

UNIVERSIDADE ESTADUAL PAULISTA JÚLIO DE MESQUITA FILHO
CAMPUS DE SÃO JOÃO DA BOA VISTA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

LÚCIO NERI BORGES

Uso de algoritmos de *clusterização* para mitigação de impedimentos em sistemas ópticos coerentes digitais DP-16QAM de alta taxa de transmissão de um único *span*

São João da Boa Vista

2021

LÚCIO NERI BORGES

Uso de algoritmos de *clusterização* para mitigação de impedimentos em sistemas ópticos coerentes digitais DP-16QAM de alta taxa de transmissão de um único *span*

Versão original

Dissertação apresentada à Universidade Estadual Paulista Júlio de Mesquita Filho - Campus de São João da Boa Vista para obtenção do título de Mestre em Engenharia Elétrica pelo Programa de Pós-graduação em Engenharia Elétrica.

Área de concentração: Automação e Sistemas Eletrônicos

Orientador: Prof. Dr. Ivan Aritz Aldaya Garde

São João da Boa Vista

2021

B732u Borges, Lúcio Neri
 Uso de algoritmos de clusterização para mitigação de impedimentos em sistemas ópticos coerentes digitais DP-16QAM de alta taxa de transmissão de um único span / Lúcio Neri Borges. -- São João da Boa Vista, 2021
 72 p. : il., tabs.

 Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Câmpus Experimental de São João da Boa Vista, São João da Boa Vista
 Orientador: Ivan Aritz Aldaya

 1. Comunicações Ópticas. 2. Fibras Ópticas. 3. Sistemas de Telecomunicação. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Câmpus Experimental de São João da Boa Vista. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: **Uso de algoritmos de clusterização para mitigação de impedimentos em sistemas ópticos coerentes digitais DP-16QAM de alta taxa de transmissão de um único span.**

AUTOR: **LÚCIO NERI BORGES**

ORIENTADOR: **IVAN ARITZ ALDAYA GARDE**

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 Eletrônica e de Telecomunicações / Câmpus de São João da Boa Vista

Prof. Dr. MARCELO LUÍS FRANCISCO ABBADE (Participação Virtual) 
Campos Experimental de São João da Boa Vista

Prof. Dr. JORGE DIEGO MARCONI (Participação Virtual) 
Centro de Engenharia, Modelagem e Ciências Sociais Aplicadas. / Universidade Federal do ABC

Sorocaba, 17 de maio de 2021

*Dedico este trabalho ao Grande Arquiteto do Universo, a minha família e à Camila Costa
por sempre me apoiarem em minhas mais loucas loucuras!*

Agradecimentos

Agradeço primeiramente ao Grande Arquiteto do Universo, que todos os dias está ao meu lado, sempre me guiando e mostrando o melhor caminho a seguir.

Agradeço aos meus pais e meus avós, que sempre me diziam que tudo daria certo. As palavras de sabedoria e apoio de vocês sempre foram meus guias na vida, vocês são a minha base.

Ao meu grande irmão Rodrigo Vinuto, que com sua garra e perseverança, torno-se um exemplo para todos e principalmente para mim. Tenho muito orgulho de ser seu irmão e poder contar com você em qualquer momento.

Á minha querida Camila Costa, que desde o início entendeu o quanto esta formação era importante para mim e apoiou isso todos dos dias. Participou com muita dedicação deste meu sonho e se consegui chegar até aqui, foi graças a você. Foi muito bom poder estudar com você e além disso, compartilhar meus dias com você, muito obrigado.

Ao Programa de Pós Graduação em Engenharia Elétrica da Unesp, que tem possibilitado a muitos a possibilidade de ampliar seus conhecimentos e assim poderem ajudar no desenvolvimento do país.

Meu agradecimento muito especial ao Prof. Dr Ivan Aldaya, que acreditou em mim mesmo sem me conhecer. Me proporcionou muito aprendizado e com muita paciência me ajudou muito neste processo. Obrigado e que você que seja sempre muito feliz.

Por fim, a todos que fizeram parte dessa jornada, meu muito obrigado.

"Sonhar o sonho impossível, sofrer a angústia implacável, pisar onde os bravos não ousam, reparar o mal irreparável, amar um amor casto à distância, enfrentar o inimigo invencível, tentar quando as forças se esvaem, alcançar a estrela inatingível: Essa é a minha busca."

(Dom Quixote - Miguel de Cervantes, 1605)

Resumo

BORGES, Lúcio Neri. **Uso de algoritmos de *clusterização* para mitigação de impedimentos em sistemas ópticos coerentes digitais DP-16QAM de alta taxa de transmissão de um único *span***. 2021. 73 f. Dissertação (Mestrado em Engenharia Elétrica) – Universidade Estadual Paulista (UNESP), Campus Experimental de São João da Boa Vista, São João da Boa Vista, 2021.

O progresso nos sistemas de comunicações ópticas possibilitou um grande aumento nas taxas de transmissão de dados. E os sistemas ópticos coerentes digitais têm uma contribuição muito importante neste processo. A possibilidade de modulação do sinal em amplitude, fase e polarização, permite que se possa atingir taxas de transmissão na ordem de centenas de giga bits por segundo com grande facilidade. Porém, estas altas taxas de transmissão tornam muito importante a mitigação de efeitos lineares e não lineares, o que é possível graças ao emprego dos processadores digitais de sinais. Neste conceito, os efeitos lineares foram amplamente estudados e, por meio do domínio digital, têm sido mitigados com bastante qualidade. Algoritmos de inteligência artificial têm contribuído no processo de mitigação destes efeitos e os resultados têm se mostrado muito promissores. Diante disso, propomos neste trabalho a utilização dos algoritmos de *clusterização* de dados não supervisionados denominados Maximização de Expectativa e K-means com inicialização K-means++, para mitigação dos impedimentos em sistemas ópticos coerentes com modulação DP-16QAM. São analisados os resultados da taxa de erros de bits (BER, do inglês *bit error ratio*) para sistemas operando em 100 Gbps e 400 Gbps. Para taxas de 100 Gbps são avaliados os comprimentos de enlaces variando de 120 km a 200 km, com potência de lançamento na fibra variando de 0 dBm a 15 dBm. Já os enlaces que operam a 400 Gbps são avaliados para comprimentos de 80 km a 150 km. Com os resultados obtidos neste trabalho, podemos verificar que os algoritmos são eficazes na mitigação de não linearidades, em ambas as taxas de transmissão analisadas, quando comparados com a detecção por máxima verossimilhança. Pode-se observar também que os algoritmos auxiliam na mitigação de efeitos lineares como o desbalanceamento da rede híbrida e do efeito da dessincronização do relógio. Por fim, os custos computacionais dos algoritmos de aprendizagem de máquinas utilizados são comparados, e fica evidenciado que o custo para o algoritmo de K-means é dezenas de vezes menor.

Palavras-chaves: Sistemas ópticos coerentes. DP-16QAM. Mitigação de impedimentos. Aprendizagem de máquinas. K-means. Maximização de Expectativa.

Abstract

Borges, Lúcio Neri. **Using clustering algorithms to mitigate nonlinear effects in DP-16QAM digital coherent optical systems operating at 100 Gbps and 400 Gbps**. 2021. 73 p. Dissertation (Master of Electric Engineering) – São Paulo State University (UNESP), Campus of São João da Boa Vista, São João da Boa Vista, 2021.

Progress in optical communications systems has made it possible to greatly increase data transmission rates. And digital coherent optical systems have a very important contribution in this process. The possibility of modulating the signal in amplitude, phase, and polarization allows it to achieve transmission rates in the order of hundreds of giga bits per second with great ease. However, these high transmission rates make it very important to mitigate linear and nonlinear effects, which is made possible by the use of digital signal processors. In this concept, linear effects have been extensively studied and mitigated quite well through the digital domain. Artificial intelligence algorithms have contributed to the process of mitigating these effects, and the results have shown great promise. In this context, we propose in this work to use the unsupervised data classification algorithms denominated Expectation Maximization and K-means with K-means++ initialization for mitigating nonlinear impairments in coherent optical systems with DP-16QAM modulation. Bit error rate results for systems operating at 100 Gbps and 400 Gbps are analyzed. For 100 Gbps rates, link lengths ranging from 120 km to 200 km are evaluated, with fiber launch power ranging from 0 dBm to 15 dBm. Links operating at 400 Gbps, on the other hand, are evaluated for link lengths ranging from 80 km to 150 km. With the results obtained in this work, we can verify that the algorithms are effective in mitigating nonlinearities, at both transmission rates analyzed, when compared to maximum likelihood detection. It can also be seen that the algorithms assist in mitigating linear effects such as hybrid imbalance I/Q and the effect of clock desynchronization. Finally, the computational cost of the machine learning algorithms used are compared, and it is evident that the cost for the K-means algorithm is tens of times lower.

Keywords: Coherent optical systems. DP-16QAM. Mitigation of impairments. Machine learning. K-means. Expectation maximization.

Lista de figuras

- Figura 1 – Diagrama de blocos de um sistema de comunicações óptico coerente digital. DAC: conversor digital/analógico; LD: diodo *laser*; DP-MZM: modulador duplo paralelo Mach–Zehnder; DA: amplificador de *drive*; PBS: divisor de feixe por polarização; PBC: combinador de feixe por polarização; EDFA: amplificadores de fibra dopada com érbio; SSMF: fibra padrão monomodo; LD_{LO}: diodo *laser* oscilador local; PD: fotodetectores e ADC: conversor analógico/digital. 23
- Figura 2 – Distorções não lineares em diferentes configurações de sistemas coerentes digitais. (a) Canal único com polarização única; (b) Canal único com dupla polarização; (c) Múltiplos canais WDM com dupla polarização. (d) Espectros dos sinais com diferentes configurações do sistema. . . . 29
- Figura 3 – Diagramas de blocos com a sequência das principais funcionalidades dos DSPs para sistemas coerentes digitais. 31
- Figura 4 – Arranjo da simulação implementada no Software VPI com os diagramas de constelação para demonstração da aplicação de cada módulo do DSP. 44
- Figura 5 – Diagrama de blocos dos módulos (a) transmissor e (b) receptor para sistemas coerentes digitais disponíveis em VPI *Transmission Maker*. . . 46
- Figura 6 – Espectros ópticos normalizados e diagramas de constelações no *front end* do receptor para sistemas com 100 km e operando em 100 Gbps. . 48
- Figura 7 – Espectros ópticos normalizados e diagramas de constelações no *front end* do receptor para sistemas com 100 km e operando em 400 Gbps. . 49
- Figura 8 – BER para enlaces com taxa de transmissão de 100 Gbps no Cenário 1. (a) BER em função da potência lançada para enlaces de 130 km, 150 km e 170 km; (b) Diagrama de contorno com os valores da BER em função da potência óptica lançada e do comprimento do enlace. . . 54
- Figura 9 – BER para enlaces com taxa de transmissão de 400 Gbps no Cenário 1. (a) BER em função da potência lançada para enlaces de 90 km, 110 km e 130 km; (b) Diagrama de contorno com os valores da BER em função da potência óptica lançada e do comprimento do enlace. 56

Figura 10 – BER para enlaces com taxa de transmissão de 100 Gbps no Cenário 2. (a) BER em função da potência lançada para enlaces de 130 km, 150 km e 170 km; (b) Diagrama de contorno com os valores da BER em função da potência óptica lançada e do comprimento do enlace.	58
Figura 11 – BER para enlaces com taxa de transmissão de 400 Gbps no Cenário 2. (a) BER em função da potência lançada para enlaces de 90 km, 110 km e 130 km; (b) Diagrama de contorno com os valores da BER em função da potência óptica lançada e do comprimento do enlace.	59
Figura 12 – Tempo médio de aprendizagem dos algoritmos de <i>clusterização</i> para taxa de transmissão de 100 Gbps.	60
Figura 13 – Tempo médio de aprendizagem dos algoritmos de <i>clusterização</i> para taxa de transmissão de 400 Gbps.	61
Figura 14 – Interação entre a SPM e a ISI para enlaces de taxas de transmissão maiores e menores em diferentes distâncias.	63

Lista de tabelas

Tabela 1 – Tabela resumo com os sistemas coerentes e os efeitos não lineares representados na Figura 2.	30
Tabela 2 – Tabela comparativa entre os algoritmos K-means e EM.	42
Tabela 3 – Tabela resumo dos parâmetros utilizados nas simulações.	52

Lista de abreviaturas e siglas

ADC	Conversor analógico/digital
ANN	Redes neurais artificiais
AP	Propagação de afinidade
ASE	Ruído de emissão espontânea amplificada
BER	Taxa de erro de bits
CD	Dispersão cromática
CW	<i>Laser</i> diodo de onda contínua
DA	Amplificador de <i>drive</i>
DAC	Conversor digital/analógico
DBP	Retro-propagação digital
DBSCAN	<i>Clusterização</i> Espacial Baseada em Densidade de Aplicações com Ruído
DP-MZM	Modulador Mach-Zehnder duplo-paralelo
DSP	Processador digital de sinais ou Processamento digital de sinais
EDFA	Amplificador fibra dopada a érbio
EM	Maximização de expectativa
FWM	Mistura de quatro ondas
Gbaud	Gigabaud
Gbps	Gigabit por segundo
GHz	Gigahertz
GSps	Gigaamostras por segundo
IM/DD	Modulação de intensidade com detecção direta
ISI	Interferência inter-simbólica

IVSTF	Função de transferência inversa da série de Volterra
km	Quilômetro
K-NN	K-ésimo vizinho mais próximo
LD	Diodo <i>laser</i>
m-PSK	Chaveamento por deslocamento de fase m-ária
ML	Máxima verossimilhança
MMA	Algoritmo de módulo múltiplo
nm	Nanômetro
OFDE	Equalização no domínio da frequência por sobreposição
OFDM	Multiplexação por divisão de frequências ortogonais
OOK	Chaveamento de ligado e desligado
OPC	Conjugação de fase óptica
PBC	Combinador de feixe por polarização
PBS	Divisor de feixe por polarização
PD	Fotodetectores
PCTW	Par de ondas gêmeas de fases conjugadas
PLL	Malha de captura de fase
PMD	Dispersão por modo de polarização
PRBS	Sequências de bits pseudo-aleatórias
QAM	Modulação de amplitude em quadratura
rms	Raiz média quadrática
SNR	Relação sinal-ruído
SoP	Estado de polarização

SPM	Automodulação de fase
SSMF	Fibra monomodo padrão
SVM	Máquina de vetores de suporte
SVR	Regressão por vetores de suporte
TDE	Equalizador no domínio do tempo
WDM	Multiplexação por divisão em comprimento de onda
XPM	Modulação de fase cruzada

Sumário

1	Introdução	17
1.1	<i>Contextualização do trabalho</i>	17
1.2	<i>Objetivos de pesquisa</i>	19
1.3	<i>Contribuição do trabalho</i>	20
1.4	<i>Organização do documento</i>	21
2	Mitigação de impedimentos em sistemas coerentes digitais . .	22
2.1	<i>Fundamentos de sistemas coerentes digitais</i>	22
2.2	<i>Impedimentos em sistemas coerentes digitais</i>	24
2.2.1	Impedimentos lineares em sistemas coerentes digitais	24
2.2.2	Impedimentos não lineares em sistemas coerentes digitais	26
2.3	<i>Mitigação de impedimentos em sistemas coerentes digitais</i>	31
2.3.1	Mitigação de impedimentos lineares em sistemas coerentes digitais .	31
2.3.2	Mitigação de impedimentos não lineares em sistemas coerentes digitais	33
2.4	<i>Mitigação de impedimentos utilizando aprendizado de máquinas</i>	35
2.4.1	K-means	37
2.4.2	Maximização de Expectativa	40
3	Arranjo de simulação e métricas de avaliação	43
3.1	<i>Arranjo de simulação</i>	43
3.1.1	Transmissor	43
3.1.2	Canal	45
3.1.3	Receptor	47
3.1.4	DSP	50
3.2	<i>Métricas de avaliação</i>	51
4	Resultados	53
4.1	<i>Análise do Cenário 1</i>	54
4.2	<i>Análise do Cenário 2</i>	57
4.3	<i>Análise da complexidade computacional</i>	59
5	Discussão dos resultados	62

6	Conclusão	65
	Referências	68

1 Introdução

1.1 Contextualização do trabalho

O uso das fibras ópticas como meio de comunicação teve início na segunda metade da década de 1970 e ficou conhecida como a primeira geração de sistemas de comunicações ópticas. Naquele momento, os sistemas operavam em aproximadamente $0,8 \mu\text{m}$, com uma taxa de transmissão de 45 Mb/s e com fibras multimodo que limitavam os enlaces a apenas 10 km (AGRAWAL, 2016). Na década seguinte, na denominada segunda geração, a utilização de fibras monomodo e a migração para a janela de comunicação de 1310 nm possibilitaram alcançar enlaces próximos a 50 km e taxas de transmissão de até 1,7 Gbps (GNAUCK *et al.*, 2008). Na terceira geração, migrou-se para comprimentos de onda ainda mais longos de 1550 nm, em que a fibra de vidro apresenta seu mínimo de atenuação, permitindo a transmissão de sinais com taxas de até 10 Gbps. No entanto, mesmo operando no comprimento de onda de mínima atenuação as limitações impostas pelo efeito da dispersão cromática (CD, do inglês *chromatic dispersion*), para estas taxas de transmissão, limitavam o alcance do enlace a menos de 30 km (KIKUCHI, 2015). Ainda neste período, deu-se início às pesquisas com sistemas heteródinos, também chamados de sistemas ópticos coerentes (LINKE; GNAUCK, 1988). Porém, o ruído de fase dos *lasers* disponíveis, em combinação com as flutuações do estado de polarização, resultavam em implementações muito complexas para serem adotadas comercialmente (AGRAWAL, 2012). Assim, as pesquisas em sistemas coerentes foram abandonadas com a irrupção nos anos 90 dos amplificadores a fibra dopada com érbio (EDFA, do inglês *erbium doped fiber amplifier*). Entre as principais vantagens destes amplificadores está a relativa larga banda de ganho, a qual permitiu a utilização de sistemas com multiplexação por divisão por comprimento de onda (WDM, do inglês *wavelength-division multiplexing*). Desta forma, os sistemas de quarta geração caracterizaram-se por sistemas *multi-span* com multiplexação WDM. Estes avanços proporcionaram um grande aumento na distância entre transmissor e receptor sem precisar de regeneradores eletro-ópticos, chegando a 11300 km (OTANI *et al.*, 1995), e também um avanço significativo da taxa de bits transmitidos, que chegou a 10 Tb/s em laboratório (AGRAWAL, 2016). Graças à WDM, aumentando o número de canais e a transmissão por canal, os sistemas de comunicações de quarta geração satisfizeram os requerimentos de capacidade até meados dos anos 2000 sem precisar modificar signifi-

cativamente o formato de modulação, que continuava sendo o tradicional chaveamento de ligado e desligado (OOK, do inglês *on-off keying*) que permitia utilizar sistemas de modulação de intensidade com detecção direta (IM/DD do inglês *intensity-modulation with direct detection*). No entanto, com a necessidade de aumento na taxa de transmissão de dados, a sincronização da amostragem no receptor e o efeito da dispersão por modo de polarização (PMD, do inglês *polarization mode dispersion*) se tornaram grandes desafios a serem superados. Para contornar estes problemas, uma das alternativas era o aumento da duração dos símbolos que, para manter uma taxa de transmissão, requer um aumento no número de bits por símbolo. A solução natural foi aumentar o número de níveis em sistemas IM/DD, mas estes eram extremadamente sensíveis ao ruído, tornando-se inviáveis (KIKUCHI, 2008). Desta forma, para que houvesse uma melhora na eficiência espectral dos sistemas WDM, seria necessária a adoção de novas dimensões na modulação do sinal (KIKUCHI, 2015).

Concomitantemente à evolução das comunicações ópticas, a eletrônica de alta frequência teve significativos avanços, viabilizando o processamento digital de sinal ¹ (DSP, do inglês *digital signal processing*) em tempo real de sinais a taxas de amostragem de várias Gigaamostras/s. Neste cenário, diferentes grupos de pesquisa identificaram a oportunidade de utilizar estes DSPs de alta velocidade para compensar os efeitos que dificultavam a detecção coerente, i.e. ruído de fase e flutuações do estado de polarização. Com isso, foram deixados de lado os sistemas baseados em IM/DD, permitindo a utilização dos sistemas com recepção coerente digital². Esta inovação deu início à quinta geração, que diferenciou-se da anterior justamente pelo emprego massivo de DSPs com alta taxa de processamento. Com isso, foi possível a aplicação de avançados sistemas de modulação, como chaveamento por deslocamento de fase (m-PSK, do inglês *phase shift keying*) e modulação de amplitude e em quadratura (QAM, do inglês *quadrature amplitude modulation*) (TSUKAMOTO *et al.*, 2005) (KIKUCHI, 2015), que permitem a modulação de um sinal em amplitude e fase. Adicionalmente, a natureza interferométrica dos sistemas coerentes digitais permite a exploração da diversidade em polarização. Outro ponto que deve ser mencionado é a sua versatilidade e eficiência na compensação ou minimização dos efeitos que degradam o sinal

¹ O termo DSP pode ser interpretado como Processador Digital de Sinais ou Processamento Digital de Sinais, neste caso cabe ao leitor avaliar o contexto de aplicação deste termo, em cada caso em que é utilizado no texto.

² Para distinguir os sistemas em que DSPs são usados para atingir a coerência daqueles que utilizam laços de controle ópticos, utilizaremos o termo *sistemas coerentes digitais*

transmitido, como o efeito da CD (GREEN; MITRA; WEGENER, 2003), da dispersão de modo de polarização (SUNNERUD *et al.*, 2002), de rotação do estado de polarização (SoP, do inglês *state of polarization*) (EKANAYAKE; H.HERATH, 2013) e ruído de fase linear (HO; KAHN, 2004). Com isso, os sistemas coerentes que outrora foram deixados de lado, neste momento começaram a ser aplicados amplamente (AGRAWAL, 2012).

