os símbolos menos energéticos tenham sido pouco afetados pelo efeito Kerr, aqueles com maior amplitude sofreram uma perceptível distorção. Deste modo, nota-se que as regiões de decisão obtidas por máxima verossimilhança deixam de ser adequadas, uma vez que sua utilização pode acarretar no aumento BER, principalmente durante a classificação de símbolos com alta distorção não linear. Por outro lado, as regiões de decisão obtidas por meio do algoritmo de ANN se mostrou consideravelmente eficaz, uma vez que as fronteiras passaram a se às distribuições (distorcidas) de cada símbolo que compõe a constelação.

Para o caso em que a potência de transmissão é de 12 mW, a constelação é completamente distorcida, inviabilizando o método de detecção convencional (máxima verossimilhança). Entretanto, assim como no caso anterior, o algoritmo de ANN foi capaz de otimizar as regiões de decisão de acordo com a distorção de fase não linear.

A Figura 8 apresenta a classificação final considerando um LOP de 12 mW. É possível notar uma diferença considerável entre os resultados obtidos utilizando os dois métodos descritos. Neste caso, a classificação por máxima verossimilhança apresenta resultados pouco satisfatórios e contrastantes aos do algoritmo de ANN, que é consideravelmente próximo ao caso ideal (Figura 8 (a)). As taxas de erro de bit obtidas utilizando detecção convencional e ANN foram $1,338 \times 10^{-2}$ e $3,152 \times 10^{-3}$, respectivamente.

