PIDT: Physics-Informed Digital Twin for Optical Fiber Parameter Estimation

PIDT: Physics-Informed Digital Twin for Optical Fiber Parameter Estimation

Zicong Jiang,[1,*] Magnus Karlsson,[2] Erik Agrell,[1] and Christian Häger[1]

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

[*] *zicongj@chalmers.se*

Abstract: We propose physics-informed digital twin (PIDT): a fiber parameter estimation approach that combines a parameterized split-step method with a physics-informed loss. PIDT improves accuracy and convergence speed with lower complexity compared to previous neural operators.

1. Introduction

In optical fiber communication systems, accurate knowledge of fiber parameters such as attenuation, dispersion and nonlinearity, and their variation along the link is essential for tasks such as impairment compensation, performance monitoring, and system optimization. Recently, learning-based parameter estimation methods have been proposed that embed physics knowledge into training. For example, physics-informed neural networks (PINNs) [1] incorporate partial differential equation constraints, such as the nonlinear Schrödinger equation (NLSE), as a physics-informed loss. PINNs have been applied to optical channel modeling [2] and estimating fiber parameters [3], but they require retraining for each new initial condition (i.e., transmitted signal). To overcome this problem, neural operator methods have been proposed [4], which learn a solution operator over a distribution of initial conditions for better generalization. Variants of these methods, called physics-informed neural operator (PINO) [5], have been applied to fiber parameter estimation [6], and channel modeling [7,8]. However, prior works mainly use simplified or short signals to simplify training. Moreover, both PINN and PINO rely on conventional neural network (NN) models that require extensive hyperparameter tuning (e.g., number of layers or neurons) and many training iterations to converge.

We propose physics-informed digital twin (PIDT): a neural operator method for accurate and fast parameter estimation with realistic signal sequences. PIDT replaces the conventional NN with an interpretable physics-based model, effectively combining the parameterized split-step Fourier method (SSFM) [9, 10] with a physics-informed loss [1]. A key technical challenge is that, unlike a NN, the SSFM produces outputs that are discrete in space and time. This is not directly compatible with the neural operator framework. To address this, we introduce a differentiable interpolation step that links SSFM outputs to the physics-informed loss. We show that PIDT significantly improves estimation accuracy and convergence speed with lower complexity compared to previous PINO approaches in [6]. Moreover, PIDT also addresses a limitation of prior SSFM-based parameter estimation approaches in, e.g., [11], which rely solely on minimizing the observation loss at the fiber output. By applying the physics-informed loss to the full t- and z-dependent optical wavefield, PIDT enhances physical consistency of the mathematical twin model at intermediate fiber locations compared to prior SSFM-based estimation methods.

2. Background and related approaches

Following the convention in the PINO literature, we define the NLSE in residual form $R(s(z,t);\theta)=0$, where

$$R(s(z,t);\theta)=\frac{\partial s(z,t)}{\partial z}+\frac{\alpha}{2}s(z,t)+\frac{\mathrm{j}\beta_2}{2}\frac{\partial^2 s(z,t)}{\partial t^2}-\mathrm{j}\gamma|s(z,t)|^2 s(z,t), \tag{1}$$

$s(z,t)\in\mathbb{C}$ is the optical baseband signal at space–time coordinate $(z,t)\in[0,L]\times\mathbb{R}$, L is the propagation distance, and $\theta=\{\alpha,\beta_2,\gamma\}$ are the attenuation, dispersion, and nonlinear parameters. The parameter estimation problem is to estimate the coefficients in θ (or their profiles along the fiber) from observed input–output signal pairs. Since our method builds on two existing approaches, we first review SSFM-based and PINO-based parameter estimation.