A geração atual de sistemas de comunicações ópticas alia os avanços obtidos ao longo dos anos e o progresso nos dispositivos opto-eletrônicos para implementar os denominados sistemas coerentes digitais, os quais permitem explorar a diversidade de amplitude, fase e polarização para atingir taxas de transmissão por canal sem precedentes. Estes sistemas foram inicialmente concebidos para aplicações de longa distância, mas progressivamente foram adotados em sistemas de menor alcance. Em particular, esta tecnologia tem sido considerada para satisfazer as altas demandas de capacidade dos enlaces que conectam *data centers* de grande escala, tal e como evidencia o recentemente publicado padrão 400ZR (OIF, 2020). Em particular, prevem-se dois modos de operação em sistemas ópticos de alta capacidade: por um lado, enlaces sem amplificação em linha com um único canal e, por outro lado, enlaces WDM com amplificação em linha. Neste trabalho, nos centramos no primeiro destes modos de operação.

1.2 *Objetivos de pesquisa*

Como ocorre em todos os sistemas ópticos de comunicação, o desempenho deste tipo de enlace é afetado pelos impedimentos induzidos pelo transmissor, receptor e fibra. Numerosos trabalhos indicam que, em sistemas com taxas de transmissão de até 56 Gbps, impedimentos lineares como a dispersão, o ruído de fase dos *lasers* e as flutuações de polarização podem ser compensados mediante DSP, sendo o desempenho do sistema limitado pela combinação do ruído aditivo e as não linearidades da fibra (GIACOUMIDIS *et al.*, 2019; PAULA *et al.*, 2020; ALDAYA *et al.*, 2020). Nestes tipos de sistemas, algoritmos de *clusterização*³, têm-se demonstrado efetivos na compensação de não linearidades e, conseqüentemente, na extensão do alcance da rede (BORGES *et al.*, 2020; ZIBAR *et al.*, 2015; GIACOUMIDIS *et al.*, 2019). Porém, não tem se avaliado se estas afirmações são

³ Por todo o documento o termo *clusterização* será utilizado como expressão que define o ato de rotular um conjunto de dados de forma não supervisionada. Entende-se o termo como processo de associar cada símbolo transmitido a um determinado *cluster*, de acordo com o símbolo do diagramas de constelação.

também aplicáveis a sistemas operando a taxas de transmissão superiores. Em outras palavras, considerando um sistema coerente digital de um único *span* utilizando formato de modulação 16-QAM de alta taxa de transmissão (100 Gbps e 400 Gbps), este trabalho tem como objetivo responder as seguintes perguntas:

1. Sobre a generalização dos sistemas. O desempenho do sistema continua sendo limitado unicamente pelos efeitos combinados do ruído aditivo e não linearidades da fibra? Pode-se continuar assumindo que todos os efeitos lineares, à exceção do ruído aditivo, são compensados completamente pelo DSP?
2. Sobre a mitigação de impedimentos usando algoritmos de *clusterização*:
 - a) No caso em que o DSP não seja capaz de compensar completamente os impedimentos lineares, podem os algoritmos de *clusterização* compensar os efeitos residuais dos mesmos?
 - b) Algoritmos de *clusterização* como K-means e Maximização de Expectativa (EM, do inglês *Expectation Maximization*) podem mitigar as não linearidades nestes sistemas, como acontece em taxas de transmissão inferiores?
3. Finalmente, uma pergunta com implicações mais práticas. Pode-se estender o alcance do sistema utilizando algoritmos de *clusterização*?

1.3 Contribuição do trabalho

Neste trabalho, é avaliada a utilização de algoritmos de inteligência artificial de *clusterização*, como K-means com inicialização K-means++ e o algoritmo EM, para mitigar os impedimentos em sistemas ópticos coerentes de um único *span* com altas taxas de transmissão. Neste contexto, a *clusterização* auxiliará na melhora dos resultados de demodulação nas regiões de decisão, em que haverá o auxílio do algoritmo na escolha de qual símbolo da constelação pertence cada símbolo recebido. Os algoritmos são aplicados em um sistema óptico digital com modulação DP-16QAM (do inglês *dual-polarization 16-ary quadrature amplitude modulation*), canal único e taxa de transmissão de dados de 100 Gbps e 400 Gbps. Os algoritmos são avaliados para enlaces de 120 km a 200 km, operando em 100 Gbps, e enlaces de 80 km a 150 km, para a taxa de 400 Gbps. A potência lançada na fibra para estes cenários é analisada entre 0 dBm e 15 dBm. Para identificar os

impedimentos limitantes, assim como a capacidade da *clusterização* para compensar estes, são analisados dois cenários distintos:

- **Cenário 1** - Considera os efeitos lineares da fibra (dispersão cromática, dispersão por modo de polarização e flutuação do estado de polarização), os efeitos não-lineares da fibra, o ruído dos fotodetectores, o desbalanceamento da rede híbrida e a dessincronização de relógio;
- **Cenário 2** - Considera apenas os efeitos lineares da fibra previamente mencionados, os efeitos não lineares e o ruído dos fotodetectores.

A comparação entre o desempenho dos algoritmos de *clusterização* nestes dois cenários permitiu-nos identificar a capacidade dos dois algoritmos de *clusterização* para compensar os diferentes impedimentos. Resultados revelaram que, para sistemas operando a taxas superiores, a *clusterização* não é tão eficiente para mitigar os efeitos das não linearidades, mas mitiga efeitos lineares como dessincronização de relógio, que é mais relevante em sistemas de altas taxas. Por fim, as simulações mostram que, efetivamente, os algoritmos de *clusterização* permitem uma extensão do máximo alcance do enlace, porém esta extensão não é atingida devido à mitigação de não linearidades, mas advém da mitigação dos efeitos residuais dos impedimentos lineares.

1.4 Organização do documento

Este documento contém cinco capítulos, sendo que no primeiro capítulo é descrita a evolução dos sistemas de comunicação até chegarem aos sistemas coerentes digitais, no qual são apresentados os limites superiores no desempenho dos sistemas coerentes. Neste capítulo também são apresentados os questionamentos que este trabalho pretende responder, a fim de verificar a mitigação de impedimentos em enlaces coerentes. No segundo capítulo, são descritos os principais efeitos lineares e não lineares nestes sistemas, quais são as maneiras mais utilizadas para a mitigação destes efeitos e, por fim, como a aprendizagem de máquinas pode ajudar a mitigá-los. No capítulo três, é descrita a configuração da simulação da rede DP-16QAM utilizada, bem como os parâmetros utilizados na mesma. Os resultados são abordados no capítulo quatro e algumas discussões sobre estes resultados são apresentadas no quinto capítulo. Finalmente, no capítulo seis, são apresentadas as conclusões e as propostas para trabalhos futuros.

2 Mitigação de impedimentos em sistemas coerentes digitais

2.1 Fundamentos de sistemas coerentes digitais

A principal característica dos sistemas coerentes que os distingue dos sistemas IM/DD anteriores é a presença de um oscilador local no receptor, de forma que o sinal recebido possa ser misturado com um tom gerado por este oscilador local, o que confere aos sistemas coerentes uma natureza interferométrica (KIKUCHI, 2008). Como já falado previamente, atingir a correlação das fases do sinal e do oscilador local é uma tarefa complexa devido ao grande ruído de fase dos *lasers* e às flutuações do estado de polarização. Desta forma, nos sistemas coerentes digitais os dois *lasers* apresentam fases não equivalentes e, inclusive, com um *offset* frequencial. Conseqüentemente, a detecção não pode ser considerada coerente e, portanto, não corresponde estritamente à detecção heteródina nem a detecção homódina, sendo então denominada como intradina. Os sistemas coerentes digitais são chamados de coerentes, pois conseguem recuperar a fase do sinal recebido no domínio digital sem a utilização da detecção coerente.

Tem-se proposto diferentes arquiteturas para os sistemas coerentes digitais. Estas podem ser classificadas entre sistemas de polarização única e sistemas de polarização dupla. Os primeiros podem se implementar fazendo uso de um rastreador de polarização dinâmico ou utilizando um receptor com diversidade de polarização. Por outro lado, a detecção pode ser simples ou balanceada, sendo esta última mais efetiva, pois amplifica o termo de batimento e cancela os outros termos (TSUKAMOTO *et al.*, 2005). Na Figura 1 mostra-se o diagrama de blocos típico de um sistema coerente digital de dupla polarização com detecção balanceada, o qual será adotado no presente trabalho.

No transmissor, a informação digital a ser enviada em cada polarização passa por um conversor serial/paralelo convertendo o *stream* binário em blocos que são mapeados na constelação desejada. Estes símbolos são convertidos ao domínio elétrico utilizando um conversor digital analógico (DAC do inglês *digital-to-analog converter*), cujas saídas (correspondentes às componentes em fase e quadratura) são filtradas usando filtros de Nyquist¹. As saídas filtradas são amplificadas utilizando amplificadores de alta potência (DA, do inglês *driver amplifiers*) para atingir os níveis de tensão requeridos para alimentar

¹ Cabe mencionar que o filtro de Nyquist pode ser implementado tanto no domínio digital como no analógico

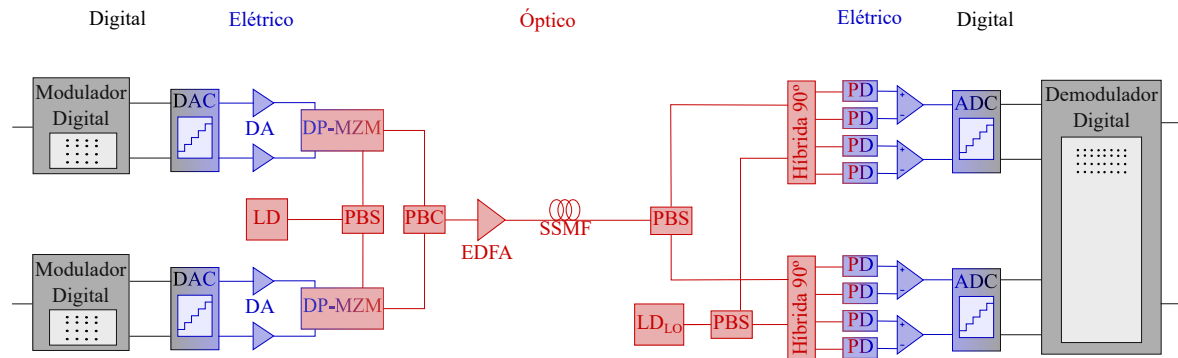


Figura 1 – Diagrama de blocos de um sistema de comunicações óptico coerente digital. DAC: conversor digital/analogico; LD: diodo *laser*; DP-MZM: modulador duplo paralelo Mach-Zehnder; DA: amplificador de *drive*; PBS: divisor de feixe por polarização; PBC: combinador de feixe por polarização; EDFA: amplificadores de fibra dopada com érbio; SSMF: fibra padrão monomodo; LD_{LO}: diodo *laser* oscilador local; PD: fotodetectores e ADC: conversor analógico/digital.

as entradas de um modulador Mach-Zehnder duplo-paralelo (DP-MZM, do inglês *Dual-parallel Mach-Zehnder modulator*) que modula tanto a componente em fase como em quadratura de cada uma das polarizações da saída de um *laser* diodo de onda contínua (CW, do inglês *continuous wave*). As duas polarizações são posteriormente combinadas usando um combinador de feixe de polarização (PBC, do inglês *polarization beam combiner*). Antes de transmitir o sinal pela fibra, pode se incluir um amplificador EDFA que compense as perdas de inserção dos diferentes dispositivos e permite o que possa ser utilizado para controlar a potência lançada.

O sinal modulado em amplitude e fase, multiplexado em duas polarizações ortogonais é transmitido pela rede de distribuição que, no caso de sistemas ópticos de um único canal e sem amplificação em linha, está composto por um único *span* de fibra monomodo padrão (SSMF, do inglês *standard single mode fiber*). Esta fibra introduz os impedimentos de CD, PMD, e rotação do estado de polarização, assim como distorção não linear causada pelo efeito Kerr.

No *front-end* do receptor, as duas polarizações ortogonais são separadas utilizando um divisor de feixe por polarização (PBS, do inglês *polarization beam splitter*) e cada uma das polarizações é misturada em uma rede híbrida de 90° com a correspondente polarização da saída de um segundo *laser* que atua como oscilador local. Posteriormente, as saídas das redes híbridas são fotodetectadas, filtradas, e amplificadas diferencialmente, obtendo-se sinais elétricos correspondentes às componentes em fase e quadratura de cada

uma das polarizações do sinal recebido. Estes sinais são digitalizados usando conversores analógicos digitais (ADCs, do inglês *analog-to-digital converters*) para serem processados no DSP, sendo os impedimentos compensados e os sinais demodulados. Vale ressaltar que as frequências dos *lasers* do transmissor e receptor não são as mesmas, o que difere os sistemas coerentes digitais atuais dos sistemas coerentes primigênicos que necessitavam de uma malha de captura de fase (PLL, do inglês *phase-locked loop*) (KIKUCHI, 2015).

2.2 Impedimentos em sistemas coerentes digitais

Os diferentes efeitos que degradam o sinal em redes ópticas coerentes podem ser classificados dependendo se satisfazem a propriedade de linearidade. Assim, os impedimentos podem ser classificados em impedimentos lineares e não lineares. Desde o ponto de vista prático, podemos definir os impedimentos lineares como aqueles que causam uma distorção no sinal que não depende da potência da mesma.

2.2.1 Impedimentos lineares em sistemas coerentes digitais

Os impedimentos lineares, por sua vez, podem ser divididos dependendo de se são causados pela fibra ou pelo transmissor/receptor. Os efeitos lineares induzidos pela fibra são a CD, a PMD e as flutuações do estado de polarização (TSUKAMOTO *et al.*, 2005). A CD é o nome dado ao fenômeno em que diferentes componentes espectrais viajam em diferentes velocidades dentro da fibra óptica. Esta distorção causa um alargamento do pulso² do sinal enviado, o que dará origem a interferência intersimbólica (ISI, do inglês *intersymbol interference*). Esta interferência pode dificultar a interpretação do sinal e, conseqüentemente, diminuir a distância entre os regeneradores do enlace, tornando-os menos eficientes e aumentando o seu custo de implementação (GREEN; MITRA; WEGENER, 2003). Por outro lado temos a PMD e as flutuações do estado de polarização. Apesar dos modernos processos de fabricação das fibras ópticas, podem haver imperfeições na circularidade dos núcleos destas, dando origem a um efeito chamado de birrefringência (AGRAWAL, 2012). Esta birrefringência pode ser ainda mais significativa devido ao estresse mecânico que

² Cabe destacar que, neste contexto, o termo pulso pode ser interpretado como um pulso na envoltória complexa, isto é, não necessariamente indica modulação de intensidade, senão também é aplicável a modulação de fase ou modulação conjunta de amplitude e fase

sofre a fibra e varia espacialmente ao longo do eixo de propagação. Esse efeito faz com que os dois modos com polarizações ortogonais que teoricamente são degenerados apresentem constantes de propagação locais diferentes, o qual causa uma dispersão do pulso denominada PMD. A PMD tem significativa importância em sistemas com taxas de transmissão de símbolos acima de 40 Gbaud e com longos comprimentos de enlaces (SUNNERUD *et al.*, 2002). Nos sistemas em que há a diversidade de polarização, a birrefringência da fibra pode causar outro efeito, a rotação do SoP. Neste caso, em sistemas coerentes com multiplexação de polarização em que diferentes informações são transmitidas em duas polarizações ortogonais do campo elétrico, a birrefringência causa uma flutuação do SoP que varia no tempo, levando a uma interferência linear entre estas informações moduladas nas duas polarizações originais. Assim, ambos os efeitos oriundos da birrefringência da fibra podem diminuir a eficiência dos enlaces e devem ser mitigados a fim de amenizar os efeitos destes impedimentos (AGRAWAL, 2012).

Quanto aos impedimentos lineares causados pelo transmissor e receptor, temos os ruídos tanto de fase como de amplitude, o *offset* frequencial entre transmissor e receptor, impedimentos causados por imperfeições de quadratura, assim como os erros na digitalização que incluem erro de quantização e sincronização de relógio. Como já mencionado, nos sistemas coerentes o sinal recebido nos sistemas coerentes digitais é combinado com a saída de um *laser* que atua como oscilador local. Idealmente, este oscilador local deveria ter a fase correlacionada com o oscilador usado no processo de modulação. No entanto, atingir isso no domínio analógico mediante controle opto-eletrônico é bastante complexo, devido aos ruídos de fase dos *lasers* de semicondutor e das flutuações do estado de polarização do sinal recebido. Assim, em sistemas coerentes digitais, o oscilador local opera em condições de *free running* que tem um controle de frequência independente do sinal recebido. Em consequência, os osciladores de sinal e local apresentam fases não correlacionadas e, inclusive, uma diferença de frequência denominada de *offset* frequencial. Tanto o ruído de fase como o *offset* frequencial causam uma rotação da constelação, mas podem ser compensados no domínio digital (RIBEIRO *et al.*, 2011). Além do ruído de fase causado pelos *lasers*, o sinal também é afetado pelo ruído aditivo dos fotodetetores. Este ruído aditivo tem duas contribuições, por um lado o ruído térmico e por outro, o ruído de disparo³ (AGRAWAL, 2012). Por sua natureza aleatória e o curto tempo

³ É bem conhecido que a densidade de potência do ruído térmico de um fotodetector é independente da intensidade óptica recebida, enquanto que a densidade de potência do ruído de disparo é proporcional

de coerência, estes ruídos não podem ser eliminados no domínio digital e representam um impedimento limitante em sistemas coerentes digitais. Outros impedimentos que ocorrem, tanto no transmissor como no receptor, são o desbalançamento e desalinhamento das quadraturas. Estes efeitos acontecem quando as componentes em fase e quadratura apresentam diferentes perdas ou quando não tem uma diferença de exatamente 90° entre elas. Como resultado, a constelação adquire uma forma romboidal. Finalmente, além dos impedimentos previamente discutidos, a conversão analógico digital prévia ao DSP também pode causar distorção no sinal recebido. Se por um lado, a quantização introduz um erro aditivo, por outro lado, em sistemas de alta taxa, nos que a frequência de amostragem excede os 20 GSps, a estabilidade da frequência dos relógios apresenta um sério desafio, tendo que ser compensado no receptor. Por exemplo, o padrão de comunicações 400ZR estabelece um desvio rms de frequência de relógio máximo de 200 MHz (OIF, 2020).

2.2.2 Impedimentos não lineares em sistemas coerentes digitais

Nos sistemas ópticos coerentes, o impacto do espalhamento estimulado Brillouin está condicionado ao formato de modulação utilizado e à taxa de bits. Neste caso, a principal contribuição está relacionada ao espalhamento de onda acústica guiada de Brillouin, que não é tão significativa devido à interação eletrostritivas (DIANOV *et al.*, 1992). Ainda nestes sistemas, o espalhamento Raman possui uma influência muito menor do que a existente em sistemas IM/DD, pois a transferência de potência induzida por Raman depende somente da potência do canal WDM. Geralmente, as não linearidades de espalhamento não tem um impacto significativo nos sistemas coerentes digitais. Assim, o efeito Kerr é comumente identificado como a principal causa dos impedimentos não lineares no enlace e estabelece um limite superior à potência máxima transmitida pelos sistemas coerentes (ELLIS *et al.*, 2017).

O efeito Kerr é um efeito eletro-óptico em que o índice de refração de um material depende da intensidade óptica, isto é, do quadrado da amplitude do campo elétrico incidente. Este efeito resulta em três distorções que são conhecidas como automodulação

a esta. Desta forma, poderia parecer que a potência de ruído de disparo dependeria da potência do sinal transmitido e, por conseguinte, não seria um efeito linear. Porém, em sistemas ópticos de um único canal sem amplificação em linha, a potência do oscilador local é significativamente maior que a potência do sinal recebido e, conseqüentemente, a potência de ruído de disparo é principalmente governada pela potência do oscilador local e não pela potência do sinal recebido.

de fase (SPM, do inglês *self-phase modulation*), modulação de fase cruzada (XPM, do inglês *cross-phase modulation*) e mistura de quatro ondas (FWM, do inglês *four-wave mixing*) (AGRAWAL, 2000). Tanto a SPM como a XPM causam uma modulação de fase num sinal propagante que depende da intensidade do mesmo e dos sinais copropagantes respectivamente. Pode-se interpretar como uma interferência multiplicativa. Por outro lado, FWM induz um interferência aditiva não linear entre os diferentes sinais.

O primeiro dos efeitos, a SPM, trata-se de um deslocamento de fase autoinduzido por um campo elétrico que ocorre durante a propagação em um meio não linear. Assim, quando se tem um segmento de comprimento igual a L , considerando uma potência constante, podemos encontrar o deslocamento de fase da seguinte forma:

$$\phi = k_0 L(n + n_2 I), \quad (1)$$

em que $k_0 = \omega/c = 2\pi/\lambda$, n simboliza o índice de refração linear de fase, n_2 o coeficiente não linear, I é a intensidade do sinal, ω é a frequência angular e λ é o comprimento de onda. Assim, o deslocamento de fase não linear resultante de SPM é (AGRAWAL, 2000):

$$\phi_{NL} = n_2 k_0 L I = \gamma L P, \quad (2)$$

quando expressamos em função da potência lançada. Sendo que $\gamma = n_2 k/A_{eff}$.

Para considerar a atenuação do sinal propagante, uma aproximação comumente utilizada considera o comprimento efetivo da fibra, L_{eff} , que leva em conta tanto a atenuação da fibra como o comprimento da mesma. O comprimento efetivo é calculado pela seguinte equação:

$$L_{eff} = \frac{1 - \exp(-\alpha L)}{\alpha}, \quad (3)$$

sendo α é o coeficiente de atenuação da fibra. Desta forma, a expressão para o deslocamento de fase não linear adquire a seguinte expressão:

$$\phi_{NL} \approx \gamma L_{eff} P. \quad (4)$$

Como pode-se observar, o deslocamento de fase não linear depende linearmente da potência óptica instantânea. Em ausência de CD, o deslocamento de fase não linear segue a forma do pulso. Já quando temos uma CD significativa, ocorre a interferência inter-simbólica, que faz com que os símbolos adjacentes interajam de maneira não linear.