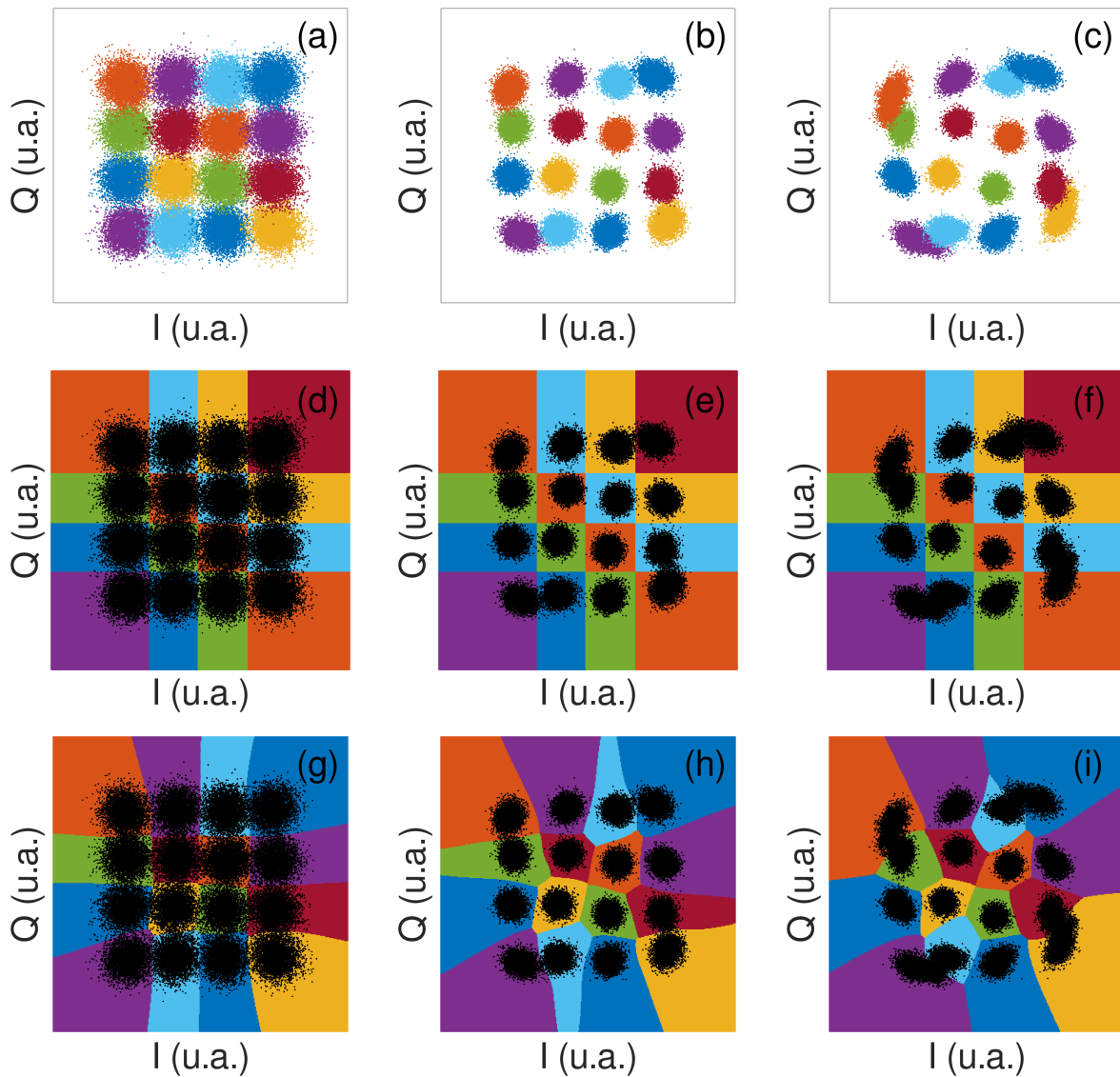


Figura 7 – Constelações recebidas (16-QAM) em sistema óptico passivo *single channel* de 100 km de comprimento para potências transmitidas de (a) 2 mW, (b) 7 mW e (c) 12 mW. Regiões de decisão considerando detecção por máxima verossimilhança para as constelações recebidas considerando LOP de (d) 2 mW, (e) 7 mW e (f) 12 mW, e regiões de decisão considerando detecção por meio de algoritmo de redes neurais artificiais para LOP de (g) 2 mW, (h) 7 mW e (i) 12 mW. Valores de amplitude em unidades arbitrárias.

Partindo dos resultados descritos, foi possível validar o funcionamento de ANNs para a classificação de símbolos. Desta forma, tornou-se necessário avaliar o desempenho do algoritmo considerando diferentes topologias e realizar a otimização de seus principais parâmetros, conciliando baixas taxas de erro de bit com o mínimo custo computacional.

5.2 OTIMIZAÇÃO DOS PARÂMETROS DO ALGORITMO DE INTELIGÊNCIA ARTIFICIAL

Dentre os parâmetros da topologia da ANN implementada, o número de neurônios que compõem a camada escondida é um dos mais importantes, uma vez que está diretamente relacionado ao custo computacional necessário para classificação de símbolos. Para otimizar este parâmetro, fez-se necessá-

rio obter as taxas de erro de bit para cada valor de potência (de 2 a 12 mW) considerado topologias com diferentes tamanhos de camada escondida. Além disso, a comparação entre os valores de BER alcançados utilizando ANN e detecção por máxima verossimilhança possibilitou identificar as faixas de potência cujo ganho de desempenho com relação à detecção convencional é mais significativo.

A Figura 9 apresenta a BER obtida utilizando detecção por máxima verossimilhança e ANNs com diferentes quantidades de neurônios na camada escondida. As ANNs foram submetidas a um treinamento realizado em 200 iterações com um *training set* de 16000 símbolos. Os valores de potência variam de 2 a 12 mW.

É possível observar que, considerando valores de potência transmitida entre 2 e 4 mW, todos os algoritmos de classificação apresentam desempenho semelhante. Nesta faixa, o aumento de potência resulta na diminuição da BER, que pode ser justificada pelo aumento da SNR do sinal. É perceptível que para estes valores de LOP o sistema é limitado por ruído, uma vez que a modulação de fase não linear não é suficientemente grande para distorcer a constelação, assim como já analisado na Figura 7 (a). Também é perceptível que, para esta faixa de potência, o desempenho de algoritmos de ANNs é consideravelmente próximo ao do método de detecção por máxima verossimilhança, sendo uma alternativa mais complexa e pouco atraente. Entretanto, ao se comparar o desempenho dos algoritmos de classificação para valores de potência iguais ou superiores a 5 mW, a redução da BER obtida mediante a ANNs faz-se perceptível. Esta diferença aumenta consideravelmente para LOP superiores a 7 mW, uma vez que o sistema passa a ser limitado por não linearidades.