2.1. SSFM-based parameter estimation: The SSFM approximates (1) using M segments of length $\Delta z=L/M$, where $\hat{\mathbf{s}}^{(m)}=H(\hat{\mathbf{s}}^{(m-1)};\theta_{\mathrm{DT}}^{(m)})=\mathbf{D}\sigma(\mathbf{D}\hat{\mathbf{s}}^{(m-1)})\in\mathbb{C}^N$ is the approximated signal at $z_m=m\Delta z$ starting with $\hat{\mathbf{s}}^{(0)}=[s(0,t_1),\dots,s(0,t_N)]^\top$, $t_n=n\Delta t$, $\Delta t=T/N$ is the sampling period, T the time window, and N the number of samples. Here, $\mathbf{D}=\mathbf{F}^{-1}\operatorname{diag}\big(e^{(-\alpha^{(m)}/2+\mathrm{j}\beta_2^{(m)}\omega_k^2/2)\Delta z/2}\big)\mathbf{F}$, $\sigma(x)=xe^{\mathrm{j}\gamma^{(m)}\Delta z|x|^2}$ is applied element-wise, $\theta_{\mathrm{DT}}^{(m)}=\{\alpha^{(m)},\beta_2^{(m)},\gamma^{(m)}\}$ are segment-wise tunable parameters, $\mathbf{F}$ is the discrete Fourier transform matrix (periodic boundary conditions are assumed), and ω_k the k-th angular frequency. We denote the overall SSFM mapping as $\hat{\mathbf{y}}(\hat{\mathbf{s}}^{(0)};\theta_{\mathrm{DT}})=\hat{\mathbf{s}}^{(M)}$, where $\theta_{\mathrm{DT}}=\{\theta_{\mathrm{DT}}^{(m)}\}_{m=1}^{M}$. For parameter estimation (see Fig. 1a), θ_{DT} is optimized by minimizing the observation loss $\mathscr{L}_{\mathrm{o}}(\theta_{\mathrm{DT}})=(1/|\mathscr{D}|)\sum_{(\mathbf{x},\mathbf{y})\in\mathscr{D}}\|\hat{\mathbf{y}}(\mathbf{x};\theta_{\mathrm{DT}})-\mathbf{y}\|^2$, where $\mathscr{D}$ is the training set consisting

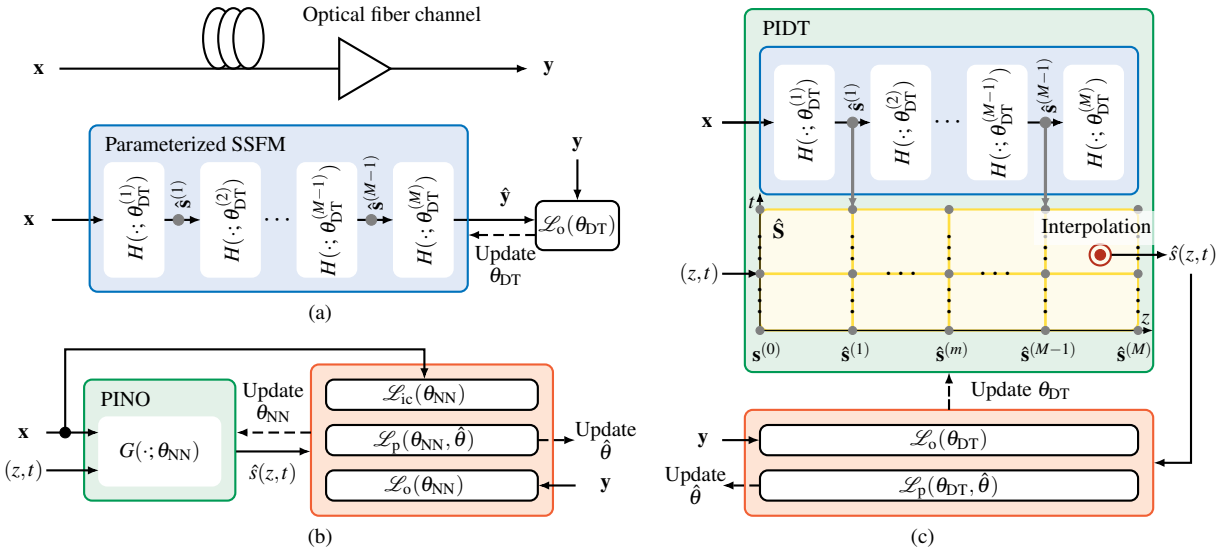

Fig. 1. Fiber-parameter estimation using (a) the parameterized SSFM, (b) PINO, and (c) PIDT (proposed).

of pairs of input and received reference signals. After training, θ_{DT} can be directly extracted from the model. However, since $\mathscr{L}_{\mathrm{o}}$ constrains only the output, it may yield parameter sets that violate physical consistency along the fiber. This motivates the use of additional loss functions that constrain the full wavefield.

