Learning to Extract Distributed Polarization Sensing Data from Noisy Jones Matrices

(article starts on next page)

Learning to Extract Distributed Polarization Sensing Data from Noisy Jones Matrices

Mohammad Farsi[1], Christian Häger[1], Magnus Karlsson[2], and Erik Agrell[1]

[1]*Dept. of Electrical Engineering, Chalmers Univ. of Technology, Sweden,*
[2]*Dept. of Microtechnology and Nanoscience, Chalmers Univ. of Technology, Sweden*

[*]farsim@chalmers.se

Abstract: We consider the problem of recovering spatially resolved polarization information from receiver Jones matrices. We introduce a physics-based learning approach, improving noise resilience compared to previous inverse scattering methods, while highlighting challenges related to model overparameterization. © 2023 The Author(s)

1. Introduction

The possibility of repurposing existing telecommunication fibers as environmental sensors has recently gained significant attention [1]. Compared to established sensing approaches like distributed acoustic sensing, extracting polarization sensing data directly from adaptive equalizers in coherent communication receivers [2–5] could have potential advantages, particularly in terms of cost, scalability, and compatibility with existing fiber infrastructure and equipment. An important open question is to what extent polarization sensing can provide distributed (i.e., spatially resolved) location information about external environmental events. Prior work makes relatively strong assumptions such as access to multiple optical paths [6], per-span readouts [7], or loop-back configurations [8].

In this paper, we study the problem of recovering spatially resolved polarization information directly from the estimated (frequency-dependent) Jones matrix at the receiver. A similar problem has been previously studied in [9–11]. In [9, 10], an inverse scattering algorithm (ISA) is considered to recover the distributed differential group delay (DGD) profile from the overall system impulse response (i.e., the time-domain Jones matrix), essentially factorizing the system into a cascade of individual responses. The ISA was extended in [11] to account for polarization-dependent loss (PDL), recovering both DGD and PDL profiles. In principle, the ISA is exact for the settings considered in [9–11], assuming that the overall impulse response is known and noise-free.

Our contributions are as follows. First, we show that the performance of the ISA quickly deteriorates in the presence of noise. We then propose and investigate a physics-based learning approach [12, 13], which directly parameterizes the underlying propagation model and jointly optimizes all parameters. We then show some examples where this approach can recover location information about time-varying polarization perturbations, even in the presence of noise. Finally, we discuss persisting limitations and challenges, in particular cases where model overparameterizations lead to ambiguities in interpreting the learned parameter configurations. We note that our learning approach is also related to work on distributed polarization-mode dispersion (PMD) compensation [14–21] and longitudinal path and PDL monitoring [22–25]. However, [14–21] do not investigate the feasibility of utilizing the optimized model representations for sensing or monitoring purposes and the approaches in [22–25] exploit additional nonlinear effects to enable spatial resolvability. The setting considered in this paper is different as we only assume access to the estimated (linear) Jones matrix. As such, the proposed approach is in principle fully compatible with existing polarization sensing approaches based on adaptive equalization.

2. System Model

We study dual-polarization optical transmission accounting for PDL, polarization rotations, and DGD. The assumed channel model is based on [11]. In particular, the overall channel response at time k and frequency ω_i is described by a 2×2 complex-valued Jones matrix resulting from N concatenated sections according to

$$H(\omega_i; \theta_k) = \prod_{n=1}^{N} \Gamma(\gamma_n) R\left(\phi_n(k), \psi_n(k)\right) T(\omega_i), \tag{1}$$

where we use the convention $\prod_{n=1}^{N} A_n = A_N A_{N-1} \cdots A_1$, $n \in \{1, \ldots, N\}$ denotes the section index,

$$\Gamma(\gamma_n) = \begin{bmatrix} e^{\gamma_n/2} & 0 \\ 0 & e^{-\gamma_n/2} \end{bmatrix}, \ R(\phi_n(k), \psi_n(k)) = \begin{bmatrix} \cos\phi_n(k) & je^{j\psi_n(k)}\sin\phi_n(k) \\ je^{-j\psi_n(k)}\sin\phi_n(k) & \cos\phi_n(k) \end{bmatrix}, \ T(\omega_i) = \begin{bmatrix} 1 & 0 \\ 0 & e^{j\omega_i\tau} \end{bmatrix}, \tag{2}$$