Nota-se que a BER mínima utilizando detecção por máxima verossimilhança é alcançada quando a potência transmitida é 5 mW, sendo de aproximadamente $6,87 \times 10^{-4}$. Considerando todas as topologias de ANNs apresentadas na Figura 9 (12, 16, 32 e 64 neurônios), o ponto mínimo é deslocado para 6 mW, e os valores ótimos de BER são, em todos os casos, inferiores ao do método convencional. Para redes neurais artificiais com 32 e 64 neurônios na camada escondida, a BER ótima é de $5,25 \times 10^{-4}$, valor que representa uma redução de 23,55% com relação à máxima verossimilhança. Essa diferença é ainda mais significativa se levarmos em conta o aumento de 1 mW na potência ótima de transmissão (de 5 para 6 mW), uma vez que este acréscimo possibilita o alongamento do enlace.

Ao se comparar os desempenhos alcançados pelas diferentes topologias de ANNs, as configurações que obtiveram os melhores resultados foram as de 32 e 64 neurônios na camada escondida. Nota-

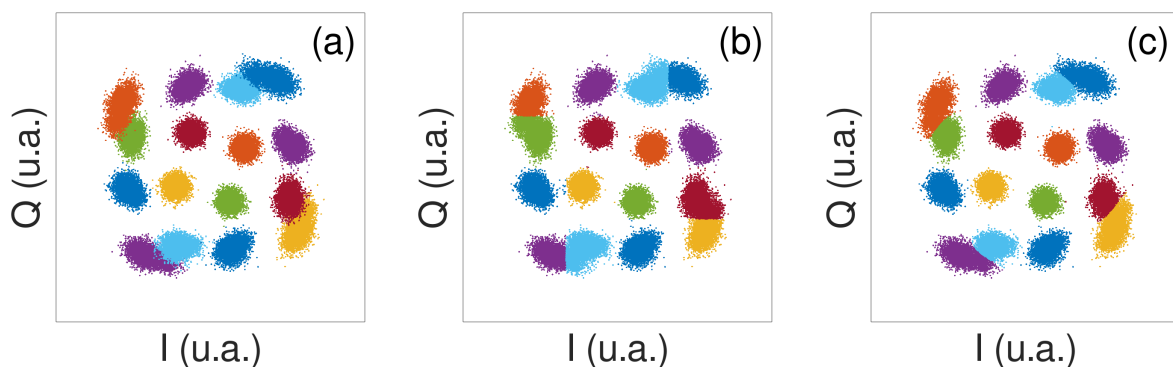


Figura 8 – (a) Constelação recebida e corretamente classificada considerando LOP de 12 mW, classificação obtida por meio de (b) detecção por máxima verossimilhança e (c) redes neurais artificiais.

se ainda que, conforme reduz-se o número de neurônios artificiais, o desempenho do algoritmo de classificação também é reduzido. Como parâmetro de comparação, a BER obtida por meio de ANN com 12 neurônios foi superior (pior) que a da detecção por máxima verossimilhança considerando LOP de 5 mW. Ainda assim, o desempenho do método convencional é inferior ao obtido em todas as topologias de ANN apresentadas na Figura 9 considerando valores de potência de transmissão iguais ou superiores a 6 mW.

Com base na discussão apresentada, a topologia que oferece melhor relação entre desempenho e complexidade dentre os casos estudados é a ANN com 32 neurônios na camada intermediária, uma vez que apresenta resultados semelhantes à configuração com 64 neurônios e apresenta menor custo computacional.

Após definir a topologia da ANN, buscou-se otimizar a quantidade de símbolos no bloco de treinamento. Para esta etapa, simularam-se resultados considerando blocos com comprimento de 2000 a 16000 símbolos e compararam-se os valores de BER obtidos na classificação dos blocos de treinamento e na etapa de teste. Assim como nos procedimentos anteriores, consideraram-se 81500 símbolos, sendo o bloco de teste composto por todos os símbolos que não constituíram o *training set*. O bloco de treinamento, por sua vez, compôs-se de símbolos selecionados aleatoriamente dentre os

