Optical Fiber Technology 90 (2025) 104047

Contents lists available at ScienceDirect

Optical Fiber Technology

journal homepage: www.elsevier.com/locate/yofte

Machine learning opportunities for integrated polarization sensing and communication in optical fibers

Andrej Rode [a,*,1], Mohammad Farsi [b], Vincent Lauinger [a], Magnus Karlsson [c], Erik Agrell [b], Laurent Schmalen [a], Christian Häger [b]

[a] Communications Engineering Lab (CEL), Karlsruhe Institute of Technology, Karlsruhe, Germany
[b] Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden
[c] Department of Microtechnology and Nanoscience, Chalmers University of Technology, Gothenburg, Sweden

ARTICLE INFO

Keywords:
End-to-end autoencoders
Machine learning
Physics-based learning
Polarization sensing
Variational autoencoders

ABSTRACT

As the bedrock of the Internet, optical fibers are ubiquitously deployed and historically dedicated to ensuring robust data transmission. Leveraging their extensive installation, recent endeavors have focused on utilizing these telecommunication fibers also for environmental sensing, exploiting their inherent sensitivity to various environmental disturbances. In this paper, we consider integrated sensing and communication (ISAC) systems that combine data transmission and sensing functionalities, by monitoring the state of polarization to detect environmental changes. In particular, we investigate various machine learning techniques to enhance the performance and capabilities of such polarization-based ISAC systems. Gradient-based techniques such as adaptive zero-forcing equalization are examined for their potential to enhance sensing accuracy at the expense of communication performance, with strategies discussed for mitigating this trade-off. Additionally, the paper reviews novel machine-learning-based approaches for blind channel estimation using variational autoencoders, aimed at improving channel estimates compared to traditional adaptive equalization methods. We also discuss the problem of distributed polarization sensing and review a recent physics-based learning approach for Jones matrix factorization, potentially enabling spatial resolution of sensed events. Lastly, we discuss the potential of leveraging dual-functional autoencoders to optimize ISAC transmitters and the corresponding transmit waveforms. Our paper underscores the potential of telecom fibers for joint data transmission and environmental sensing, facilitated by advancements in digital signal processing and machine learning.

1. Introduction

With the emergence of optical fiber communications as the backbone of the Internet, optical fibers are deployed virtually everywhere. While their primary use is to ensure reliable data transmission, many recent studies have investigated the possibility of repurposing such fibers into environmental sensors [1–3]. Indeed, optical fibers have been harnessed for dedicated sensing applications for several decades, playing an important role in environments that require sensitive, precise, and rapid monitoring [4–9].

Traditional fiber sensing works by launching a probing signal into the fiber and examining the subsequent backscattered light. This backscattered light carries the imprint of various external influences such as changes in temperature, pressure, or mechanical strain. By translating time-of-flight information into distances, one can further pinpoint where a specific disturbance occurred along the fiber's length. This feature, known as distributed optical fiber sensing (DOFS), has revolutionized various industries with applications ranging from monitoring pressure variations in pipelines to overseeing structural integrity of critical infrastructures [10–12].

Whereas dedicated fiber sensing applications utilize fibers not intended for simultaneous data transmission, telecom fibers can be buried deep beneath the ground, lie on the ocean floor surrounded by massive amounts of protective shielding, or simply hang in the air suspended from poles. Even though these are not exactly scenarios envisioned for sensing, it has been shown that DOFS techniques applied on telecom fibers are sensitive enough to pick up urban traffic patterns [13,14], seismic activities [15–18] or even whale movements [19]. See also [20–22] for an overview of related work.

* Corresponding author.
E-mail address: rode@kit.edu (A. Rode).

[1] A. Rode conducted this research while he was visiting the Communication Systems Group at the Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden.

https://doi.org/10.1016/j.yofte.2024.104047
Received 1 July 2024; Accepted 12 November 2024
Available online 11 December 2024

While the backscatter-based approaches used in the above papers can provide accurate and spatially-resolved sensing information, they suffer from multiple drawbacks. In particular, they require extra equipment, have limited range, are often incompatible with existing telecom infrastructure such as optical amplifiers, and sacrifice valuable communication bandwidth. A different, less intrusive, approach is to unify both functionalities into an integrated sensing and communication (ISAC) system. Prior work has shown that sensing functionality can indeed be built on top of standard digital signal processing (DSP) blocks that are already widely implemented in current coherent optical communication receivers such as phase estimation [23,24] or adaptive equalization [25,26]. This paper focuses on the latter approach, which aims at the extraction of the signal's state of polarization (SOP). The SOP is known to be sensitive to the surrounding environment [27–30] and has to be estimated continuously in communication receivers to compensate the observed signal.

The idea of using polarization information for sensing purposes was investigated for example in [31], where SOP changes of an out-of-band (non-information-carrying) laser are tracked and processed using a Markov state model. The authors in [25] show that SOP information can be extracted directly from a standard coherent telecom receiver and propose an approach to monitor certain mechanical stresses applied to the fiber. In recent field trials over submarine links [26,32], it was shown that SOP monitoring is sufficiently accurate to allow for seismic and water wave sensing. Subsequent work has shown that similar approaches are also applicable in urban settings [33] and over aerial fiber links [34,35]. See also [36–39] for related works.

In this paper, we explore opportunities to apply techniques from machine learning to optimize the performance of optical ISAC systems, focusing on polarization sensing. Indeed, the workhorse that enables polarization sensing in current coherent telecom transponders is an adaptive multiple-input multiple-output (MIMO) equalizer [40,41]. From a machine learning perspective, such equalizers perform gradient-based *online learning* via least mean square (LMS) filtering [42]. With this perspective, the paper covers the following topics:

- *Adaptive Zero-Forcing Equalization and Trade-offs between Communication and Sensing:* Adaptive equalizers in optical receivers are optimized using purely communication-centric loss functions, in particular minimizing the error between transmitted and received (or estimated) data symbols. As a result, the acquired equalizer taps do not necessarily converge to the inverse of the (estimated) Jones matrix of the channel. Here, we explore the use of adaptive zero-forcing (ZF) equalization [43,44] and show some examples where such equalizers can provide more accurate estimates of the Jones matrix, which may translate into improved sensing performance and higher sensitivity. While adaptive ZF comes at the expense of a loss in communication performance (due to possible noise enhancement), we also show that by weighting (or interpolating) the regression vectors of LMS and ZF (defined in Eqs. (15) and (16) below), one can achieve flexible trade-offs between communication and sensing performance.
- *Dedicated Blind Channel Estimation using Variational Autoencoders:* In the communication context, a novel blind equalization algorithm called the VAE was recently introduced in [45]. It utilizes state-of-the-art machine learning techniques and inherently separates the machine learning models for communication (i.e., equalization) and sensing (i.e., channel estimation) without required labeled training data. We review our recent works in [46–50] on VAE-based equalization for optical systems and show that this approach can be used to explicitly learn the MIMO channel impulse responses and provide a more accurate channel estimate compared to adaptive LMS algorithms without the need for pilot symbols which reduce spectral efficiency.
- *Distributed Polarization Sensing via Jones Matrix Factorization:* A major issue with current polarization sensing is that it does not allow for event localization, i.e., distributed sensing. This is because adaptive equalizers only acquire SOP information that is integrated over the entire propagation distance. Here, we review the recently proposed physics-based machine learning (PML) approach in [51], which aims at recovering spatially resolved polarization information by factorizing the estimated Jones matrix into individual responses. We show that this approach leads to improved noise resilience compared to previous inverse scattering algorithms (ISAs) [52–54], while highlighting challenges related to model overparameterization.
- *ISAC Transmitter Trade-Offs and Waveform Optimization:* In addition to the receiver DSP algorithms discussed so far, we also examine communication–sensing trade-offs associated with ISAC transmitter design. Furthermore, we highlight the potential of leveraging advanced machine learning methods such as dual-functional autoencoders (AEs) to optimize transmit waveforms and enhance both communication and sensing capabilities.

Lastly, we remark that machine learning has already been extensively studied to improve communication performance of coherent optical transmission systems, particularly in nonlinear settings, see, e.g., [55–65] and references therein. Moreover, there has also been significant work on utilizing machine learning to process raw sensing data, e.g., for event detection and classification [66–76]. On the other hand, the learning-based ISAC approaches discussed in this paper operate at a lower level, targeting instead the accurate *acquisition* of raw sensing data, while simultaneously ensuring reliable data transmission.

The remainder of this paper is structured as follows. First, in Section 2 we introduce the system model that is used in the remainder of this paper. The system model captures key features of the optical channel as it is observed in the DSP and highlights parts of the system where we believe that machine learning can provide a benefit for ISAC. In Section 3, we discuss the application of adaptive ZF and VAEs, and numerically investigate the associated trade-offs between communication and sensing performance. While we discuss the retrieval of channel parameters up to this point, in Section 4 we discuss how we can apply machine learning to infer location information about the physical channel from the estimated overall channel impulse response. Transmitter trade-offs and optimization opportunities using machine learning techniques are discussed in Section 5. Finally, the paper is concluded in Section 6.

2. System model

In Fig. 1, we show the considered system model. At the transmitter, we choose $2N_{\text{sym}}$ transmit symbols $\boldsymbol{X} \in \mathbb{C}^{2\times N_{\text{sym}}}$ according to a desired probability mass function $P(\mathcal{M})$ over a constellation alphabet $\mathcal{M}$. The transmit symbols $\boldsymbol{X}$ are upsampled T times and a root-raised-cosine (RRC) pulse-shaping filter is applied to obtain the transmit signal $\boldsymbol{S} = \left(\boldsymbol{s}_1\ \boldsymbol{s}_2\right)^{\mathsf{T}} \in \mathbb{C}^{2\times N_{\text{sym}}T}$. Then, we apply the optical channel transfer matrix $\boldsymbol{H}(\omega,\boldsymbol{\theta}) \in \mathbb{C}^{2\times 2}$ in the frequency domain, add additive white Gaussian noise (AWGN) $\boldsymbol{N} = \left(\boldsymbol{n}_1\ \boldsymbol{n}_2\right)^{\mathsf{T}}$ where each component of $\boldsymbol{N}$ is independent and identically distributed according to $\mathcal{N}\left(0,\sigma_{\text{w}}^2\right)$, and obtain the received signal

$$\boldsymbol{Z} = \mathcal{F}^{-1}\left(\boldsymbol{H}(\omega,\boldsymbol{\theta})\,\mathcal{F}(\boldsymbol{S})\right) + \boldsymbol{N}, \tag{1}$$

where $\boldsymbol{\theta}$ is a vector of channel parameters, and $\mathcal{F}$ and $\mathcal{F}^{-1}$ are the Fourier transform and inverse Fourier transform respectively. The parameter vector $\boldsymbol{\theta}$ and the optical channel model used in our work are described in more detail below in Section 2.1.

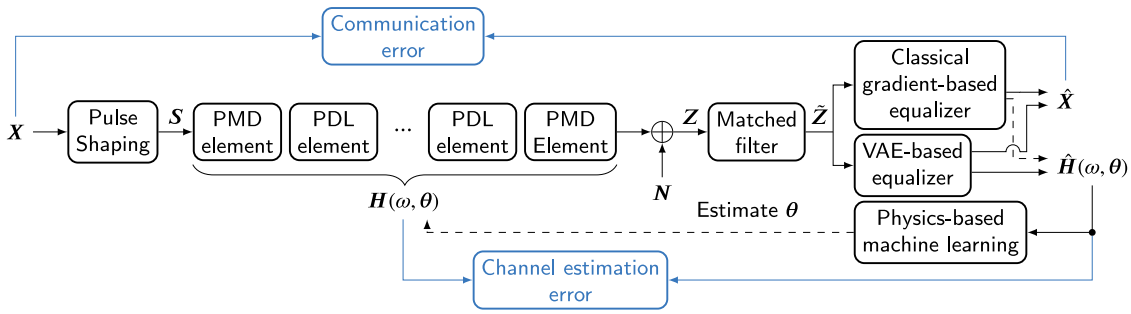

Fig. 1. Considered system model to study integrated polarization sensing and communication. Transmit symbols $\boldsymbol{X}$ are sent through an optical channel model with PMD, PDL, and additive noise. Gradient-based equalizers and VAE-based equalizers obtain the communication symbols estimate $\hat{\boldsymbol{X}}$ and channel estimate $\hat{\boldsymbol{H}}(\omega,\boldsymbol{\theta})$ (Section 3). Physics-based machine learning aims to infer the underlying channel parameters $\boldsymbol{\theta}$ from the channel estimate (Section 4).

