

---

# Fourier Neural Operators for Fast Simulation and Inverse Design of Second-Harmonic Generation in TFLN Waveguide

---

Valentin Duruisseau<sup>1\*</sup> Robert M. Gray<sup>2\*</sup> Siyuan Jiang<sup>3\*</sup> Selina Zhou<sup>1</sup>  
Robert J. George<sup>1</sup> Kamyar Azizzadenesheli Alireza Marandi<sup>1</sup> Anima Anandkumar<sup>1</sup>  
<sup>1</sup>California Institute of Technology, <sup>2</sup>ETH Zürich, <sup>3</sup>University of Pennsylvania  
vduruiss@caltech.edu, rogray@ethz.ch, siyuanj@seas.upenn.edu

## Abstract

Advances in nanophotonics require increasingly fast and accurate modeling tools for nonlinear optical phenomena, yet conventional numerical solvers are constrained by high computational cost and limited scalability. Thin-film lithium niobate (TFLN), combining strong nonlinearities and low-loss integration, enables efficient second-harmonic generation (SHG) for telecommunications, frequency combs, and quantum photonics with stable, broadband, low-noise signals. In this paper, we present an Fourier Neural Operator (FNO) based framework for simulating SHG in TFLN waveguides. Trained on high-fidelity reference simulations from a Fourier split-step solver, the FNO achieves a relative  $L^2$  error around 4% on the test set while delivering an approximate 550,000× speedup over the solver, when executed in batch on GPUs. This computational efficiency enables extensive sampling of large design spaces, accelerating the discovery of configurations with desired properties. By further exploiting the differentiability of the FNO, we develop an inverse design pipeline that applies gradient-based optimization to directly adjust waveguide geometry, poling period mismatch and pump pulse energy to maximize conversion efficiency. The pipeline identifies high-efficiency device configurations very rapidly, reducing design cycles from days or weeks to seconds. Across 1000 random initializations, 75% of designs exceeded 85% second-harmonic conversion efficiency within 100 optimization iterations, underscoring the potential of FNO-driven methods to revolutionize nonlinear photonic device design.

## 1 Introduction and Background

Optical systems underpin modern technologies in telecommunications, spectroscopy, sensing, computing, and information processing. The rise of photonic integration has reduced the size and cost of essential components while enabling dense, low-loss on-chip interconnects. In particular, integrated waveguides provide strong mode confinement for flexible signal routing, dispersion engineering, and enhanced nonlinear interactions, yielding a scalable toolbox of linear and nonlinear building blocks for photonic circuits. Nonlinearity is central to this landscape: spectral translation via frequency conversion extends laser access to otherwise challenging spectral windows [1, 2], and nonlinear dynamics improve sources through driven dissipative solitons and supercontinuum generation [3–7], while also empowering advanced sensing and information processing [8–10]. Thin-film lithium niobate (TFLN) has emerged as a dominant integrated platform by combining strong  $\chi^{(2)}$  and  $\chi^{(3)}$  responses with a large electro-optic coefficient and wafer-scale, low-loss waveguide fabrication [11, 12].

---

\*Equal contribution.

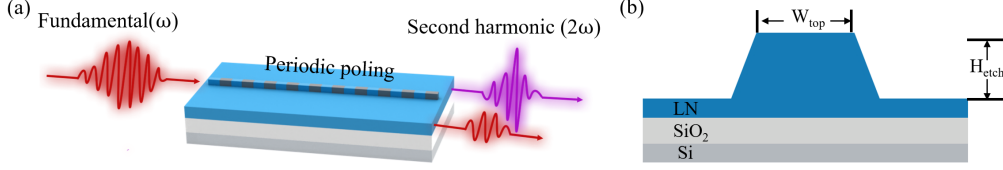


Figure 1: (a) Schematic of SHG in a periodically poled TFLN. The red pulse at the fundamental frequency ( $\omega$ ) drives a dispersion-engineered nanophotonic waveguide designed for quasi-phase matched SHG, producing the purple pulse at second-harmonic ( $2\omega$ ). (b) Cross-sectional view of the device, where the top width ( $W_{\text{top}}$ ) and etch depth ( $H_{\text{etch}}$ ) indicate waveguide geometry parameters.

The design space for integrated nonlinear devices (e.g., geometry, poling, and driving conditions for second-harmonic generation) is high-dimensional. Accurate modeling typically relies on physics-based solvers such as FDTD/FEM or split-step Fourier methods, which can become computational bottlenecks that hinder large-scale parameter sweeps and inverse design [13]. To alleviate this, machine learning has been increasingly adopted to accelerate simulations by learning relationships directly from data [14, 15]. However, conventional neural networks learn mappings between finite-dimensional vectors, which may overfit to training discretizations and limit out-of-distribution generalization in systems more naturally described by functions. Neural operators address this limitation by learning mappings between functions, e.g., solution operators of differential equations [15, 16]. They offer universal approximation results at the operator level [17], support queries at arbitrary resolutions, and promote discretization-agnostic generalization. Neural operators also integrate naturally with physics-informed training to reduce data requirements [18–20].