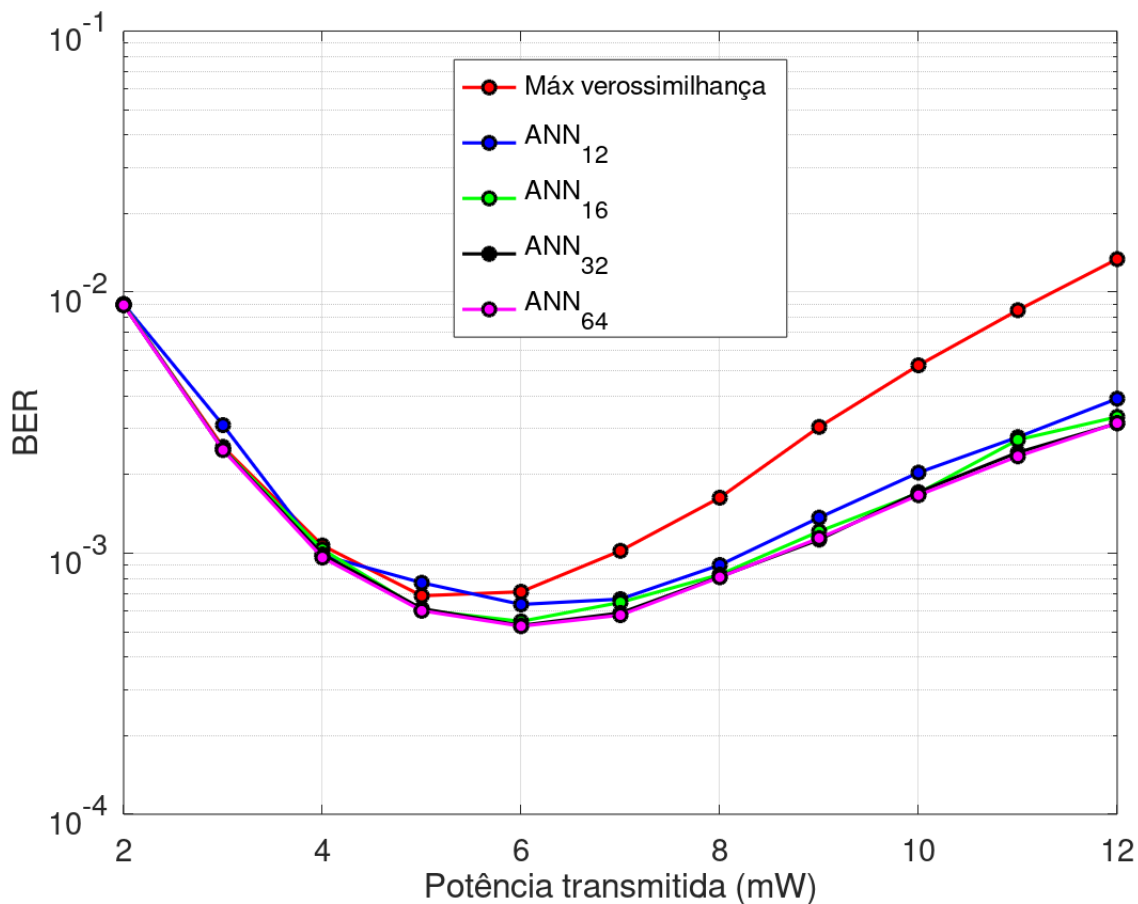


Figura 9 – Taxa de erro de bit em função da potência transmitida considerando detecção por máxima verossimilhança e ANNs com diferentes quantidades de neurônios na camada escondida, todas submetidas a um treinamento com 16000 símbolos e 200 iterações.

símbolos do bloco total, sendo único para cada etapa de treinamento e classificação.

Devido à aleatoriedade dos símbolos do *training set*, realizaram-se 10 simulações para cada caso, uma vez que blocos de treinamento estocásticos podem gerar desempenhos diferentes, mesmo que possuam tamanhos iguais. Sendo assim, após as simulações foram considerados os valores médios de BER obtidos para cada configuração, assim como apresentado na Figura 10.

É possível notar que para quantidades pequenas de símbolos de treinamento, os valores de BER obtidos nas duas etapas são contrastantes. Para estes casos, o tamanho reduzido do *training set* ocasiona baixos valores de BER na classificação do bloco de treinamento e altas taxas de erro no bloco de teste. Nestes casos, é perceptível que a quantidade de símbolos utilizados na etapa de aprendizagem não é suficientemente grande para treinar corretamente o sistema.