2.2. PINO-based parameter estimation: The PINO in Fig. 1b learns a solution operator $G(\mathbf{x},z,t;\theta_{\mathrm{NN}})$ that maps input signals $\mathbf{x}\in\mathbb{C}^N$ and space-time coordinates (z,t) to the approximated signal $\hat{s}(z,t)\in\mathbb{C}$. In prior works [6,7], G is implemented as a black-box NN[1] with parameters θ_{NN} (weights and biases), so fiber parameters cannot be directly extracted from it. Instead, fiber parameters $\hat{\theta}=\{\hat{\alpha},\hat{\beta}_2,\hat{\gamma}\}$ are estimated by minimizing the physics-informed loss $\mathscr{L}_{\mathrm{p}}(\theta_{\mathrm{NN}},\hat{\theta})=(1/(|\mathscr{X}||\mathscr{C}|))\sum_{\mathbf{x}\in\mathscr{X}}\sum_{(z,t)\in\mathscr{C}}|R(G(\mathbf{x},z,t;\theta_{\mathrm{NN}});\hat{\theta})|^2$, where $\mathscr{X}=\{\mathbf{x}\,|\,(\mathbf{x},\mathbf{y})\in\mathscr{D}\}$ and $\mathscr{C}$ is a set of predefined space-time coordinates. However, training only with $\mathscr{L}_{\mathrm{p}}$ may yield incorrect $\hat{\theta}$, as the operator G can learn to satisfy the NLSE residual R with mismatched parameters. To address this, $\mathscr{L}_{\mathrm{p}}$ is combined with an initial-condition loss $\mathscr{L}_{\mathrm{ic}}(\theta_{\mathrm{NN}})=(1/|\mathscr{X}|)\sum_{\mathbf{x}\in\mathscr{X}}\|\hat{\mathbf{x}}(\mathbf{x};\theta_{\mathrm{NN}})-\mathbf{x}\|^2$ and an observation loss $\mathscr{L}_{\mathrm{o}}(\theta_{\mathrm{NN}})=(1/|\mathscr{D}|)\sum_{(\mathbf{x},\mathbf{y})\in\mathscr{D}}\|\hat{\mathbf{y}}(\mathbf{x};\theta_{\mathrm{NN}})-\mathbf{y}\|^2$, where $\hat{\mathbf{x}}(\cdot;\theta_{\mathrm{NN}})=[G(\cdot,0,t_1;\theta_{\mathrm{NN}}),\ldots,G(\cdot,0,t_N;\theta_{\mathrm{NN}})]^\top$ and $\hat{\mathbf{y}}(\cdot;\theta_{\mathrm{NN}})=[G(\cdot,L,t_1;\theta_{\mathrm{NN}}),\ldots,G(\cdot,L,t_N;\theta_{\mathrm{NN}})]^\top$. The total loss is $\mathscr{L}(\theta_{\mathrm{NN}},\hat{\theta})=\lambda_{\mathrm{o}}\mathscr{L}_{\mathrm{o}}(\theta_{\mathrm{NN}})+\lambda_{\mathrm{ic}}\mathscr{L}_{\mathrm{ic}}(\theta_{\mathrm{NN}})+\lambda_{\mathrm{p}}\mathscr{L}_{\mathrm{p}}(\theta_{\mathrm{NN}},\hat{\theta})$, where $\lambda_{\mathrm{o}},\lambda_{\mathrm{ic}},\lambda_{\mathrm{p}}$ are adjustable weights.

3. Proposed physics-informed digital twin (PIDT) for fiber parameter estimation

While neural operator methods like PINO can enable parameter estimation across different input signals, learning the full NLSE solution operator remains challenging: conventional NN models are sensitive to architecture choices and need many training iterations to converge. To overcome these issues, we propose PIDT, which embeds the parameterized SSFM as a differentiable, physics-based twin model within the operator-learning framework.

The main challenge is that PINO requires predictions at arbitrary coordinates, whereas the SSFM produces outputs on a discrete space–time grid. We resolve this interface mismatch through a differentiable interpolation step, as illustrated in Fig. 1c. In PIDT, the proposed neural operator has two inputs: the discretized transmitted signal $\mathbf{x}$ and a querying coordinate (z,t), consistent with the PINO structure in Fig. 1b. Leveraging the SSFM, we obtain the output signal vector $\hat{\mathbf{s}}^{(m)}$ at the end of each segment m and combine them with the initial condition $\hat{\mathbf{s}}^{(0)}=\mathbf{x}$ to form an intermediate signal matrix $\hat{\mathbf{S}}=[\hat{\mathbf{s}}^{(0)},\ldots,\hat{\mathbf{s}}^{(m)},\ldots,\hat{\mathbf{s}}^{(M)}]\in\mathbb{C}^{N\times(M+1)}$. The approximated signal value $\hat{s}(z,t)$ at arbitrary coordinates (z,t) is then obtained by appropriately interpolating the matrix $\hat{\mathbf{S}}$, where the choice of interpolation method is nontrivial in general. For example, a simple piecewise linear interpolation in time yields a zero second-order derivative, preventing gradient flow to the dispersion parameter $\hat{\beta}_2$. In this paper, we employ natural bicubic splines [12], which ensure continuous second-order derivatives across intervals.