Building on these capabilities, we propose an operator-learning framework for strongly nonlinear guided-wave optics in integrated TFLN waveguides. Specifically, we employ Fourier Neural Operators (FNOs) [21] to model nonlinear propagation and to enable fast, differentiable simulation of second-harmonic generation. By achieving speedups up to  $\sim 550,000\times$  compared to the solver while maintaining high fidelity to the solver ( $L^2$  relative error around 4%), the proposed framework overcomes long-standing barriers to large-scale design exploration, enabling millions of evaluations and the generation of extensive datasets. Through the differentiability of the FNO, gradient-based inverse design becomes possible, over geometry, poling period mismatch, and pump conditions. In this paper, we leverage these new possibilities to optimize for designs with high second-harmonic conversion efficiency. By combining resolution-agnostic surrogates with photonic design priors (dispersion engineering, phase matching, and waveguide loss), the methodology scales to large design spaces and accelerates discovery of high-performance nonlinear photonic devices.

## 2 Nonlinear Propagation in TFLN Waveguides

**TFLN Waveguides Design.** Figure 1 illustrates the class of nonlinear waveguides under study. A pump pulse of duration 35 fs at central angular frequency  $\omega_0$  (corresponding to a wavelength of 2090 nm) propagates along a 5-mm-long waveguide oriented along the crystalline  $y$ -axis of X-cut TFLN on a silica substrate. The thin-film thickness is fixed at 700 nm and only the fundamental transverse electric mode is considered. Quasi-phase matching for the  $\chi^{(2)}$  nonlinearity is achieved by periodic poling along the propagation direction. The design variables are the waveguide etch depth  $H_{\text{etch}}$ , the top width  $W_{\text{top}}$ , the poling period  $\Lambda$ , and the pump pulse energy  $E_{\text{in}}$ . Varying the top width and etch depth modifies the dispersion and confinement of the guided mode while tuning the poling period controls the phase mismatch between fundamental and second-harmonic waves. We focus on periodically poled TFLN ridge waveguides, a device architecture that has become a leading candidate for high-efficiency second-harmonic generation (SHG). Such waveguides combine subwavelength modal confinement with low-loss integration, enabling strong quadratic nonlinear interactions over centimeter-scale propagation lengths. Optimizing their geometry and poling parameters directly impacts the conversion efficiency and bandwidth of SHG, making them highly relevant for compact sources of visible and near-infrared light. These devices find wide application in frequency conversion for telecommunications, frequency-comb stabilization, quantum photonics, and precision metrology. We construct a dataset of device geometries, poling conditions, pump parameters, and input optical waveforms, together with the corresponding output complex optical fields. From this optical field, we can later derive SHG metrics of interest. Details on the dataset generation are given in Appendix B.

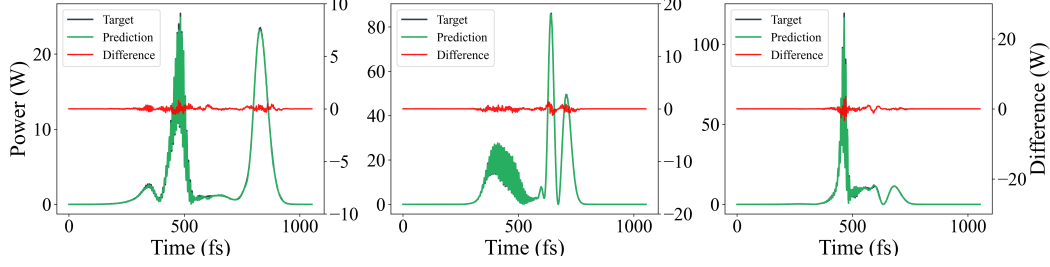


Figure 2: Examples of temporal profiles obtained from FNO predictions, against the reference ones.

**Nonlinear Propagation using a Fourier Neural Operator.** To efficiently simulate nonlinear propagation in TFLN waveguides, we employ the Fourier Neural Operator (FNO) as a surrogate forward model. See Appendices A.1 and A.2 for a presentation of the FNO architecture and the precise hyperparameters used. The input scalar design parameters (i.e. the top width, etch depth, and poling period mismatch of the waveguide) are encoded as amplitudes of sinusoidal signals that are stacked with the input complex optical waveform. Together, these are passed through the FNO to generate a prediction for the complex optical waveform at the end of the waveguide. After training on reference simulations, the FNO achieves substantial speedups of up to 550,000 $\times$  while maintaining high physical fidelity with a relative  $L^2$  error around 4% on previously unseen designs. Illustrative examples in Figures 2 and 6 further demonstrate that the FNO predictions are nearly indistinguishable from the reference simulations, highlighting the high reliability of the neural operator. A detailed analysis of runtime performance is also provided in Appendix A.4.