γ_n is the PDL extinction parameter, $\phi_n(k)$ and $\psi_n(k)$ are (possibly time-varying) rotation angles, τ is the DGD per section, and $\theta_k = \{\gamma_n, \phi_n(k), \psi_n(k)\}_{n=1}^{N}$ denotes all channel parameters. In addition to [11], we assume that an environmental perturbation introduces time variations in the rotation angles $\phi_n(k)$ and $\psi_n(k)$. The time evolution is

modeled as a random walk according to $\phi_n(k) = \phi_n(k-1) + \Phi_n(k)$ and $\psi_n(k) = \psi_n(k-1) + \Psi_n(k)$ where $\Phi_n(k) \sim \mathcal{N}(0, \sigma_n^2(k))$ and $\Psi_n(k) \sim \mathcal{N}(0, \rho_n^2(k))$. Here, $\sigma_n^2(k)$ and $\rho_n^2(k)$ are emulating the intensity of the environmental change, i.e., larger values correspond to more severe changes in the environment. Note that the model (1) assumes that the time scale of the dynamic polarization perturbation is much slower than the DGD parameter τ.

In coherent receivers, the Jones matrix (1) is typically estimated through adaptive equalization. We therefore assume that we have access to a noisy version of the Jones matrix $\tilde{H}(\omega_i; \theta_k) = H(\omega_i; \theta_k) + Z_k(\omega_i)$, where $Z_k(\omega_i)$ is a 2×2 matrix with zero-mean independent complex normal elements with variance σ_z^2 modeling the estimation error, i.e., $\mathbb{E}[Z_k(\omega_i) \circ Z_k(\omega_i)^\dagger] = \sigma_z^2 \left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, where $\circ$ is the Hadamard product. The estimation noise variance σ_z^2 depends on various factors such as the choice of estimation algorithm and the channel signal-to-noise ratio (SNR), where better algorithms and higher SNR lead to lower σ_z^2.

3. Distributed Parameter Estimation Techniques

In this section, we describe the ISA and the proposed learning approach, both aiming at recovering the unknown parameters θ_k from $\tilde{H}(\omega_i, \theta_k)$.[1]

Inverse Scattering Algorithm: The ISA proposed in [11] operates under the assumption that the overall channel response is noiseless. The algorithm analytically calculates the values of the unknown parameters θ_k and is performed iteratively in the time domain by first taking the inverse Fourier transform of $\tilde{H}(\omega_i, \theta_k)$. The algorithm's procedure then calculates the unknown parameter values for section N, i.e., γ_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.

Machine Learning Approach: We propose a physics-based machine learning approach, which directly parameterizes the propagation model (1) and jointly optimizes all its associated parameters. Denoting the set of trainable parameters as $\hat{\theta}_k = \{\hat{\gamma}_n, \hat{\phi}_n(k), \hat{\psi}_n(k)\}_{n=1}^N$, we define the frequency-averaged cost function as $\mathscr{L}(\hat{\theta}_k) = \sum_{i=1}^L \left\| \tilde{H}(\omega_i; \theta_k) - H(\omega_i; \hat{\theta}_k) \right\|^2$, where L is the number of frequency samples. This loss function is then iteratively minimized by applying M iterations of a gradient-descent optimizer with learning rate α.

4. Numerical Results and Discussion

For the numerical results, we consider $N = 5$ sections, where, within each section, we sample the extinction parameter γ_n from a uniform distribution within the range of $[0.07, 0.17]$ (equivalent to $[0.3, 0.7]$ dB PDL). Furthermore, the initial rotation parameters $\phi_n(0)$ and $\psi_n(0)$ are sampled uniformly from $[-\pi, \pi]$. The DGD parameter τ is normalized to 1 and assumed to be known and constant across all sections. Note that this assumption is required for the ISA, whereas the learning approach can potentially be extended to also learn unknown DGD sections, similar to, e.g., [18]. For the learning approach, the Adam optimizer is used with $M = 300$ and $\alpha = 0.05$, where all trainable parameters $\hat{\theta}_k$ are initialized to zero at $k = 0$ and then tracked for $k > 0$. The frequency samples in the loss function correspond to $1/\tau$ sampling distance in the frequency domain with $L = N$, where we verified that increasing the number of frequency samples does not affect the results. Lastly, we consider a time-varying scenario where an environmental change within section $n = 2$ at $k \in [15, 35]$ affects parameters $\phi_2(k)$ and $\psi_2(k)$ with $\sigma_2^2(k) = \rho_2^2(k) = 0.1$ for $k \in [15, 35]$ and $\sigma_2^2(k) = \rho_2^2(k) = 0$ elsewhere.

Fig. 1 compares the tracking capability of the ISA and the learning approach by plotting $|\cos \phi_n(k)|$ for all sections in four different scenarios. The absolute value is used since two negative signs can potentially commute between different sections in (1), effectively canceling each other and leading to sign ambiguity of the estimated angle parameters.