Since PIDT follows the neural-operator formalism, the loss functions are identical to those in Sec. 2.2. However, owing to the physics-based model and interpolation, the initial condition is inherently satisfied at all coordinates $(z=0,t_n)$, and no additional $\mathscr{L}_{\mathrm{ic}}$ term is required. Moreover, after training, physically meaningful and interpretable parameters appear both in the SSFM twin model via θ_{DT} (where they govern signal propagation), and in the physics-informed loss via $\hat{\theta}$ (where they define the NLSE residual for consistency enforcement).

[1]The mapping is typically divided into a branch and a trunk network, with complex signals separated into real and imaginary parts.

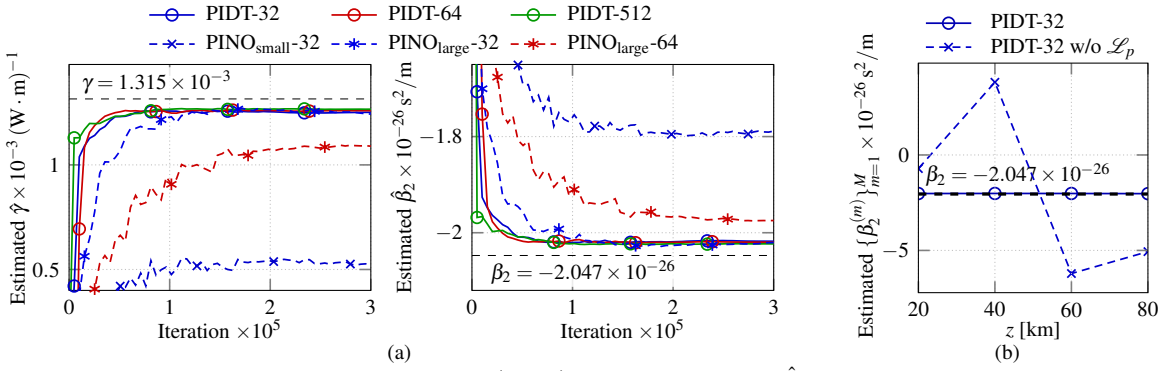

Fig. 2. Estimation results comparing PIDT-N_{sym} ($M = 4$) with (a) PINO-N_{sym} for $\hat{\theta}$ and (b) parameterized SSFM for θ_{DT}. Both PINO$_{\text{small}}$ (3 layers, 200 neurons/layer) and PINO$_{\text{large}}$ (7 layers, 900 neurons/layer) use `tanh` activation.

4. Numerical Results

To enable comparison with prior work, we adopt a similar setup as in [6]: N_{sym} 16-QAM symbols at 14 Gbaud are pulse-shaped (root-raised cosine, roll-off 0.1, oversampling $\rho = 2$, launch power $P_0 = 0\,\text{dBm}$) and transmitted over a fiber of length $L = 80$ km, resulting in $N = \rho N_{\text{sym}}$ samples. Attenuation is neglected ($\alpha = 0$) and only β_2 and γ are estimated, following [6]. Reference signals are generated using the SSFM with $M = 800$, where noise with power $P_{\text{n}} = -20\,\text{dBm}$ is added at $L = 80\,\text{km}$ in each training iteration, giving $\text{SNR} = 10\log_{10}(P_0/P_{\text{n}}) = 20\,\text{dB}$. For the neural operator in PIDT, we set $M = 4$. For PINO, several NN architectures for G are evaluated, as discussed in the following. Since no initialization scheme was specified in [6], we initialize the parameters in $\theta_{\text{DT}}, \theta_{\text{NN}}$ as $\mathscr{N}(0,1)$ (following [4]) and parameters in $\hat{\theta}$ to 0 (following [3]). Loss weights $\lambda_{\text{ic}}, \lambda_{\text{o}}, \lambda_{\text{p}}$ are adaptively updated using the balancing scheme in [13]. Other implementation details are provided in the code [14].

Figure 2a compares parameter estimation accuracy for PINO and the proposed PIDT. For $N_{\text{sym}} = 32$ (the longest symbol sequence considered in [6]), PINO performance strongly depends on the NN architecture: a smaller model (PINO$_{\text{small}}$) performs poorly, while larger configurations improve accuracy but require careful tuning. Even with an optimized architecture (PINO$_{\text{large}}$), PIDT achieves slightly higher accuracy and converges in fewer iterations. Increasing N_{sym} to 64, the performance of PINO$_{\text{large}}$ degrades, and further enlarging the NN in our simulations did not yield any improvement. In contrast, PIDT–64 maintains high accuracy without any tuning, and continues to perform reliably even for $N_{\text{sym}} = 512$, highlighting its potential for handling longer signal sequences.