O segundo efeito, a XPM, só ocorre quando múltiplos sinais são multiplexados e transmitidos simultaneamente pelo meio. Desta forma, a XPM não acontece unicamente

em sistemas com multiplexação em comprimento de onda mas também em sistemas com multiplexação em polarização, como os enlaces ópticos estudados neste trabalho. A expressão do deslocamento de fase não linear induzido em um sinal por outro sinal de potência P_2 por meio de XPM é dada por:

$$\phi_{NL1} = 2\gamma L_{eff} P_2. \quad (5)$$

Nota-se que a única diferença entre as equações para XPM (4) e SPM (5) é um fator 2, adicionado na XPM. Esta inclusão tem como objetivo considerar a degenerescência presente na XPM (AGRAWAL, 2000).

Já o terceiro efeito, a FWM, quando é considerado o caso degenerado, até quatro ondas interatuam e transferem energia entre si. Como acontece com XPM, FWM pode acontecer tanto em sistemas com multiplexação de comprimento de onda como em sistemas com multiplexação de polarização. Quando se é verificado o princípio da conservação de energia, as frequências dos sinais estão relacionadas por $\omega_4 = \omega_1 \pm \omega_2 \pm \omega_3$, em que ω_i é a frequência do i -ésimo sinal. A expressão pode ser substituída para o caso parcialmente degenerado ($\omega_3 = 2\omega_1 \pm \omega_2$) e para o caso completamente degenerado ($\omega_2 = 3\omega_1$) (AGRAWAL, 2012). É preciso notar também que a eficiência do processo de FWM depende do casamento de fase entre as diferentes componentes que interagem. Em sistemas coerentes, em que a rede de distribuição geralmente não inclui fibra compensadora de dispersão, considera-se que a FWM entre os diferentes canais WDM não é tão eficiente como em sistemas que utilizam detecção direta (KIKUCHI, 2015). Porém, em sistemas com multiplexação de polarização, as duas polarizações praticamente apresentam a mesma constante de propagação e, em consequência, apresentam um casamento de fase muito alto. Desta forma, é esperado que o processo de FWM entre as duas polarizações de um canal seja altamente eficiente (AGRAWAL, 2000).

O peso relativo e o impacto de SPM, XPM e FWM no desempenho do sistema está governado por diferentes parâmetros operacionais e estruturais. Além do formato de modulação e o filtro utilizado, o efeito das não linearidades dependerá se são utilizados múltiplos canais WDM ou se apenas um único canal é transmitido. Vale lembrar também que os impactos de cada um destes efeitos podem variar de acordo com a polarização única ou dupla do sinal. Conforme estes parâmetros, podemos agrupar os sistemas em três classes: sistemas mono-canal de uma única polarização (por exemplo, sistemas auto-homódinos (PUTTNAM *et al.*, 2014)), sistemas mono-canal de dupla polarização (como os

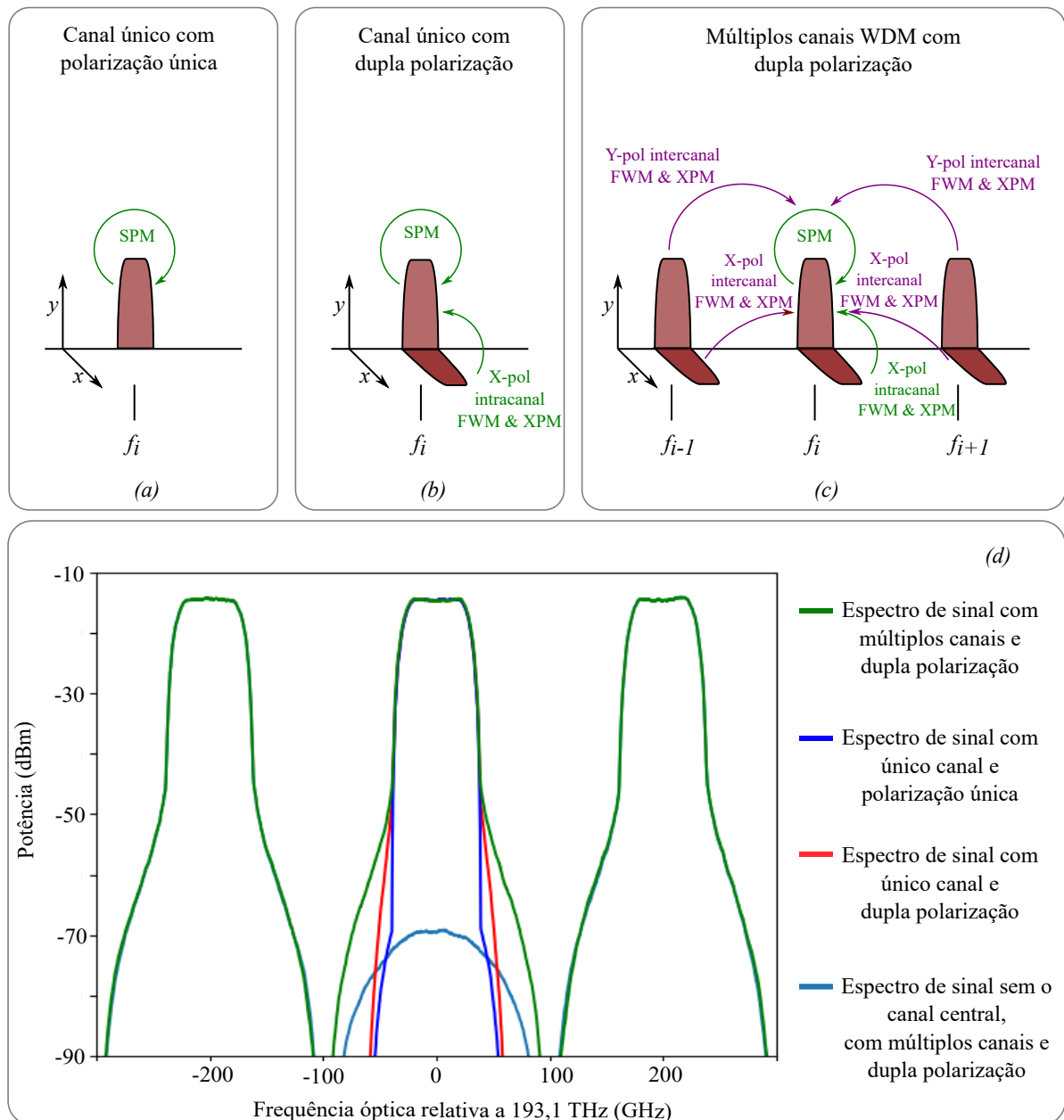


Figura 2 – Distorções não lineares em diferentes configurações de sistemas coerentes digitais. (a) Canal único com polarização única; (b) Canal único com dupla polarização; (c) Múltiplos canais WDM com dupla polarização. (d) Espectros dos sinais com diferentes configurações do sistema.

sistemas ópticos estudados neste trabalho) e sistemas multi-canal de dupla polarização. Os sistemas multi-canal de uma única polarização não foram incluídos por não representarem uma alternativa utilizada atualmente.⁴

⁴ Cabe mencionar que estes sistemas foram utilizados para minimizar FWM em sistemas WDM com detecção direta mas, no caso de sistemas coerentes, não são utilizados (SECONDINI; FORESTIERI, 2014).

A Figura 2 mostra as diferentes fontes de distorção nos três tipos de sistemas mencionados. A Figura 2(a) ilustra a situação que se tem apenas um canal em um sistema com polarização única. Neste primeiro cenário, temos somente o efeito da SPM no enlace. Quando se tem uma multiplexação com duas polarizações perpendiculares de um sinal aliado a utilização de um canal único, como mostrado na Figura 2(b), além do efeito da SPM, temos o efeito da XPM e FWM intracanal causados pela polarização cruzada. A situação fica mais complexa em sistemas WDM com múltiplos canais e dupla polarização. Neste caso, as não linearidades podem acontecer devido às seguintes influências: SPM causada pelo próprio canal; XPM e FWM intracanal devido a polarização cruzada; XPM e FWM intercanal causadas pela polarização cruzada entre diferentes canais com a mesma polarização; XPM e FWM intercanal causadas pela polarização cruzada dos outros canais WDM (EKANAYAKE; H.HERATH, 2013).

Tipo de Sistema	Distorção intracanal		Distorção intercanal	
	Y-polarização	X-polarização	Y-polarização	X-polarização
Monocanal com polarização única	SPM	—	—	—
Monocanal com dupla polarização	SPM	XPM & FWM	—	—
Multi-canal com dupla polarização	SPM	XPM & FWM	XPM & FWM	XPM & FWM

Tabela 1 – Tabela resumo com os sistemas coerentes e os efeitos não lineares representados na Figura 2.

Assim, a Tabela 1 resume estes cenários e o impacto de cada um destes para os sistemas coerentes digitais. Para poder ter uma ideia do peso relativo de cada fonte de distorção, na Figura 2(d) mostramos os espectros para diferentes configurações: canal único com polarização única, canal único com dupla polarização e múltiplos canais WDM com dupla polarização. Neste último caso, o sistema foi simulado com e sem o sinal de interesse para poder identificar a contribuição da interferência causada pela FWM. Em todos os casos, considerou-se uma potência por polarização em cada canal de 10 dBm, taxa de transmissão de símbolo de 56 Gbaud, um filtro de Nyquist de fator de *roll-off* de 0,2 e uma banda de resolução de 1,25 GHz. No espectro do sinal com canal único e única polarização (curva azul escuro), apenas a SPM está presente. Pode-se observar que o alargamento espectral é relativamente pequeno. A curva em vermelho representa o espectro

de um sinal com um único canal e dupla polarização, no qual há a presença da SPM, devido à própria polarização, e da XPM e FWM intracanalais, devido aos efeitos gerados pela outra polarização do sinal. Como se pode observar, o alongamento espectral é maior do que quando uma única polarização é considerada. O espectro na cor azul claro mostra a potência da interferência intrabanda causada pelos outros canais (ambas as polarizações). Por fim, a curva em verde, que exibe o espectro do sinal com múltiplos canais e dupla polarização, mostra o espectro do sinal afetado pela SPM, XPM e FWM intracanal e XPM e FWM intercanal. Comparando estes espectros podemos concluir que no caso de múltiplos canais WDM, a distorção não linear dominante é a XPM intercanal. Já no caso de canal único, as fontes de distorção não linear mais significativas são as XPM e FWM intracanal causadas pela polarização cruzada.

2.3 Mitigação de impedimentos em sistemas coerentes digitais

2.3.1 Mitigação de impedimentos lineares em sistemas coerentes digitais

A mitigação dos efeitos lineares nos sistemas ópticos coerentes digitais foi estudada amplamente nos últimos anos. O resultado destes estudos permitiu a implementação de algoritmos de mitigação dos impedimentos altamente efetivos em sistemas coerentes por meio de DSP. Várias estruturas, sequências e algoritmos foram propostos visando otimizar a performance do sistema (SAVORY, 2008). A Figura 3 mostra uma estrutura que é comumente utilizada.

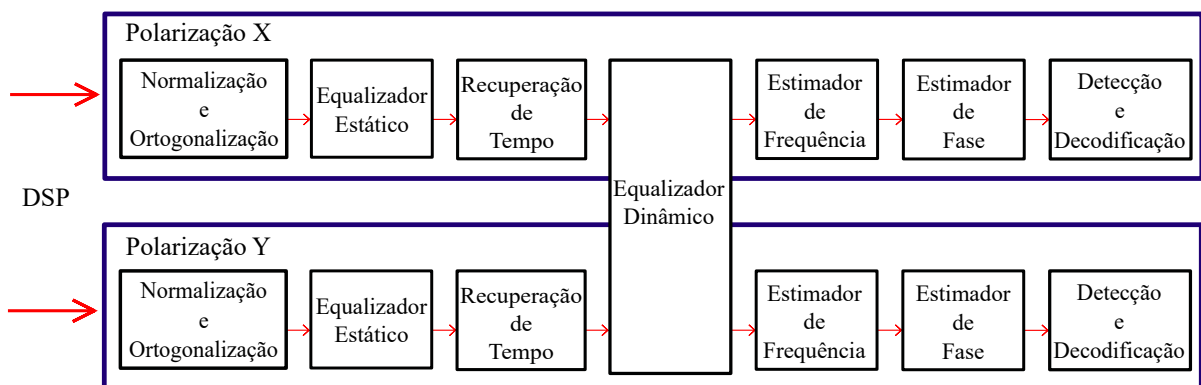


Figura 3 – Diagramas de blocos com a sequência das principais funcionalidades dos DSPs para sistemas coerentes digitais.

Como indicado na seção anterior, um dos impedimentos lineares que afeta a qualidade do sinal é a imperfeição na rede híbrida de 90° que resulta em desbalanço de potência das componentes em fase e quadratura, assim como uma quebra da ortogonalidade entre estas. Para compensar estes impedimentos, é necessário um bloco denominado *Normalização e Ortogonalização* que opera em cada polarização independentemente. Os algoritmos mais amplamente utilizados neste bloco são o algoritmo de Gram-Schmidt e o de Löwdin (PETROU *et al.*, 2009).

O segundo bloco, denominado *Equalizador Estático*, realiza a compensação da CD, podendo esta correção ser realizada no domínio do tempo ou no domínio da frequência (RIBEIRO *et al.*, 2011). As implementações tanto no domínio temporal como frequencial requerem uma taxa de amostragem de ao menos duas amostras por símbolo para satisfazerem ao critério de amostragem de Nyquist. Assim como no caso anterior, este bloco tipicamente é projetado para operar em cada polarização independentemente. Vale ressaltar que a CD pode também ser compensada no domínio óptico, porém este método introduz não linearidades e perdas ao sistemas e além disso, causa um atraso na transmissão do sinal, devido ao aumento do enlace quando se utiliza da fibra compensadora. Estas desvantagens, fazem com que a utilização de blocos de compensação no domínio elétrico sejam mais vantajosos.

O terceiro bloco, *Recuperação de Tempo de Relógio*, compensa a incerteza temporal (*timing jitter*) causada pela diferença de frequência entre os relógios do transmissor e receptor. Neste bloco são aplicados algoritmos que possibilitam a recuperação da fase e da frequência do relógio, os quais podem ser divididos em: algoritmos auxiliados por dados (MUELLER; MULLER, 1976) e algoritmos não auxiliados por dados (GARDNER, 1986), sendo estes últimos mais utilizados por apresentarem uma operação dinâmica sem comprometer a capacidade do sistema. Podemos também dividir estes algoritmos de acordo com o tipo de amostragem, podendo esta ser síncrona ou assíncrona. Na amostragem síncrona, o amostrador é sincronizado com o sinal que está sendo recebido, ao passo que na amostragem assíncrona a taxa de amostragem pode ser diferente da taxa de símbolos, não havendo dependência do relógio local fixo com o sinal recebido (MENGALI, 2013).

A compensação da interferência inter-polarização e da PMD é realizada no quarto bloco, denominado *Equalizador Dinâmico*, em que o algoritmo tem como objetivo encontrar um filtro adaptativo que possua uma resposta inversa a estes impedimentos do canal. Entre as técnicas comumente utilizadas estão os filtros adaptativos em borboleta com

otimização de pesos por minimização do erro quadrático médio. Este bloco, diferentemente dos anteriores, opera sobre as duas polarizações simultaneamente. Posteriormente se inclui um bloco de *Subamostragem* que reduz a frequência de amostragem a uma amostra por símbolo. Além de subamostrar, este bloco é responsável pela sincronização temporal do receptor, pois seleciona o instante ótimo de amostragem para otimizar o desempenho do receptor. O algoritmo geralmente utiliza algum tipo de métrica baseada em multi-módulo. Isto é, o algoritmo procura o instante que minimize a variação da amplitude dentro de cada nível.

Nos blocos *Estimador de Frequência* e *Estimador de Fase* são estimados os valores do desvio de frequência entre o *laser* de sinal e o *laser* local (*offset* de frequências) e de fase da portadora (SAVORY, 2008). O primeiro, pode ser implementado tanto no domínio temporal como no frequencial, sendo o primeiro mais utilizado. Enquanto ao algoritmo de estimação da fase, pode ser utilizada a denominada busca cega de fase.

Por fim, no bloco *Decisão e Decodificação* são aplicados algoritmos que auxiliam na decisão entre cada símbolo enviado de acordo com a modulação utilizada, e assim, há a decodificação para conclusão do processamento do sinal.

2.3.2 Mitigação de impedimentos não lineares em sistemas coerentes digitais

Diferentemente do que ocorre com os impedimentos lineares, a compensação de efeitos não lineares não tem sido utilizada comercialmente, ainda estando limitada à pesquisa. Alguns pesquisadores propõem a utilização de métodos ópticos para a mitigação das não linearidades. Dentre eles, podemos citar os trabalhos de (DU; MORSHED; LOWERY, 2012), (HO, 2005) e (LIU *et al.*, 2013). (DU; MORSHED; LOWERY, 2012) propõe conjugação de fase óptica (OPC, do inglês *optical phase conjugation*) para a compensação de distorções não lineares em supercanais multi-portadora. No segundo, (HO, 2005), é estendido o estudo a N conjugadores ópticos, de forma que os resultados ótimos são alcançados quando estes estão espaçados na fração de $2/(2N+1)$ do enlace. Na técnica proposta por (LIU *et al.*, 2013), há um cancelamento das não linearidades por meio da transmissão de um par de ondas gêmeas conjugadas em fase (PCTW, do inglês *phase conjugated twin waves*), que são sobrepostas simultaneamente. Todos estes métodos, com características comuns de

utilizar meios ópticos para mitigação das não linearidades, são complexos, e este fato dificulta sua implementação em campo.

Por outro lado, nos últimos tempos tem-se proposto uma grande variedade de técnicas digitais para mitigar o efeito das distorções. Estas ferramentas aproveitam-se de toda a flexibilidade que a eletrônica digital oferece e podem ser divididas em métodos baseados em inversão de modelo e métodos que usam de algoritmos fundamentados em inteligência artificial. Quanto às ferramentas de inversão de modelo, podemos citar o modelo Wiener-Hammerstein, que é capaz de mitigar os efeitos não lineares por meio de uma combinação entre um elemento não linear estático com dois elementos lineares dinâmicos. Este modelo apresenta uma grande simplicidade de implementação e já foi analisado em sistemas de redes metropolitanas (MHATLI *et al.*, 2015), em enlaces com formatos multiníveis de modulação (MILLAR *et al.*, 2010) e também para redes de acesso (TORRES-ZUGAIDE *et al.*, 2017)(TORRES-ZUGAIDE *et al.*, 2016). No entanto, verificou-se que o modelo de Wiener-Hammertein não possui boa eficiência em redes *multispan*, em que a distorção está presente em locais distintos dentro do enlace. Outra técnica amplamente utilizada é o algoritmo de retro-propagação digital (DBP, do inglês *digital back-propagation*) (IP; KAHN, 2008) (PAKALA; SCHMAUSS, 2015). Este algoritmo reverte os efeitos de propagação na fibra de maneira digital. Tendo em vista os parâmetros da fibra e do sinal, o algoritmo é capaz de inverter a distorção sofrida pelo sinal durante a transmissão no canal óptico. Uma outra técnica de inversão de modelo estudada por vários grupos utiliza da função de transferência inversa da série Volterra (IVSTF, do inglês *inverse volterra series transfer function*), que é aplicada para construir um equalizador não linear reconfigurável (GIACOUMIDIS *et al.*, 2013). Todavia, nos casos que se utiliza de inversão de modelo de alta eficiência, i.e. DBP e IVSTF, o custo computacional é muito elevado, o que dificulta sua aplicação em sistemas reais.

Recentemente, alguns modelos que se utilizam de algoritmos de aprendizagem de máquinas foram propostos. Com a capacidade de processamento dos atuais DSPs, algoritmos sofisticados têm sido estudados a fim de se mitigar os efeitos das não linearidades em sistemas ópticos coerentes digitais. Estas técnicas são apresentadas e discutidas na seguinte seção.

2.4 Mitigação de impedimentos utilizando aprendizado de máquinas

Desde 1959, quando o pesquisador americano Arthur Samuel (1901-1990) definiu o termo aprendizado de máquinas como o novo ramo da engenharia que estudaria a capacidade dos computadores aprenderem, sem terem sido programados para esse fim, e os inseriu dentro do campo da inteligência artificial (MCCARTHY; FEIGENBAUM, 1990), o conceito despertou o interesse de empresas e pesquisadores de todo o mundo. Atualmente, está presente nas mais diversas áreas do conhecimento e tem sido amplamente utilizado em pesquisas e desenvolvimento de soluções para melhorias nas redes de comunicação óptica, principalmente envolvendo sistemas coerentes digitais (ZIBAR *et al.*, 2015). A ciência explora a construção de algoritmos que aprendem com seus erros e podem fazer previsões sobre dados e resultados, a partir de três métodos de aprendizagem: supervisionada, não supervisionada e por reforço (ALPAYDIN, 2014).