Conforme aumenta-se o comprimento do *training set*, o desempenho da etapa de teste melhora gradativamente, e o da etapa de treinamento é reduzido. Este processo se dá pelo fato de que, conforme as regiões de decisão se aproximam do caso ótimo, os desempenhos de treinamento e teste são cada vez mais próximos, uma vez que ambos os blocos de símbolos são gerados pelo mesmo sistema e estão sujeitos às mesmas adversidades. Mesmo com a aproximação dos valores obtidos de BER nos dois casos, o desempenho obtido no treinamento sempre tende a ser superior se comparado ao do teste,

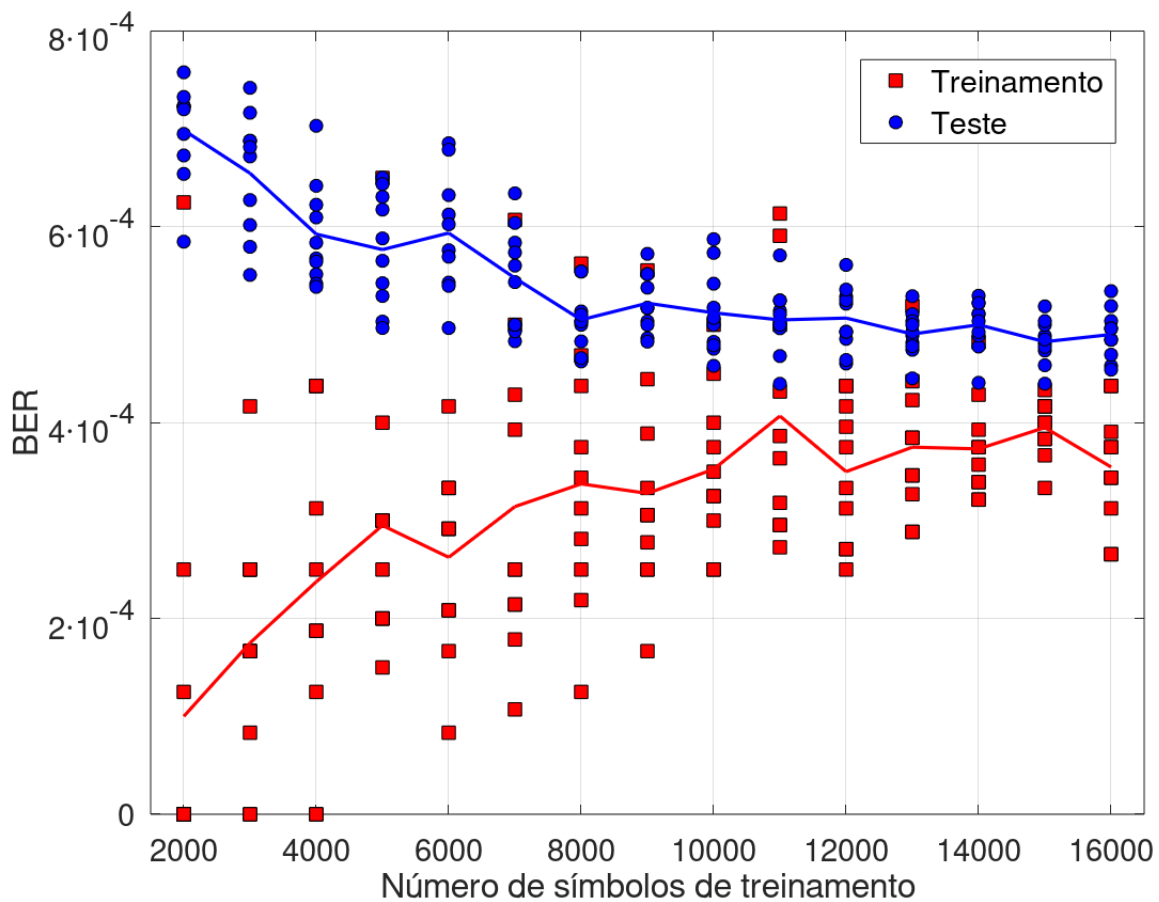


Figura 10 – Taxa de erro de bit em função do número de símbolos de treinamento considerando rede neural artificial com 32 neurônios na camada escondida e treinamento realizado em 200 iterações.

uma vez que o processo de aprendizagem é aplicado ao *training set* e busca reduzir os custos para classificar símbolos que estão especificamente contidos neste bloco.

Ainda com relação à Figura 10, é perceptível que as variações de desempenho não são uniformes, e tendem a se estabilizar para blocos de treinamento cada vez mais compridos. Após esta estabilização, o desempenho de ambas as etapas deixa de sofrer variações significativas, mesmo com o aumento considerável do número de símbolos de treinamento. Para o caso estudado, considerando 32 neurônios na camada escondida, esta saturação ocorre próxima aos 12000 símbolos. Sendo assim, podemos adotar este como sendo o comprimento ideal do *training set*, uma vez que a partir deste valor o aumento do tamanho do bloco de treinamento não ocasiona reduções significativas na BER.