### 3 Inverse Design of Nonlinear Photonic Devices

The speedup achieved by the FNO removes long-standing barriers to large-scale design exploration, making it possible to perform millions of evaluations in a few minutes on a single GPU as opposed to years on a single CPU. One can now densely explore the design parameter space and evaluate the corresponding output waveforms to identify the best designs. Alternatively, the inference speed and differentiability of the FNO enable gradient-based optimization, typically unavailable in conventional solvers due to their slow forward evaluations and lack of differentiability with respect to design parameters. In this framework, the formulation of an appropriate loss function is especially important yet nontrivial: poorly chosen objectives can create artificial local minima that impede convergence to solutions with the intended properties. To ensure effective optimization, the loss must emphasize the relevant physical features and encode the essential design targets, while the parameterization, loss, and forward model must all remain differentiable with respect to the design variables.

As a demonstration of how the trained FNO can facilitate nonlinear photonic device design, we consider the inverse design of devices targeting high second-harmonic conversion efficiency in this paper. We carry out an optimization task aimed at maximizing the second-harmonic conversion efficiency by jointly tuning the waveguide geometry, the poling period mismatch, and the pump pulse energy of the input pulse. The overall inverse design architecture is illustrated in Figure 3, where a gradient-based optimization framework is employed. At each iteration, the four design parameters are initialized from random seeds and mapped into their feasible ranges using a sigmoid function to ensure valid inputs within the training distribution of the FNO. More precisely, the parameter ranges considered are  $W_{\text{top}} \in [1000 \text{ nm}, 3000 \text{ nm}]$ ,  $H_{\text{etch}} \in [120 \text{ nm}, 400 \text{ nm}]$ , poling period mismatch  $\Delta\Lambda \in [-50 \text{ nm}, 50 \text{ nm}]$ , and pump pulse energy  $E_{\text{in}} \in [3 \text{ pJ}, 7 \text{ pJ}]$ . The optimization is driven by the Adam optimizer with a learning rate of 0.05. The objective is defined as the ratio between the integrated spectral power within the SHG frequency band and the total spectral power. Specifically, the SHG band is taken to be centered at 286 THz with a bandwidth of 100 THz (i.e., a  $\pm 50$  THz window).

Across 1000 random initializations, the gradient-based optimization was run for 100 iterations per sample. The results demonstrate that 75% of the parameter combinations exceeded 85% second-harmonic conversion efficiency after optimization, highlighting the capacity of the FNO-driven framework to rapidly identify high-performance SHG device configurations.

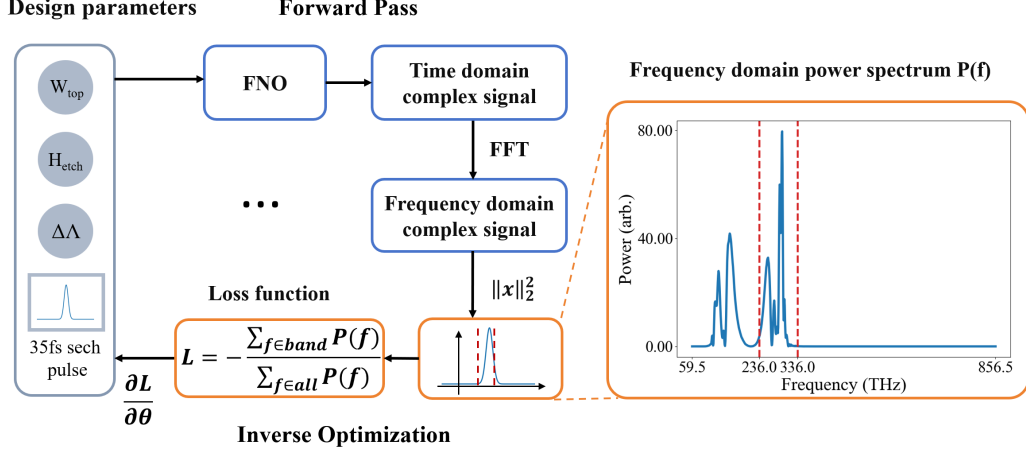


Figure 3: The overall inverse design pipeline. The design parameters include the waveguide top width ( $W_{\text{top}}$ ), etch depth ( $H_{\text{etch}}$ ), poling period mismatch ( $\Delta\Lambda$ ), and pump pulse energy. The FNO surrogate forward model predicts the complex time-domain signal, which is transformed into the frequency domain by FFT, followed by a modulus squared operation to obtain the power spectrum  $P(f)$ . The loss function is defined as the ratio of the integrated spectral power within the SHG frequency band to the total spectral power. The inset shows a representative frequency-domain power spectrum, with the SHG band highlighted between 236 THz and 336 THz.