No método de aprendizagem supervisionado, o algoritmo é capaz de rotular novos dados não rotulados a partir de um treinamento com dados que possuem rótulos conhecidos (ALPAYDIN, 2014). Este tipo de aprendizado pode implementar uma classificação ou uma regressão. A classificação associa um rótulo de saída discreto para um dos elementos de entrada, rotulando os dados em classes. Já a técnica baseada em regressão possui uma saída contínua. Assim, podemos sintetizar que, se o rótulo é um número real, temos uma regressão. Caso o rótulo seja um conjunto finito e não ordenado, temos a classificação (MARSLAND, 2015). Dentre os algoritmos de aprendizagem supervisionada, podemos citar: K-NN (*k-nearest neighbors*) (FUKUNAGA; NARENDRA, 1975), máquina de vetores de suporte (SVM, do inglês *support vector machine*) (CORTES; VAPNIK, 1995), redes neurais artificiais (ANN, do inglês *artificial neural networks*) (HASSOUN *et al.*, 1995) e regressão por vetores de suporte (SVR, do inglês *support vector regression*) (MARSLAND, 2015)

A aprendizagem por reforço está entre a aprendizagem supervisionada e a não supervisionada. O ponto central entre estes dois termos é o conceito de que o algoritmo recebe a informação se sua resposta está correta, no entanto, não recebe a informação de como melhorá-la. A aprendizagem por reforço é pautada na interação entre algum agente e o seu ambiente (MARSLAND, 2015). O agente, neste caso, é o que está aprendendo, e o seu ambiente é que traz a informação de como ele está aprendendo e o quanto este aprendizado

tem sido eficaz, por meio de recompensas. O algoritmo baseado em reforço experimenta estratégias diferentes e verifica qual destas alternativas maximiza o resultado, sendo esta etapa conhecida como fase de pesquisa. Os algoritmos mais populares que utilizam este tipo de aprendizado são *Q-Learning* (WATKINS; DAYAN, 1992), SARSA (SUTTON; BARTO *et al.*, 1998) e DYNA (SUTTON, 1991).

Por fim, a aprendizagem não supervisionada dispensa qualquer necessidade de intervenção prévia para que se tenha o aprendizado, não sendo necessário que nenhum rótulo seja inicialmente estipulado. Neste método, o algoritmo só pode explorar as semelhanças nos dados para agrupá-los a fim de encontrar as similaridades e anomalias comuns aos dados (ALPAYDIN, 2014). Este método de aprendizagem vem sendo muito utilizado, principalmente em problemas que possuem um conjunto muito grande de dados a serem classificados. Isso ocorre justamente pela não necessidade da etapa de sequência de treinamento, o que dispenderia um custo computacional muito elevado. São exemplos de aprendizado não supervisionado para *clusterização*, os algoritmos K-means, EM, *clusterização* espacial baseada em densidade de aplicações com ruído (DBSCAN, do inglês *Density-based spatial clustering of applications with noise*) e propagação de afinidade (AP, do inglês *Affinity Propagation*) (MARSLAND, 2015).

Devido ao baixo custo computacional e a flexibilidade dos algoritmos que envolvem inteligência artificial, muitos estudos têm aplicado estas ferramentas para mitigação de não linearidades em sistemas ópticos coerentes digitais. Com a alta performance dos atuais DSPs, a velocidade no processamento tem possibilitado a implementação dos mais sofisticados algoritmos em sistemas com taxas de transmissão cada vez maiores. Como abordado anteriormente, os métodos de mitigação no domínio óptico possuem uma alta complexidade e são inviáveis para aplicações comerciais (ROS *et al.*, 2014). As técnicas de compensação eletrônicas baseadas em inversão de modelo requerem um alto número de iterações, o que acarreta um elevado custo computacional aos dispositivos eletrônicos, gerando alta dissipação de potência e um elevado tempo de latência (LIN *et al.*, 2012) (GIACOU MIDIS *et al.*, 2013) (PAN *et al.*, 2011). Como alternativa, foi proposta em alguns trabalhos a utilização de algoritmos de aprendizagem de máquinas para mitigar o efeito das distorções não lineares em sistemas coerentes digitais. Considerando um algoritmo de aprendizagem supervisionada, (WANG *et al.*, 2015) utilizou SVM para tratar do ruído de fase não linear em um enlace de 480 km com modulação 8-PSK. (JARAJREH *et al.*, 2015) propõe um equalizador baseado em ANN para a mitigação dos efeitos em

um enlace com multiplexação por divisão de frequências ortogonais (OFDM, do inglês *orthogonal frequency-division multiplexing*). (WANG *et al.*, 2016) e (PAULA *et al.*, 2020) utilizam do algoritmo K-NN com o intuito de diminuir os efeitos destas distorções. Todos estes estudos são baseados em algoritmos supervisionados e, conseqüentemente, exigem uma sequência de treinamento para ajuste do modelo de compensação, o que gera um aumento no processamento necessário.

Quando se trata do aprendizado não supervisionado, os algoritmos empregados na mitigação dos ruídos de fase não lineares são baseados em processos de *clusterização*. (ZIBAR *et al.*, 2012) utiliza da maximização de expectativa para contornar os efeitos das não linearidades em um enlace com modulação 16-QAM. (ALDAYA *et al.*, 2020) propõe um novo algoritmo baseado na classificação por meio de um histograma que mitiga o ruído de fase não linear causado pela SPM. Alguns trabalhos como: (ZHANG WEI CHEN; SHEN, 2017), (FERNANDEZ *et al.*, 2016) e (BORGES *et al.*, 2020) se basearam no algoritmo K-means para amenizar os efeitos nocivos nos enlaces ópticos coerentes, sempre utilizando-os em sistemas de polarização única e para taxas de transmissão de dados abaixo de 100 Gbps. Ainda nesta linha de pesquisa, (LIU; YANG; XU, 2019) propõe o uso do algoritmo DBSCAN para a mitigação de alguns efeitos em sistemas ópticos coerentes. Por outro lado, em (ZIBAR *et al.*, 2012), os autores propõem a utilização de EM para combater os efeitos da automodulação de fase em sistemas de um único *span*. Desta forma, os algoritmos de *clusterização* se mostraram ferramentas muito poderosas para compensação de não linearidades em sistemas de uma única polarização e taxas não superiores a 56 Gbps. Porém, não está claro se estes algoritmos continuam apresentando esta eficiência em sistemas mais sofisticados nos quais utiliza-se dupla polarização e a taxa de transmissão é alta. Entre os diferentes algoritmos de *clusterização*, selecionamos K-means e EM pela simplicidade do primeiro e o alto potencial do segundo. Além disso, ambos os algoritmos não precisam de sintonização de parâmetros como DBSCAN e AP, processo este que pode ser desafiador em sistemas em que as condições variam.

2.4.1 K-means

O algoritmo de *clusterização* K-means é uma ferramenta de aprendizagem de máquinas não supervisionada que possui como qualidades a simplicidade e a eficiência. O

método proposto por (LLOYD, 1982) classifica as entradas em K grupos automaticamente, minimizando a distância média entre cada dado e um conjunto de K pontos, denominados centroides. No método K-means, cada *cluster* é associado a um destes centroides, sendo este centroide o ponto de verificação que se mensura a distância entre todos os símbolos a serem classificados, e sua classificação ocorrerá de acordo com a distância média mínima encontrada. Na versão clássica de K-means, o algoritmo inicializa aleatoriamente a escolha dos K centroides e, posteriormente, irá refinando a localização de cada centroide de acordo com os seguintes passos (MARSLAND, 2015):

- **Passo 1:** Cada dado é classificado, associando estes ao centroide que possui menor distância, obtida pela equação:

$$d_i = \min\{d(x_i, \mu_j)\} \quad \forall i \text{ e } 1 \leq j \leq K, \quad (6)$$

sendo $d(\cdot)$ representa o operador de distância (tipicamente é usada a distância euclidiana). Os pontos que são associados a um dado centroide formam um *cluster*;

- **Passo 2:** É calculada a localização do novo centroide, baseando-se no novo aglomerado formado pelos novos pontos. A posição do novo centroide é calculada pela seguinte equação:

$$\mu'_j = \frac{1}{N_j} \sum_{i=1}^{N_j} x_{i,j} \quad \text{para } 1 \leq j \leq K, \quad (7)$$

sendo, N_j e $x_{i,j}$ são, respectivamente, o número de elementos associados ao *cluster* j e os elementos desse *cluster*;

- **Passo 3:** Com os centroides calculados μ'_j , é verificado se estão nas mesmas localizações que os centroides anteriores μ_j ou com distância menor do que a precisão estipulada para o algoritmo. Se esta condição é atingida, o algoritmo é encerrado. Caso não seja atendida, são atualizados os centroides, igualando $\mu_j = \mu'_j$ e retoma-se ao passo 1.

Seguidos todos estes passos, no final do procedimento os dados são subdivididos em k agrupamentos que foram ajustados de acordo com os próprios dados de entrada, sem a necessidade de um rotulamento inicial. Apesar de ter uma implementação relativamente simples, a complexidade geral do algoritmo aumenta de acordo com a quantidade de dados a serem classificados e com o número de agrupamentos (*clusters*) a serem considerados. Outro ponto que merece destaque é a natureza estocástica do método, pois o resultado final

depende da inicialização dos centroides. Assim, os centroides podem convergir para um mínimo local, especialmente quando se trata de um número elevado de *clusters* (ARTHUR; VASSILVITSKII, 2006). Uma das soluções para esse problema é a repetição do processo com diferentes valores iniciais aleatórios dos centroides, mas essa abordagem aumenta significativamente o custo computacional.

Diante dessa dificuldade, alguns métodos de inicialização mais inteligentes foram propostos, dos quais podemos citar: K-means hierárquico (ARAI; BARAKBAH, 2007), método global de K-means (LIKAS; VLASSIS; VERBEEK, 2003), algoritmo de inicialização do centro de *cluster* (KHAN; AHMAD, 2004), método de busca de *cluster* simplificado (TOU, 1979) e K-means++ (ARTHUR; VASSILVITSKII, 2006).

Proposto em 2006, K-means++ é amplamente adotado pois, além de possuir uma boa eficiência, permite uma grande flexibilidade em diferentes tipos de aplicações. Os autores desenvolveram o algoritmo buscando maximizar a probabilidade de que os centroides iniciais estivessem o mais próximo possível dos locais ótimos. Desta maneira, o consumo computacional e, conseqüentemente, a latência do algoritmo seriam reduzidos. Nesta técnica de inicialização, os centroides iniciais são escolhidos um a um, até que os k centroides são obtidos, e só assim a classificação dos dados iniciaria. Para isso, o processo possui os seguintes passos (ARTHUR; VASSILVITSKII, 2006):

- **Passo 1:** É feita uma escolha aleatória do primeiro centroide μ_1 dentro do conjunto de dados que serão classificados x_i ;
- **Passo 2:** É encontrada a distância d_i de todos os pontos de dados x_i para o(s) centroide(s) já escolhido(s);
- **Passo 3:** O próximo centroide μ_j é escolhido de forma aleatória com uma probabilidade proporcional à distância até os centroides já selecionados conforme a:

$$P = \frac{[d_i(x)]^2}{\sum_{\forall i} [d_i(x)]^2}; \quad (8)$$

- **Passo 4:** São repetidas as etapas 2 e 3 até encontrar as posições iniciais dos k centroides.

Com a implementação deste novo método de inicialização, há uma redução significativa no número de iterações necessárias para que o algoritmo convirja ao ponto ótimo e, além disso, os erros de classificação causados pela má inicialização aleatória dos centroides iniciais podem ser diminuídos consideravelmente. Estas características tornam

a ferramenta K-means com a inicialização K-means++ muito utilizada nas mais diversas áreas do conhecimento.

2.4.2 Maximização de Expectativa

O algoritmo de EM, proposto por (DEMPSTER; LAIRD; RUBIN, 1977), é um método que visa encontrar os parâmetros de um conjunto de distribuições gaussianas para maximizar a expectativa. Esta ferramenta é amplamente utilizada em situações em que os dados do problema não são inteiramente conhecidos. Com isso, a EM permite a associação de um problema com dados incompletos a um problema com dados completos, reformulando o problema em termos dos dados completos, mediante a aplicação da verossimilhança entre os dois problemas (DEMPSTER; LAIRD; RUBIN, 1977). O algoritmo lida com os dados de um problema de dados incompletos, substituindo as inconsistências por valores estimados. Vale ressaltar que a EM pode ser aplicada não só em problemas em que a ausência de dados é clara, mas é utilizada também numa grande variedade de situações nas quais a falta de dados não é tão clara. Isso faz com a solução do problema seja calculada de uma maneira mais simples, estimando os valores por meio da verossimilhança dos dados e, posteriormente, sendo maximizados os resultados. (DEMPSTER; LAIRD; RUBIN, 1977) definiram os dados conhecidos como informações que são fornecidas diretamente (y) e os dados incompletos como aqueles que não podem ser observados diretamente (x). Para melhor entendimento destas definições, podemos analisar dois espaços amostrais X e Y e um mapeamento de X para Y . Neste exemplo, temos uma relação direta entre os dados observados y para com Y , já os dados correspondentes de x , não são diretamente ligados a X , sendo estes correlacionados indiretamente por meio de y . Assume-se que o mapeamento entre as variáveis é $x \rightarrow y(x)$ de X para Y e que x é conhecido somente pela função $X(y)$. O subconjunto de X é estipulado pela equação $y = y(x)$, sendo que y são os dados conhecidos. A função equivalente dos dados incompletos $f(x|\theta)$ e dos dados completos $g(y|\theta)$, relacionam-se pela seguinte equação:

$$g(y|\theta) = \int_{X(y)} f(x|\theta) dx. \quad (9)$$

Sendo assim, o algoritmo EM busca encontrar o valor de θ que maximiza $g(y|\theta)$ por meio da função $f(x|\theta)$ e dos dados conhecidos y . Com isso, podemos definir a função esperança condicional da log-verossimilhança $Q(\theta|\theta^{(t)})$ da seguinte maneira:

$$Q(\theta|\theta^{(t)}) = E[\ln(\theta|y, \theta^{(t)})] \quad (10)$$

A função esperança condicional da log-verossimilhança $Q(\theta|\theta^{(t)})$ é calculada com o intuito de estimar os dados não observados, permitindo criar uma expectativa sobre os mesmos. Num segundo momento, é implementada a estimativa de máxima verossimilhança para a maximização dos parâmetros, usando os dados pré estimados, de acordo com a seguinte equação:

$$\theta^{(t+1)} = \text{arg}g_{\theta}^{\text{max}}Q(\theta|\theta^{(t)}) \quad (11)$$

Conforme demonstrado, o algoritmo iterativo EM se estrutura em duas principais etapas: a expectativa, conhecida como etapa E , e a maximização, chamada de etapa M . Na primeira delas, equacionada pela Equação 10, as variáveis desconhecidas são estimadas utilizando-se dos parâmetros existentes até o momento. No segundo passo, o algoritmo maximiza os parâmetros e produz, assim, recursos para uma nova expectativa, conforme demonstra a Equação 11. Vale ressaltar que a utilização do estimador de máxima verossimilhança para a maximização dos parâmetros é necessário para que se encontre os parâmetros do algoritmo de EM, diferentemente da aplicação do método de detecção por máxima verossimilhança que é um método de referência consolidado na literatura. O algoritmo realiza iterações destes dois passos até que a estimativa do parâmetro tenha convergido à um determinado critério, sendo que, a partir deste ponto, não é necessária a estimativa de novos parâmetros (MARSLAND, 2015).

A convergência do método varia de acordo com o número de dados ausentes e, além disso, não se pode garantir que o algoritmo convirja para ótimos globais, pois o algoritmo pode convergir para os ótimos locais, dependendo da escolha dos parâmetros iniciais do algoritmo. No entanto, devido a sua alta estabilidade numérica, possui a característica de aumentar a probabilidade de dados observados a cada iteração. Por este motivo, o algoritmo de EM é amplamente utilizado em diversas áreas, principalmente na classificação de dados e no reconhecimento de padrões (ALPAYDIN, 2014).

Portanto, enquanto o algoritmo de *clusterização* K-means utiliza de verificações dos pontos médios de cada *cluster* para fixar novos centróides, o algoritmo EM utiliza de

	K-means	EM
Inicialização	Aleatório / K-means++	Inicialização com k distribuições Gaussianas
Definição dos novos centróides	Média	Média, peso e variância
Expectativa	Atribui cada ponto ao centróide mais próximo	Atribui cada ponto a uma distribuição Gaussiana
Maximização	Calcula a média dos novos clusters	Calcula os novos parâmetros para as novas Gaussianas

Tabela 2 – Tabela comparativa entre os algoritmos K-means e EM.

parâmetros de distribuições Gaussianas para verificação de novos centróides e consequentemente a formação de novos *clusters*. Após este ponto, os algoritmos convergem até o ponto em que todos os dados sejam *clusterizados*, e assim o algoritmo seja finalizado. A Tabela 2 sintetiza as principais diferenças entre as características dos algoritmos de clusterização K-means e EM.

3 Arranjo de simulação e métricas de avaliação

3.1 Arranjo de simulação

As simulações foram realizadas de maneira cooperativa, sendo utilizado o software VPI *Transmission Maker* para geração, transmissão, recepção e detecção dos sinais conforme mostra o arranjo da simulação presente na Figura 4. Por outro lado, a linguagem de alto nível Python foi utilizada na implementação dos algoritmos de *clusterização* e cálculo da BER. O diagrama de blocos do esquemático de VPI *Transmission Maker* mostrado na Figura 4 permite recuperar e armazenar as constelações para posterior *clusterização*. É importante notar que tanto a potência lançada na fibra como o comprimento do enlace estão parametrizados e podem ser variados por meio de um duplo *sweep*. Foram realizadas simulações com taxas de transmissão de dados de 100 Gbps e 400 Gbps em dois cenários, sempre operando em um único canal, sendo eles:

- **Cenário 1** - Considera além da CD, PMD, rotação da polarização, o efeito Kerr da fibra e o ruído dos fotodetectores, impedimentos como o desbalanceamento da rede híbrida a dessincronização de relógio e o ruído de fase dos *laser* assim como seu *offset* frequencial;
- **Cenário 2** - Considera os efeitos de CD, PMD, rotação do estado de polarização e ruído aditivo dos fotodetectores, assim como o efeito Kerr da fibra. Porém, não considera os efeitos do ruído de fase dos *lasers*, desbalanço da rede híbrida e dessincronização de relógio.

O arranjo da simulação pode ser dividido em quatro partes: transmissor, canal, receptor e DSP.¹

3.1.1 Transmissor

Inicialmente no transmissor, utiliza-se o módulo de geração de sinais ópticas QAM para sistemas coerentes digitais, o qual é configurado para gerar duas sequências de bits pseudo-aleatórias (PRBSs, do inglês *pseudorandom bit sequences*) independentes, correspondentes às informações a serem transmitidas nas duas polarizações da portadora

¹ Estritamente, o DSP faz parte do receptor, porém, o autor prefere considerar de forma independente para que seja detalhado de forma mais clara a sua operação.

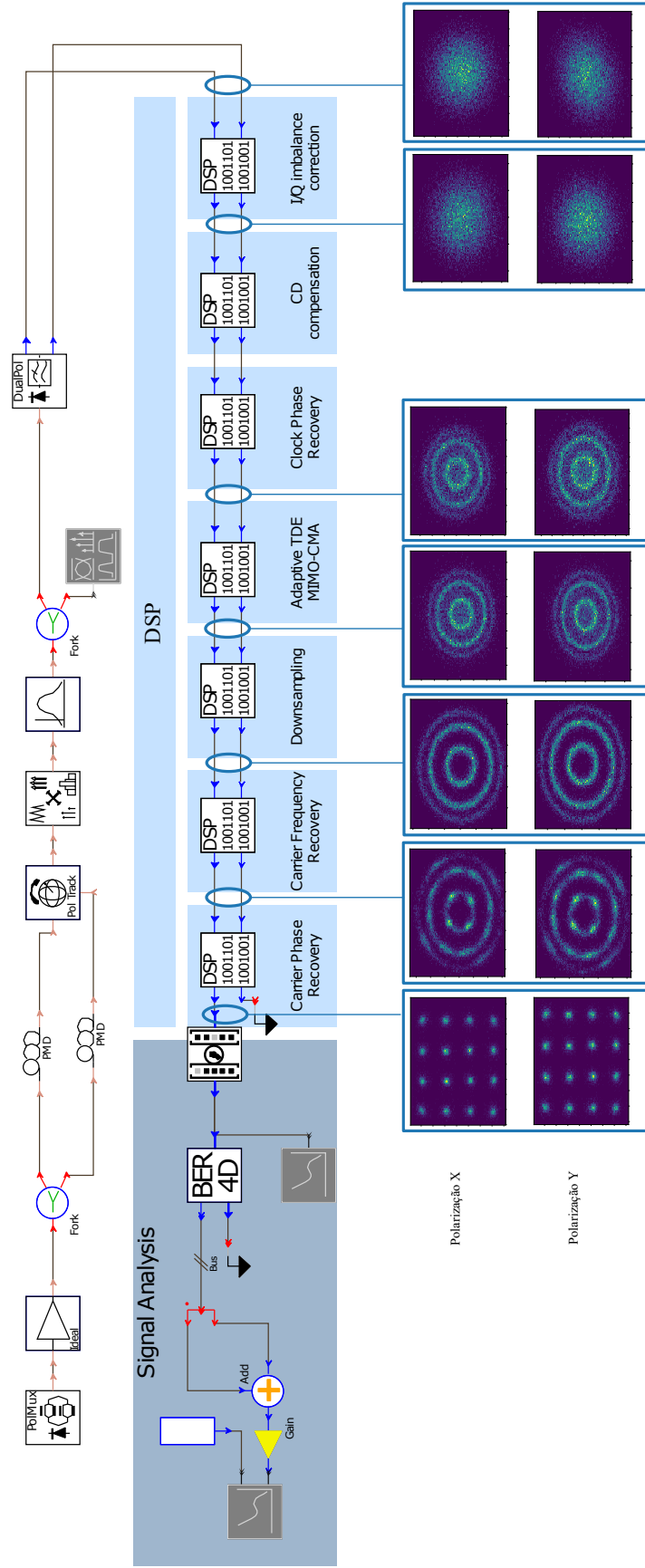


Figura 4 – Arranjo da simulação implementada no Software VPI com os diagramas de constelação para demonstração da aplicação de cada módulo do DSP.