At the receiver, we first apply a matched filter to the received signal $\boldsymbol{Z}$ separately for each polarization to obtain $\tilde{\boldsymbol{Z}}$.[2] The signal $\tilde{\boldsymbol{Z}} = (\tilde{\boldsymbol{z}}_1\ \tilde{\boldsymbol{z}}_2)^{\intercal}$ is then processed by a linear 2×2 MIMO equalizer with a set of complex-valued parameter vectors $\tilde{\boldsymbol{c}}_k = \{\boldsymbol{c}_{1,1,k}, \boldsymbol{c}_{1,2,k}, \boldsymbol{c}_{2,1,k}, \boldsymbol{c}_{2,2,k}\}$. The subscript k indicates that at each time-step an updated parameter vector is used by the equalizer. Each parameter vector $\boldsymbol{c}_{p,q,k} \in \mathbb{C}^{2L+1\times 1}$ contains the $2L+1$ complex taps of the equalizer to apply to the signal on polarization q, to obtain part of the signal on polarization p at time step k. At the input of the equalizer we only use a part of the received signal $\tilde{\boldsymbol{z}}_{p,k} = (\tilde{z}_{p,k-L}\ \dots\ \tilde{z}_{p,k}\ \dots\ \tilde{z}_{p,k+L})^{\intercal} \in \mathbb{C}^{2L+1\times 1}$ on each polarization p, which corresponds to the number of equalizer taps, to return the equalized signal

$$\tilde{x}_{p,k} = \sum_{i=1}^{2} \tilde{\boldsymbol{z}}_{i,k}^{\intercal} \boldsymbol{c}_{p,i,k} \tag{2}$$

on polarization p at time step k. The equalized signal $\tilde{\boldsymbol{x}}_p = (\tilde{x}_{p,0}\ \dots\ \tilde{x}_{p,N_{\mathrm{sym}}T})^{\intercal}$ on each polarization p is then further downsampled by a factor T to get $(\hat{\boldsymbol{x}}_1\ \hat{\boldsymbol{x}}_2)^{\intercal} = \hat{\boldsymbol{X}}$. We visualize the operation of the linear MIMO equalizer in Fig. 2.

Simultaneously, we use the equalizers to find the channel estimate $\hat{\boldsymbol{H}}(\omega,\boldsymbol{\theta})$. The channel transfer matrix estimate $\hat{\boldsymbol{H}}(\omega,\boldsymbol{\theta})$ is either derived from the parameter vectors $\tilde{\boldsymbol{c}}_k$ of the equalizers or explicitly learned and provided by VAE-based equalizers. The precise operation of the considered equalizers and channel estimation techniques is described in Section 3. In Section 4, we discuss how to further process the (noisy) channel estimate to find the underlying channel parameters $\boldsymbol{\theta}$.

2.1. Optical fiber channel model

We model the optical fiber in terms of the linear polarization-dependent effects which can observed in the DSP, in particular polarization mode dispersion (PMD) and polarization dependent loss (PDL). We assume that chromatic dispersion (CD) can be sufficiently removed with a static frequency domain equalizer and therefore we are not including it in the system model. For simplicity, we also do not consider nonlinear channel effects on the transmission signal in our investigation.

We use the standard retarder plate model, where each fiber realization is modeled as N independent polarization-maintaining fiber plates [77] with varying differential group delay (DGD) and PDL. At each fiber plate, a random state of polarization (SOP) rotation is introduced to the signal and combined with DGD this models the PMD. The frequency-dependent channel transfer matrix is given by[3]

$$\boldsymbol{H}(\omega,\boldsymbol{\theta}) = \prod_{n=1}^{N} \boldsymbol{R}_n \boldsymbol{Q}_n(\omega)\, \boldsymbol{\Gamma}_n, \tag{3}$$

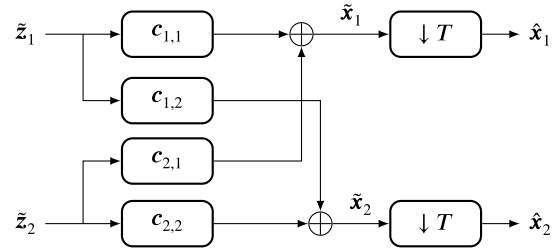

Fig. 2. Block diagram of a 2×2 MIMO equalizer, where we use $\tilde{\boldsymbol{Z}} = (\tilde{\boldsymbol{z}}_1\ \tilde{\boldsymbol{z}}_2)^{\intercal}$, $\hat{\boldsymbol{X}} = (\hat{\boldsymbol{x}}_1\ \hat{\boldsymbol{x}}_2)^{\intercal}$, and complex valued parameter vectors $\boldsymbol{c}_{(\cdot)}$. The equalizer also performs a T-fold downsampling.

where, in each segment n, $\boldsymbol{R}_n \in \mathbb{C}^{2\times 2}$ is a random polarization rotation matrix, the matrix $\boldsymbol{Q}_n(\omega) \in \mathbb{C}^{2\times 2}$ describes the DGD, and the matrix $\boldsymbol{\Gamma}_n \in \mathbb{R}^{2\times 2}$ describes the PDL. In particular, the polarization rotation matrix is

$$\boldsymbol{R}_n = \boldsymbol{I}_2 \cos(\vartheta_n) - \mathrm{j}\, \boldsymbol{a}_n^{\intercal} \vec{\boldsymbol{\sigma}} \sin(\vartheta_n), \tag{4}$$

where $\boldsymbol{I}_2$ is a 2×2 identity matrix,

$$\vec{\boldsymbol{\sigma}} = \left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -\mathrm{j} \\ \mathrm{j} & 0 \end{pmatrix} \right) \tag{5}$$

is the tensor of Pauli spin matrices, ϑ_n is a rotation angle, and $\boldsymbol{a}_n \in \mathbb{C}^{3\times 1\times 1}$ is a unit vector $\|\boldsymbol{a}_n\|^2 = 1$. The operation $\boldsymbol{a}_n^{\intercal}\vec{\boldsymbol{\sigma}}$ should be interpreted as a linear combination of the Pauli spin matrices. The accumulated DGD τ_n at each segment is modeled with

$$\boldsymbol{Q}_n(\omega) = \begin{pmatrix} \mathrm{e}^{-\mathrm{j}\frac{\tau_n}{2}\frac{\Delta z}{L_c}\omega} & 0 \\ 0 & \mathrm{e}^{\mathrm{j}\frac{\tau_n}{2}\frac{\Delta z}{L_c}\omega} \end{pmatrix}, \tag{6}$$

where L_c is the PMD correlation length and Δz is the length of each segment. The PDL for each segment k is modeled by a matrix

$$\boldsymbol{\Gamma}_n = \begin{pmatrix} \mathrm{e}^{\frac{\gamma_n}{2}\Delta z} & 0 \\ 0 & \mathrm{e}^{-\frac{\gamma_n}{2}\Delta z} \end{pmatrix}. \tag{7}$$

This matrix applies a differential loss term $\mathrm{e}^{\gamma_n \Delta z}$ to a given segment n. While this does not correctly represent the effects physically since it amplifies the signal on one polarization and attenuates the signal on the other polarization, it correctly captures the effects of the PDL in the DSP. The matrices $\boldsymbol{R}_n$ and $\boldsymbol{Q}_n(\omega)$ are unitary, while the PDL matrix $\boldsymbol{\Gamma}_n$ in general is non-unitary. The vector $\boldsymbol{\theta}$ comprises the underlying channel parameters τ_n, γ_n, ϑ_n and $\boldsymbol{a}_n$ across all N segments.

[2] In practical systems, the matched filter is not applied separately from the equalizer, rather it is contained within the equalizer. Here, we apply a dedicated matched filter to ensure that the subsequent equalizer only reverts the channel effects and not the additional inter-symbol interference introduced by the RRC pulse-shaping filter, which allows for easier performance assessment.

[3] We use the convention $\prod_{n=1}^{N} \boldsymbol{A}_n = \boldsymbol{A}_N \boldsymbol{A}_{N-1} \cdots \boldsymbol{A}_1$.

2.2. Performance metrics

To assess the communication performance of the equalizers, we use the signal-to-noise ratio (SNR)

$$\mathrm{SNR} = 10\log_{10}\left(\frac{\mathbb{E}\left(\|\boldsymbol{X}\|_{\mathrm{F}}^2\right)}{\mathbb{E}\left(\left\|\boldsymbol{X}-\hat{\boldsymbol{X}}\right\|_{\mathrm{F}}^2\right)}\right), \tag{8}$$

where $\|\cdot\|_{\mathrm{F}}^2$ denotes the Frobenius norm.

Optical fiber sensing has the goal of measuring environmental quantities that induce a change in the properties of the optical fiber. We can measure this change in the properties in optical fiber communication as a change of the channel impulse response or equivalently as a change in the channel transfer matrix $\boldsymbol{H}(\omega,\boldsymbol{\theta})$. Together with the channel transfer matrix estimate $\hat{\boldsymbol{H}}(\omega,\boldsymbol{\theta})$, we assess sensing performance via the mean squared error

$$\mathrm{MSE} = \frac{1}{L_{\mathrm{f}}}\sum_{i=1}^{L_{\mathrm{f}}}\left\|\hat{\boldsymbol{H}}\left(\omega_i,\boldsymbol{\theta}\right)-\boldsymbol{H}\left(\omega_i,\boldsymbol{\theta}\right)\right\|_{\mathrm{F}}^2 \tag{9}$$

over L_{f} discrete frequencies ω_i within the usable signal bandwidth. In our system model, this usable signal bandwidth corresponds approximately to the symbol rate.

3. Adaptive equalization for integrated polarization sensing and communication

In this section, we discuss how adaptive equalizers can be used to obtain the channel estimate $\hat{\boldsymbol{H}}(\omega,\boldsymbol{\theta})$, while performing their intended purpose of equalizing the received communication signal. In Section 3.1, we will discuss the difference between the adaptive LMS and ZF in terms of communication and sensing performance. While the LMS algorithm has seen wide adoption in adaptive equalizers, we will discuss opportunities and challenges to use the ZF algorithm for ISAC. In Section 3.2 we will discuss the recently-proposed VAE equalizer and how we can exploit its structure for ISAC systems.

3.1. Classical gradient-based adaptive equalization

To enable communications over unknown linear channels, gradient-descent-based adaptive equalizer algorithms have been employed in communications engineering at least since the introduction of the LMS algorithm by Widrow in [78,79] and the adaptive ZF equalizer by Lucky in [43,44]. For 2×2 MIMO systems, the equalization operation of the adaptive LMS and adaptive ZF equalizer can be realized with a butterfly filter structure, depicted in Fig. 2.

