
A Modeling Framework and LSTM Nonlinear Dynamics Predictor for Accelerating Data Generation in Ultrafast Optics and Laser System Design

Jack Hirschman^{1,2,*}, Minyang Wang^{3,4}, Erfan Abedi⁵, Sergio Carbajo^{2,5,6}

¹ Department of Applied Physics, Stanford University, Stanford, CA 94305

² SLAC National Accelerator Laboratory, Menlo Park, CA 94025

³ Department of Mathematics, UCLA, Los Angeles, CA 90095

⁴ Department of Statistics and Data Science, UCLA, Los Angeles, CA 90095

⁵ Department of ECE, UCLA, Los Angeles, CA 90095

⁶ Physics and Astronomy Department, UCLA, Los Angeles, CA 90095

⁷ California NanoSystems Institute, Los Angeles, CA 90095

*jhirschm@stanford.edu

Abstract

Ultrafast optics and high-power laser systems are the backbone of nearly every major industry from semiconductor manufacturing and telecommunications to advanced medical procedures and next-generation energy and defense solutions. With the increasing integration of machine learning (ML) into laser system design, there is a growing demand for efficient data generation to aid experiment design, perform in-situ optimization, and improve the efficacy of system-wide digital twins. Toward this end, we present a novel start-to-end (S2E) modeling framework for complex laser systems that can be tailored to specific applications for large data generation. The simulation output can then be used in a wide variety of ML tasks from predicting pulse propagation behavior and laser system controls optimization to diagnostic characteristics extraction. However, the models, by necessity, can involve solving several coupled partial and ordinary differential equations including the Nonlinear Schrödinger Equation (NLSE), gain dynamics, and Maxwell's equations, depending on the framework configuration. These complex cascaded nonlinear systems of equations become a significant time bottleneck in generating large quantities of data. To demonstrate a broad impact application of ML enhancing ultrafast optics simulations, we aim our studies on using long short-term memory (LSTM) networks to replace solving the NLSE for sum-frequency generation, a nonlinear optical process involving the interaction of three fields. We show how these models can provide significant speed-up for large data generation and can ultimately enable an S2E framework to be applied broadly across applications in the ultrafast optics field. Finally, we discuss a metric for assessing the performance of these networks in the context of the optics domain.

1 Introduction

Since the development of chirped pulse amplification (CPA) of optical pulses and the subsequent rise of ultrafast optics [1], CPA systems have driven the rapid advancements in high power laser systems, ultrashort pulse lasers, and nonlinear optical (NLO) systems that underlie almost every major technology in the modern world. CPA and NLO systems are required for semiconductor manufacturing and precision machining [2, 3], telecommunications [4], biomedical imaging [5], fusion energy [6], particle accelerators [7, 8], a wide array of defense and security measures [9],

and many other sectors. The impending era of optics and photonics will focus on machine-assisted optimization, intelligent ground-up design, and inverse design, which will all take advantage of the current push for integrating machine learning (ML) [10, 11]. However, the considerable data requirement to navigate the complexities of cascaded components in laser systems, each unique, presents a significant challenge due to the notoriously time-consuming data collection process.

We have developed a start-to-end (S2E) software model for simulating laser systems involving CPA and NLO devices [12]. This model renders time and frequency-resolved electromagnetic fields alongside essential physical characteristics of energy, fluence, and spectral distribution, providing a substantial data source for ML models and applications. However, the data generation process is hindered by bottlenecks when solving coupled nonlinear equations, like the generalized Nonlinear Schrödinger Equation (NLSE). To overcome this, we employ a long short-term memory (LSTM) model, trained on a large dataset generated by our S2E framework, to replace the simulation blocks solving the NLSE. Here, we showcase how our framework can be used to generate critical, application-specific datasets; a tailored dataset; how we use such a dataset for developing an LSTM to further accelerate future optical system simulations; and an evaluation metric. This paper outlines the specific application area, simulation setup, data reduction and preprocessing methods, and investigates the LSTM’s performance in solving the NLSE for a three-field mixing process.

2 Background

Showcasing the robustness of the S2E framework requires demonstrating it on a sufficiently complex laser system. Here, we choose the drive laser system at SLAC National Accelerator Laboratory’s LCLS-II, the world’s most powerful X-ray free-electron laser (XFEL) [13]. This laser system involves linked CPA and NLO systems (fig. 1), specifically incorporating a mode-locked oscillator, a pulse shaper, a CPA regenerative amplifier, and NLO upconversion. The gain dynamics of the CPA system are modeled using the modified Franz-Nodvik equations [14] and the pulse shaper is modeled with a frequency-domain transfer function. For the NLO upconversion, we use dispersion-controlled nonlinear synthesis (DCNS), which employs noncollinear sum-frequency generation (SFG) by interacting two pulsed laser beams at a set angle to render temporally-shaped flat-top upconverted pulses [15]. The nonlinear process takes place in a medium with a nonlinear response—here, a crystal—and is governed by coupled equations for each of the three fields

$$\frac{dA_i}{dz} = \frac{2i \cdot d_{\text{eff}} \cdot \omega}{k_i c^2} \cdot A_j \cdot A_k \cdot e^{-i\Delta k z} \quad (1)$$

where z is the distance along the crystal, d_{eff} represents the nonlinearity tensor, i, j, k represent the three mixing fields, Δk represents phase matching, and ω is the frequency.