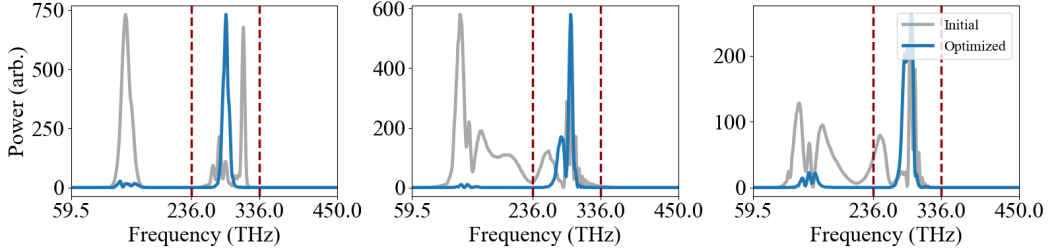


Figure 4: Examples of second-harmonic conversion efficiency optimization tasks. The light gray curves represent the model output under randomly initialized parameters, while the blue curves show the output after gradient-based optimization. The two red dashed lines indicate the frequency band used to evaluate and optimize the second-harmonic conversion efficiency.

## 4 Discussion

We proposed leveraging FNOs to simulate strongly nonlinear guided-wave optics in integrated TFLN waveguides for second-harmonic generation. By achieving speedups up to  $\sim 550,000\times$  compared to the solver while maintaining high accuracy, the proposed framework enables extensive sampling of photonic devices. The differentiability of the FNO further enables gradient-based optimization, which we leveraged to identify designs with high second-harmonic conversion efficiency.

The same framework could be employed to systematically tune and propose novel promising photonic devices through accessible controls, for instance for spectral translation, broadening, and pulse shaping. The speedup provided by the FNO allows rapid exploration of large design spaces, and when coupled with experiments, can expand the range of realizable nonlinear integrated photonic devices. Further improvements include enhancing efficiency with physics-informed constraints, enabling additional inputs (e.g., arbitrary pulse profiles) without greatly increasing model size, and integrating more advanced mode solvers to provide accurate dispersion and handle higher-order modes.

## References

- [1] Valdas Pasiskevicius, Shunhua Wang, Jens A. Tellefsen, Fredrik Laurell, and Håkan Karlsson. Efficient nd:yag laser frequency doubling with periodically poled ktp. *Appl. Opt.*, 37(30): 7116–7119, Oct 1998. doi: 10.1364/AO.37.007116. URL <https://opg.optica.org/ao/abstract.cfm?URI=ao-37-30-7116>.
- [2] Tingting Yu, Jianan Fang, Kun Huang, and Heping Zeng. Widely tunable mid-infrared fiber-feedback optical parametric oscillator. *Photon. Res.*, 12(10):2123–2129, Oct 2024. doi: 10.1364/PRJ.531271. URL <https://opg.optica.org/prj/abstract.cfm?URI=prj-12-10-2123>.
- [3] T. Herr, V. Brasch, J. D. Jost, C. Y. Wang, N. M. Kondratiev, M. L. Gorodetsky, and T. J. Kippenberg. Temporal solitons in optical microresonators. *Nature Photonics*, 8(2):145–152, Feb 2014. ISSN 1749-4893. doi: 10.1038/nphoton.2013.343. URL <https://doi.org/10.1038/nphoton.2013.343>.
- [4] Xu Yi, Qi-Fan Yang, Ki Youl Yang, Myoung-Gyun Suh, and Kerry Vahala. Soliton frequency comb at microwave rates in a high-q silica microresonator. *Optica*, 2(12):1078–1085, Dec 2015. doi: 10.1364/OPTICA.2.001078. URL <https://opg.optica.org/optica/abstract.cfm?URI=optica-2-12-1078>.
- [5] Yunxiang Song, Yaowen Hu, Xinrui Zhu, Kiyoul Yang, and Marko Lončar. Octave-spanning kerr soliton frequency combs in dispersion- and dissipation-engineered lithium niobate microresonators. *Light: Science & Applications*, 13(1):225, Sep 2024. ISSN 2047-7538. doi: 10.1038/s41377-024-01546-7. URL <https://doi.org/10.1038/s41377-024-01546-7>.
- [6] Christian Rosenberg Petersen, Uffe Møller, Irnis Kubat, Binbin Zhou, Sune Dupont, Jacob Ramsay, Trevor Benson, Slawomir Sujecki, Nabil Abdel-Moneim, Zhuoqi Tang, David Furniss, Angela Seddon, and Ole Bang. Mid-infrared supercontinuum covering the 1.4–13.3 $\mu$ m molecular fingerprint region using ultra-high na chalcogenide step-index fibre. *Nature Photonics*, 8(11):830–834, Nov 2014. ISSN 1749-4893. doi: 10.1038/nphoton.2014.213. URL <https://doi.org/10.1038/nphoton.2014.213>.
- [7] Camille-Sophie Brès, Alberto Della Torre, Davide Grassani, Victor Brasch, Christian Grillet, and Christelle Monat. Supercontinuum in integrated photonics: generation, applications, challenges, and perspectives. *Nanophotonics*, 12(7):1199–1244, 2023. doi: doi:10.1515/nanoph-2022-0749. URL <https://doi.org/10.1515/nanoph-2022-0749>.
- [8] Myoung-Gyun Suh, Xu Yi, Yu-Hung Lai, S. Leifer, Ivan S. Grudinin, G. Vasisht, Emily C. Martin, Michael P. Fitzgerald, G. Doppmann, J. Wang, D. Mawet, Scott B. Papp, Scott A. Diddams, C. Beichman, and Kerry Vahala. Searching for exoplanets using a microresonator astrocomb. *Nature Photonics*, 13(1):25–30, Jan 2019. ISSN 1749-4893. doi: 10.1038/s41566-018-0312-3. URL <https://doi.org/10.1038/s41566-018-0312-3>.
- [9] Scott B. Papp, Katja Beha, Pascal Del’Haye, Franklyn Quinlan, Hansuek Lee, Kerry J. Vahala, and Scott A. Diddams. Microresonator frequency comb optical clock. *Optica*, 1(1):10–14, Jul 2014. doi: 10.1364/OPTICA.1.000010. URL <https://opg.optica.org/optica/abstract.cfm?URI=optica-1-1-10>.
- [10] Pablo Marin-Palomo, Juned N. Kemal, Maxim Karpov, Arne Kordts, Joerg Pfeifle, Martin H. P. Pfeiffer, Philipp Trocha, Stefan Wolf, Victor Brasch, Miles H. Anderson, Ralf Rosenberger, Kovendhan Vijayan, Wolfgang Freude, Tobias J. Kippenberg, and Christian Koos. Microresonator-based solitons for massively parallel coherent optical communications. *Nature*, 546(7657):274–279, Jun 2017. ISSN 1476-4687. doi: 10.1038/nature22387. URL <https://doi.org/10.1038/nature22387>.
- [11] Yifan Qi and Yang Li. Integrated lithium niobate photonics. *Nanophotonics*, 9(6):1287–1320, 2020.
- [12] Di Zhu, Linbo Shao, Mengjie Yu, Rebecca Cheng, Boris Desiatov, C\_J Xin, Yaowen Hu, Jeffrey Holzgrafe, Soumya Ghosh, Amirhassan Shams-Ansari, et al. Integrated photonics on thin-film lithium niobate. *Advances in Optics and Photonics*, 13(2):242–352, 2021.