The interesting part of the adaptive LMS and ZF equalizer is the set of parameter update equations. Together with the error

$$\epsilon_{p,k} = x_{p,k} - \hat{x}_{p,k}. \tag{10}$$

at the current time step k and polarization p between a known pilot symbol $x_{p,k}$ and the output of the equalizer $\hat{x}_{p,k}$, and a regression vector $\boldsymbol{y}_{p,k}\in\mathbb{C}^{2L-1\times 1}$, an updated set of parameters can be calculated for the next time step. We will introduce the used regression vectors in (15) and (16) for the adaptive LMS and adaptive ZF equalizer respectively. In [44,80], the adaptive parameter update was derived for a single-input single-output (SISO) channel. In [40], update equations for a decision-directed LMS are given for a dual-polarization 2×2 MIMO equalizer, which we can use to formulate the update of the parameter vectors

$$\boldsymbol{c}_{1,1,k+1} \leftarrow \boldsymbol{c}_{1,1,k} + 2\mu\epsilon_{1,k}\boldsymbol{y}_{1,k}^* \tag{11}$$

$$\boldsymbol{c}_{1,2,k+1} \leftarrow \boldsymbol{c}_{1,2,k} + 2\mu\epsilon_{1,k}\boldsymbol{y}_{2,k}^* \tag{12}$$

$$\boldsymbol{c}_{2,2,k+1} \leftarrow \boldsymbol{c}_{2,2,k} + 2\mu\epsilon_{2,k}\boldsymbol{y}_{2,k}^* \tag{13}$$

$$\boldsymbol{c}_{2,1,k+1} \leftarrow \boldsymbol{c}_{2,1,k} + 2\mu\epsilon_{2,k}\boldsymbol{y}_{1,k}^*. \tag{14}$$

The variable μ is a user-defined step size, which is the only free parameter of the adaptive update equations. The proper selection of the step size is crucial to control both robustness and speed of convergence. The calculation of the error signal $\epsilon_{p,k}$ requires prior knowledge of a pilot sequence. After initial convergence with a known pilot sequence, the operation is commonly switched to a decision-directed mode, where a (delayed) feedback is used after a hard decision on a symbol from the constellation alphabet $\mathcal{M}$ has been performed with a sufficiently small symbol error rate. The parameter vectors $\tilde{\boldsymbol{c}}_k$ are then used to estimate the sent information symbol according to the principle of the 2×2 butterfly filter. Downsampling the received signal by a factor of T is achieved by advancing the received signal and the known pilot sequence by T samples for processing in the next time step.

This general description of the update equations is the same for the adaptive LMS and the adaptive ZF algorithm. The difference between them lies in the choice of regression vectors. The regression vector in case of the LMS equalizer is

$$\boldsymbol{y}_{p,k}^{(\mathrm{LMS})} = \begin{pmatrix}\tilde{z}_{p,k-L} & \dots & \tilde{z}_{p,k} & \dots & \tilde{z}_{p,k+L}\end{pmatrix}^{\mathsf{T}} \tag{15}$$

and

$$\boldsymbol{y}_{p,k}^{(\mathrm{ZF})} = \begin{pmatrix}x_{p,k-L} & \dots & x_{k,l} & \dots & x_{p,k+L}\end{pmatrix}^{\mathsf{T}} \tag{16}$$

in case of the ZF equalizer. The regression vector of the LMS equalizer can be derived from minimizing the mean squared error (MSE) between the estimated symbol and the known pilot symbol. For the ZF equalizer, the regression vector can be derived from the minimization of the peak distortion [80].

It is known that the application of the ZF algorithm leads to noise enhancement, while the LMS algorithm leads to solutions more robust to noise. For ISAC systems, we want to explore equalizer algorithms which do not only show a good communication performance and robustness in noisy environments but simultaneously provide a more accurate channel estimate. Therefore, we propose to combine the regression vectors of the LMS and ZF algorithms to form a weighted regression vector

$$\boldsymbol{y}_{p,k}^{(\mathrm{LMS-ZF})} = \sqrt{\nu}\boldsymbol{y}_{p,k}^{(\mathrm{LMS})} + \sqrt{1-\nu}\boldsymbol{y}_{p,k}^{(\mathrm{ZF})}, \tag{17}$$

where $\nu\in[0,1]$ is the weight given to the adaptive LMS algorithm over the adaptive ZF algorithm.

The parameter update equations only help us to obtain parameter vectors $\tilde{\boldsymbol{c}}_k$, but they do not directly give $\hat{\boldsymbol{H}}(\omega,\boldsymbol{\theta})$. The parameter vectors $\tilde{\boldsymbol{c}}_k$ define the equalizer in the time domain. Following the concept of zero-forcing, we can calculate the ZF inverse to $\tilde{\boldsymbol{c}}_k$ and use that as a channel estimate. We follow the approach in [81] to calculate the inverse in the context of a multiuser MIMO channel. First, we expand each of the parameter vectors $\boldsymbol{c}_{p,q,k}$ contained in $\tilde{\boldsymbol{c}}_k$ to its corresponding Toeplitz form $\boldsymbol{C}_{p,q,k}\in\mathbb{C}^{2L+2\times 2(2L+1)}$. Then we combine the four Toeplitz form parameter matrices to a larger compound matrix $\boldsymbol{C}_{\mathrm{T},k}\in\mathbb{C}^{2(2L+2)\times 2(2L+1)}$, which corresponds to the operation [81, Eq. (4)]. Now, this compound matrix fully describes the operation of the equalizer when the received and filtered signals $\tilde{\boldsymbol{Z}}_k$ on the two polarizations are stacked and multiplied from the right

$$\begin{pmatrix}\check{\boldsymbol{x}}_{1,k}\\ \check{\boldsymbol{x}}_{2,k}\end{pmatrix} = \boldsymbol{C}_{\mathrm{T},k}\begin{pmatrix}\dot{\boldsymbol{z}}_{1,k}\\ \dot{\boldsymbol{z}}_{2,k}\end{pmatrix}, \tag{18}$$

where $\dot{\boldsymbol{z}}_{p,k} = (\tilde{z}_{p,k-(2L+1)} \ \dots \ \tilde{z}_{p,k+(2L)})^{\mathsf{T}}$ and $\check{\boldsymbol{x}}_{p,k}\in\mathbb{C}^{2L+2\times 1}$. With the compound parameter matrix, we can calculate an estimated channel impulse response in the time-domain

$$\hat{\boldsymbol{h}}_{\varrho,p,k} = \left(\boldsymbol{C}_{\mathrm{T},k}\boldsymbol{C}_{\mathrm{T},k}^{\mathrm{H}}\right)^{-1}\boldsymbol{C}_{\mathrm{T},k}\boldsymbol{e}_{\varrho,p} \tag{19}$$

for output polarization p and time delay ϱ. The vector $\boldsymbol{e}_{\varrho,p}$ is a unit vector with a single one entry. This selects the channel impulse responses for polarization p at a time delay ϱ. In theory, the delay ϱ can be chosen freely, however in practice it is beneficial to select $\varrho = L+1$ to center the channel impulse response around the center tap. The

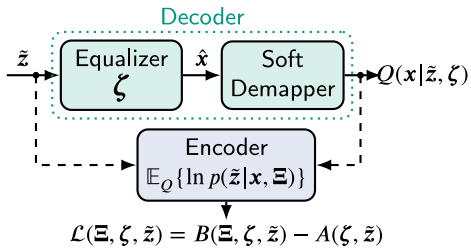

Fig. 3. Block diagram of the VAE-based equalizer. Note that, in contrast to end-to-end AEs, both encoder and decoder (colored blocks) are parts of the receiver only.

vector $\hat{\boldsymbol{h}}_{\varrho,p} = (\hat{\boldsymbol{h}}_{\varrho,1,p}\hat{\boldsymbol{h}}_{\varrho,2,p})^{\intercal}$ contains the stacked version of the channel impulse response $\hat{\boldsymbol{h}}_{\varrho,1,p}$ observed between polarization 1 and p and the channel impulse response $\hat{\boldsymbol{h}}_{\varrho,2,p}$ observed between polarization 2 and p. This is the inverse to the MIMO ZF solution described in more detail in [81, Eq.(27)]. Equipped with the set of channel impulse responses $\{\hat{\boldsymbol{h}}_{\varrho,1,1}, \hat{\boldsymbol{h}}_{\varrho,1,2}, \hat{\boldsymbol{h}}_{\varrho,2,1}, \hat{\boldsymbol{h}}_{\varrho,2,2}\}$ we can assemble the 2×2 Jones matrix and apply the discrete Fourier transform to obtain an estimate of the channel transfer matrix $\hat{\boldsymbol{H}}_{\varrho}(\omega,\boldsymbol{\theta})$. With the aforementioned selection of $\varrho = L+1$ we get a channel estimate $\hat{\boldsymbol{H}}(\omega,\boldsymbol{\theta}) = \hat{\boldsymbol{H}}_{L+1}(\omega,\boldsymbol{\theta})$.

3.2. Variational autoencoders for integrated sensing and communications

Ideally, both DSP algorithms of an ISAC system, i.e., the channel estimation and equalization, should be derived from a common criterion and optimized jointly to ensure the best overall performance. This is the case for the novel equalization algorithm based on a VAE. It combines a classical adaptive equalizer with an adaptive channel estimator and jointly optimizes both blindly, i.e., without pilot symbols as required for the algorithms in Section 3.1 which reduce the spectral efficiency. Originally introduced as maximum likelihood approximating channel equalizer in [45], we extended the concept to higher-order modulation formats and probabilistic constellation shaping and applied it to a linear dispersive optical channel model in [46]. In [48] we showed that the VAE equalizer can be used to reliably start-up blind equalizers in coherent optical communications at working points where other state-of-the-art blind equalization algorithms fail. The robust start-up has further been leveraged for a blind joint channel estimation and symbol detection on time-variant linear inter-symbol-interference channels [82].

Using the observation of the received sequence $\tilde{z}$ after the transmission of some unknown data sequence $\boldsymbol{x}$ over an unknown channel, we want to estimate the parameters $\Xi = (\boldsymbol{h}_{11}, \boldsymbol{h}_{12}, \boldsymbol{h}_{21}, \boldsymbol{h}_{22}, \hat{\sigma}_{\mathrm{w}})$ of the transmission channel, i.e., the noise variance $\hat{\sigma}_{\mathrm{w}}$ and the channel impulse responses composing a 2×2 butterfly estimator. Specifically, we want to find the best channel estimate for the given (observed) received sequence $\tilde{z}$, i.e., the maximization of the (log-) likelihood function

$$\hat{\Xi} = \arg\max_{\Xi} \ln\left(p(\tilde{z}|\Xi)\right) \tag{20}$$

which, however, is usually intractable since its computation would require the summation over all possible transmit sequences.

Instead, the log-likelihood can be re-written using Bayes' theorem as

$$\ln p(\tilde{z}|\Xi) = \ln \frac{p(\tilde{z}|\boldsymbol{x},\Xi)\cdot p(\boldsymbol{x})}{p(\boldsymbol{x}|\tilde{z})}\,. \tag{21}$$