Após definir o número de neurônios na camada escondida e a quantidade de símbolos do bloco de treinamento, buscou-se otimizar a quantidade de iterações de *backpropagation* necessárias para treinar adequadamente o sistema. Neste processo, simularam-se resultados considerando números de iterações que variaram entre 25 e 250 e compararam-se os valores de BER obtidos na classificação do bloco de teste. Novamente foram utilizados blocos de treinamentos com símbolos aleatoriamente escolhidos dentre os 81500 possíveis, considerando um *training set* de 12000 símbolos. Este procedimento foi realizado 5 vezes e os resultados obtidos podem ser analisados na Figura 11.

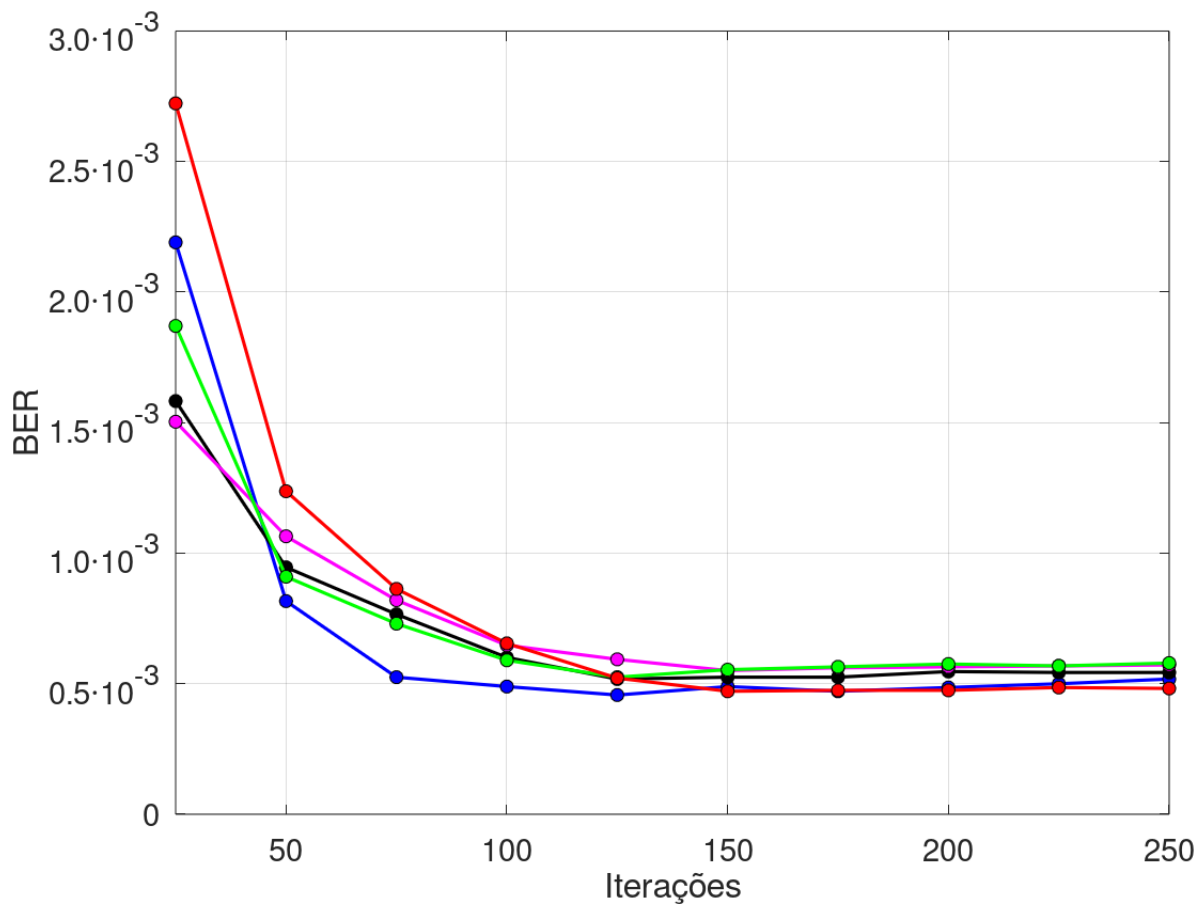


Figura 11 – Taxa de erro de bit em função do número de iterações na etapa de treinamento considerando rede neural artificial com 32 neurônios e 12000 símbolos de *training set*. Processo realizado considerando blocos de treinamentos (símbolos) diferentes, porém de mesmo tamanho