- [13] Yuheng Chen, Alexander Montes McNeil, Taehyuk Park, Blake A. Wilson, Vaishnavi Iyer, Michael Bezick, Jae-Ik Choi, Rohan Ojha, Pravin Mahendran, Daksh Kumar Singh, Geetika Chitturi, Peigang Chen, Trang Do, Alexander V. Kildishev, Vladimir M. Shalaev, Michael Moebius, Wenshan Cai, Yongmin Liu, and Alexandra Boltasseva. Machine-learning-assisted photonic device development: a multiscale approach from theory to characterization. *Nanophotonics*, July 2025. ISSN 2192-8614. doi: 10.1515/nanoph-2025-0049. URL <http://dx.doi.org/10.1515/nanoph-2025-0049>.
- [14] Steven L. Brunton and J. Nathan Kutz. Machine learning for partial differential equations, 2023. URL <https://arxiv.org/abs/2303.17078>.
- [15] Kamyar Azizzadenesheli, Nikola Kovachki, Zongyi Li, Miguel Liu-Schiaffini, Jean Kossaifi, and Anima Anandkumar. Neural operators for accelerating scientific simulations and design. *Nature Reviews Physics*, pages 1–9, 2024.
- [16] Julius Berner, Miguel Liu-Schiaffini, Jean Kossaifi, Valentin Duruisseaux, Boris Bonev, Kamyar Azizzadenesheli, and Anima Anandkumar. Principled approaches for extending neural architectures to function spaces for operator learning, 2025. URL <https://arxiv.org/abs/2506.10973>.
- [17] Nikola Kovachki, Samuel Lanthaler, and Siddhartha Mishra. On universal approximation and error bounds for Fourier neural operators. *J. Mach. Learn. Res.*, 22(1), 2021. ISSN 1532-4435.
- [18] Zongyi Li, Hongkai Zheng, Nikola Kovachki, David Jin, Haoxuan Chen, Burigede Liu, Kamyar Azizzadenesheli, and Anima Anandkumar. Physics-informed neural operator for learning partial differential equations. *ACM/JMS Journal of Data Science*, 1(3):1–27, 2024.
- [19] Ryan Y. Lin, Julius Berner, Valentin Duruisseaux, David Pitt, Daniel Leibovici, Jean Kossaifi, Kamyar Azizzadenesheli, and Anima Anandkumar. Enabling automatic differentiation with mollified graph neural operators, 2025.
- [20] Adarsh Ganeshram, Haydn Maust, Valentin Duruisseaux, Zongyi Li, Yixuan Wang, Daniel Leibovici, Oscar Bruno, Thomas Hou, and Anima Anandkumar. Fc-pino: High precision physics-informed neural operators via fourier continuation, 2025.
- [21] Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895*, 2020.
- [22] Jean Kossaifi, Nikola Kovachki, Zongyi Li, David Pitt, Miguel Liu-Schiaffini, Valentin Duruisseaux, Robert Joseph George, Boris Bonev, Kamyar Azizzadenesheli, Julius Berner, and Anima Anandkumar. A library for learning neural operators, 2025. URL <https://arxiv.org/abs/2412.10354>.
- [23] Luis Ledezma. Snow: Simulator for nonlinear optical waveguides. <https://github.com/ledezmaluism/snow>, 2025. [Online; accessed 21-July-2025].