We note that the transmit sequence $\boldsymbol{x}$ is independent from the parameters Ξ, i.e., $p(\boldsymbol{x}|\tilde{z},\Xi) = p(\boldsymbol{x}|\tilde{z})$ and $p(\boldsymbol{x}|\Xi) = p(\boldsymbol{x})$. While we can consider the prior distribution $p(\boldsymbol{x})$ as known (e.g., uniform or Maxwell–Boltzmann distribution), the true posterior distribution $p(\boldsymbol{x}|\tilde{z})$ is unknown. VAEs are based on *variational inference*, which introduces a parametric function to find the best variational approximation $Q(\boldsymbol{x}|\tilde{z},\zeta) \in \mathcal{Q}$ from a family of possible distributions $\mathcal{Q}$ by varying the parameters ζ [83,84]. Then, the log-likelihood can be re-written as

$$\begin{aligned}\ln p(\tilde{z}|\Xi) &= \mathrm{D}_{\mathrm{KL}}\left(Q(\boldsymbol{x}|\tilde{z},\zeta)\,\middle\|\,p(\boldsymbol{x}|\tilde{z})\right) + \mathrm{ELBO}(Q)\\ &\geq \mathrm{ELBO}(Q)\end{aligned} \tag{22}$$

with the Kullback–Leibler divergence $\mathrm{D}_{\mathrm{KL}}(u \parallel v)$ as a measure of similarity between two densities (u and v), which is non-negative and zero for $u = v$. Since $p(\tilde{z}|\Xi)$ is also called *evidence*, the term on the right hand side of (22) is called *evidence lower bound (ELBO)*, with

$$\begin{aligned}\mathrm{ELBO}(Q) = &\underbrace{\mathbb{E}_Q\{\ln p(\tilde{z}|\boldsymbol{x},\Xi)\}}_{=:\mathrm{B}}\\ &\underbrace{-\mathrm{D}_{\mathrm{KL}}\left(Q(\boldsymbol{x}|\tilde{z},\boldsymbol{\phi})\,\middle\|\,P(\boldsymbol{x})\right)}_{=:\mathrm{A}},\end{aligned} \tag{23}$$

where $\mathbb{E}_Q\{\cdot\}$ is the expectation with respect to the approximated posterior density $Q(\boldsymbol{x}|\tilde{z},\zeta)$. The ELBO can then be used as a proxy to maximize the evidence by jointly optimizing the parameters of the channel estimation Ξ and ζ of the variational approximation $Q(\boldsymbol{x}|\tilde{z},\zeta)$, e.g., by using an *Adam* optimizer [85]. The latter updates the parameters based on gradient-descent with momentum for a batch of N_{batch} samples with a learning rate μ. The ELBO corresponds to the same loss as derived for the blind equalizer in [46] and can be implemented as sketched in Fig. 3 as a VAE at the receiver. In fact, the variational approximation can be interpreted as a joint equalization and (soft) demapping, and could be any parametric function including neural networks. In order to have a fair comparison with the other algorithms, we use a classical 2×2 butterfly finite impulse response (FIR) structure and a Gaussian soft demapper with variance $\sigma^2_{\mathrm{demap}}$ (which can be either a hyperparameter or updated by the optimizer as well), so the parameters ζ correspond to the equalizer taps $\tilde{c}_k$ as introduced in Section 3.1. The main differences of the VAE approach compared to the other algorithms is that, firstly, the channel is estimated directly and not via filter inversion, and, secondly, no pilot symbols are required for the parameter update. While prior work mainly focuses on either the communication or the estimation performance of the VAE approach, we highlight the ISAC characteristic in this paper. Future work could further analyze the hyperparameter tuning with respect to ISAC and may optimize the filter structures to extract more interpretable information.

In [47], we further generalized the concept for channel estimators based on memory polynomials and linked it to a regularized decision-directed LMS, which we used to update frequency-domain equalizers [49] and low-complexity non-linear equalizers [50].

Table 1
Simulation parameters for the numerical results in Section 3.3.

Constellation	$\mathcal{M}$	64-QAM
Probability mass function	$p(\mathcal{M})$	uniform
AWGN	$\frac{1}{\sigma_{\mathrm{w}}^2}$	25 dB
Fiber length	L_{fiber}	1000 km
PMD elements	N	10 000
PMD coefficient	D	$0.3\,\mathrm{ps}/\sqrt{\mathrm{km}}$
PDL elements		5
PDL attenuation	γ	$\mathcal{U}(0.1,\ 0.3)$ dB
Channel realizations		1000
Transmit symbols		20 000
Equalizer length	$2L+1$	25
LMS and LMS-ZF step size	μ	5×10^{-4}
ZF step size	μ	1×10^{-3}
VAE step size	μ	3×10^{-3}
Symbol rate		100 GBd
Samples per symbol	T	2
Simulation bandwidth		800 GHz

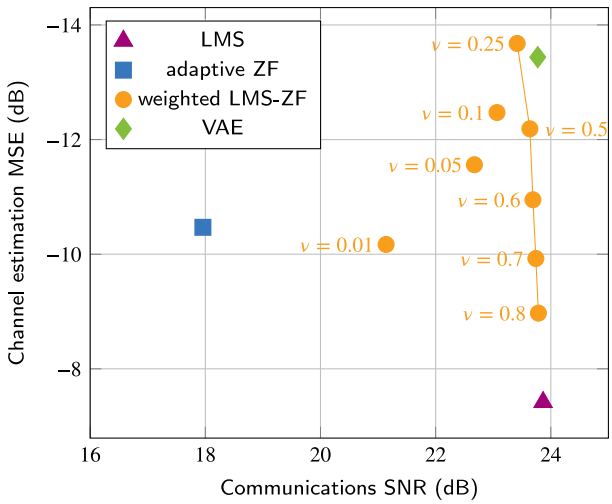

Fig. 4. Adaptive equalizer sensing–communication trade-off. For the weighted LMS-ZF algorithm the chosen parameter ν is noted at the corresponding mark.

3.3. Numerical results

We evaluate the communication and sensing performance of the ZF, LMS, LMS-ZF, and VAE equalizers according to the introduced performance metrics in Section 2.2.[4] We give the set of default simulation parameters for the numerical results in this section in Table 1. While the number of PMD segments is $N = 10\,000$, we chose to limit the number of active PDL elements to 5, where these active elements are equally spaced along the full fiber length.

For each channel realization, we initialize ϑ_n and $\boldsymbol{a}_n \in \mathbb{C}^{3\times1\times1}$ according to [86, Summary]. To model the DGD for an optical fiber of length L_{fiber}, we split the fiber span into $N = \lfloor L_{\text{fiber}}/L_{\text{c}} \rfloor$ segments. The DGD per segment is then distributed according to

$$\tau_n \sim \mathcal{N}\left(\frac{\tau}{\sqrt{N}}, \left(\frac{\tau}{5\sqrt{N}}\right)^2\right), \tag{24}$$

where

$$\tau = \sqrt{\frac{8L_{\text{fiber}}}{3\pi}}\,D \tag{25}$$

is the mean DGD over the fiber span with a diffusion constant D in $\text{ps}/\sqrt{km}$ [87].

We study the performance of the different equalizers in the steady-state operation. This means that we simulate a time-invariant channel with PMD, PDL, and AWGN and provide a sufficient number of known pilot symbols such that the equalizers can converge.

The LMS, ZF, and LMS-ZF equalizers use the first 19 000 transmit symbols as pilots to converge their parameter vectors $\tilde{\boldsymbol{c}}_k$. Due to known start-up convergence problems,[5] of the adaptive ZF equalizer, the LMS equalizer is used to initialize the parameter vector with the first 2000 symbols to provide a stable working point.

In Fig. 4, we show the average communication signal-to-noise ratio (SNR) and the average channel estimation MSE for 1000 channel realizations. We observe that the adaptive ZF algorithm exhibits the worst communication SNR of all tested algorithms. It is especially interesting to see that the interpolation between the LMS and ZF regression vectors does not only interpolate the performance in a region between the performance of the ZF and the LMS algorithm. But already a small weight $\nu = 0.25$ shows significant improvement in the communication performance from 18 dB to 23 dB while simultaneously reducing the MSE in the channel estimation compared to the ZF algorithm. This is an unexpected result, as one would think that interpolating the weights would also interpolate the performance. A possible explanation for this behavior is that interpolation of the ZF regression vector with the LMS regression vector has a stabilizing effect on the convergence of the adaptive ZF algorithm. In [88], the SISO version of the adaptive ZF algorithm has been studied. In particular, it has been highlighted that for certain worst-case sequences, the adaptive ZF algorithm diverges. In a communication setting, where the transmitted sequence cannot be controlled, certain worst-case sequences may drive the adaptive ZF out of convergence, albeit slowly.

Further increase of the weight parameter ν shows an interpolation between $\nu = 0.25$ and $\nu = 1.0$, where the latter corresponds to the LMS algorithm, since the ZF regression vector values are not used at all. For the selected channel parameters, the useful range of parameters for the LMS-ZF equalizer is between $\nu = 0.25$ and $\nu = 1.0$, where it is not beneficial to select ν smaller than 0.25. We expect this region to change depending on the channel parameterization, especially the experienced noise at the receiver.

The VAE equalizer was simulated with 200 000 transmit symbols. This is because our implementation of the VAE algorithm performs batch-wise updates with a batch-size of 200 and is a blind algorithm. Both factors contribute to a slower convergence rate compared to the symbol-wise updates performed by the LMS and ZF equalizers. It has been shown that the convergence rate of the VAE can be improved by tuning the hyperparameters or using a flexible step size [46]. Additionally, due to the lack of pilot symbols, the VAE equalizer may converge to a solution that has a flipped polarization and a constant phase offset of a multiple of $\pi/2$ compared to the transmit symbols $\boldsymbol{X}$. We correct this polarization flip and phase offset for both the signal $\hat{\boldsymbol{X}}$ and the channel estimate $\hat{\boldsymbol{H}}(\omega, \boldsymbol{\theta})$. Furthermore, we can observe that the VAE provides an improved channel estimation while having an average communication performance similar to LMS-ZF with a weight $\nu = 0.5$.

4. Physics-based machine learning for distributed polarization sensing

In, e.g., [25,26], the extraction and processing of sensing information is limited to tracking the accumulated SOP at the channel center frequency over the entire propagation distance. However, this approach does not provide distributed (i.e., spatially resolved) location information about external environmental events. Coarse localization information can be obtained by assuming per-span readouts of polarization data [89] or access to multiple optical paths [38,90]. To improve localization accuracy, time-resolved backscattering approaches for distributed polarization sensing similar to traditional DOFS techniques have been recently investigated in [91–93]. However, since backscatter is blocked by optical amplifiers, applying such methods in amplified multi-span systems requires a bidirectional link topology in combination with (typically high-loss) loop-back reflectors in order for the backscatter signal to reach the transmitter, thereby limiting applicability.

Here, we review the approach recently proposed in [51], which employs physics-based machine learning [57,64] to estimate spatially resolved parameters of the optical fiber channel by factorizing the full channel state information in the form of the noisy received Jones matrix at the receiver. Properly estimating PMD and PDL enhances spatial resolution due to the non-commutative nature of their corresponding Jones matrices. This enables a more accurate characterization of spatial dependencies, as demonstrated in [51]. Tracking PMD however, requires a sufficient bandwidth in the adaptive receiver filter. The PDL is likely to be more effective for estimating the spatial dependencies,

[4] Open-source implementations are provided at https://github.com/kit-cel/pol_sensing_communication.

[5] In [80] it is pointed out that a sufficiently large eye-opening must be already available in the received signal for robust convergence of the adaptive ZF algorithm.

as it typically arises from localized components, such as reconfigurable optical add-drop multiplexers or amplifiers, whereas the PMD effects are often distributed over the transmission fiber. Note that high PDL does not straightforwardly enhance distributed sensing performance. While it may introduce greater degrees of freedom, high PDL can also lead to situations where the received signal power becomes comparable to the noise level, ultimately degrading sensing accuracy. Compared to [89–93], this approach does not require a special link topology and only assumes access to the estimated (linear) Jones matrix at the receiver. It is therefore compatible with the standard DSP approaches based on adaptive equalization or VAEs described in the previous section. We note that this approach is also closely related to work on distributed PMD compensation [87,94–100] and longitudinal path and PDL monitoring [101–104].

4.1. Modified channel model

The problem of recovering spatially resolved polarization information from estimated Jones matrices has been previously studied via the so-called ISA [52–54]. To allow for direct comparisons to the ISA, we consider a slight modification of the channel model in (3), where the rotation matrices are parametrized in the same way as in [54]. Moreover, we additionally assume that the underlying rotation parameters are time-varying to account for external environmental disturbances. In particular, we redefine (4) with a rotation matrix defined at segment n as

$$\boldsymbol{R}_{n,k} = \begin{bmatrix} \cos\phi_{n,k} & \mathrm{j}\,\mathrm{e}^{\mathrm{j}\psi_{n,k}}\sin\phi_{n,k} \\ \mathrm{j}\,\mathrm{e}^{-\mathrm{j}\psi_{n,k}}\sin\phi_{n,k} & \cos\phi_{n,k} \end{bmatrix}, \tag{26}$$

where $\phi_{n,k}$ and $\psi_{n,k}$ are time-varying rotation angles and k refers to the time index. Substituting $\boldsymbol{R}_{n,k}$ from (26) in (3) gives a time-varying channel transfer matrix. Note that unlike the rotation matrices described in (4), which produce a uniform distribution over the Poincaré sphere, the model defined in (26) represents a special case that does not necessarily result in a uniform distribution over the Poincaré sphere.