óptica. São gerados 65.536 (2^{16}) bits aleatórios para a geração da PRBS com taxa de transmissão de 100 Gbps e 262.144 (2^{18}) bits para a taxa de transmissão de 400 Gbps. Posteriormente, cada sequência é processada por um conversor serial/paralelo que agrupa os bits em blocos de 4 elementos, sendo cada um destes conjuntos mapeado em um símbolo 16-QAM utilizando código Gray. Com isso, o gerador de sinais QAM é ajustado para filtrar o sinal com um filtro de Nyquist com formato de cosseno levantado com fator de *roll-off* de 0,2 e são utilizadas 8 amostras por símbolo, que tem a função de simular o sinal analógico em que o filtro é utilizado. Esta sequência de amostras complexas é dividida nas suas partes real e imaginária e, após serem pré-distorcidas para compensar a resposta não linear do modulador, são utilizadas para alimentar as duas entradas de um DP-MZM com perdas de inserção de 6 dB e uma razão de extinção de 35 dB. Por outro lado, este módulo de geração de sinal considera uma fonte quase-monocromática que simula o *laser* de sinal, o qual é configurado para operar a 1550 nm, apresentando uma potência de emissão de 16 dBm e uma largura de linha de 200 kHz para o Cenário 1 e 0 kHz para o Cenário 2.

A saída do *laser* é dividida em dois sinais por meio de um divisor de feixe por polarização, sendo cada saída modulada com o DP-MZM previamente descrito. Os dois sinais ópticos são posteriormente multiplexados em polarização para gerar a saída do módulo. O sinal modulado é amplificado por um EDFA com figura de ruído de 4,5 dB, cuja potência de saída é controlada para estabelecer uma potência de transmissão entre 0 dBm e 15 dBm, em passos de 1 dBm. Estes parâmetros da simulação atribuídos ao transmissor são utilizados tanto no Cenário 1, quanto no Cenário 2, à exceção da largura de linha do *laser*, conforme descrito. É importante notar que, como a potência de entrada no EDFA é relativamente alta (tendo o valor de -17,12 dBm para as duas polarizações) a potência do ruído agregado pelo amplificador via emissão espontânea amplificada (ASE, do inglês *amplified spontaneous emission*) é menor que $0,1887 \mu\text{W}$ (AGRAWAL, 2012) e, portanto, não causa uma penalidade significativa. A Figura 5(a) descreve os elementos do módulo transmissor de dupla polarização utilizado nas simulações.

3.1.2 Canal

Quanto ao canal, a propagação na fibra é simulada usando o método de passo alternado de Fourier (*split-step Fourier method*) vetorial de passo adaptativo, o qual

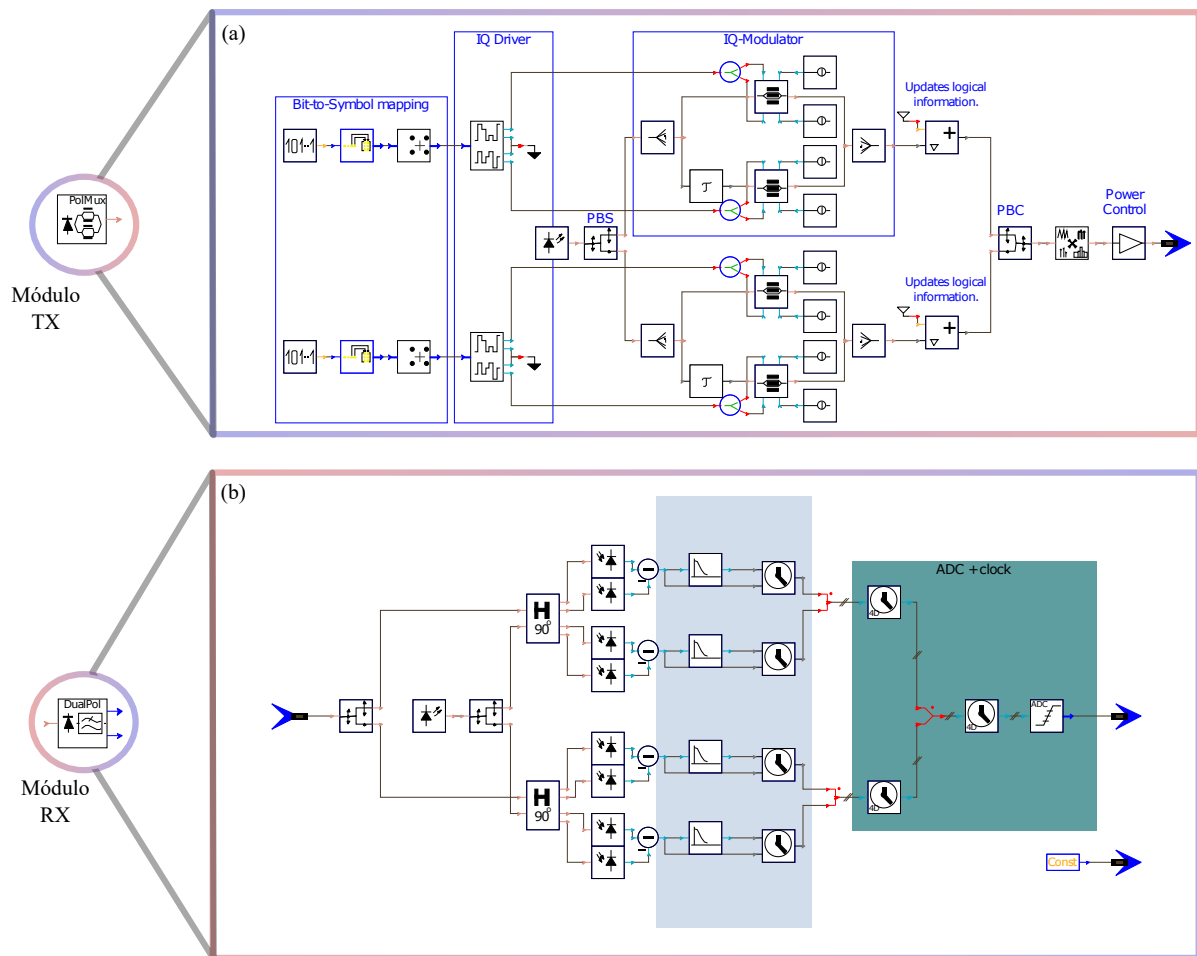


Figura 5 – Diagrama de blocos do módulos (a) transmissor e (b) receptor para sistemas coerentes digitais disponíveis em VPI *Transmission Maker*.

permite a integração das equações de Manakov com birrefringência estocástica para ambos os cenários propostos. O módulo que simula a fibra foi configurado para apresentar um parâmetro de dispersão de 16 ps/nm/km, uma área modal efetiva de $80 \mu\text{m}^2$ e um índice de refração não linear de $2,6 \cdot 10^{-20} \text{ m}^2$. O coeficiente de atenuação é fixado em 0,2 dB/km, o coeficiente de dispersão do modo de polarização é ajustado em $3,16 \text{ ps}/\sqrt{\text{km}}$ e o comprimento de correlação da birrefringência é configurados a 50 m. Para as simulações em 100 Gbps, o comprimento do enlace utilizado foi variado de 120 km a 200 km, considerando um passo de 10 km, resultando em perdas na fibra entre 24 dB e 40 dB. O comprimento da fibra, para as simulações em 400 Gbps, foi variado de 80 km até 150 km, resultando em perdas na fibra variando entre 16 dB e 30 dB.

3.1.3 Receptor

A Figura 5(b) detalha os componentes do módulo receptor utilizado na simulação. No *front-end* do receptor, a polarização do sinal recebido é primeiramente controlada por um rastreador de polarização dinâmica para minimizar o desalinhamento de polarização entre o sinal recebido e o *laser* do oscilador local. Em seguida, um filtro óptico Gaussiano de primeira ordem de 90 GHz é utilizado a fim de remover o ruído de emissão espontânea amplificada presente fora da banda de sinal². Posteriormente, foi utilizado o módulo de recepção de sinais coerente de dupla polarização já implementado no VPI *Transmission Maker*. Neste módulo, o primeiro passo é a demultiplexação das duas polarizações utilizando um divisor de feixe por polarização, que é configurado para ter uma perda de inserção de 1 dB. O sinal associado a cada polarização é misturado com a correspondente polarização do *laser* do oscilador local numa rede híbrida de 90°, que considera uma perda de inserção de 6 dB, com um desbalanço entre I e Q de 0,5 dB e um desbalanço de fase de 10°. Como ocorre no transmissor, o *laser* é modelado como uma fonte quase monocromática com uma largura de linha de 200 kHz e uma frequência de operação que difere em 50 MHz daquela do oscilador local. A potência de emissão do *laser* é estabelecida em 16 dBm, como indicam vários padrões (TSUKAMOTO *et al.*, 2005). As saídas das redes híbridas utilizadas para as duas polarizações são convertidas ao domínio elétrico utilizando fotodetectores balanceados com responsividade de 0,7 A/W e considerando ruídos de disparo e térmico com uma densidade espectral de ruído de corrente de 21 pA/ \sqrt{Hz} (potência equivalente de ruído de 30 pW/ \sqrt{Hz}). A resposta em frequência combinada dos fotodetectores e do filtro subsequente é modelada por um filtro equivalente passa baixa de Bessel de 4^a ordem. O amplificador diferencial de transimpedância é considerado ideal pois o ruído agregado é muito menor que o ruído dos fotodetectores. A última etapa do módulo de recepção é a conversão analógico-digital, configurada para utilizar uma resolução de 8 bits e uma frequência de amostragem próxima a duas amostras por símbolo, com um erro de relógio de 5 MHz. Estes parâmetros são referentes ao Cenário 1. Para o Cenário 2, são utilizados os mesmos parâmetros, apenas são desconsiderados os parâmetros de desbalanço entre I e Q da rede híbrida, o erro de frequência do relógio e os ruídos de fase dos *lasers*, assim como o *offset* frequencial.

² Este filtro é importante em enlaces *multi-span* limitados por ruído e não tanto em enlaces não amplificados, porém este foi incluído devido ao fato de que muitos sistemas incluem este filtro.

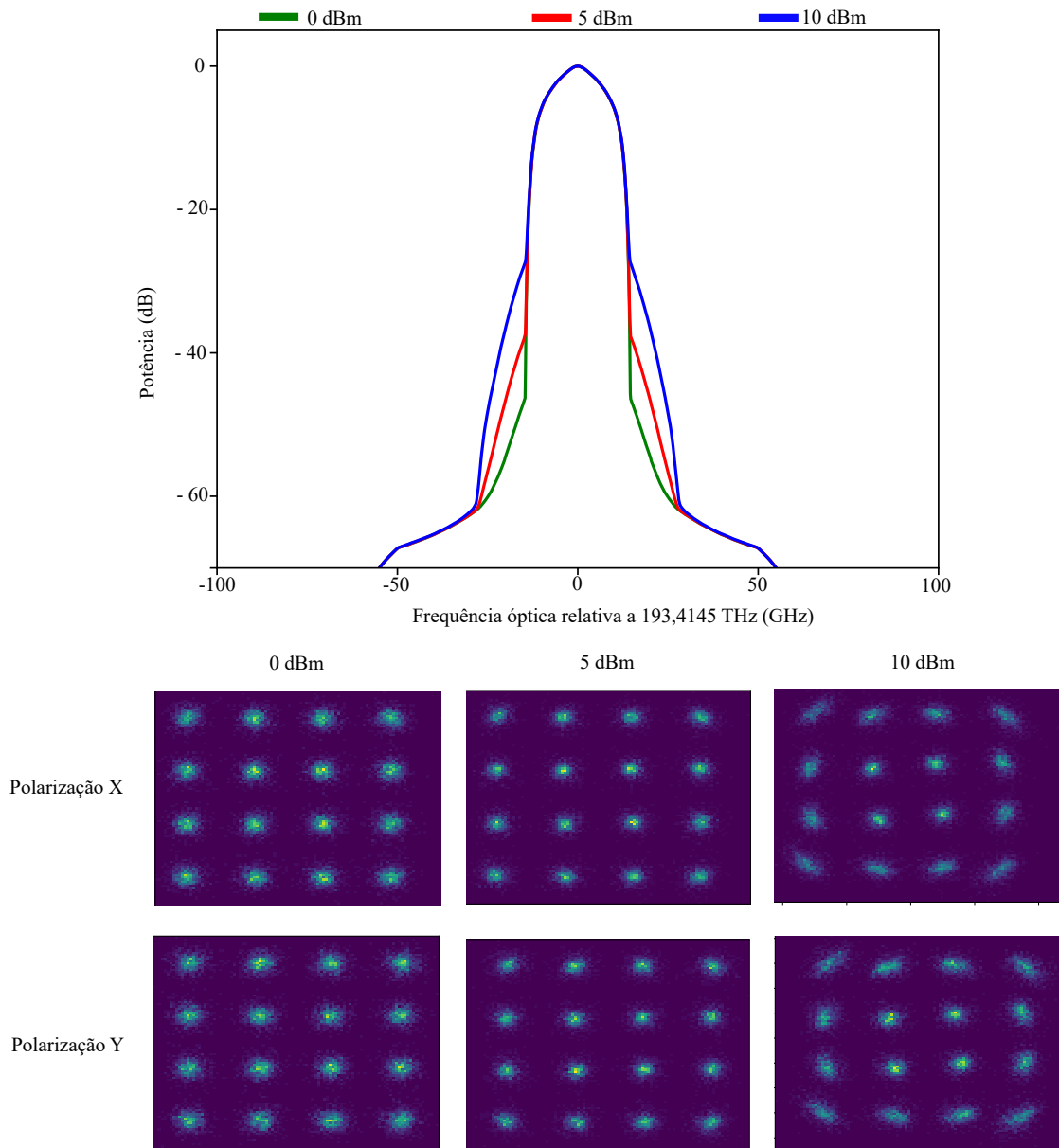


Figura 6 – Espectros ópticos normalizados e diagramas de constelações no *front end* do receptor para sistemas com 100 km e operando em 100 Gbps.

Nas constelações mostradas na Figura 4 podemos observar que os diferentes blocos do DSP realizam as operações previstas. Para fins de ilustração, utilizamos um enlace curto de 100 km para assegurar uma alta relação sinal-ruído (SNR, do inglês *signal to noise ratio*) incluso com uma potência de lançamento baixa. Por outro lado, a fim de comprovar que as não linearidades efetivamente estão sendo modeladas na fibra, na Figura 6 mostramos os espectros ópticos e os diagramas de constelação na saída do DSP (sem aplicação de algoritmos de *clusterização*) para um enlace de 100 km operando a 100 Gbps com potências

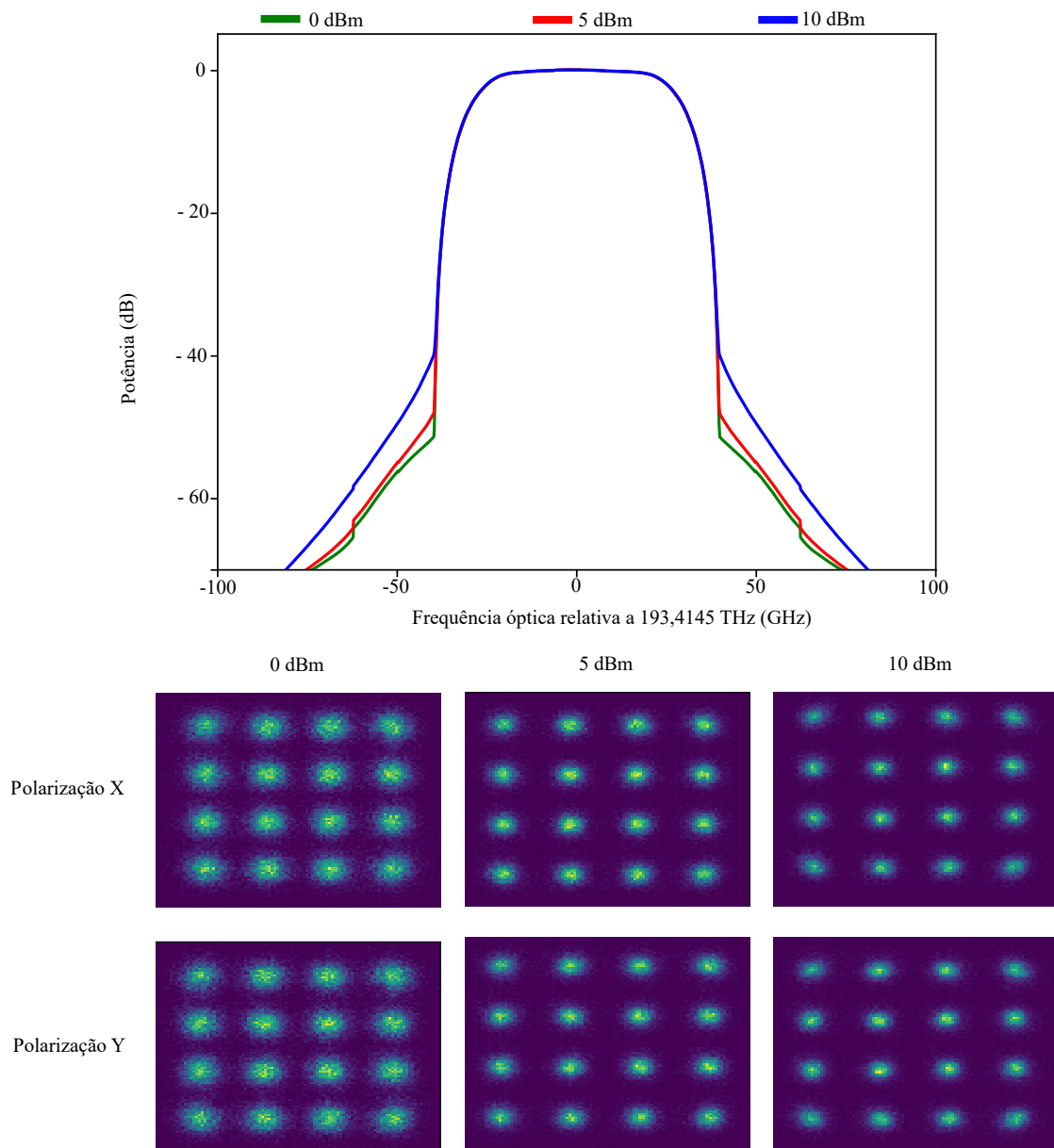


Figura 7 – Espectros ópticos normalizados e diagramas de constelações no *front end* do receptor para sistemas com 100 km e operando em 400 Gbps.

de lançamento de 0 dBm, 5dBm e 10 dBm. Como pode se observar, o espectro apresenta um alargamento que aumenta com a potência lançada na fibra. Sobre as constelações, pode-se observar que para altas potências, adquirem uma forma característica em que os pontos da periferia apresentam uma dispersão angular maior. Já a Figura 7 ilustramos as mesmas figuras para o sistema operando a 400 Gbps. Podemos ver que mesmo tendo um comportamento similar àquele de 100 Gbps, a constelação para altas potências não apresenta um padrão de dispersão tão claro. Como veremos ao longo deste trabalho, este

comportamento é consequência da iteração entre CD e efeitos não lineares, e acarreta importantes consequências em relação ao desempenho de algoritmos de *clusterização*.

3.1.4 DSP

No DSP, o sinal digitalizado passa por algoritmos capazes de mitigar ou até mesmo compensar alguns impedimentos do sinal enviado. Além disso, o processamento deste sinal é necessário para a demodulação do sinal. A Figura 4 mostra o efeito de cada bloco do DSP que implementa um determinado algoritmo na constelação. Neste modelo de simulação utilizado, o sinal é processado inicialmente pelo bloco de correção de desequilíbrio de fase e amplitude I/Q, que baseia-se no método de correção proposto por Gram-Schmidt (SAVORY, 2008). No próximo passo, é realizada a correção da CD, na qual o algoritmo é baseado na equalização no domínio da frequência por sobreposição (OFDE, do inglês *overlap frequency domain equalization*) (KUDO *et al.*, 2009), que compensa a distorção em $16 \mu\text{s}/\text{m}^2$. Posteriormente, é realizada a recuperação da fase do relógio, no qual é utilizado o estimador de Oerder-Meyr, que se trata de uma variante digital do estimador quadrático (OERDER; MEYR, 1988), em que o sinal de entrada no algoritmo é elevado ao quadrado e a componente espectral resultante na taxa de símbolo é extraída por filtragem. O sinal passa por um equalizador no domínio do tempo (TDE, do inglês *time-domain equalizer*) adaptativo, de múltiplas entradas e múltiplas saídas, e que usa um algoritmo de módulo múltiplo (MMA, do inglês *multi-modulus algorithm*) responsável por realizar a rotação para compensar o desalinhamento entre o sinal e o SoP do receptor. Neste caso foram utilizados 25 *taps* para o filtro TDE. Assim, ocorre a subamostragem do sinal, em que as polarizações são subamostradas independentemente. Há então a recuperação da frequência e fase da portadora, de acordo com o método proposto por (LOUCHET; KUZMIN; RICHTER, 2008), em que o tamanho do bloco adotado foi de 5 símbolos.

Após estes algoritmos, a programação em Python é utilizada para aplicação dos algoritmos de *clusterização*. Estes algoritmos são implementados para auxiliar na detecção e na classificação dos símbolos que se encontram na região de decisão entre cada símbolo da constelação. Inicialmente, o sinal sem distorções é classificado e, assim, cada símbolo é rotulado de acordo com o símbolo da constelação a que pertence. Assim, são aplicados os algoritmos de *clusterização* ao sinal que já passou pelos blocos de detecção e mitigação de

impedimentos citados acima. O algoritmo de ML é o primeiro a ser avaliado, e os símbolos transmitidos são rotulados após a aplicação do algoritmo. Posteriormente, é aplicado o algoritmo de classificação K-means com o método de inicialização chamado de K-means++ e o algoritmo de maximização de expectativa. Estes algoritmos também rotulam os símbolos de acordo com a classificação verificada para cada método. Para fins de equidade na utilização dos métodos de *clusterização*, alguns parâmetros foram fixados em ambos os casos, sendo estes: o número máximo de iterações em 300; a tolerância de distância para convergência do algoritmo em 10^{-4} , observando a norma de Frobenius (POWELL, 2004), em que deve haver duas iterações consecutivas com resultados iguais para ser declarada a convergência. Por fim, os rótulos de cada algoritmo são comparados ao rotulamento do sinal original transmitido e, assim, contabilizado o número de erros de símbolos, que possibilita o levantamento da BER para cada caso específico. Os passos referentes ao DSP são realizados para cada um dos cenários propostos.