## A The Fourier Neural Operator (FNO) for Forward Modeling

### A.1 The FNO Architecture

Neural operators compose linear integral operators  $\mathcal{K}$  with pointwise non-linear activation functions  $\sigma$  to approximate highly non-linear operators. More precisely, we define the neural operator

$$\mathcal{G}_\theta := \mathcal{Q} \circ \sigma(W_L + \mathcal{K}_L + b_L) \circ \cdots \circ \sigma(W_1 + \mathcal{K}_1 + b_1) \circ \mathcal{P} \quad (1)$$

where  $\mathcal{P}$ ,  $\mathcal{Q}$  are the pointwise neural networks that encode the lower dimension function into higher dimensional space and vice versa. The model stacks  $L$  layers of  $\sigma(W_l + \mathcal{K}_l + b_l)$  where  $W_l$  are pointwise linear operators (matrices),  $\mathcal{K}_l$  are integral kernel operators,  $b_l$  are bias terms, and  $\sigma$  are fixed activation functions. The parameters  $\theta$  consists of all the parameters in  $\mathcal{P}$ ,  $\mathcal{Q}$ ,  $W_l$ ,  $\mathcal{K}_l$ ,  $b_l$ . Kossaifi et al. [22] maintain a comprehensive open-source library for learning neural operators in PyTorch, which serves as the foundation for our implementation.

Throughout our numerical experiments, we employ Fourier neural operators to approximate the solution operators of PDEs. A **Fourier neural operator (FNO)** [21] is a neural operator using Fourier integral operator layers, which are defined via

$$(\mathcal{K}(\phi)v_t)(x) = \mathcal{F}^{-1} \left( R_\phi \cdot (\mathcal{F}v_t) \right)(x) \quad (2)$$

where  $R_\phi$  is the Fourier transform of a periodic function  $\kappa$  parameterized by  $\phi$ . On a uniform mesh, the Fourier transform  $\mathcal{F}$  can be implemented using the fast Fourier transform (FFT). Note that Fourier neural operators are differentiable. Here is a depiction of the Fourier Neural Operator:

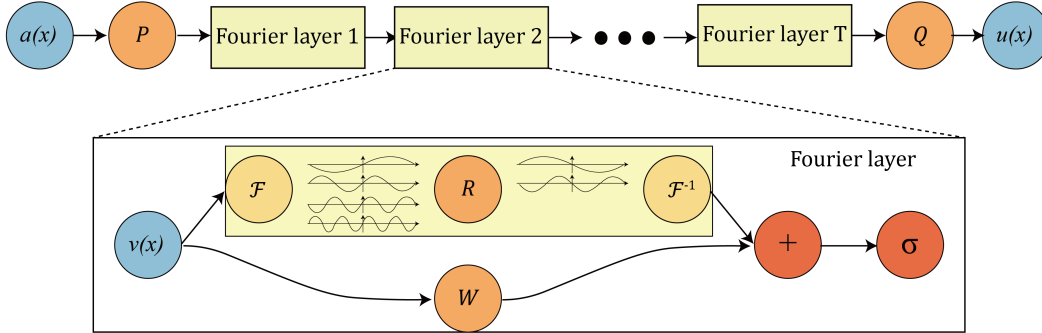


Figure 5: The FNO architecture (extracted from [21]).

### A.2 FNO Hyperparameters Used

The neural operator used as a surrogate for the SNOW solver in this paper is a 1D FNO with 18 layers, each with 22 hidden channels and 256 Fourier modes. This FNO possesses 4.5M trainable parameters and was trained in PyTorch for 1200 epochs to minimize the absolute  $L^2$  error using the Adam optimizer with learning rate 0.001, and the ReduceLROnPlateau scheduler with factor 0.4 and patience 6.

The input scalar parameters (i.e. the top width, etch depth, and poling period mismatch of the waveguide) are encoded as amplitudes of sinusoidal signals that are stacked with the input complex optical waveform. Together, these are passed through the FNO to generate a prediction for the complex optical waveform at the end of the waveguide.

### A.3 FNO Predictive Performance

The trained FNO achieves a relative  $L^2$  error of 4.39% on the test set (3.63% on the training set). We display below two examples of inputs, FNO predictions, and reference solutions for comparison.