É possível notar que, conforme aumenta-se o número de iterações no intervalo de [25,100], a BER decai rapidamente e se estabiliza para quantidades iguais ou superiores a 150 iterações. Embora os desempenhos tenham sido diferentes para cada uma das 5 simulações, os valores finais de BER (após a estabilização) são próximos, e as pequenas diferenças estão relacionadas com a utilização de blocos de treinamento que, embora possuam o mesmo comprimento (12000 símbolos), são diferentes entre si.

Nota-se que os valores finais de BER são próximos à $0,5 \times 10^{-3}$ de modo que, para valores acima de 150 iterações, o desempenho do sistema não sofre variações significativas. Sendo assim, o número ideal de iterações de *backpropagation* a serem realizadas na etapa de treinamento considerando uma ANN com 32 neurônios na camada escondida e *training set* de 12000 símbolos é de 150.

5.3 ANÁLISE DE COMPLEXIDADE

Após otimizar os principais parâmetros que compõem as ANNs implementadas neste trabalho, foi possível analisar a complexidade do algoritmo. Para tal procedimento, foi necessário computar a quantidade de operações de ponto flutuante necessárias para classificar cada símbolo inserido no sistema.

De acordo com a Equação 3.1 apresentada na Seção 3.2, o número de operações realizadas por cada neurônio para o cálculo do potencial de ativação é igual a duas vezes a quantidade de entradas que possui, sendo metade deste montante operações de soma e a outra metade multiplicação. Além disso, a aplicação deste potencial na função de ativação (equação 3.2) requer operações normalmente não implementadas em unidades lógicas aritméticas, mas que podem ser calculadas por meio de *look up tables*. Deste modo, considerando uma ANN com A neurônios de entrada, B na camada escondida e C na saída, é possível computar o número de operações realizadas pelo algoritmo para classificar cada símbolo de acordo com a equação 5.1.

$$Num = 2AB + 2BC = 2B(A + C) \quad (5.1)$$

Particularizando essa expressão para uma ANN cujos comprimentos da entrada, camada escondida e saída são respectivamente 2, 32 e 16 neurônios, obtém-se que o número total de operações necessárias para classificar cada símbolo é de 1152 operações de ponto flutuante.

6 CONSIDERAÇÕES FINAIS

Neste trabalho analisaram-se os impactos de impedimentos não lineares em sistemas de comunicações ópticas digitais coerentes com um único canal e implementaram-se algoritmos baseados em redes neurais artificiais com a finalidade de compensar estas adversidades. Para isso, foram desenvolvidas ANNs *feedforward* com diferentes topologias para classificar símbolos por meio da obtenção de regiões de decisão otimizadas. Estas redes neurais foram submetidas à processo de treinamento com diferentes comprimentos de *training set* e número de iterações, e seus resultados foram comparados ao método de detecção por máxima verossimilhança.

Os algoritmos ANN foram otimizados de modo a oferecer desempenho satisfatório (em termos de BER) com o baixo custo computacional, e apresentaram reduções significativas nas taxas de erro, principalmente se comparados à forma de detecção convencional. As diferenças obtidas foram significativas considerando casos em que o sistema é limitado por não linearidades, demonstrando a aplicabilidade dos algoritmos implementados para a compensação deste tipo de efeito.

Foi perceptível que, para sistemas limitados por ruído, os algoritmos de ANNs apresentaram desempenho extremamente próximo à classificação por máxima verossimilhança, mostrando-se pouco viáveis especificamente para estes casos, uma vez que são consideravelmente mais complexos e não apresentam melhorias consideráveis com relação à detecção convencional.

Além disso, os resultados foram obtidos considerando um sistema de comunicação óptica operando com polarização única e um único canal. Sendo assim, este estudo pode ser futuramente estendido para sistemas WDM (múltiplos canais) e/ou com dupla polarização. Trabalhos futuros também poderão incluir a implementação de algoritmos de ANN para compensação de não linearidades por meio de equalização não linear do canal, estabelecendo uma comparação entre os desempenhos obtidos por meio de algoritmos de regressão com relação aos de classificação de símbolos implementados no presente trabalho.