A Tabela 3 sintetiza os parâmetros utilizados nas simulações.

3.2 Métricas de avaliação

A qualidade do sinal será avaliada utilizando a taxa de erro de bits (BER, do inglês *bit error ratio*), a qual é calculada por contagem de erros, isso devido ao fato de que, outras métricas como a magnitude do vetor de erro (EVM, do inglês *error vector magnitude*) não apresentam uma relação analítica simples com a BER, devido à distorção não linear.

Já o desempenho dos algoritmos de *clusterização* será quantificado comparando a BER após a aplicação do algoritmo de *clusterização* com a BER de referência obtida utilizando máxima verossimilhança. Cabe mencionar que para coerência entre as comparações, devem ser utilizados os valores de BER nas potências de lançamento ótimas. Adicionalmente, utilizaremos o máximo alcance do enlace para uma métrica de avaliação secundária obtida da comparativa anterior.

Parâmetros Gerais	Número de canais	1	canal	
	Janela de tempo da simulação	1,17	μs	
	Taxa de amostragem no domínio elétrico	28 / 112	GSps	
	Taxa de amostragem no domínio óptico	112 / 448	GSps	
Sinal	Formato de modulação	16-QAM	-	
	Fator de roll-off	0,2	-	
	Forma do filtro de Nyquist	Cosseno levantado	-	
	Taxa de bits	100 / 400	Gbps	
Transmissor	PRBS	$2^{16} / 2^{18}$	bits	
	Potência do <i>Laser</i>	16	dBm	
	Potência óptica lançada	0-15	dBm	
	Figura de ruído do amplificador	4,5	dB	
	Razão de extinção do MZM	35	dB	
	Inserção de perdas MZM	6	dB	
	Largura de linha do <i>laser</i>	Cenário 1 Cenário 2	200 0	kHz kHz
Canal	Dispersão na fibra	16	ps/nm/km	
	Área modal	80	μm^2	
	Índice de refração linear	$2,6 \times 10^{-20}$	m^2	
	Coefficiente de atenuação	0,2	dB/km	
	Coefficiente de PMD	3,16	ps/ \sqrt{km}	
	Comprimento dos enlaces	120-200 / 80-150	km	
	Perdas na fibra	24-40 / 16-30	dBm	
Receptor	Filtro Gaussiano (1 ^a ordem)	90	GHz	
	Perdas de inserção no PBS	1	dB	
	Perdas de inserção na rede híbrida	6	dB	
	Desbalanço I/Q na rede híbrida	Cenário 1 Cenário 2	0,5 0	dB dB
	Desbalanço de fase na rede híbrida	Cenário 1 Cenário 2	10 0	graus graus
	Largura de linha <i>laser</i> Lo	200	kHz	
	Potência do <i>laser</i> Lo	16	dBm	
	Responsividade dos fotodetectores	0,7	A/W	
	Densidade de ruído do fotodiodo	21	pA/ \sqrt{Hz}	
	Ordem do filtro elétrico em RX	4 ^a	-	
	Erro de relógio	Cenário 1 Cenário 2	5 0	MHz MHz

Tabela 3 – Tabela resumo dos parâmetros utilizados nas simulações.

4 Resultados

Este capítulo aborda os resultados de simulação obtidos, comparando a qualidade do sinal quando consideramos máxima verosimilhança e quando aplicamos os algoritmos de *clusterização* K-means com inicialização K-means++ e EM na mitigação de impedimentos em um sistema DP-16QAM coerente digital de um único *span* com taxas de transmissão de 100 Gbps e 400 Gbps.

Além do uso dos métodos de *clusterização* K-means com inicialização K-means++ e a EM, outros dois algoritmos foram inicialmente testados, DBSCAN e K-means Mini Batch (SCULLEY, 2010). O algoritmo DBSCAN é amplamente utilizado em várias aplicações, inclusive já teve seu uso comprovado para compensação de ruídos em sistemas ópticos (GIACOMIDIS *et al.*, 2019). No entanto, devido ao seu método necessitar de ajustes constantes nos parâmetros *número mínimo de amostras* e *eps*, seu uso não foi eficaz para este trabalho, cujo intuito era fazer uma varredura entre as distâncias e potências propostas na análise. Os resultados obtidos com o uso de DBSCAN foram muito piores, em termos da BER, que os encontrados ao aplicar algoritmos discutidos neste trabalho, e por isso foi descartada a hipótese de seu uso. Quanto ao algoritmo K-means Mini Batch, que tem como principal vantagem a redução do tempo de processamento do algoritmo K-means convencional em detrimento à eficiência, este não trouxe bons resultados para a situação proposta neste trabalho. Além de resultar em valores de BER superiores aos encontrados pelos métodos analisados neste trabalho, o algoritmo exigiu um custo computacional bem superior, que tornou inviável seu uso neste sistema proposto. Assim, diante dos resultados expostos, para este tipo de sistema analisado, decidiu-se por não apresentar detalhadamente os resultados destes métodos neste trabalho.

Como já mencionado, o sistema foi avaliado utilizando dois cenários distintos em que diferentes impedimentos foram considerados. Nesta seção, primeiro analisamos o Cenário 1 que inclui todos os impedimentos e, posteriormente, analisamos o Cenário 2 em que se inclui unicamente o efeito Kerr, a CD e o ruído aditivo dos fotodetectores¹. Por meio da análise destes dois cenários utilizando taxas de transmissão de 100 Gbps e 400 Gbps, podemos identificar as configurações em que os algoritmos de *clusterização* efetivamente melhoram o desempenho do sistema e que impedimentos são compensados.

¹ É importante mencionar que a CD também foi considerada no Cenário 2 devido à interação entre efeito Kerr e dispersão, o qual não acontece com outros efeitos como ruído de fase.

4.1 Análise do Cenário 1

O Cenário 1, conforme citado no Capítulo 3, considera todos os impedimentos do sistema de comunicações ópticas. Neste cenário, os efeitos lineares e não lineares da fibra, o ruído dos fotodetectores, o desbalanceamento da rede híbrida e a dessincronização de relógio, são considerados. Em primeiro lugar, foi simulado um sistema operando a uma taxa de transmissão de 100 Gbps, com 2^{14} símbolos, cuja potência óptica de lançamento na fibra variou de 0 dBm a 15 dBm, e o comprimento do enlace foi varrido de 120 km a 200 km, sempre operando em um único canal. Todas as combinações entre potência de lançamento na fibra e distância foram simuladas e processadas para obtenção da BER. Os resultados obtidos, após o processamento dos algoritmos de *clusterização* e da detecção de ML, estão presentes na Figura 8.

A Figura 8(a) mostra as curvas de BER em termos da potência lançada na fibra para três diferentes distâncias de enlace, 130 km, 150 km e 170 km, e três esquemas de detecção, máxima verossimilhança, K-means e EM. Analisando esta figura, podemos constatar que, para a distância de 170 km, as curvas que representam a BER para cada algoritmo (Máxima verossimilhança: preto, K-means: vermelho e EM: verde), possuem valores muito parecidos até aproximadamente 8 dBm. Neste caso, os algoritmos de *clusterização* conseguem

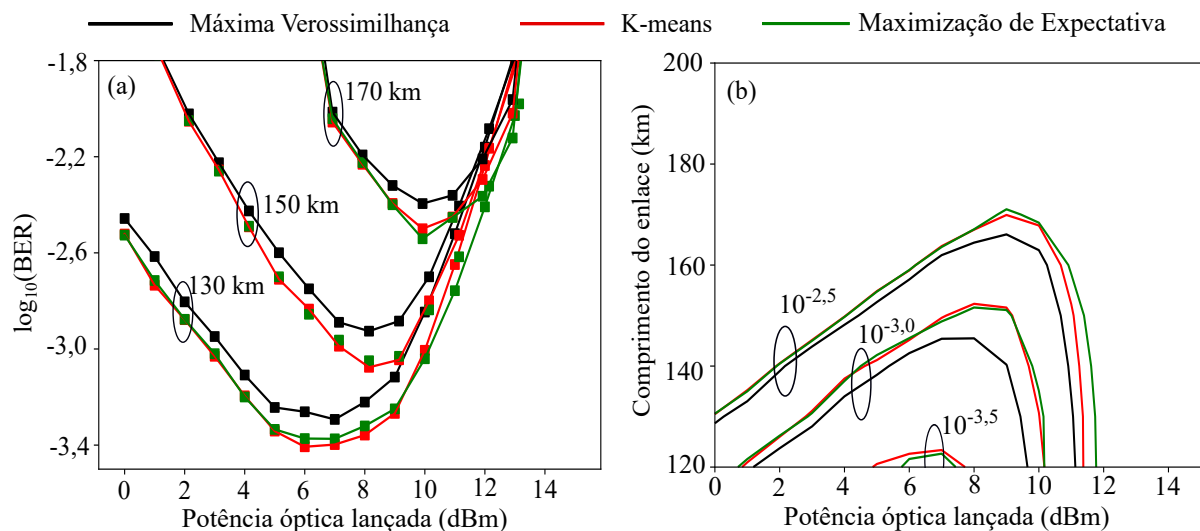


Figura 8 – BER para enlaces com taxa de transmissão de 100 Gbps no Cenário 1. (a) BER em função da potência lançada para enlaces de 130 km, 150 km e 170 km; (b) Diagrama de contorno com os valores da BER em função da potência óptica lançada e do comprimento do enlace.

diminuir o valor da BER de $10^{-2,35}$ para $10^{-2,53}$ na potência ótima de 10 dBm. Quando avaliamos a curva característica para o enlace de 150 km, nota-se que os valores de BER, para as diferentes técnicas de detecção, são muito semelhantes para potências inferiores a 4 dBm. Acima desta potência, temos uma redução da BER que pode chegar a $10^{-3,08}$ na potência de 8 dBm, sendo a BER para ML de $10^{-2,91}$. Este comportamento é similar na curva para 130 km, mas a *clusterização* começa a mitigar os impedimentos a partir de 0 dBm, com uma diminuição média de $10^{-1,5}$. Neste caso, a redução do valor da BER ocorre até 5 dBm, em que há um pequeno aumento na diferença entre os valores de BER, sendo encontrada no valor ótimo de potência de 6 dBm, $10^{-3,40}$, enquanto que para ML a BER é de $10^{-3,25}$. Comparando o desempenho dos algoritmos de *clusterização* para as diferentes curvas de BER, podemos observar que estes apresentam um melhor desempenho para potências médias quando a distância aumenta. Além disso, observa-se que para potências acima de 10 dBm, os resultados obtidos com a aplicação de EM são melhores que com a aplicação de outras ferramentas.

Estas observações indicam que os algoritmos de *clusterização* não conseguem compensar o ruído aditivo dos fotodetetores, que é o efeito dominante para longas distâncias e baixas potências, nem os efeitos das não linearidades. Consequentemente, podemos visualizar que os algoritmos de *clusterização* estão compensando algum outro impedimento que é significativo na região de médias potências. Por outro lado, comparando o desempenho dos algoritmos de K-means e EM, podemos concluir que para baixas e médias potências, tem um comportamento similar. Já para altas potências, EM apresenta um desempenho melhor que K-means. Realizando varreduras similares às apresentadas na Figura 8(a) para enlaces de comprimento entre 120 km e 200 km, podemos analisar a dependência da BER em termos da potência de lançamento e o comprimento de enlace. Para facilitar a comparação com a detecção por máxima verossimilhança, na Figura 8(b), representamos os diagramas de contorno da BER para $10^{-2,5}$, $10^{-3,0}$ e $10^{-3,5}$. Nesta figura, podemos verificar que os algoritmos de *clusterização* conseguem mitigar os impedimentos para as diferentes distâncias e potências consideradas, permitindo estender o alcance do enlace por 10 km para uma BER de $10^{-3,0}$ e aproximadamente 5 km para uma BER de $10^{-2,5}$. Também podemos observar um aumento na margem de potência no sistema.

Com o intuito de se verificar a eficiência das ferramentas propostas na mitigação de impedimentos em sistemas operando com maiores taxas de transmissão, foram realizadas simulações com taxa de transmissão de 400 Gbps. O arranjo sistêmico é o mesmo utilizado

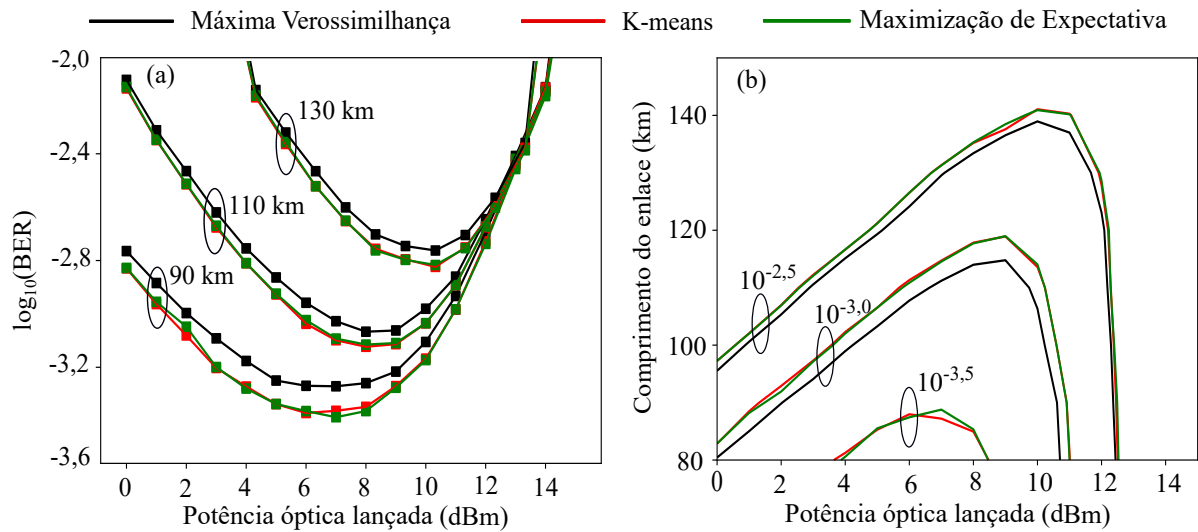


Figura 9 – BER para enlaces com taxa de transmissão de 400 Gbps no Cenário 1. (a) BER em função da potência lançada para enlaces de 90 km, 110 km e 130 km; (b) Diagrama de contorno com os valores da BER em função da potência óptica lançada e do comprimento do enlace.

na simulação para 100 Gbps, porém o número de símbolos simulados agora é de 2^{18} . Na Figura 9, apresentamos os resultados obtidos para 400 Gbps, com a potência de lançamento na fibra também variando de 0 dBm e 15 dBm e o comprimento do enlace variando de 80 km a 150 km. A Figura 9(a) mostra a BER calculada para comprimentos de enlace de 90 km, 110 km e 130 km em relação à potência lançada. Para 130 km, verifica-se que a contribuição referente à aplicação dos algoritmos de *clusterização*, está na redução de $10^{-2,76}$ para K-means e EM, enquanto que para máxima verossimilhança tem o valor de $10^{-2,68}$, todos estes para a potência ótima de 10 dBm. Para 110 km, o comportamento é similar ao anterior, tendo como potência ótima o valor de 8 dBm e BER de $10^{-3,07}$ para ML e $10^{-3,15}$ para K-means e EM. Já o comportamento da curva da BER em função da potência para 90 km mostra que, para valores de potência maiores do que 2 dBm, há a compensação de impedimentos que chegam, na potência de 7 dBm, ao valor $10^{-3,47}$ com o uso dos algoritmos de *clusterização*, enquanto que para ML é de $10^{-3,33}$. Podemos observar que a o comportamento das curvas é similar a aquele descrito para 100 Gbps, com uma exceção. Como pode-se observar, para altas potências o desempenho de ML, EM e K-means é praticamente o mesmo, indicando que a eficiência dos algoritmos de *clusterização* para compensar as não linearidade é ainda menor que a 100 Gbps.

O diagrama de contorno que representa a BER em função do comprimento do enlace e da potência do sinal para 400 Gbps, dentro do Cenário 1, é mostrado na Figura 9(b). Nesta situação, os resultados obtidos com K-means e com EM são praticamente os mesmos. Observando esta figura, podemos verificar que é possível obter uma BER de $10^{-3,5}$, apenas em enlaces menores que 90 km e para potências de lançamento entre 4 dBm e 8 dBm. Além disso, fica claro que a possibilidade de aumento do comprimento do enlace, mantendo a mesma BER e a mesma potência, cai em relação ao uso do sistema 100 Gbps. Para a potência 9 dBm, e com a BER de $10^{-3,0}$, o comprimento do enlace pode ser 5 km maior utilizando os algoritmos de *clusterização* do que usando da detecção por ML (114 km para ML e 119 km para K-means e EM). Considerando uma BER de $10^{-2,5}$ e potência de 9 dBm, a diferença passa a ser de apenas 2 km, sendo que, para ML, é limitado em 139 km, e para os algoritmos de *clusterização*, em 141 km.

4.2 Análise do Cenário 2

Como visto no Cenário 1, a utilização de algoritmos de *clusterização* para mitigação de impedimentos em sistemas coerentes DP-16QAM parece que não consegue mitigar a distorção causada pelo efeito Kerr. Para verificar esta afirmação, utilizamos o Cenário 2, no qual impedimentos como o ruído de fase dos *lasers*, o desbalanço da rede híbrida e a dessincronização do relógio não são considerados. A Figura 10 apresenta os resultados obtidos da BER para 100 Gbps considerando o segundo cenário proposto.

Na Figura 10(a) apresentamos as curvas de BER para enlaces de 130 km, 150 km e 170 km em função da potência lançada na fibra. Observando a figura, pode-se observar que para baixas potências, inferiores a 9 dBm, não se vê uma diferença entre os valores da BER obtidos pela detecção da máxima verossimilhança e dos algoritmos de *clusterização* utilizados. Podemos observar também que, a medida que a potência ultrapassa valores de 10 dBm, a utilização dos algoritmos de *clusterização* proporciona uma pequena compensação em relação à detecção por máxima verossimilhança. Vale ressaltar que, para o enlace de 130 km, os valores de BER são nulos ou muito próximos de zero e, por isso, a curva característica desta distância só aparece após 10 dBm.

Na Figura 10(b) apresentamos as curvas de contorno obtidas de varreduras de potência para diferentes comprimentos de enlace. Podemos constatar nesta figura que as

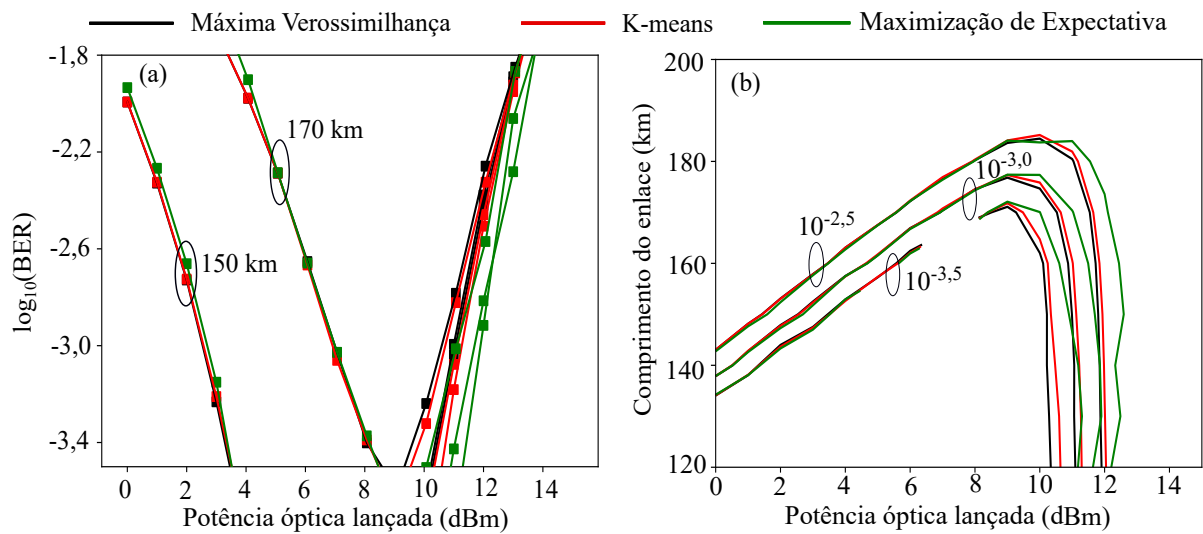


Figura 10 – BER para enlaces com taxa de transmissão de 100 Gbps no Cenário 2. (a) BER em função da potência lançada para enlaces de 130 km, 150 km e 170 km; (b) Diagrama de contorno com os valores da BER em função da potência óptica lançada e do comprimento do enlace.

curvas para máxima verossimilhança, K-means e EM são muito próximas para baixas potências, o que confirma as observações feitas na Figura 10(a). Para altas potências, podemos observar que os algoritmos de *clusterização* continuam compensando ligeiramente os efeitos da distorção não linear. Porém, para potências ótimas de lançamento, estes algoritmos não conseguem melhorar a qualidade do sinal, resultando em um aumento de alcance marginal.

Na Figura 11 mostramos os valores da BER para o Cenário 2, operando a 400 Gbps. Como ocorre nos casos anteriores, primeiro mostramos na Figura 11(a) as curvas com os valores da BER em função da potência enviada na fibra para três comprimentos de enlace diferentes, neste caso 90 km, 110 km e 130 km. Nela, os valores para a BER com potência inferior a 12 dBm são zero para o comprimento de 90 km, e somente após este valor de potência que a curva é apresentada. Para as potências de 110 km e 130 km, vemos que os valores de BER após o processamento dos algoritmos e da detecção por máxima verossimilhança são muito próximos inclusive para altas potências, revelando que para este caso, a *clusterização* não é capaz de compensar os efeitos das não linearidades.