In terms of complexity, we compare the number of trainable parameters in G and the approximate number of real multiplications per symbol to evaluate G. For the considered setting, PIDT has $|\theta_{\text{DT}}| = 2M = 8$ trainable parameters in G, whereas PINO$_{\text{small}}$ and PINO$_{\text{large}}$ contain, respectively $|\theta_{\text{NN}}| \approx 4 \times 10^5$ and $|\theta_{\text{NN}}| \approx 2 \times 10^7$ weights alone (excluding biases), making training substantially heavier. For each input signal, PIDT requires $C_{\text{PIDT}} = (M(8N\log_2 N + cN) + C_{\text{int}})/N_{\text{sym}}$ real multiplications per symbol, where $cN \approx 24N$ covers $8N$ nonlinear and $16N$ linear frequency-domain operations, and $C_{\text{int}} = 52$ accounts for the interpolation (assuming precalculated interpolation coefficients). In contrast, the number of multiplications per symbol for NN-based PINO, approximated by $C_{\text{PINO}} = |\theta_{\text{NN}}|/N_{\text{sym}}$, is roughly $C_{\text{PINO}}/C_{\text{PIDT}} \approx 20$ times higher for PINO$_{\text{small}}$ and $N_{\text{sym}} = 32$.

To assess physical consistency, Fig. 2b shows the estimated $\{\beta_2^{(m)}\}_{m=1}^M$ extracted from PIDT after training with and without the physics-informed loss $\mathscr{L}_{\text{p}}$. The latter corresponds to training the parameterized SSFM only on the observation loss $\mathscr{L}_{\text{o}}$. While this achieves a smaller output prediction error after training ($\mathscr{L}_{\text{o}} \approx 6 \times 10^{-4}$) compared to PIDT ($\mathscr{L}_{\text{o}} \approx 1.1 \times 10^{-3}$), the resulting parameter profiles deviate notably from the ground truth. In contrast, the inclusion of $\mathscr{L}_{\text{p}}$ in PIDT yields more accurate profiles, confirming the benefit of enforcing physical consistency.

Acknowledgements: The work of C. Häger and E. Agrell was supported by the Swedish Research Council under grants no. 2020-04718 and 2021-03709, resp. The simulation resources were provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS), partially funded by the Swedish Research Council through grant agreement no. 2022-06725.

References

1. M. Raissi *et al.*, "Physics-informed neural networks: A deep learning framework for solving [...]," Journal Computational physics **378**, 686–707 (2019).
2. Y. Zang *et al.*, "Principle-driven fiber transmission model based on pinn neural network," Journal Lightwave Technol. **40**, 404–414 (2021).
3. X. Jiang *et al.*, "Physics-informed neural network for optical fiber parameter estimation [...]," Journal Lightwave Technol. **40**, 7095–7105 (2022).
4. L. Lu *et al.*, "Learning nonlinear operators via deeponet based on the universal approximation theorem [...]," Nature machine intelligence **3**, 218–229 (2021).
5. Z. Li *et al.*, "Physics-informed neural operator for learning partial differential equations," ACM/IMS Journal Data Science **1**, 1–27 (2024).
6. Y. Song *et al.*, "Physics-informed neural operator-based full [...]," in *Optical Fiber Communication Conference,* (Optica Publishing Group, 2023), pp. M2F–5.
7. Y. Song *et al.*, "Physics-informed neural operator for fast [...]," in *2022 European Conference on Optical Communication (ECOC),* (IEEE, 2022), pp. 1–4.
8. X. Luo *et al.*, "Fast and generalizable light field propagation [...]," in *2024 Conference on Lasers and Electro-Optics (CLEO),* (IEEE, 2024), pp. 1–2.
9. C. Häger *et al.*, "Nonlinear interference mitigation via [...]," in *Optical Fiber Communication Conference,* (Optica Publishing Group, 2018), p. W3A.4.
10. C. Häger *et al.*, "Physics-based deep learning for fiber-optic communication systems," IEEE Journal on Selected Areas Communications **39**, 280–294 (2020).
11. T. Sasai *et al.*, "Digital longitudinal monitoring of optical fiber communication link," Journal Lightwave Technol. **40**, 2390–2408 (2021).
12. C. Hall, "Natural cubic and bicubic spline interpolation," SIAM Journal on Numerical Anal. **10**, 1055–1060 (1973).
13. S. Wang *et al.*, "An expert's guide to training physics-informed neural networks," arXiv preprint arXiv:2308.08468 (2023).
14. Available at https://github.com/ZicongJiang/Physics-informed-digital-twin.