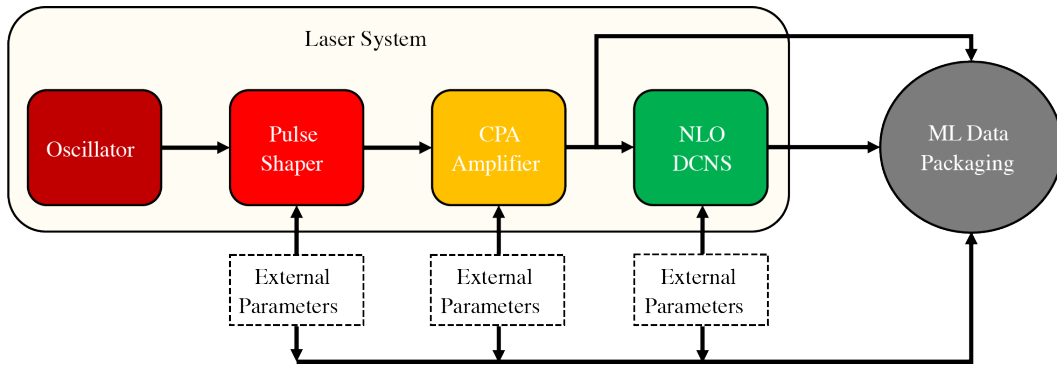


Figure 1: Laser system model incorporating initial oscillator, pulse shaper, amplifier, and upconversion process where each module has externally loaded parameters. Here, we package data from before and after the NLO DCNS upconversion process for use in an LSTM.

Fig. 2 illustrates the input/output structure of this crystal. The SFG process, specifically, involves the interaction of these two laser pulses, or fields, within this nonlinear medium [16] culminating in a third field produced at a frequency equivalent to the sum of the two input fields (labeled as output 2 or SFG). As the process is noncollinear, with the two input fields entering the medium at an angle,

two additional outputs emerge, corresponding to a process called second-harmonic generation (SHG) applied to each input independently (input 1 yields output 3 and input 3 yields output 1). For our work, these two input fields are derived from the same initial field but are subject to equal yet opposite second- and third-order phase modulation [15] in order to achieve a specific output pulse shape.

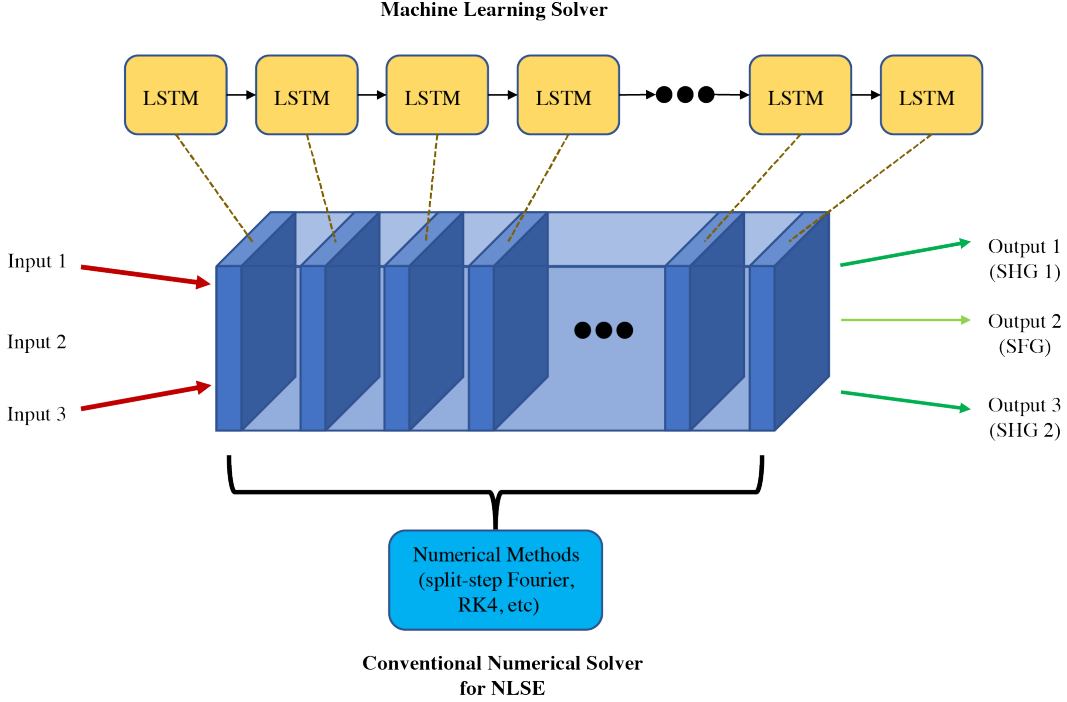


Figure 2: DCNS setup as a three-input, three-output system that propagates the fields from start to end where identical LSTM layers can replace each slice in the medium.