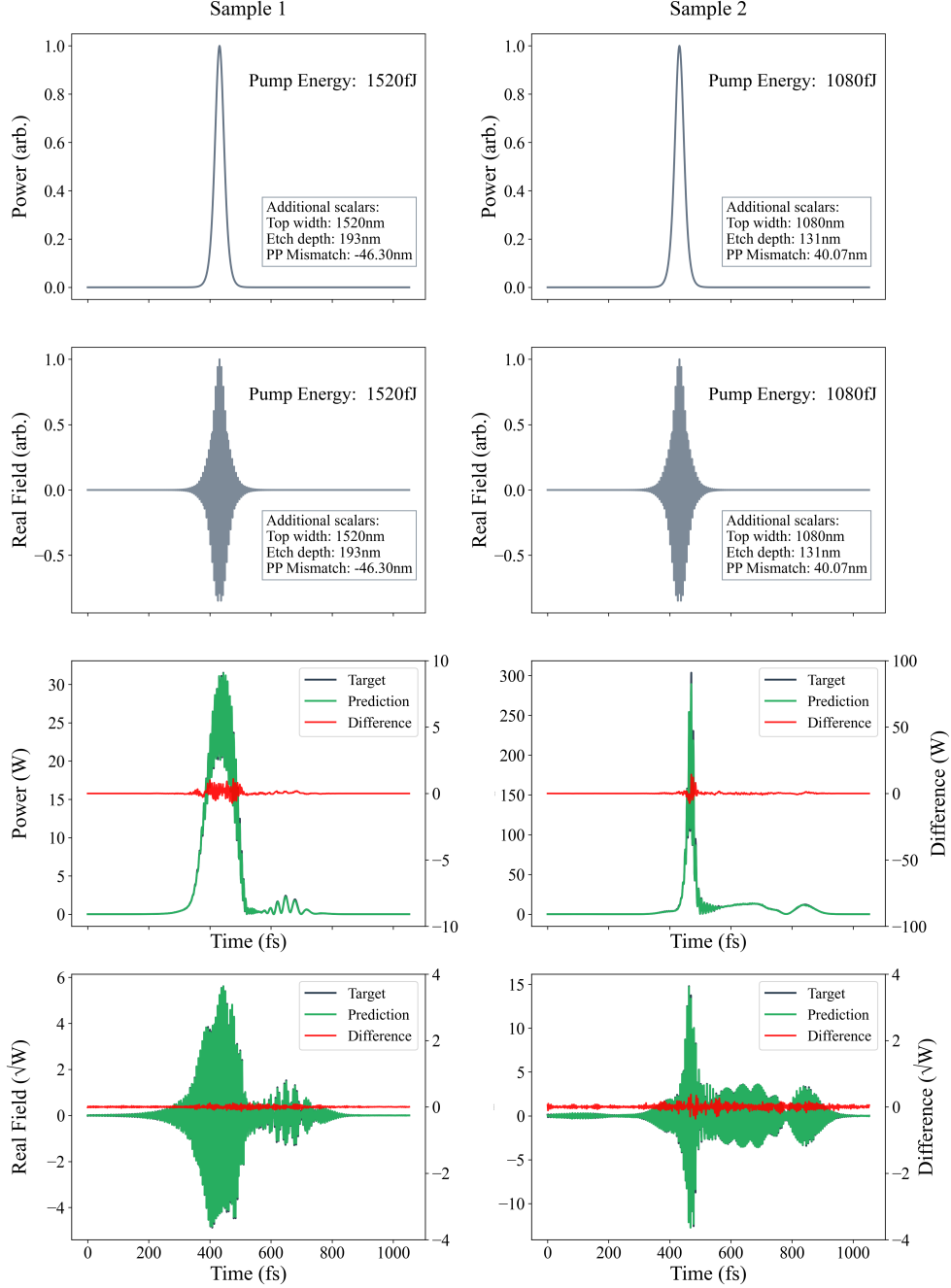


Figure 6: Examples of FNO predictions on previously unseen design parameters. The top two rows show the model inputs, composed of a complex optical waveform (determined by the pump pulse energy) and input scalar design parameters (the top width, etch depth, and poling period mismatch) of the waveguide. The bottom two rows show the FNO predictions against the reference optical waveforms. We display both the real parts and temporal profiles of the waveforms. We also display in red the mismatch between the predictions and reference simulations (on a separate scale since these are small, relatively).



#### A.4 Speedups enabled by the FNO

The trained FNO generates predictions significantly faster than the reference numerical solver. We record the time necessary to generate predictions on a workstation equipped with a single NVIDIA RTX 4090 GPU (24GB VRAM), an AMD Ryzen 9 7900X CPU, and 64GB of system RAM. The numerical solver only generates a single prediction at a time, and takes roughly 52 seconds per prediction (averaged over 2,000 different samples, with a standard deviation of 16 seconds). On the other hand, doing one inference at a time with the trained FNO (i.e. batch size of 1), takes roughly 6.7 milliseconds (averaged over 10,000 different samples), i.e. it gives a speedup of  $\sim 7,700\times$  compared to the solver. In addition, the trained FNO can generate multiple simultaneous predictions on a single GPU by making predictions in parallel. In particular, with a batch size of 92 (which we found to be optimal for speed), the trained FNO takes roughly 0.095 milliseconds per prediction (averaged over 100,000 different samples), i.e. a speedup of  $\sim 550,000\times$  compared to the solver. Furthermore, unlike the numerical solver, the FNO takes the same amount of time for inference regardless of the inputs. These speedups should be interpreted with caution, as the solver was not optimized for speed and could be made substantially faster. Similarly, smaller and more efficient models with comparable accuracy might have been obtained via further hyperparameter tuning and training.