To model the time-dependence, we assume that the rotation angles remain static when the fiber is undisturbed and become dynamic under disturbances. Specifically, we assume that environmental disturbances cause temporal fluctuations in the rotation angles $\phi_{n,k}$ and $\psi_{n,k}$, which are modeled as Wiener processes, characterized by

$$\phi_{n,k} = \phi_{n,k-1} + \Phi_{n,k}, \tag{27}$$

$$\psi_{n,k} = \psi_{n,k-1} + \Psi_{n,k}. \tag{28}$$

Here, $\Phi_{n,k}$ and $\Psi_{n,k}$ are normally distributed with a mean of zero and standard deviations $\sigma_{\Phi_{n,k}}$ and $\sigma_{\Psi_{n,k}}$, respectively, representing the intensity of environmental fluctuations. Larger values of $\sigma_{\Phi_{n,k}}$ and $\sigma_{\Psi_{n,k}}$ indicate more significant environmental changes. The model assumes a timescale for the dynamic polarization perturbation that is significantly slower than the DGD parameters τ_n.

Let $\boldsymbol{\theta}_k = \{\gamma_n, \phi_{n,k}, \psi_{n,k}, \tau_n\}_{n=1}^{N}$ represent all the underlying channel parameters, where γ_n denotes the PDL ratio parameter, and τ_n represents the per-section DGD parameters. The channel matrix (3) at time k and frequency ω can be fully described using $\boldsymbol{\theta}_k$. Therefore, extracting $\boldsymbol{\theta}_k$ from the estimated channel matrix at the receiver enables the receiver to pinpoint the location of the environmental disturbance by capturing the time-dependency of the parameters in $\boldsymbol{\theta}_k$. In the following, we review two algorithms that try to reconstruct $\boldsymbol{\theta}_k$ from the noisy estimate of the channel $\hat{\boldsymbol{H}}(\omega, \boldsymbol{\theta}_k)$ obtained, e.g., from the channel estimation techniques introduced in Section 3.

4.2. Inverse scattering algorithm

The use of an ISA to extract the DGD profile from the overall impulse response (i.e., the time-domain Jones matrix) was first explored in [52,53]. This technique was further advanced in [54] to handle PDL, enabling the recovery of both DGD and PDL profiles. The ISA method is theoretically accurate for the scenarios discussed in these references, assuming that the overall impulse response is completely known and free from noise.

The ISA algorithm analytically calculates the values of the unknown underlying channel parameters $\boldsymbol{\theta}_k$ containing the SOP rotation, PDL, and DGD parameters of each section at time k and is performed iteratively in the time domain by first taking the inverse Fourier transform of the noisy estimate of the Jones matrix $\hat{\boldsymbol{H}}(\omega, \boldsymbol{\theta}_k)$. Then it calculates the unknown parameter values for the last section N, i.e., a_N, $\phi_{N,k}$, and $\psi_{N,k}$, and subsequently employs these values to construct the equivalent total channel response for sections 1 to $N-1$, effectively removing the last section from the model. By repeating these steps, all the subsequent parameters will be obtained. We emphasize that in the ISA algorithm proposed by [54], the DGD parameters τ_k are assumed to be identical and known for all sections, eliminating the need for their estimation.

4.3. Physics-based machine learning (PML) approach

As previously mentioned, the ISA algorithm [54] is accurate when the channel matrix is known and noise-free; however, as we show in Section 4.4, its performance deteriorates in the presence of noise. To address the noise sensitivity of ISA, a PML method has been recently proposed in [51] based on the approach in [57,64]. This method directly parameterizes the propagation model and simultaneously optimizes all its underlying parameters, $\boldsymbol{\theta}_k$, improving robustness to noise.

Let the set of trainable parameters be denoted as $\hat{\boldsymbol{\theta}}_k = \{\hat{\gamma}_n, \hat{\phi}_{n,k}, \hat{\psi}_{n,k}\}_{n=1}^{N}$. Then, we define the frequency-averaged cost function as

$$\mathcal{L}(\hat{\boldsymbol{\theta}}_k) = \sum_{i=1}^{L_\mathrm{f}} \left\| \hat{\boldsymbol{H}}(\omega_i; \boldsymbol{\theta}_k) - \boldsymbol{H}(\omega_i; \hat{\boldsymbol{\theta}}_k) \right\|_\mathrm{F}^2, \tag{29}$$

where L_f represents the number of frequency samples. This loss function is then minimized iteratively using a gradient-descent optimizer with a learning rate of μ and M number of iterations. Note that we only have access to the overall noisy channel response $\hat{\boldsymbol{H}}(\omega_i; \boldsymbol{\theta}_k)$ without knowing its underlying parameters $\boldsymbol{\theta}_k$.

4.4. Numerical results

For the numerical results[6], we consider $N = 5$ fiber segments. Within each segment, the PDL parameter γ_n is sampled from a uniform distribution in the range $[0.07, 0.17]$ (equivalent to $[0.3, 0.7]$ dB PDL). Additionally, the initial rotation parameters $\phi_{n,0}$ and $\psi_{n,0}$ are uniformly sampled from $[-\pi, \pi)$. The DGD parameter $\tau_n = \frac{L_\mathrm{c}}{\Delta z}$ and is assumed to be known and constant across all segments. This assumption is necessary for the ISA method, while the PML approach could potentially be extended to learn unknown DGD elements, as demonstrated in [105].

For the PML approach, the Adam optimizer is used with $M = 300$ iterations and a learning rate of $\mu = 0.05$. All trainable parameters $\hat{\boldsymbol{\theta}}_k$ are initialized to zero at $k = 0$ and then tracked for $k > 0$. The frequency samples in the loss function correspond to a $1/\tau_n$ sampling distance in the frequency domain, with $L_\mathrm{f} = N$. It was verified that increasing the number of frequency samples does not affect the results.

Finally, to model an environmental disturbance, it is assumed that an incidence occurs within segment $n = 2$ at $k \in [15, 35]$ that only affects parameters $\phi_{2,k}$ and $\psi_{2,k}$ with $\sigma_{\Phi_{2,k}} = \sigma_{\Psi_{2,k}} = 0.1$ for $k \in [15, 35]$ and $\sigma_{\Phi_{2,k}} = \sigma_{\Psi_{2,k}} = 0$ elsewhere.

[6] Open-source implementations are provided at https://github.com/Mohammadfarsi1994/physics-based-distributed-polarization-sensing.

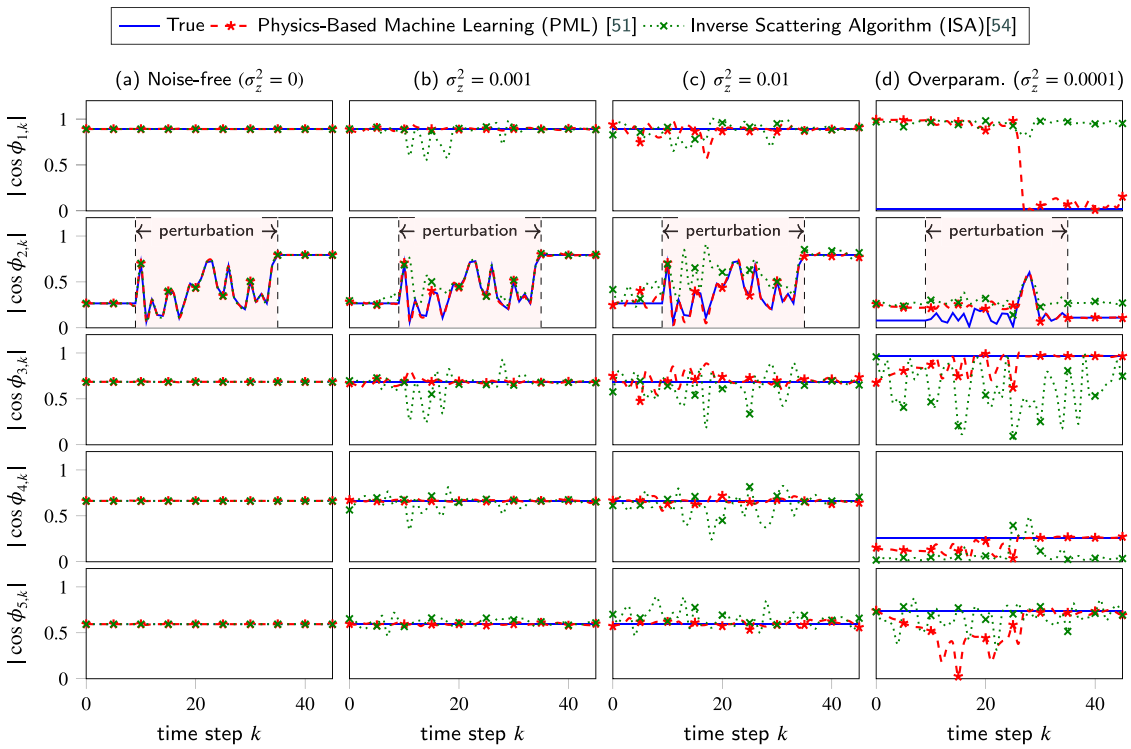

Fig. 5. Numerical results: (a)–(c) trackable channel realization with varying amounts of noise; (d) non-trackable realization [51, Fig. 1].

4.4.1. Trackable cases

Fig. 5 compares the tracking capability of the ISA and the learning approach by plotting $|\cos\phi_{n,k}|$ for all sections in four different scenarios. As depicted in Fig. 5(a), when the channel estimation is perfect (noise-free), both approaches are able to track the perturbation, pinpointing its location and timing. However, when the channel response is subject to estimation noise (Fig. 5(b)–(c)), the PML approach exhibits improved noise resilience by successfully tracking and identifying the perturbation's location and timing, whereas the ISA looses its ability to track or precisely determine the affected section.

4.4.2. Non-trackable cases

For certain realizations of the channel parameters $\boldsymbol{\theta}_k$, the channel response $\boldsymbol{H}(\omega_i;\boldsymbol{\theta}_k)$ could be described with a smaller number of independent channel parameters. In this case, multiple different $\boldsymbol{\theta}_k$ can produce nearly identical channel responses $\boldsymbol{H}(\omega_i;\boldsymbol{\theta}_k)$. An example of this phenomenon is illustrated in Fig. 5(d). In such cases, the two algorithms exhibit different behaviors. For the ISA, noise-induced numerical instabilities lead to mismatches between the overall response and the estimated parameters, causing them to deviate from the actual underlying channel parameters $\boldsymbol{\theta}_k$. On the other hand, PML achieves a channel response $\boldsymbol{H}(\omega_i;\hat{\boldsymbol{\theta}}_k)$ that remains closely aligned with the true response $\boldsymbol{H}(\omega_i;\boldsymbol{\theta}_k)$, due to the specific loss function used. However, the model is effectively overparameterized because various parameter configurations can yield nearly identical responses, which prevents finding a stable and unique solution. A potential avenue for future work involves modeling the actual channel response using a reduced number of sections, which may offer a solution to the overparameterization problem by trading off localization resolution.

5. Transmitter trade-offs and optimization opportunities

So far, we have considered the design of polarization-based ISAC systems by discussing various receiver DSP algorithms, assuming the transmission of standard communication signals. In this section, we briefly discuss opportunities to optimize also the transmitter and underlying waveforms through the lens of dual-functional system design. Previous research on dual-functional transmitters for ISAC over optical fibers has focused on backscatter-based DOFS sensing with the intention of improving spectral efficiency and/or reducing hardware cost compared to utilizing separate sensing and communication systems. For example, the authors in [106] showed that probing signals used for distributed acoustic sensing (DAS) can be multiplexed together with data-carrying digital subcarriers in the frequency domain. Other proposed integration schemes merge DAS sensing probes and data signals by replacing the optical carrier with linear frequency modulation (LMF) [107] or by inserting LMF sensing probes into the training signals of communication data frames [108], thereby sharing both the same spectrum and transmitter hardware for sensing and communication.