Do mesmo modo, a Figura 11(b) mostra as curvas de contorno correspondentes aos valores de BER para as distintas potências e comprimento de enlace. As curvas da

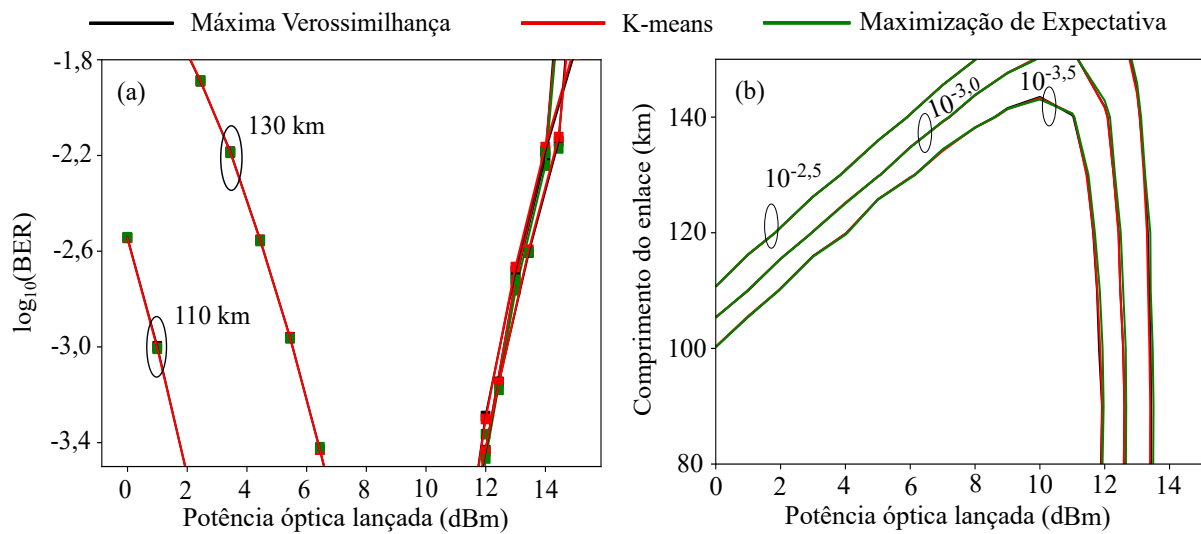


Figura 11 – BER para enlaces com taxa de transmissão de 400 Gbps no Cenário 2. (a) BER em função da potência lançada para enlaces de 90 km, 110 km e 130 km; (b) Diagrama de contorno com os valores da BER em função da potência óptica lançada e do comprimento do enlace.

BER para máxima verossimilhança, K-means e EM são muito próximas, resultando numa extensão de alcance nula.

4.3 Análise da complexidade computacional

Como mencionado anteriormente, os algoritmos de aprendizagem de máquinas empregues neste trabalho utilizam da aprendizagem não supervisionada para o processo de classificação dos dados. Este processo de aprendizagem pode variar, dependendo da particularidade de cada algoritmo. Diante disso, podemos comparar o desempenho de cada ferramenta, analisando tanto a eficiência na compensação de impedimentos, quanto o tempo necessário para a aprendizagem da tarefa. A primeira destas comparações foi realizada nas seções anteriores, em que o desempenho de cada algoritmo de compensação foi analisado para cada configuração do enlace. Neste aspecto, a eficiência de ambos os algoritmos foi comprovada para a mitigação de impedimentos em um enlace coerente, mas verifica-se também que os resultados entre si foram muito similares. Na grande maioria das vezes, tanto K-means, quanto EM, tinham valores muito parecidos de BER.

Desta forma, faz-se necessária uma avaliação do segundo ponto comparativo, que seria o tempo necessário para o processamento de dados. Este tempo destinado ao proces-

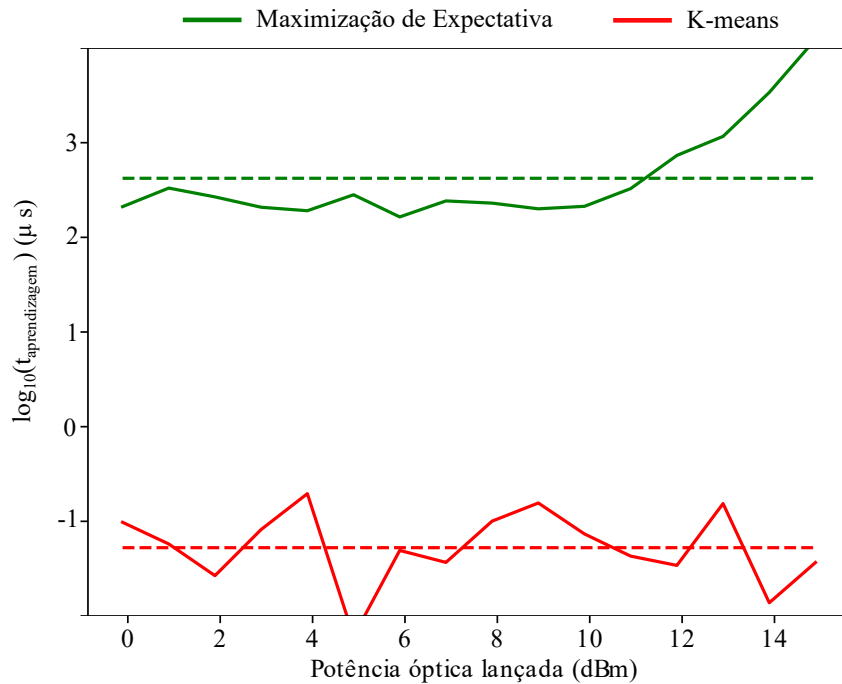


Figura 12 – Tempo médio de aprendizagem dos algoritmos de *clusterização* para taxa de transmissão de 100 Gbps.

samento da aprendizagem do algoritmo é diretamente proporcional ao custo computacional requerido, sendo este um importante parâmetro para a efetividade da aplicação da ferramenta. Vale lembrar que toda a operação de transmissão de dados, processamento, detecção e decodificação do sinal leva algum tempo, e o tempo necessário para o aprendizado do algoritmo seria somado a este tempo. Para efeitos comparativos entre os dois algoritmos implementados, foram medidos os tempos médios de aprendizagem requeridos por uma mesma máquina para cada um dos algoritmos. Além disso, como a complexidade do algoritmo depende do número de símbolos a serem classificados, foi normalizado o tempo requerido pelo número de símbolos. Em todas as combinações de potência lançada na fibra e o comprimento do enlace, o tempo de aprendizado foi medido, sempre se atentando para que os mesmos critérios fossem mantidos. Também foram analisados os tempos de processamento para cada taxa de transmissão.

Na Figura 12 foi mostrado o tempo médio necessário para a aprendizagem do algoritmo em função da potência lançada na fibra, para a taxa de 100 Gbps. A linha vermelha sólida mostra o tempo médio quando utilizamos o algoritmo K-means com inicialização K-means++ para cada potência e nos variados comprimentos de enlace, e a linha vermelha tracejada informa a média do tempo necessário no processo de aprendizagem retratado

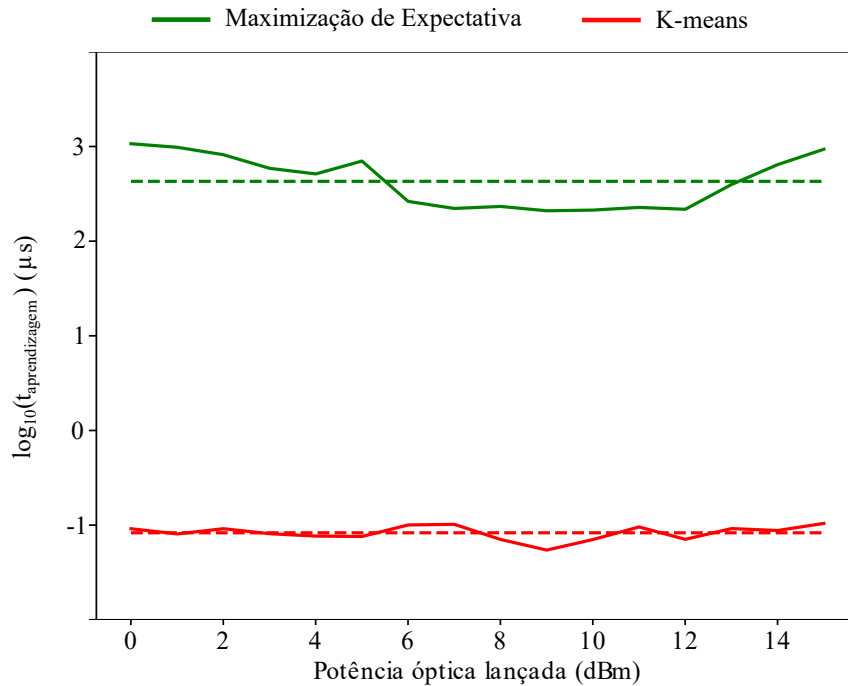


Figura 13 – Tempo médio de aprendizagem dos algoritmos de *clusterização* para taxa de transmissão de 400 Gbps.

pela linha sólida. As linhas verdes representam o tempo médio para o processamento da EM, sendo a linha sólida o valor do tempo médio para os distintos comprimentos de enlace e a tracejada a média deste tempo. Assim, podemos verificar que, para a taxa de transmissão de 100 Gbps, o algoritmo K-means necessita de um tempo médio de $0,055 \mu s$ e, para a mesma taxa, o algoritmo de EM necessita de $2,61 \mu s$. Estes valores demonstram que o algoritmo de EM leva, em média, um tempo 47 vezes maior de aprendizagem do que o algoritmo K-means, mesmo utilizando da inicialização K-means++.

Na Figura 13, os mesmos resultados são mostrados, havendo apenas a mudança da taxa de transmissão de dados e, conseqüentemente, a quantidade de símbolos processados. Nesta condição, o tempo de processamento médio para a EM é de $2,65 \mu s$ e, para aprendizagem de K-means, o tempo médio é de $0,085 \mu s$. Sendo assim, o tempo de aprendizagem para a EM ainda é 30 vezes maior do que o algoritmo K-means, neste tipo de aplicação.

5 Discussão dos resultados

Os resultados apresentados na seção anterior revelam que os algoritmos de *clusterização* conseguem estender o alcance do enlace tanto para 100 Gbps como para 400 Gbps. Porém, a comparação entre os dois cenários analisados revelam que esta extensão não é graças a compensação das não linearidades mas de impedimentos lineares residuais. Neste capítulo discutimos porque os algoritmos de *clusterização* não conseguem compensar as não linearidades em sistemas de dupla polarização com altas taxas de transmissão.

Em primeiro lugar, como mostrado na Figura 2, a distorção não linear em sistemas de um único canal e dupla-polarização não está limitado à SPM. Como se observou no espectro mostrado na Figura 2(d), a XPM entre as duas polarizações também causa uma distorção significativa no sinal. De fato, conforme à (AGRAWAL, 2000), a XPM causa uma modulação de fase equivalente a $2/3$ da causada por SPM. Isto é, ao implementar a *clusterização* em cada polarização separadamente, a distorção XPM não pode ser compensada, o qual já resulta em uma degradação do desempenho do esquema de compensação. Neste sentido, parece natural pensar em um algoritmo de *clusterização* que contemple as duas polarizações simultaneamente, porém, para o caso de DP-16QAM, isso resultaria em 256 agrupamentos que seria extremamente difícil de se realizar a *clusterização*.

Por outro lado, precisamos considerar a interação entre a dispersão cromática e o efeito Kerr. É bem conhecido que a dispersão tem uma influência crítica no impacto que os efeitos não lineares causam na transmissão de sinal. O exemplo mais claro é a eficiência de FWM que é altamente dependente da condição de casamento de fase, isto é, da dispersão do guia de onda, tipicamente da dispersão cromática (AGRAWAL, 2016). A interação de dispersão e SPM, por outro lado, também é amplamente estudada pois pode-se conseguir que a distorção da forma de pulso de ambos os efeitos se compensem, nos denominados *solitóns*. Esta interação não é trivial pois apresenta uma importante dependência espacial e da taxa de transmissão. Para entender esta dependência, na Figura 14 mostramos graficamente, a evolução de três pulsos adjacentes por meio de uma fibra. Devido ao decaimento da potência, os efeitos não lineares concentram-se no começo da fibra enquanto que a ISI aumenta ao longo da distância de propagação. Desta forma, no início da fibra, temos um grande peso da SPM e XPM e uma ISI baixa. No final da fibra esta situação é invertida tendo SPM e XPM baixa e uma ISI alta. A primeira

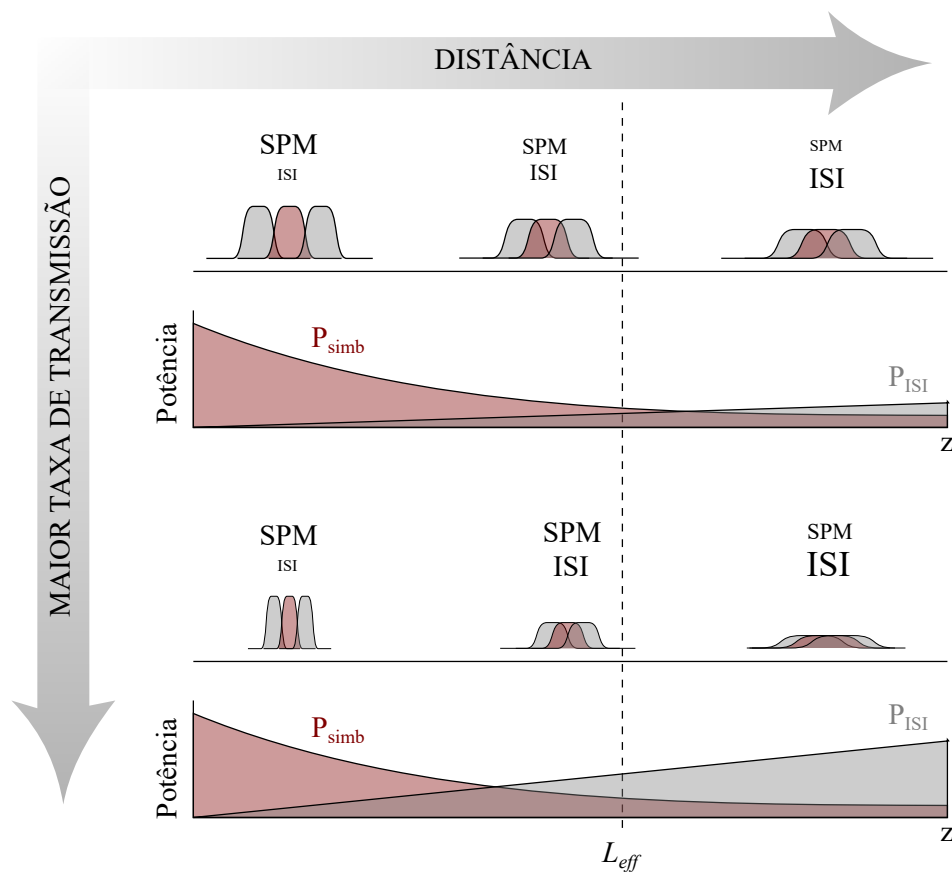


Figura 14 – Interação entre a SPM e a ISI para enlaces de taxas de transmissão maiores e menores em diferentes distâncias.

situação pode ser eficientemente compensada pois a distorção depende do próprio símbolo a classificar. Por outro lado, a distorção causada na última parte da fibra em que a SPM e XPM é baixa e a ISI é alta, é compensada mediante o módulo de compensação de CD. Na parte central do enlace, temos simultaneamente SPM e ISI causando interferência não linear entre os símbolos. Como sabemos, o ISI aumenta com a banda do sinal, isto é com a taxa de transmissão. Desta forma, para sinais com uma taxa de transmissão baixa, a ISI continua baixa na parte do enlace em que a distorção por SPM e XPM é alta, i.e. $z < L_{\text{eff}}$. Neste caso, o efeito da interferência intersimbólica não linear não é muito significativo. Ao aumentar a taxa de transmissão, a ISI aumenta de forma mais rápida a medida que o sinal é transmitido pela fibra, causando a presença de ISI inclusive na parte em que a SPM e XPM é significativa. Conseqüentemente, o impacto da interferência intersimbólica não linear é maior a medida que aumentamos a taxa de transmissão.

O peso relativo entre distorção não linear gerado pelo próprio símbolo e por símbolos adjacentes, pode-se observar de forma qualitativa na Figura 14. Como os algoritmos

de *clusterização* utilizados somente consideram um único símbolo, não são capazes de compensar a interferência intersimbólica não linear, o qual explica o baixo desempenho destes algoritmos para altas taxas de transmissão. É importante notar que esta interferência intersimbólica não linear é causada em sistemas de dupla polarização principalmente por causa da SPM e XPM entre as duas polarizações. Estas, a princípio, introduzem unicamente uma modulação de fase e não causam modulação de amplitude. Porém, é importante considerar que esta modulação de fase é convertida a modulação de amplitude pela própria dispersão da fibra. Como acontece com os sistemas de dupla polarização, a solução seria contemplar um conjunto de símbolos adjacentes de forma que o sistema tenha informação dos diferentes símbolos que interagem de forma não linear. Porém isto aumentaria de forma excessiva o número de *clusters* ($N_{clusters} = 16^{N_{simbolos}}$ para o caso de 16QAM e polarização simples e $N_{clusters} = 16^{2N_{simbolos}}$ para dupla polarização), o que impossibilitaria a sua implementação.

Assim, como os algoritmos de *clusterização* utilizados somente podem compensar impedimentos que dependem unicamente do símbolo e polarização considerada, não são adequados para compensar efeitos lineares em sistemas de altas taxas com dupla polarização. No entanto, outros impedimentos intrapolarização e que não tenham uma grande dependência de símbolos adjacentes em princípio, podem ser mitigados por meio destes algoritmos. Em particular, a dessincronização de relógio é um impedimento que apresenta estas características e que tem um impacto no desempenho do sistema para níveis de potência médias em que o efeito combinado do ruído dos fotodetectores e das não linearidades da fibra é menor.

6 Conclusão

Neste trabalho foi estudada a utilização de algoritmos de aprendizagem de máquinas, que realizam a classificação de dados não supervisionada, para auxiliar na mitigação de impedimentos em sistemas ópticos coerentes. O sistema analisado é um sistema de um único span e utiliza modulação DP-16QAM, operando em 100 Gbps e 400 Gbps. Para isso, foram estudados os algoritmos não supervisionados K-means com o método de inicialização K-means++ e o método de EM, além da detecção por ML. Após as simulações dos sistemas realizadas no VPI *Transmission Maker*, utilizou-se da programação em Python para implementação dos algoritmos e para verificação da eficácia dos mesmos na mitigação de impedimentos em sistemas coerentes.

Com os resultados obtidos, podemos responder aos questionamentos levantados na Seção 1.2, que são os objetivos deste trabalho. A primeira questão se refere ao contexto dos sistemas coerentes digitais, que consideram que após a aplicação de algoritmos de processamento digital, apenas são encontrados os efeitos não lineares e o ruído aditivo. Quanto a esta pergunta, os resultados obtidos demonstram que os algoritmos utilizados no DSP das simulações não são capazes de compensar todos os efeitos lineares. Como podemos observar nos resultados, o uso dos algoritmos auxilia na mitigação de impedimentos lineares residuais e não lineares para uma baixa potência lançada, quando se trata de enlaces com menores comprimentos. Nos resultados, para 100 Gbps, podemos observar que para distâncias inferiores a 140 km, a mitigação dos impedimentos para potências entre 0 dBm e 4 dBm ocorre de maneira constante. Para a taxa de 400 Gbps, a mitigação dos efeitos do desbalanceamento da rede híbrida e da dessincronização do relógio ocorre nos enlaces com distâncias inferiores a 100 km. Quando comparamos os resultados obtidos para os Cenários 1 e 2, concluímos que o uso destes algoritmos de *clusterização* auxiliam na mitigação do desbalanceamento da rede híbrida e da dessincronização do relógio, ou seja, os efeitos lineares da dessincronização do relógio e desbalanço da rede híbrida não são completamente compensados pelos algoritmos de DSP, são apenas mitigados.

O segundo questionamento, que se divide em duas hipóteses, se refere ao uso de algoritmos de *clusterização* para mitigação de impedimentos em sistemas coerentes. A primeira hipótese sugere que os algoritmos de *clusterização* podem corrigir os efeitos lineares que não tenham sido compensados pelos algoritmos do DSP. Neste caso, conforme

citado anteriormente, para taxas de 100 Gbps com comprimentos de enlace inferiores a 140 km, podemos melhorar o valor de BER, enquanto que para taxas de 400 Gbps, isso acontece para enlaces inferiores a 100 km. Portanto, os algoritmos de *clusterização* aplicados permitem a mitigação dos efeitos lineares que não sejam totalmente compensados pelos algoritmos no DSP. A segunda hipótese, trata da possibilidade de mitigação de efeitos não lineares com o uso dos algoritmos de *clusterização* K-means e EM. Neste aspecto, diferentemente do que ocorre a taxas de transmissão inferiores, os algoritmos propostos permitem uma correção de não linearidades, mas de uma maneira não tão eficiente. Este fato é verificado para ambos os algoritmos testados neste trabalho e para todos os comprimentos de enlace entre 120 km e 200 km em taxas de 100 Gbps. Para a taxa de 400 Gbps, a mitigação dos efeitos não lineares diminui, quando comparamos com 100 Gbps. A redução de BER observada em maiores taxas de transmissão ocorre devido aos efeitos da distorção nos efeitos não lineares, ou seja, devido ao aumento da ISI em enlaces que operem com maiores taxas de transmissão.