Trackable cases: As depicted in Fig. 1(a), when the noiseless channel response is accessible, both methods are able to track the perturbation, pinpointing its location and timing. However, the situation changes when we examine Figs. 1(b) and (c), in which the channel response is subject to varying amounts of noise. In these cases, our proposed approach demonstrates improved noise resilience by successfully tracking and identifying the perturbation's location and timing, whereas the ISA loses its capability to track or precisely determine the section affected by the environmental change.

Non-trackable cases: For certain realizations of the channel parameters θ_k, there are situations where simultaneous changes in multiple parameters result in nearly identical channel responses $H(\omega_i; \theta_k)$ implying that it is possible to approximately describe $H(\omega_i; \theta_k)$ using fewer independent parameters. An example is shown in Fig. 1(d). Such cases manifest different behavior for the two algorithms. For the ISA, numerical instabilities arise due to noise and neither the overall response nor the estimated parameters align with the actual model. For the learning method, the learned overall channel response $H(\omega_i; \hat{\theta}_k)$ remains close to the actual response $H(\omega_i; \theta_k)$ due to the employed loss function. However, since there are many parameter configurations that produce approximately the same response, the model is effectively overparameterized and no stable, unique solution can be found.

---

[1] Open-source implementations are provided at https://github.com/Mohammadfarsi1994/physics-based-distributed-polarization-sensing.
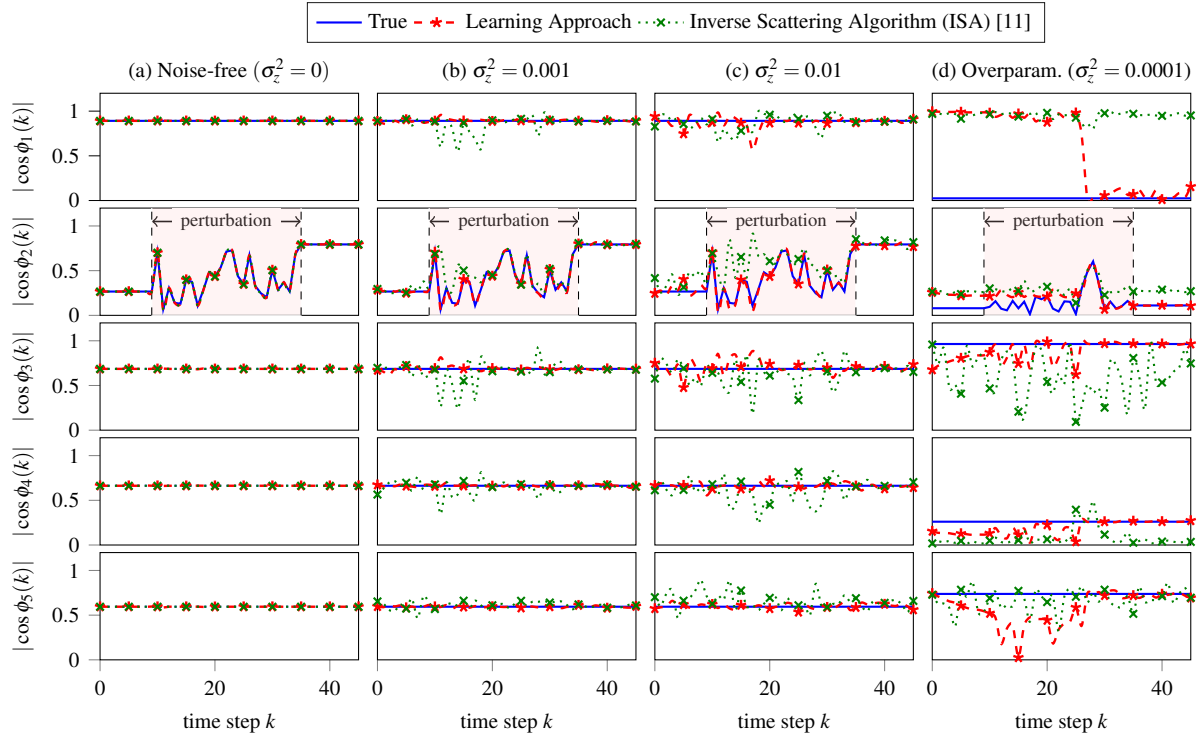
Fig. 1: Numerical results: (a)-(c) trackable channel realization with varying amounts of noise; (d) non-trackable realization.

To address this issue, one possible approach would be to reduce the number of trainable parameters, for example by reducing the number of sections in the learned model, potentially sacrificing some localization accuracy. In this case, however, the ISA is no longer applicable and it becomes nontrivial to assess how well the learned distributed response matches the propagation model because it is not possible to compare on a parameter-by-parameter basis.