In general, the transmitted signal in ISAC systems serves a dual role as it may be chosen to either maximize communication performance or aid with channel estimation, thereby enhancing sensing functionality. In [109], it was shown that this dual role gives rise to fundamental trade-offs that, for example, involve balancing the "degree of randomness" of the transmitted signal, where sensing tends to prefer deterministic probing signals while communication prefers signals that are maximally random. Applying such insights to the considered system setup, transmitter trade-offs between communication and polarization sensing could for example be achieved by increasing the frequency of pilot transmissions or suitably adjusting the modulation format, thereby improving the tracking capabilities of adaptive equalizers. While this may lead to a temporary decrease in communication performance (e.g., sacrificing data throughput or accepting higher error rates), it could provide a more granular and accurate picture of an unfolding environmental event. After the event ends, normal communication operations can be restored.

A possible avenue for future work to tackle the problem of designing dual-functional ISAC waveforms for polarization sensing could rely on dual-functional autoencoders (AEs). Indeed, treating a communication system as a denoising AE can facilitate transmitter optimizations, and,

more generally, joint transmitter and receiver design [110]. AE approaches have already been widely studied for optical systems assuming many different system setups [111–120]. However, these papers consider purely communication-centric design, where the optimization is based on metrics like error rates, or suitable surrogate loss functions such as cross-entropy. Dual-functional autoencoders AEs have been recently studied in the context of wireless ISAC [121–124], where weighting loss functions for sensing and communication can be used to optimize dual-functional beamformers and corresponding receiver algorithms. This approach allows for flexible optimization of both sensing and communication capabilities within the same framework. Similar approaches could be applicable in optical domains as well, for example by devising sensing-centric optimization metrics similar to the one used in adaptive ZF.[7] However, we leave this as an open avenue for future research.

6. Conclusion and future work

In this paper, we studied the communication performance and channel sensing utility of different adaptive equalizer algorithms in the context of ISAC for optical fiber communication. In a numerical simulation we found that the adaptive ZF algorithm shows the previously reported convergence problems at startup. When interpolating the regression vector between the adaptive ZF and LMS algorithm we could find significant improvement in channel estimation, exceeding the performance of the adaptive ZF algorithm. Similarly, the recently introduced pilot-less VAE equalizer with its dedicated channel parameter estimation maintains a significantly improved channel estimation compared to the LMS equalizer, while maintaining a very good communication performance. Further investigation of the presented LMS-ZF and VAE equalizer on a time-varying optical channel is required to test the tracking performance for channels with polarization drift. In particular, the tracking accuracy may be of significance to enable novel sensing applications.

A critical gap in the literature is the lack of realistic time-varying channel models that accurately describe how various environmental perturbations lead to variations in the model parameters over time. Here and in [51], it was assumed that environmental disturbances affect only the SOP rotation elements $(\phi_{k,t}, \psi_{k,t})$ and we modeled their time dependence using random walks. This simplifying assumption was made due to the lack of more realistic channel models. In [32], it was shown how dynamic polarization fluctuations induced by hydrostatic pressure resulting from, e.g., earthquakes, relate to relative birefringence fluctuations in long-haul submarine cables . However, it is not immediately clear how to integrate the insights from [32] into the model (3). Moreover, other types of events or deployment scenarios lack corresponding models that describe the environmental impact on the underlying fiber channel parameters. Developing such models therefore presents a significant opportunity to facilitate further study of polarization-based ISAC systems, develop new data-driven algorithms, and accurately assess their performance trade-offs.

CRediT authorship contribution statement

Andrej Rode: Investigation, Methodology, Writing – original draft, Writing – review & editing. **Mohammad Farsi:** Investigation, Methodology, Writing – original draft, Writing – review & editing. **Vincent Lauinger:** Investigation, Methodology, Writing – original draft, Writing – review & editing. **Magnus Karlsson:** Funding acquisition, Methodology, Supervision, Writing – review & editing. **Erik Agrell:** Funding acquisition, Methodology, Supervision, Writing – review & editing. **Laurent Schmalen:** Funding acquisition, Methodology, Supervision, Writing – review & editing. **Christian Häger:** Funding acquisition, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Andrej Rode reports financial support and travel were provided by Karlsruhe Institut of Technology Karlsruhe House of Young Scientists. Andrej Rode reports financial support was provided by European Research Council. Laurent Schmalen reports financial support was provided by European Research Council. Christian Haeger reports financial support was provided by Swedish Research Council. Vincent Lauinger reports financial support was provided by German Federal Ministry of Education and Research. Mohammad Farsi reports financial support was provided by Knut and Alice Wallenberg Foundation. Erik Agrell reports financial support was provided by Knut and Alice Wallenberg Foundation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The work of A. Rode and L. Schmalen was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 101 001899).

The research visit of A. Rode at the Chalmers University of Technology was supported by the Karlsruhe House of Young Scientists (KHYS).

The work of C. Häger was supported by the Swedish Research Council under grant no. 2020-04718.

The work of M. Farsi and E. Agrell was supported by the Knut and Alice Wallenberg Foundation, Sweden under grant no. 2018.0090.

Parts of this work was carried out in the framework of the CELTIC-NEXT project AI-NET-ANTILLAS (C2019/3-3) and was funded by the German Federal Ministry of Education and Research (BMBF) under grant agreement 16KIS1316.

Data availability

Data will be made available on request.

References

[1] N.J. Lindsey, E.R. Martin, Fiber-optic seismology, Annu. Rev. Earth Planet. Sci. 49 (1) (2021) 309–336.
[2] E. Ip, F. Ravet, H. Martins, M.-F. Huang, T. Okamoto, S. Han, C. Narisetty, J. Fang, Y.-K. Huang, M. Salemi, E. Rochat, F. Briffod, A. Goy, M. Del Rosario Fernandez-Ruiz, M.G. Herraez, Using global existing fiber networks for environmental sensing, Proc. IEEE 110 (11) (2022) 1853–1888.
[3] G.A. Wellbrock, T.J. Xia, M.-F. Huang, S. Han, Y. Chen, T. Wang, Y. Aono, Explore benefits of distributed fiber optic sensing for optical network service providers, J. Lightw. Technol. 41 (12) (2023) 3758–3766.
[4] M.K. Barnoski, S.M. Jensen, Fiber waveguides: a novel technique for investigating attenuation characteristics, Appl. Opt. 15 (9) (1976) 2112–2115.
[5] T. Horiguchi, K. Shimizu, T. Kurashima, M. Tateda, Y. Koyamada, Development of a distributed sensing technique using Brillouin scattering, J. Lightw. Technol. 13 (7) (1995) 1296–1302.
[6] M.A. Farahani, T. Gogolla, Spontaneous Raman scattering in optical fibers with modulated probe light for distributed temperature Raman remote sensing, J. Lightw. Technol. 17 (8) (1999) 1379–1391.
[7] A. Rogers, Distributed optical-fibre sensing, Meas. Sci. Technol. 10 (8) (1999) R75–R99.
[8] S. Russell, K. Brady, J. Dakin, Real-time location of multiple time-varying strain disturbances, acting over a 40-km fiber section, using a novel dual-Sagnac interferometer, J. Lightw. Technol. 19 (2) (2001) 205–213.

[7] Note that the ZF optimization criterion is not given in the form of a traditional loss function, but in terms of a system of equations which is then solved using gradient descent [43,80]

[9] Q. Liu, Z. He, T. Tokunaga, Sensing the earth crustal deformation with nano-strain resolution fiber-optic sensors, Opt. Express 23 (11) (2015) A428–A436.

[10] Y. Lu, T. Zhu, L. Chen, X. Bao, Distributed vibration sensor based on coherent detection of phase-OTDR, J. Lightw. Technol. 28 (22) (2010) 3243–3249.

[11] L. Thévenaz, Brillouin distributed time-domain sensing in optical fibers: state of the art and perspectives, Front. Optoelectron. China 3 (1) (2010) 13–21.

[12] X. Bao, L. Chen, Recent progress in distributed fiber optic sensors, Sensors 12 (7) (2012) 8601–8639.

[13] S. Liang, X. Zhao, R. Liu, X. Zhang, L. Wang, X. Zhang, Y. Lv, F. Meng, Fiber-optic auditory nerve of ground in the suburb: For traffic flow monitoring, IEEE Access 7 (2019) 166704–166710.

[14] M.F. Huang, P. Ji, T. Wang, Y. Aono, M. Salemi, Y. Chen, J. Zhao, T.J. Xia, G.A. Wellbrock, Y.K. Huang, G. Milione, E. Ip, First field trial of distributed fiber optical sensing and high-speed communication over an operational telecom network, J. Lightw. Technol. 38 (1) (2020) 75–81.

[15] N.J. Lindsey, E.R. Martin, D.S. Dreger, B. Freifeld, S. Cole, S.R. James, B.L. Biondi, J.B. Ajo-Franklin, Fiber-optic network observations of earthquake wavefields, Geophy. Res. Lett. 44 (23) (2017) 11792–11799.

[16] P. Jousset, T. Reinsch, T. Ryberg, H. Blanck, A. Clarke, R. Aghayev, G.P. Hersir, J. Henninges, M. Weber, C.M. Krawczyk, Dynamic strain determination using fibre-optic cables allows imaging of seismological and structural features, Nat. Commun. 9 (1) (2018).

[17] H.F. Martins, M.R. Fernández-Ruiz, L. Costa, E. Williams, Z. Zhan, S. Martin-Lopez, M. Gonzalez-Herraez, Monitoring of remote seismic events in metropolitan area fibers using distributed acoustic sensing (DAS) and spatiotemporal signal processing, in: Proc. OFC, San Diego, CA, 2019.

[18] M.R. Fernández-Ruiz, M.A. Soto, E.F. Williams, S. Martin-Lopez, Z. Zhan, M. Gonzalez-Herraez, H.F. Martins, Distributed acoustic sensing for seismic activity monitoring, APL Photonics 5 (3) (2020) 030901.

[19] M. Landrø, L. Bouffaut, H.J. Kriesell, J.R. Potter, R.A. Rørstadbotnen, K. Taweesintananon, S.E. Johansen, J.K. Brenne, A. Haukanes, O. Schjelderup, F. Storvik, Sensing whales, storms, ships and earthquakes using an Arctic fibre optic cable, Sci. Rep. 12 (1) (2022) 19226.

[20] G.A. Wellbrock, T.J. Xia, M.-F. Huang, J. Fang, Y. Chen, C. Narisetty, D. Peterson, J.M. Moore, A. Scarpaci, P. Westbrook, J. Li, R. Lingle, T. Wang, Y. Aono, Perimeter intrusion detection with backscattering enhanced fiber using telecom cables as sensing backhaul, in: Proc. OFC, San Diego, CA, 2022.

[21] S. Guerrier, K. Benyahya, C. Dorize, E. Awwad, H. Mardoyan, J. Renaudier, Vibration detection and localization in buried fiber cable after 80km of SSMF using digital coherent sensing system with co-propagating 600Gb/s WDM channels, in: Proc. OFC, San Diego, CA, 2022.

[22] T. Wang, M.-F. Huang, S. Han, C. Narisetty, Employing fiber sensing and on-premise AI solutions for cable safety protection over telecom infrastructure, in: Proc. OFC, San Diego, CA, 2022.

[23] E. Ip, J. Fang, Y. Li, Q. Wang, M.-F. Huang, M. Salemi, Y.-K. Huang, Distributed fiber sensor network using telecom cables as sensing media: technology advancements and applications, J. Opt. Commun. Netw. 14 (1) (2022).

[24] E. Ip, Y.-K. Huang, G. Wellbrock, T. Xia, M.-F. Huang, T. Wang, Y. Aono, Vibration detection and localization using modified digital coherent telecom transponders, J. Lightw. Technol. 40 (5) (2022) 1472–1482.