A terceira pergunta que norteia este estudo traz um contexto mais prático para ser analisado. Nela é indagado se o uso destes algoritmos de *clusterização* auxiliam na extensão do comprimento do enlace para uma mesma BER e mesma potência lançada. Sobre esta questão, pode-se concluir que isso ocorre para ambas as taxas de transmissão. Para a taxa de 100 Gbps, há a possibilidade de se obter a mesma BER alcançada pela detecção de ML em uma distância de até 10 km maior, quando aplicados os algoritmos de *clusterização*. Na taxa de 400 Gbps, como é de se esperar, o aumento do alcance diminuiu para apenas 5 km, para um mesmo valor de BER de quando utilizamos ML, sendo este valor $10^{-3,0}$.

Quando analisamos a eficiência dos dois algoritmos aplicados, notamos que ambos tem resultados muito parecidos. Por isso, uma análise do custo computacional também foi realizada, com intuito de verificar qual das ferramentas possui melhor aplicabilidade para este tipo de problema. Os resultados mostram que o custo computacional do algoritmo K-means com inicialização K-means++ é bem mais baixo. Esta redução foi verificada entre 30 a 47 vezes menor, quando comparados os tempos de processamento no processo de aprendizagem do algoritmo.

Com isso, concluímos que apesar de terem um resultado positivo em relação a mitigação de impedimentos, os valores obtidos não são tão vantajosos quanto aos obtidos em sua utilização em menores taxas e com polarização única, conforme demonstram outros

estudos. Quanto mais aumentamos as taxas de transmissão, piores são os resultados obtidos com o uso dos algoritmos de *clusterização*. Isso ocorre por dois pontos, o primeiro deles é a grande ISI presentes nos sistemas com taxas de transmissão maiores, como nos casos estudados. Outro ponto a ser observado é o fato de que a aplicação dos algoritmos de *clusterização* é realizada independentemente em cada polarização. Este fato faz com que a aplicação dos algoritmos não seja capaz de mitigar os efeitos da FWM e XPM referentes à outra polarização, sendo este efeitos não mitigados. Portanto, conclui-se que o uso de algoritmos de *clusterização* podem auxiliar na mitigação das não linearidades e dos efeitos lineares causados pelo desbalanceamento da rede híbrida e da dessincronização do relógio, porém tem limitações quando se aumentam as taxas de transmissão de dados.

Como sequência deste trabalho, propõe-se a aplicação dos mesmos algoritmos para outros formatos de modulação, como o 32-QAM e 64-QAM. Da mesma forma, sugere-se que outros algoritmos de aprendizagem de máquinas sejam utilizados com o mesmo intuito de auxiliar na mitigação dos impedimentos em sistemas coerentes digitais. Além disso, o estudo da aplicabilidade destes algoritmos em sistemas práticos é uma parte muito importante a ser estudada no futuro.

Referências

- AGRAWAL, G. P. Nonlinear fiber optics. In: *Nonlinear Science at the Dawn of the 21st Century*. [S.l.]: Springer, 2000. p. 195–211. Citado 3 vezes nas páginas 27, 28 e 62.
- AGRAWAL, G. P. *Fiber-optic communication systems*. [S.l.]: John Wiley & Sons, 2012. v. 222. Citado 6 vezes nas páginas 17, 19, 24, 25, 28 e 45.
- AGRAWAL, G. P. Optical communication: its history and recent progress. In: *Optics in Our Time*. [S.l.]: Springer, Cham, 2016. p. 177–199. Citado 2 vezes nas páginas 17 e 62.
- ALDAYA, I.; GIACOUMIDIS, E.; OLIVEIRA, G. de; WEI, J.; PITA, J. L.; MARCONI, J. D.; FAGOTTO, E. A. M.; BARRY, L.; ABBADE, M. L. F. Histogram based clustering for nonlinear compensation in long reach coherent passive optical networks. *Applied Sciences*, Multidisciplinary Digital Publishing Institute, v. 10, n. 1, p. 152, 2020. Citado 2 vezes nas páginas 19 e 37.
- ALPAYDIN, E. *Introduction of Machine Learning*. [S.l.]: The MIT Press, 2014. ISBN 9780262028189. Citado 3 vezes nas páginas 35, 36 e 41.
- ARAI, K.; BARAKBAH, A. R. Hierarchical k-means: an algorithm for centroids initialization for k-means. *Reports of the Faculty of Science and Engineering*, v. 36, n. 1, p. 25–31, 2007. Citado na página 39.
- ARTHUR, D.; VASSILVITSKII, S. *k-means++: The advantages of careful seeding*. [S.l.], 2006. Citado na página 39.
- BORGES, L. N.; COSTA, C. B.; PAULA, R. de; ABBADE, M. L.; ALDAYA, I. K-means clustering for mitigation of nonlinear phase noise in digital coherent optical systems using 16-qam modulation format. In: *Simpósio Brasileiro de Telecomunicações - SBrT*. [S.l.: s.n.], 2020. Citado 2 vezes nas páginas 19 e 37.
- CORTES, C.; VAPNIK, V. Support-vector networks. *Machine learning*, Springer, v. 20, n. 3, p. 273–297, 1995. Citado na página 35.
- DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, Wiley Online Library, v. 39, n. 1, p. 1–22, 1977. Citado na página 40.
- DIANOV, E.; LUCHNIKOV, A.; PILIPETSKII, A.; PROKHOROV, A. Long-range interaction of picosecond solitons through excitation of acoustic waves in optical fibers. *Applied Physics B*, Springer, v. 54, n. 2, p. 175–180, 1992. Citado na página 26.
- DU, L. B.; MORSHED, M. M.; LOWERY, A. J. Fiber nonlinearity compensation for ofdm super-channels using optical phase conjugation. *Optics Express*, Optical Society of America, v. 20, n. 18, p. 19921–19927, 2012. Citado na página 33.
- EKANAYAKE, N.; H.HERATH. Effect of nonlinear phase noise on the performance of M-ary PSK signals in optical fiber links. *Journal Lightwave Technology*, v. 31, n. 3, p. 447–454, 2013. Citado 2 vezes nas páginas 19 e 30.

- ELLIS, A.; MCCARTHY, M.; KHATEEB, M. A.; SOROKINA, M.; DORAN, N. Performance limits in optical communications due to fiber nonlinearity. *Advances in Optics and Photonics*, Optical Society of America, v. 9, n. 3, p. 429–503, 2017. Citado na página 26.
- FERNANDEZ, E. A.; TORRES, J. J. G.; SOTO, A. M. C.; GONZÁLEZ, N. G. Demodulation of m-ary non-symmetrical constellations using clustering techniques in optical communication systems. In: IEEE. *2016 IEEE Latin American Conference on Computational Intelligence (LA-CCI)*. [S.l.], 2016. p. 1–6. Citado na página 37.
- FUKUNAGA, K.; NARENDRA, P. M. A branch and bound algorithm for computing k-nearest neighbors. *IEEE transactions on computers*, IEEE, v. 100, n. 7, p. 750–753, 1975. Citado na página 35.
- GARDNER, F. A bpsk/qpsk timing-error detector for sampled receivers. *IEEE Transactions on communications*, IEEE, v. 34, n. 5, p. 423–429, 1986. Citado na página 32.
- GIACOUMIDIS, E.; ALDAYA, I.; VGENOPOULOU, V.; DORAN, N.; JAOUËN, Y. Impact of reduced complexity inverse volterra series transfer function-based nonlinear equalizer in coherent OFDM systems for next-generation core networks. In: IEEE. *2013 15th International Conference on Transparent Optical Networks (ICTON)*. [S.l.], 2013. p. 1–4. Citado 2 vezes nas páginas 34 e 36.
- GIACOUMIDIS, E.; LIN, Y.; JARAJREH, M.; O'DUILL, S.; MCGUINNESS, K.; WHELAN, P. F.; BARRY, L. P. A blind nonlinearity compensator using dbscan clustering for coherent optical transmission systems. *Applied Sciences*, Multidisciplinary Digital Publishing Institute, v. 9, n. 20, p. 4398, 2019. Citado 2 vezes nas páginas 19 e 53.
- GNAUCK, A. H.; TKACH, R.; CHRAPLYVY, A.; LI, T. High-capacity optical transmission systems. *Journal of Lightwave Technology*, IEEE, v. 26, n. 9, p. 1032–1045, 2008. Citado na página 17.
- GREEN, A.; MITRA, P.; WEGENER, L. Effect of chromatic dispersion on nonlinear phase noise. *Optics Letters*, Optical Society of America, v. 28, n. 24, p. 2455–2457, 2003. Citado 2 vezes nas páginas 19 e 24.
- HASSOUN, M. H. *et al. Fundamentals of artificial neural networks*. [S.l.]: MIT press, 1995. Citado na página 35.
- HO, K.-P. Mid-span compensation of nonlinear phase noise. *Optics Communications*, Elsevier, v. 245, n. 1-6, p. 391–398, 2005. Citado na página 33.
- HO, K.-P.; KAHN, J. M. Electronic compensation technique to mitigate nonlinear phase noise. *Journal of Lightwave Technology*, IEEE, v. 22, n. 3, p. 779, 2004. Citado na página 19.
- IP, E.; KAHN, J. M. Compensation of dispersion and nonlinear impairments using digital backpropagation. *Journal of Lightwave Technology*, IEEE, v. 26, n. 20, p. 3416–3425, 2008. Citado na página 34.

- JARAJREH, M. A.; GIACOU MIDIS, E.; ALDAYA, I.; LE, S. T.; TSOKANOS, A.; GHASSEMLOOY, Z.; DORAN, N. J. Artificial neural network nonlinear equalizer for coherent optical OFDM. *IEEE Photonics Technology Letters*, v. 27, n. 4, p. 387–390, 2015. Citado na página 36.
- KHAN, S. S.; AHMAD, A. Cluster center initialization algorithm for k-means clustering. *Pattern Recognition Letters*, Elsevier, v. 25, n. 11, p. 1293–1302, 2004. Citado na página 39.
- KIKUCHI, K. Coherent optical communication systems. In: *Optical Fiber Telecommunications VB*. [S.l.]: Elsevier, 2008. p. 95–129. Citado 2 vezes nas páginas 18 e 22.
- KIKUCHI, K. Fundamentals of coherent optical fiber communications. *Journal of Lightwave Technology*, IEEE, v. 34, n. 1, p. 157–179, 2015. Citado 4 vezes nas páginas 17, 18, 24 e 28.
- KUDO, R.; KOBAYASHI, T.; ISHIHARA, K.; TAKATORI, Y.; SANO, A.; MIYAMOTO, Y. Coherent optical single carrier transmission using overlap frequency domain equalization for long-haul optical systems. *Journal of Lightwave Technology*, IEEE, v. 27, n. 16, p. 3721–3728, 2009. Citado na página 50.
- LIKAS, A.; VLASSIS, N.; VERBEEK, J. J. The global k-means clustering algorithm. *Pattern Recognition Letters*, Elsevier, v. 36, n. 2, p. 451–461, 2003. Citado na página 39.
- LIN, C.-Y.; ASIF, R.; HOLTMANN SPOETTER, M.; SCHMAUSS, B. Nonlinear mitigation using carrier phase estimation and digital backward propagation in coherent QAM transmission. *Optics Express*, Optical Society of America, v. 20, n. 26, p. B405–B412, 2012. Citado na página 36.
- LINKE, R. A.; GNAUCK, A. H. High-capacity coherent lightwave systems. *Journal of Lightwave Technology*, IEEE, v. 6, n. 11, p. 1750–1769, 1988. Citado na página 17.
- LIU, X.; CHRAPLYVY, A.; WINZER, P.; TKACH, R.; CHANDRASEKHAR, S. Phase-conjugated twin waves for communication beyond the kerr nonlinearity limit. *Nature Photonics*, Nature Publishing Group, v. 7, n. 7, p. 560–568, 2013. Citado na página 33.
- LIU, Y.; YANG, B.; XU, T. Machine learning for fiber nonlinearity mitigation in long-haul coherent optical transmission systems. In: IEEE. *2019 IEEE 11th International Conference on Advanced Infocomm Technology (ICAIT)*. [S.l.], 2019. p. 124–127. Citado na página 37.
- LLOYD, S. P. Leats squares quantization in pcm. *IEEE Transactions on Information Theory*, v. 28, n. 2, p. 129–136, 1982. Citado na página 38.
- LOUCHET, H.; KUZMIN, K.; RICHTER, A. Improved dsp algorithms for coherent 16-qam transmission. In: IEEE. *2008 34th European Conference on Optical Communication*. [S.l.], 2008. p. 1–2. Citado na página 50.
- MARSLAND, S. *Machine Learning: an Algorithmic Perspective*. [S.l.]: CRC Press, 2015. ISBN 9781466583337. Citado 4 vezes nas páginas 35, 36, 38 e 41.

- MCCARTHY, J.; FEIGENBAUM, E. A. In memoriam: Arthur samuel: Pioneer in machine learning. *AI Magazine*, v. 11, n. 3, p. 10–10, 1990. Citado na página 35.
- MENGALI, U. *Synchronization techniques for digital receivers*. [S.l.]: Springer Science & Business Media, 2013. Citado na página 32.
- MHATLI, S.; NSIRI, B.; JARAJREH, M. A.; CHANNOUFI, M.; ATTIA, R. Analysis of Wiener Hammerstein equalizer for optical OFDM modem. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. *Photonics, Devices, and Systems VI*. [S.l.], 2015. v. 9450, p. 94500X. Citado na página 34.
- MILLAR, D. S.; MAKOVEJS, S.; BEHRENS, C.; HELLERBRAND, S.; KILLEY, R. I.; BAYVEL, P.; SAVORY, S. J. Mitigation of fiber nonlinearity using a digital coherent receiver. *IEEE Journal of Selected Topics in Quantum Electronics*, v. 16, n. 5, p. 1217–1226, 2010. Citado na página 34.
- MUELLER, K.; MULLER, M. Timing recovery in digital synchronous data receivers. *IEEE transactions on communications*, IEEE, v. 24, n. 5, p. 516–531, 1976. Citado na página 32.
- OERDER, M.; MEYR, H. Digital filter and square timing recovery. *IEEE Transactions on communications*, IEEE, v. 36, n. 5, p. 605–612, 1988. Citado na página 50.
- OIF. *Implementation Agreement 400ZR, OIF-400ZR-01.0*. [S.l.], 2020. Citado 2 vezes nas páginas 19 e 26.
- OTANI, T.; GOTO, K.; ABE, H.; TANAKA, M.; YAMAMOTO, H.; WAKABAYASHI, H. 5.3 gbit/s 11300 km data transmission using actual submarine cables and repeaters. *Electronics Letters*, IET, v. 31, n. 5, p. 380–381, 1995. Citado na página 17.
- PAKALA, L.; SCHMAUSS, B. Non-linear mitigation using carrier phase estimation and k-means clustering. In: VDE. *Photonic Networks; 16. ITG Symposium*. [S.l.], 2015. p. 1–5. Citado na página 34.
- PAN, Z.; CHÂTELAIN, B.; CHAGNON, M.; PLANT, D. V. Volterra filtering for nonlinearity impairment mitigation in DP-16QAM and DP-QPSK fiber optic communication systems. In: OPTICAL SOCIETY OF AMERICA. *National Fiber Optic Engineers Conference*. [S.l.], 2011. p. JThA040. Citado na página 36.
- PAULA, R. de; BORGES, L. N.; ABBADE, M. L. F.; ALDAYA, I. Mitigation of nonlinear phase noise in coherent 16-qam long-reach pons by k-nearest neighbors-based classification. In: *Simpósio Brasileiro de Telecomunicações - SBrT*. [S.l.: s.n.], 2020. Citado 2 vezes nas páginas 19 e 37.
- PETROU, C.; VGENIS, A.; ROUDAS, I.; RAPTIS, L. Quadrature imbalance compensation algorithms for coherent pdm qpsk systems. In: IEEE. *2009 IEEE LEOS Annual Meeting Conference Proceedings*. [S.l.], 2009. p. 642–643. Citado na página 32.
- POWELL, M. J. Least frobenius norm updating of quadratic models that satisfy interpolation conditions. *Mathematical Programming*, Springer, v. 100, n. 1, p. 183–215, 2004. Citado na página 51.

PUTTNAM, B. J.; LUÍS, R. S.; MENDINUETA, J. M. D.; SAKAGUCHI, J.; KLAUS, W.; KAMIO, Y.; NAKAMURA, M.; WADA, N.; AWAJI, Y.; KANNO, A. *et al.* Self-homodyne detection in optical communication systems. In: MULTIDISCIPLINARY DIGITAL PUBLISHING INSTITUTE. *Photonics*. [S.l.], 2014. v. 1, n. 2, p. 110–130. Citado na página 28.

RIBEIRO, V. B.; DINIZ, J. C. M.; PARAHYBA, V. E. S.; ROSA, E.; RANZINI, S.; SILVA, F.; OLIVEIRA, J. Processamento de sinais para redes ópticas coerentes digitais de alta velocidade. *Cadernos CPqD tecnologia*, 2011. Citado 2 vezes nas páginas 25 e 32.

ROS, F. D.; SACKKEY, I.; ELSCHNER, R.; RICHTER, T.; MEUER, C.; NÖLLE, M.; JAZAYERIFAR, M.; PETERMANN, K.; PEUCHERET, C.; SCHUBERT, C. Kerr nonlinearity compensation in a 5×28 -GBd PDM 16-QAM WDM system using fiber-based optical phase conjugation. In: IEEE. *The European Conference on Optical Communication (ECOC)*. [S.l.], 2014. p. 1–3. Citado na página 36.

SAVORY, S. J. Digital filters for coherent optical receivers. *Optics express*, Optical Society of America, v. 16, n. 2, p. 804–817, 2008. Citado 3 vezes nas páginas 31, 33 e 50.

SCULLEY, D. Web-scale k-means clustering. In: *Proceedings of the 19th international conference on World wide web*. [S.l.: s.n.], 2010. p. 1177–1178. Citado na página 53.

SECONDINI, M.; FORESTIERI, E. On xpm mitigation in wdm fiber-optic systems. *IEEE Photonics Technology Letters*, IEEE, v. 26, n. 22, p. 2252–2255, 2014. Citado na página 29.

SUNNERUD, H.; XIE, C.; KARLSSON, M.; SAMUELSSON, R.; ANDREKSON, P. A. A comparison between different PMD compensation techniques. *Journal of Lightwave Technology*, IEEE, v. 20, n. 3, p. 368, 2002. Citado 2 vezes nas páginas 19 e 25.

SUTTON, R. S. Dyna, an integrated architecture for learning, planning, and reacting. *ACM Sigart Bulletin*, ACM New York, NY, USA, v. 2, n. 4, p. 160–163, 1991. Citado na página 36.

SUTTON, R. S.; BARTO, A. G. *et al.* *Introduction to reinforcement learning*. [S.l.]: MIT press Cambridge, 1998. v. 135. Citado na página 36.

TORRES-ZUGAIDE, J.; ALDAYA, I.; CAMPUZANO, G.; CASTANON, G. Hammerstein-based equalizer for nonlinear compensation in coherent OFDM long-reach PONs. In: IEEE. *18th International Conference on Transparent Optical Networks (ICTON)*. [S.l.], 2016. p. 1–3. Citado na página 34.

TORRES-ZUGAIDE, J.; ALDAYA, I.; CAMPUZANO, G.; GIACOUMIDIS, E.; BEAS, J.; CASTAÑÓN, G. Range extension in coherent OFDM passive optical networks using an inverse Hammerstein nonlinear equalizer. *Journal of Optical Communications and Networking*, Optical Society of America, v. 9, n. 7, p. 577–584, 2017. Citado na página 34.

TOU, J. T. DYNOC-A dynamic optimal cluster-seeking technique. *International Journal of Computer & Information Sciences*, Springer, v. 8, n. 6, p. 541–547, 1979. Citado na página 39.

TSUKAMOTO, S.; LY-GAGNON, D.-S.; KATOH, K.; KIKUCHI, K. Coherent demodulation of 40-Gbit/s polarization-multiplexed QPSK signals with 16-GHz spacing after 200-km transmission. In: OPTICAL SOCIETY OF AMERICA. *Optical Fiber Communication Conference*. [S.l.], 2005. p. PDP29. Citado 4 vezes nas páginas 18, 22, 24 e 47.

WANG, D.; ZHANG, M.; FU, M.; CAI, Z.; LI, Z.; HAN, H.; CUI, Y.; LUO, B. Nonlinearity mitigation using a machine learning detector based on k -Nearest Neighbors. *IEEE Photonics Technology Letters*, IEEE, v. 28, n. 19, p. 2102–2105, 2016. Citado na página 37.

WANG, D.; ZHANG, M.; LI, Z.; CUI, Y.; LIU, J.; YANG, Y.; WANG, H. Nonlinear decision boundary created by a machine learning-based classifier to mitigate nonlinear phase noise. In: *European Conference on Optical Communication (ECOC)*. [S.l.: s.n.], 2015. p. 3–16. Citado na página 36.

WATKINS, C. J.; DAYAN, P. Q-learning. *Machine learning*, Springer, v. 8, n. 3-4, p. 279–292, 1992. Citado na página 36.

ZHANG WEI CHEN, M. G. J.; SHEN, G. K-means-clustering-based fiber nonlinearity equalization techniques for 64-QAM coherent optical communication system. *Optics Express*, v. 25, n. 22, p. 27570–27580, 2017. Citado na página 37.

ZIBAR, D.; PIELS, M.; JONES, R.; SCHÄEFFER, C. G. Machine learning techniques in optical communication. *Journal of Lightwave Technology*, IEEE, v. 34, n. 6, p. 1442–1452, 2015. Citado 2 vezes nas páginas 19 e 35.

ZIBAR, D.; WINTHER, O.; FRANCESCHI, N.; BORKOWSKI, R.; CABALLERO, A.; ARLUNNO, V.; SCHMIDT, M. N.; GONZALES, N. G.; MAO, B.; YE, Y. Nonlinear impairment compensation using expectation maximization for dispersion managed and unmanaged PDM 16-QAM transmission. *Optics Express*, Optical Society of America, v. 20, n. 26, p. B181–B196, 2012. Citado na página 37.