REFERÊNCIAS

- AGRAWAL, G. P. Nonlinear fiber optics. In: **Nonlinear Science at the Dawn of the 21st Century**. [S.l.]: Springer, 2000. p. 195–211.
- AGRAWAL, G. P. **Fiber-optic communication systems**. [S.l.]: John Wiley & Sons, 2012. v. 222.
- ALPAYDIN, E. **Introduction to machine learning**. [S.l.]: MIT press, 2009.
- BECKER, P. M.; OLSSON, A. A.; SIMPSON, J. R. **Erbium-doped fiber amplifiers: fundamentals and technology**. [S.l.]: Elsevier, 1999.
- BRITO, S.; LUIS, J. et al. Pentes de frequências ópticas baseados em moduladores eletro-ópticos e fibras altamente não lineares. [sn], 2015.
- ERIKSSON, T. A.; BÜLOW, H.; LEVEN, A. Applying neural networks in optical communication systems: possible pitfalls. **IEEE Photonics Technology Letters**, IEEE, v. 29, n. 23, p. 2091–2094, 2017.
- ESSIAMBRE, R.-J. et al. Capacity limits of optical fiber networks. **Journal of Lightwave Technology**, IEEE, v. 28, n. 4, p. 662–701, 2010.
- HAYKIN, S. **Redes neurais: princípios e prática**. [S.l.]: Bookman Editora, 2007.
- JARAJREH, M. A. et al. Artificial neural network nonlinear equalizer for coherent optical OFDM. **IEEE Photonics Technology Letters**, IEEE, v. 27, n. 4, p. 387–390, 2014.
- KOGELNIK, H. High-capacity optical communications: personal recollections. **IEEE Journal of Selected Topics in Quantum Electronics**, IEEE, v. 6, n. 6, p. 1279–1286, 2000.
- LATHI, B. P. **Modern Digital and Analog Communication Systems 3e Osece**. [S.l.]: Oxford University Press, Inc., 1998.
- LAVERY, D. et al. A long-reach ultra-dense 10 gbit/s wdm-pon using a digital coherent receiver. **Optics Express**, Optical Society of America, v. 18, n. 25, p. 25855–25860, 2010.
- LIU, L. et al. Intrachannel nonlinearity compensation by inverse volterra series transfer function. **Journal of Lightwave Technology**, IEEE, v. 30, n. 3, p. 310–316, 2011.
- MATA, J. et al. Artificial intelligence (AI) methods in optical networks: A comprehensive survey. **Optical Switching and Networking**, Elsevier, v. 28, p. 43–57, 2018.
- MECOZZI, A.; ESSIAMBRE, R.-J. Nonlinear shannon limit in pseudolinear coherent systems. **Journal of Lightwave Technology**, IEEE, v. 30, n. 12, p. 2011–2024, 2012.
- MITRA, P. P.; STARK, J. B. Nonlinear limits to the information capacity of optical fibre communications. **Nature**, Nature Publishing Group, v. 411, n. 6841, p. 1027, 2001.
- RAFIQUE, D. et al. Compensation of intra-channel nonlinear fibre impairments using simplified digital back-propagation algorithm. **Optics Express**, Optical Society of America, v. 19, n. 10, p. 9453–9460, 2011.
- SANFERRARE, R. J. Terrestrial lightwave systems. **AT&T Technical Journal**, Nokia Bell Labs, v. 66, n. 1, p. 95–107, 1987.

SHANNON, C. E. A mathematical theory of communication. **Bell System Technical Journal**, Wiley Online Library, v. 27, n. 3, p. 379–423, 1948.

SILVA, I. d.; SPATTI, D. H. Redes neurais artificiais para engenharia e ciências aplicadas. **São Paulo: Artliber**, v. 23, n. 5, p. 33–111, 2010.

TORRES, J. J. G. et al. Adaptive nonsymmetrical demodulation based on machine learning to mitigate time-varying impairments. In: IEEE. **2016 IEEE Avionics and Vehicle Fiber-Optics and Photonics Conference (AVFOP)**. [S.l.], 2016. p. 289–290.

WIDROW, B.; HOFF, M. E. **Adaptive Switching Circuits**. [S.l.], 1960.

WU, X. et al. Applications of artificial neural networks in optical performance monitoring. **Journal of Lightwave Technology**, IEEE, v. 27, n. 16, p. 3580–3589, 2009.