[25] F. Boitier, V. Lemaire, J. Pesic, L. Chavarría, P. Layec, S. Bigo, E. Dutisseuil, Proactive fiber damage detection in real-time coherent receiver, in: Proc. ECOC, 2017.

[26] Z. Zhan, M. Cantono, V. Kamalov, A. Mecozzi, R. Müller, S. Yin, J.C. Castellanos, Optical polarization–based seismic and water wave sensing on transoceanic cables, Science 371 (6532) (2021) 931–936.

[27] M. Karlsson, J. Brentel, P.A. Andrekson, Long-term measurement of PMD and polarization drift in installed fibers, J. Lightw. Technol. 18 (7) (2000) 941–951.

[28] D. Waddy, Ping Lu, Liang Chen, Xiaoyi Bao, Fast state of polarization changes in aerial fiber under different climatic conditions, IEEE Photonics Technol. Lett. 13 (9) (2001) 1035–1037.

[29] J. Wuttke, P. Krummrich, J. Rosch, Polarization oscillations in aerial fiber caused by wind and power-line current, IEEE Photonics Technol. Lett. 15 (6) (2003) 882–884.

[30] D. Charlton, S. Clarke, D. Doucet, M. O'Sullivan, D.L. Peterson, D. Wilson, G. Wellbrock, M. Bélanger, Field measurements of SOP transients in OPGW, with time and location correlation to lightning strikes, Opt. Express 25 (9) (2017) 9689.

[31] J. Pesic, E. Le Rouzic, N. Brochier, L. Dupont, Proactive restoration of optical links based on the classification of events, in: Proc. ONDM, 2011.

[32] A. Mecozzi, M. Cantono, J.C. Castellanos, V. Kamalov, R. Muller, Z. Zhan, Polarization sensing using submarine optical cables, Optica 8 (6) (2021) 788.

[33] R. Bratovich, A. D'Amico, F. Aquilino, R. Pastorelli, V. Curri, Surveillance of metropolitan anthropic activities by WDM 10G optical data channels, 2022.

[34] M. Mazur, D. Wallberg, L. Dallachiesa, E. Börjeson, R. Ryf, M. Bergroth, B. Josefsson, N.K. Fontaine, H. Chen, D.T. Neilson, J. Schröder, P. Larsson-Edefors, M. Karlsson, Field trial of FPGA-based real-time sensing transceiver over 524 km of live aerial fiber, in: Proc. OFC, San Diego, CA, 2023.

[35] M. Mazur, L. Dallachiesa, R. Ryf, D. Wallberg, E. Börjesson, M. Bergroth, B. Josefsson, N.K. Fontaine, H. Chen, D.T. Neilson, J. Schröder, P. Larsson-Edefors, M. Karlsson, Real-time monitoring of cable break in a live fiber network using a coherent transceiver prototype, in: Proc. OFC, San Diego, CA, 2024.

[36] M. Mazur, J.C. Castellanos, R. Ryf, E. Borjeson, T. Chodkiewicz, V. Kamalov, S. Yin, N.K. Fontaine, H. Chen, L. Dallachiesa, S. Corteselli, P. Copping, J. Gripp, A. Mortelette, B. Kowalski, R. Dellinger, D.T. Neilson, P. Larsson-Edefors, Transoceanic phase and polarization fiber sensing using real-time coherent transceiver, in: Proc. Optical Fiber Communication Conf., OFC, San Diego, CA, 2022.

[37] K.S.Y. Skarvang, S. Bjørnstad, R.A. Rørstadbotnen, K. Bozorgebrahimi, D.R. Hjelme, Observation of local small magnitude earthquakes using state of polarization monitoring in a 250km passive arctic submarine communication cable, in: Proc. Optical Fiber Communication Conf., OFC, San Diego, CA, 2023.

[38] H. Awad, F. Usmani, E. Virgillito, R. Bratovich, R. Proietti, S. Straullu, F. Aquilino, R. Pastorelli, V. Curri, Environmental surveillance through machine learning-empowered utilization of optical networks, Sensors 24 (10) (2024) 3041.

[39] S. Bjørnstad, K.S.Y. Skarvang, D.R. Hjelme, A. Tunheim, F. Fjermestad, E. Østerli, First impact movement characterization of shallow buried live subsea-cable, in: Proc. Optical Fiber Communication Conf., OFC, San Diego, CA, 2024.

[40] S.J. Savory, Digital filters for coherent optical receivers, Opt. Express 16 (2) (2008) 804–817.

[41] S.J. Savory, Digital coherent optical receivers: Algorithms and subsystems, IEEE J. Sel. Top. Quantum Electron. 16 (5) (2010) 1164–1179.

[42] P.S.R. Diniz, Adaptive Filtering, Springer US, Boston, MA, 2013, Publication Title: Adaptive Filtering.

[43] R.W. Lucky, Automatic equalization for digital communication, Bell Syst. Tech. J. 44 (4) (1965) 547–588.

[44] R.W. Lucky, Techniques for adaptive equalization of digital communication systems, Bell Syst. Tech. J. 45 (2) (1966) 255–286.

[45] A. Caciularu, D. Burshtein, Unsupervised linear and nonlinear channel equalization and decoding using variational autoencoders, IEEE Trans. Cogn. Commun. Netw. 6 (3) (2020) 1003–1018.

[46] V. Lauinger, F. Buchali, L. Schmalen, Blind equalization and channel estimation in coherent optical communications using variational autoencoders, IEEE J. Sel. Areas Commun. 40 (9) (2022) 2529–2539.

[47] J. Song, V. Lauinger, C. Häger, J. Schröder, A. Graell i Amat, L. Schmalen, H. Wymeersch, Blind frequency-domain equalization using vector-quantized variational autoencoders, in: Proc. European Conf. Optical Communication, ECOC, Glasgow, UK, 2023.

[48] V. Lauinger, F. Buchali, L. Schmalen, Improving the bootstrap of blind equalizers with variational autoencoders, in: Proc. Optical Fiber Communication Conf., OFC, San Diego, CA, 2023.

[49] J. Song, V. Lauinger, Y. Wu, C. Häger, J. Schröder, A. Graell i Amat, L. Schmalen, H. Wymeersch, Blind channel equalization using vector-quantized variational autoencoders, 2023, arXiv:2302.11687 [eess.SP].

[50] V. Lauinger, P. Matalla, J. Ney, N. Wehn, S. Randel, L. Schmalen, Fully-blind neural network based equalization for severe nonlinear distortions in 112 Gbit/s passive optical networks, in: Proc. Optical Fiber Communication Conf., OFC, San Diego, CA, 2024.

[51] M. Farsi, C. Häger, M. Karlsson, E. Agrell, Learning to extract distributed polarization sensing data from noisy jones matrices, in: Proc. Optical Fiber Communication Conf., OFC, San Diego, CA, 2024.

[52] S.E. Harris, E.O. Ammann, I.C. Chang, Optical network synthesis using birefringent crystals. I synthesis of lossless networks of equal-length crystals, J. Opt. Soc. Am. 54 (10) (1964) 1267.

[53] L. Moller, Filter synthesis for broad-band PMD compensation in WDM systems, IEEE Photon. Technol. Lett. 12 (9) (2000) 1258–1260.

[54] R. Noé, B. Koch, D. Sandel, V. Mirvoda, Polarization-dependent loss: New definition and measurement techniques, J. Lightw. Technol. 33 (10) (2015) 2127–2138.

[55] T. Shen, A. Lau, Fiber nonlinearity compensation using extreme learning machine for DSP-based coherent communication systems, in: Proc. OECC, 2011.

[56] D. Wang, M. Zhang, Z. Li, C. Song, M. Fu, J. Li, X. Chen, System impairment compensation in coherent optical communications by using a bio-inspired detector based on artificial neural network and genetic algorithm, Opt. Commun. 399 (2017) 1–12.

[57] C. Häger, H.D. Pfister, Nonlinear interference mitigation via deep neural networks, in: Proc. Optical Fiber Communication Conf., OFC, San Diego, CA, 2018.

[58] V. Kamalov, L. Jovanovski, V. Vusirikala, S. Zhang, F. Yaman, K. Nakamura, T. Inoue, E. Mateo, Y. Inada, Evolution from 8QAM live traffic to PS 64-QAM with neural-network based nonlinearity compensation on 11000 km open subsea cable, in: Proc. OFC, 2018.

[59] O. Sidelnikov, A. Redyuk, S. Sygletos, Equalization performance and complexity analysis of dynamic deep neural networks in long haul transmission systems, Opt. Express 26 (25) (2018) 32765–32776.

[60] S. Zhang, F. Yaman, K. Nakamura, T. Inoue, V. Kamalov, L. Jovanovski, V. Vusirikala, E. Mateo, Y. Inada, T. Wang, Field and lab experimental demonstration of nonlinear impairment compensation using neural networks, Nature Commun. 10 (1) (2019) 1–8, ISBN: 4146701910.

[61] F.N. Khan, Q. Fan, C. Lu, A.P.T. Lau, An optical communication's perspective on machine learning and its applications, J. Lightw. Technol. 37 (2) (2019) 493–516.

[62] T. Koike-Akino, Y. Wang, D. Millar, K. Kojima, K. Parsons, Neural turbo equalization: Deep learning for fiber-optic nonlinearity compensation, J. Lightw. Technol. 38 (11) (2020) 3059–3066.

[63] S. Deligiannidis, A. Bogris, C. Mesaritakis, Y. Kopsinis, Compensation of fiber nonlinearities in digital coherent systems leveraging long short-term memory neural networks, J. Lightw. Technol. 38 (21) (2020) 5991–5999.

[64] C. Häger, H.D. Pfister, Physics-based deep learning for fiber-optic communication systems, IEEE J. Sel. Areas Commun. 39 (1) (2021) 280–294.

[65] P.J. Freire, A. Napoli, B. Spinnler, N.M.S. da Costa, S.K. Turitsyn, J. Prilepsky, Neural networks-based equalizers for coherent optical transmission: Caveats and pitfalls, IEEE J. Sel. Top. Quantum Electron. (2022) 1–23.

[66] L. Shiloh, A. Eyal, R. Giryes, Efficient processing of distributed acoustic sensing data using a deep learning approach, J. Lightw. Technol. 37 (18) (2019) 4755–4762.

[67] P. Westbrook, Big data on the horizon from a new generation of distributed optical fiber sensors, APL Photonics 5 (2) (2020).

[68] M. Wada, Y. Maeda, H. Shimabara, T. Aihara, Manhole locating technique using distributed vibration sensing and machine learning, in: Proc. OFC, Washington, DC, 2021.

[69] K. Abdelli, F. Azendorf, H. Grieser, C. Tropschug, S. Pachnicke, Gated recurrent unit based autoencoder for optical link fault diagnosis in passive optical networks, in: Proc. European Conf. Optical Communication, ECOC, Bordeaux, France, 2021.

[70] Y. Lu, Y. Tian, S. Han, E. Cosatto, S. Ozharar, Y. Ding, Automatic fine-grained localization of utility pole landmarks on distributed acoustic sensing traces based on bilinear resnets, in: Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing, ICASSP, Toronto, Canada, 2021.

[71] K. Abdelli, H. Grieser, C. Tropschug, S. Pachnicke, Optical fiber fault detection and localization in a noisy OTDR trace based on denoising convolutional autoencoder and bidirectional long short-term memory, J. Lightw. Technol. 40 (8) (2022) 2254–2264.

[72] D. Wang, C. Zhang, W. Chen, H. Yang, M. Zhang, A.P.T. Lau, A review of machine learning-based failure management in optical networks, Sci. China Inf. Sci. (2022).

[73] A. Bukharin, S. Han, Y. Chen, M.-F. Huang, Y.-K. Huang, Y. Xie, T. Wang, Ambient noise-based weakly supervised manhole localization methods over deployed fiber networks, Opt. Express 31 (6) (2023) 9591.

[74] L. Sadighi, S. Karlsson, C. Natalino, M. Furdek, Machine learning-based polarization signature analysis for detection and categorization of eavesdropping and harmful events, in: Proc. Optical Fiber Communication Conf., OFC, San Diego, CA, 2024.