5. Conclusion

In this paper, we considered the problem of estimating the distributed channel response from noisy Jones matrices. We showed that the previously proposed ISA is relatively susceptible to noise. To address this limitation, we proposed a physics-based learning approach which exploits the propagation model and optimizes all parameters simultaneously, yielding promising results. We demonstrated successful location retrieval of a time-varying polarization perturbation in noisy conditions. Persisting challenges relate to cases of model overparameterization, making it ambiguous to relate the learned parameters to the actual perturbation in channel parameters. 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.

References
1. E. Ip *et al.*, "Using Global Existing Fiber Networks for Environmental Sensing," Proc. IEEE **110**, 1853–1888 (2022).
2. F. Boitier *et al.*, "Proactive Fiber Damage Detection in Real-time Coherent Receiver," in *Proc. ECOC*, (2017).
3. Z. Zhan *et al.*, "Optical polarization–based seismic and water wave sensing on transoceanic cables," Science **371**, 931–936 (2021).
4. A. Mecozzi *et al.*, "Polarization sensing using submarine optical cables," Optica **8**, 788–795 (2021).
5. M. Mazur *et al.*, "Field Trial of FPGA-Based Real-Time Sensing Transceiver over 524 km of Live Aerial Fiber," in *Proc. OFC*, (2023).
6. J. C. Castellanos *et al.*, "Optical polarization-based sensing and localization of submarine earthquakes," in *Proc. OFC*, (2022).
7. J. E. Simsarian and P. J. Winzer, "Shake Before Break: Per-Span Fiber Sensing with In-Line Polarization Monitoring," in *Proc. OFC*, (2017).
8. D. Charlton *et al.*, "Field measurements of SOP transients in OPGW, with time and location [...]," Opt. Express **25**, 9689–9696 (2017).
9. S. E. Harris *et al.*, "Optical Network Synthesis Using Birefringent Crystals [...]," J. Opt. Soc. Am. **54**, 1267–1279 (1964).
10. L. Moller, "Filter synthesis for broad-band PMD compensation in WDM systems," IEEE Photon. Technol. Lett. **12**, 1258–1260 (2000).
11. R. Noé *et al.*, "Polarization-Dependent Loss: New Definition and Measurement Techniques," J. Light. Technol. **33**, 2127–2138 (2015).
12. C. Häger and H. D. Pfister, "Nonlinear Interference Mitigation via Deep Neural Networks," in *Proc. OFC*, (2018).
13. C. Häger and H. D. Pfister, "Physics-Based Deep Learning for Fiber-Optic [...]," IEEE J. Sel. Areas Commun. **39**, 280–294 (2021).
14. K. Goroshko *et al.*, "Overcoming performance limitations of digital back propagation due to polarization mode dispersion," in *Proc. ICTON*, (2016).
15. C. B. Czegledi *et al.*, "Digital backpropagation accounting for polarization-mode dispersion," Opt. Express **25**, 1903–1915 (2017).
16. G. Liga *et al.*, "A PMD-adaptive DBP receiver based on SNR optimization," in *Proc. OFC*, (2018).
17. B. I. Bitachon *et al.*, "Deep learning based digital backpropagation demonstrating SNR gain at [...]," Opt. Express **28**, 29318–29334 (2020).
18. R. M. Bütler *et al.*, "Model-Based Machine Learning for Joint Digital Backpropagation and [...]," J. Light. Technol. **39**, 949–959 (2021).
19. P. Jain *et al.*, "Joint PMD Tracking and Nonlinearity Compensation With Deep Neural Networks," J. Light. Technol. **41**, 3957–3966 (2023).
20. M. Abu-romoh *et al.*, "Equalization in dispersion-managed systems using learned digital back-propagation," Opt. Continuum **2**, 2088 (2023).
21. K. Liu *et al.*, "FPGA Implementation of Multi-Layer Machine Learning Equalizer with On-Chip Training," in *Proc. OFC*, (2023).
22. T. Sasai *et al.*, "Digital Longitudinal Monitoring of Optical Fiber Communication Link," J. Light. Technol. **40**, 2390–2408 (2022).
23. M. Eto *et al.*, "Location-resolved PDL Monitoring with Rx-side Digital Signal Processing in [...]," in *Proc. OFC*, (2022).
24. A. May *et al.*, "Receiver-Based Localization and Estimation of Polarization Dependent Loss," in *Proc. OECC*, (2022).
25. M. Takahashi *et al.*, "DSP-based PDL Estimation and Localization in Multi-Span Optical Link Using [...]," in *Proc. OECC*, (2023).