## B Data Generation

### B.1 SNOW: Simulator for Nonlinear Optical Waveguides

SNOW (Simulator for Nonlinear Optical Waveguides) is an open-source computational toolbox designed for the modeling and design of integrated nonlinear photonic devices. The package was originally developed to simulate nonlinear optical waveguides in TFLN (thin-film lithium niobate), but its framework is extensible to other materials and photonic platforms. SNOW provides a self-contained set of tools that cover the key aspects of nonlinear integrated optics, including material dispersion database, waveguide mode simulation and nonlinear pulse propagation.

### B.2 Nonlinear Envelope Equation

The propagation of the broadband electric field envelope  $U(z, t)$  along the waveguide is described by the nonlinear envelope equation (NEE). Defining the Fourier transform of  $U$  by  $\tilde{U}(z, \omega) = \mathcal{F}\{U(z, t)\}$  and the envelope frequency offset  $\Omega = \omega - \omega_0$ , the NEE in the frequency domain reads

$$\frac{\partial}{\partial z} \tilde{U} = -i(\beta - \beta_0 - \frac{\Omega}{v_0} - i\frac{\alpha}{2})\tilde{U} - \frac{i\omega\epsilon_0 X_0}{8P} \mathcal{F}_\Omega \{ \bar{d}U^2 e^{i\phi} + 2\bar{d}^*|U|^2 e^{-i\phi} \}. \quad (3)$$

where  $\beta(\omega)$  is the wavenumber of the guided mode,  $\beta_0 = \beta(\omega_0)$  is the reference phase constant,  $v_0$  is the group velocity at  $\omega_0$ , and  $\alpha(\omega)$  is a frequency-dependent loss. The nonlinear coefficient  $\chi_0$  incorporates the effective  $\chi^{(2)}$  susceptibility and mode overlap;  $P$  is a normalisation constant taken as 1 W. The poling function  $d(z)$  alternates sign with a period  $\Lambda$  to realise quasi-phase matching, and  $\phi(z, t) = \omega_0 t - [\beta_0 - \omega_0/v_{\text{ref}}]z$  accounts for the phase accrued in the moving frame. The operators  $\mathcal{F}_\Omega$  and  $\mathcal{F}_\Omega^{-1}$  denote forward and inverse Fourier transforms with respect to  $\Omega$ .

The waveguide dispersion and loss are obtained by solving for the effective index of the fundamental TE mode using a finite-element eigenmode solver at a temperature of 25°C. The phase matching condition for SHG is

$$\Lambda_{\text{pm}} = \frac{2\pi}{|\beta(2\omega_0) - 2\beta(\omega_0)|}, \quad (4)$$

and the implemented poling period is  $\Lambda = \Lambda_{\text{pm}} + \Delta\Lambda$ . The poling function is modelled as the sign of a sinusoid,

$$d(z) = \text{sgn}[\sin(2\pi z/\Lambda)], \quad z \in [0, L_{\text{wg}}], \quad (5)$$

which inverts the nonlinear coefficient every half period.

### B.3 Dataset Generation

Training and testing data are obtained by numerically solving Equation (3) with the Fourier split-step algorithm implemented in the SNOW package [23]. The propagation is advanced along  $z$  in discrete steps: at each step, the nonlinear contribution is calculated using the Runge–Kutta–Fehlberg scheme, after which a frequency-domain operator is applied to capture both dispersion and loss during the linear evolution.

The dataset contains 100,000 distinct samples. For each sample  $i$ , the parameters include the input pump pulse energy  $E_{\text{in},i}$ , the poling period  $\Lambda_i = \Lambda_{\text{pm},i} + \Delta\Lambda_i$ , the waveguide top width  $W_{\text{top},i}$ , and the etch depth  $H_{\text{etch},i}$ . The pump pulse energy  $E_{\text{in},i}$  and the poling period mismatch  $\Delta\Lambda_i$  are drawn from uniform distributions, with  $\Delta\Lambda_i \in \mathcal{U}[-50 \text{ nm}, 50 \text{ nm}]$  and  $E_{\text{in},i} \in \mathcal{U}[10 \text{ fJ}, 10,000 \text{ fJ}]$ . The structural parameters are sampled over a finite set of discrete pairs:  $W_{\text{top},i}$  and  $H_{\text{etch},i}$  take values only from 51 evenly spaced points within  $[1000 \text{ nm}, 3000 \text{ nm}]$  and  $[120 \text{ nm}, 400 \text{ nm}]$ , respectively. These parameter ranges are consistent with practical fabrication limits and tolerances of the TFLN platform.