[75] K. Abdelli, M. Lonardi, J. Gripp, D. Correa, S. Olsson, F. Boitier, P. Layec, Anomaly detection and localization in optical networks using vision transformer and SOP monitoring, in: Proc. Optical Fiber Communication Conf., OFC, San Diego, CA, 2024.

[76] F. Usmani, H. Awad, E. Virgillito, R. Bratovich, S. Straullu, F. Aquilino, R. Proietti, R. Pastorelli, V. Curri, Earthquake early warning through terrestrial optical networks: A Bi-GRU attention model approach on SOP data, in: Proc. Optical Fiber Communication Conf., OFC, San Diego, CA, 2024.

[77] A. Dal Forno, A. Paradisi, R. Passy, J. von der Weid, Experimental and theoretical modeling of polarization-mode dispersion in single-mode fibers, IEEE Photon. Techn. Lett. 12 (3) (2000) 296–298.

[78] B. Widrow, M.E. Hoff, Adaptive Switching Circuits, Technical Report, Defense Technical Information Center, Fort Belvoir, VA, 1960.

[79] B. Widrow, J. Glover, J. McCool, J. Kaunitz, C. Williams, R. Hearn, J. Zeidler, J. Eugene Dong, R. Goodlin, Adaptive noise cancelling: Principles and applications, Proc. IEEE 63 (12) (1975) 1692–1716.

[80] J.G. Proakis, M. Salehi, Digital Communications, Fifth ed., McGraw-Hill, Boston, 2008.

[81] M. Rupp, Robust design of adaptive equalizers, IEEE Trans. Signal Process. 60 (4) (2012) 1612–1626.

[82] L. Schmid, T. Raviv, N. Shlezinger, L. Schmalen, Blind channel estimation and joint symbol detection with data-driven factor graphs, 2024, arXiv:2401.12627.

[83] D.P. Kingma, M. Welling, Auto-encoding variational Bayes, in: Proc. Internat. Conf. Learn. Represent., ICLR, 2014.

[84] D.M. Blei, A. Kucukelbir, J.D. McAuliffe, Variational inference: A review for statisticians, J. Amer. Statist. Assoc. 112 (518) (2017) 859–877.

[85] D.P. Kingma, J. Ba, Adam: A method for stochastic optimization, 2014, Preprint arXiv:1412.6980.

[86] C.B. Czegledi, M. Karlsson, E. Agrell, P. Johannisson, Polarization drift channel model for coherent fibre-optic systems, Sci. Rep. 6 (1) (2016) 21217.

[87] C.B. Czegledi, G. Liga, D. Lavery, M. Karlsson, E. Agrell, S.J. Savory, P. Bayvel, Digital backpropagation accounting for polarization-mode dispersion, Opt. Express 25 (3) (2017) 1903–1915.

[88] M. Rupp, Convergence properties of adaptive equalizer algorithms, IEEE Trans. Signal Process. 59 (6) (2011) 2562–2574.

[89] J.E. Simsarian, P.J. Winzer, Shake before break: Per-span fiber sensing with in-line polarization monitoring, in: Proc. Optical Fiber Communication Conf., OFC, Los Angeles, CA, 2017.

[90] J.C. Castellanos, Z. Zhan, V. Kamalov, M. Cantono, S. Yin, A. Mecozzi, S. Bhattacharya, R.M. Allen, Optical polarization-based sensing and localization of submarine earthquakes, in: Proc. Optical Fiber Communication Conf., OFC, San Diego, CA, 2022.

[91] L. Costa, S. Varughese, P. Mertz, V. Kamalov, Z. Zhan, Localization of seismic waves with submarine fiber optics using polarization-only measurements, Comms. Eng. 2 (1) (2023).

[92] F. Yaman, Y. Li, S. Han, T. Inoue, E. Mateo, Y. Inada, Polarization sensing using polarization rotation matrix eigenvalue method, in: Proc. Optical Fiber Communication Conf., OFC, San Diego, CA, 2023.

[93] A. Mecozzi, Sensing with submarine optical cables, 2024, arXiv:2403.18448 [physics].

[94] K. Goroshko, H. Louchet, A. Richter, Overcoming performance limitations of digital back propagation due to polarization mode dispersion, in: Proc. Int. Conf. Transparent Optical Networks, ICTON, Trento, Italy, 2016.

[95] G. Liga, C. Czegledi, P. Bayvel, A PMD-adaptive DBP receiver based on SNR optimization, in: Proc. OFC, 2018.

[96] B.I. Bitachon, A. Ghazisaeidi, M. Eppenberger, B. Baeurle, M. Ayata, J. Leuthold, Deep learning based digital backpropagation demonstrating {SNR} gain at low complexity in a 1200 km transmission link, Opt. Express 28 (20) (2020) 29318–29334.

[97] R.M. Bütler, C. Häger, H.D. Pfister, G. Liga, A. Alvarado, Model-based machine learning for joint digital backpropagation and PMD compensation, J. Lightw. Technol. 39 (4) (2021) 949–959.

[98] P. Jain, L. Lampe, J. Mitra, Joint PMD tracking and nonlinearity compensation with deep neural networks, J. Lightw. Technol. 41 (12) (2023) 3957–3966.

[99] M. Abu-romoh, N. Costa, Y. Jaouën, A. Napoli, J. Pedro, B. Spinnler, M. Yousefi, Equalization in dispersion-managed systems using learned digital back-propagation, Optics Continuum 2 (10) (2023) 2088.

[100] K. Liu, E. Börjeson, C. Häger, P. Larsson-Edefors, FPGA implementation of multi-layer machine learning equalizer with on-chip training, in: Proc. Optical Fiber Communication Conf., OFC, San Diego, CA, 2023.

[101] T. Sasai, M. Nakamura, E. Yamazaki, S. Yamamoto, H. Nishizawa, Y. Kisaka, Digital longitudinal monitoring of optical fiber communication link, J. Lightw. Technol. 40 (8) (2022) 2390–2408.

[102] M. Eto, K. Tajima, S. Yoshida, S. Oda, T. Hoshida, Location-resolved PDL monitoring with Rx-side digital signal processing in multi-span optical transmission system, in: Proc. Optical Fiber Communication Conf., OFC 2022, San Diego, CA, 2022.

[103] A. May, E. Awwad, P. Ramantanis, P. Ciblat, Receiver-based localization and estimation of polarization dependent loss, in: Proc. Opto-Electronics and Commun. Conf., OECC, Toyama, Japan, 2022.

[104] M. Takahashi, T. Sasai, E. Yamazaki, Y. Kisaka, DSP-based PDL estimation and localization in multi-span optical link using least squares-based longitudinal power monitoring, in: Proc. Opto-Electronics and Commun. Conf., OECC, Shanghai, China, 2023.

[105] R.M. Bütler, C. Häger, H.D. Pfister, G. Liga, A. Alvarado, Model-based machine learning for joint digital backpropagation and PMD compensation, J. Lightw. Technol. 39 (4) (2020) 949–959.

[106] Z. Hu, C. Zhao, Y. Chen, M. Zhang, J. Chen, W. Li, M. Tang, Simultaneous distributed acoustic sensing and communication in digital subcarrier multiplexing systems, in: Proc. European Conf. Optical Communication, ECOC, Glasgow, UK, 2024.

[107] H. He, L. Jiang, Y. Pan, A. Yi, X. Zou, W. Pan, A.E. Willner, X. Fan, Z. He, L. Yan, Integrated sensing and communication in an optical fibre, Light Sci. Appl. 12 (1) (2023) 25.

[108] J. Wang, L. Lu, L. Wang, Y. Yan, A.P.T. Lau, C. Lu, High-efficiency ISAC to enable sub-meter level vibration sensing for coherent fiber networks, in: Proc. Optical Fiber Communication Conf., OFC, San Diego, CA, 2024.

[109] Y. Xiong, F. Liu, Y. Cui, W. Yuan, T.X. Han, G. Caire, On the fundamental tradeoff of integrated sensing and communications under Gaussian channels, IEEE Trans. Inform. Theory 69 (9) (2023) 5723–5751.

[110] T. O'Shea, J. Hoydis, An introduction to deep learning for the physical layer, IEEE Trans. Cogn. Commun. Netw. 3 (4) (2017) 563–575.

[111] S. Li, C. Häger, N. Garcia, H. Wymeersch, Achievable information rates for nonlinear fiber communication via end-to-end autoencoder learning, in: Proc. European Conf. Optical Communication, ECOC, Rome, Italy, 2018.

[112] B. Karanov, M. Chagnon, F. Thouin, T.A. Eriksson, H. Bülow, D. Lavery, P. Bayvel, L. Schmalen, End-to-end deep learning of optical fiber communications, J. Lightw. Technol. 36 (20) (2018) 4843–4855.

[113] R.T. Jones, T.A. Eriksson, M.P. Yankov, D. Zibar, Deep learning of geometric constellation shaping including fiber nonlinearities, in: Proc. European Conf. Optical Communication, ECOC, Rome, Italy, 2018.

[114] R.T. Jones, M.P. Yankov, D. Zibar, End-to-end learning for GMI optimized geometric constellation shape, in: Proc. European Conf. Optical Communication, ECOC, Dublin, Ireland, 2019.

[115] K. Gümüs, A. Alvarado, B. Chen, C. Häger, E. Agrell, End-to-end learning of geometrical shaping maximizing generalized mutual information, in: Proc. Optical Fiber Communication Conf., OFC, San Diego, CA, 2020.

[116] T. Uhlemann, S. Cammerer, A. Span, S. Dörner, S. ten Brink, Deep-learning autoencoder for coherent and nonlinear optical communication, in: IEEE/ITG Symp. Photon. Netw., SPN, Leipzig, Germany, 2020.

[117] S. Gaiarin, F. Da Ros, R.T. Jones, D. Zibar, End-to-end optimization of coherent optical communications over the split-step Fourier method guided by the nonlinear Fourier transform theory, J. Lightw. Technol. 39 (2) (2021) 418–428.

[118] M. Srinivasan, J. Song, A. Grabowski, K. Szczerba, H.K. Iversen, M.N. Schmidt, D. Zibar, J. Schröder, A. Larsson, C. Häger, H. Wymeersch, End-to-end learning for VCSEL-based optical interconnects: State-of-the-art, challenges, and opportunities, J. Lightw. Technol. 41 (11) (2023) 3261–3277.

[119] O. Jovanovic, F. Da Ros, D. Zibar, M.P. Yankov, Geometric constellation shaping for fiber-optic channels via end-to-end learning, J. Lightw. Technol. 41 (12) (2023) 3726–3736.

[120] A. Rode, B. Geiger, S. Chimmalgi, L. Schmalen, End-to-end optimization of constellation shaping for Wiener phase noise channels with a differentiable blind phase search, J. Lightw. Technol. 41 (12) (2023) 3849–3859.

[121] J.M. Mateos-Ramos, J. Song, Y. Wu, C. Häger, M.F. Keskin, V. Yajnanarayana, H. Wymeersch, End-to-end learning for integrated sensing and communication, in: Proc. IEEE Int. Conf. Communications, ICC, Seoul, South Korea, 2022.

[122] J.M. Mateos-Ramos, C. Häger, M.F. Keskin, L.L. Magoarou, H. Wymeersch, Model-driven end-to-end learning for integrated sensing and communication, in: Proc. IEEE Int. Conf. Communications, ICC, Rome, Italy, 2023.

[123] C. Muth, L. Schmalen, Autoencoder-based joint communication and sensing of multiple targets, in: Proc. ITG Conf. Systems, Communications, Coding, SCC, Braunschweig, Germany, 2023.

[124] J.M. Mateos-Ramos, B. Chatelier, C. Häger, M.F. Keskin, L.L. Magoarou, H. Wymeersch, Semi-supervised end-to-end learning for integrated sensing and communications, in: Proc. Int. Conf. Mach. Learning for Communication and Networking, ICMLCN, Stockholm, Sweden, 2024.