Simulating the entire series of cascaded processes in fig. 1 is essential for characterizing and exploring system-wide capabilities as well as performing setup optimizations based on large datasets generated from the cascaded models. The dataset generated from our S2E framework for this study scans through different pulse shaper parameters to inspect the downstream effects on the NLO DCNS process. These S2E simulations require accurate models for each sub-component. Traditionally, DCNS upconversion is simulated by solving the coupled equations numerically with split-step Fourier and fourth-order Runge-Kutta (RK4) methods, which discretize the crystal and propagate a solution through the medium [17, 18, 19]. Here, these computationally intensive procedures account for 1.85s of the 1.98s, or over 93%, for one iteration of the simulator, creating a significant bottleneck. This can be detrimental in cases where fast system exploration is required, such as in future research linking simulation and lab-collected data for controls. To overcome this, we propose reframing the problem as a three-input, three-output network with one input field (input 2 in fig. 2) set to identically zero. We then employ an LSTM network to simulate each slice in the crystal, effectively learning the propagation function within the medium, offering a more efficient alternative, and, simultaneously, showcasing one such use-case for the dataset we generated from the S2E framework.

Our approach is inspired by Salmela, et al. [20], who demonstrated the efficacy of LSTMs in predicting ultrafast dynamics in optical fiber simulations. While their study focused on a single input single output process, tracking only spectral intensity, it presented good alignment with numerical solutions to the NLSE and achieved a speedup factor of ~ 270 times for single-mode NLSE simulations and a speedup factor of ~ 11400 times for multi-mode NLSE simulations where, in both cases, the NLSE was run on CPU and the LSTM on GPU. The same group similarly explored feed-forward neural networks as a nonlinear dynamics integrator for the specific application of supercontinuum generation [21]. Others have since explored convolutional neural networks [22], various extensions with recurrent neural network architectures [23, 24], and physics-informed neural networks [25],

predominantly in wave guides. Our study extends this concept to a more complex system, addressing intensity and phase tracking in noncollinear sum-frequency generation in our DCNS method.

While the equations that govern our system and the one for Salmela, et al. [20] are similar, our system solves for 3 full fields—intensity and phase—as opposed to solely the intensity of a single field. Furthermore, we modulate the fields upstream of the DCNS process and keep the parameters of the DCNS NLO section stationary, exploring how the prior pulse shapes are propagated in free-space and the crystal. Our system incorporates the parameters for each component along the cascaded simulation chain, as shown in fig. 1. These parameters influence various aspects of the process, including the initial pulse shaper parameters altering input pulse shape, the CPA parameters altering the spectrum, and the DCNS internal parameters altering the propagated pulse (see Hirschman, et al. [12]). With this model, we generate a large data repository for an LSTM to learn the impact of spectral phase and amplitude shaping from the pulse shaper on NLO upconversion outcomes.

3 Data Generation and Processing

Initial dataset generation runs the S2E simulator using the standard numerical methods for each module. Since the NLO DCNS simulation uses a split-step Fourier method, the input data must have sufficient sampling in both the frequency and temporal domains to capture the short laser pulses’ electric field oscillations during the forward and inverse Fourier transforms. On the contrary, using an LSTM in the time or frequency domain relaxes this high sampling requirement since it can operate in just one domain, allowing for significant data reduction. For the preparation of the data for the LSTM, we transform the data from simulation through two rounds of preprocessing.

The first round downsamples and cuts vectors for one domain from all three fields in the simulation, packages the fields into one vector, and stores this data (see supplementary material). Table 1 shows the vector sizes from initial simulation, after downsampling, and after cutting.

Table 1: Vector sizes of three fields from initial simulation, downsampling, and cutting.

	SHG 1	SHG 2	SFG
Initial	32768	32768	32768
Downsampled	10589	10589	4681
Downsampled & Cut	1892	1892	348

Fig. 3a–c show these corresponding waveforms for one example input in the frequency domain, decomposing the full field into intensity (solid) and phase (dashed). The first row shows the SFG signal and the second and third rows show the SHG 1 and SHG 2 signals. To check that the downsampling and cutting of the frequency domain field does not adversely affect the represented information, we transform the data into the time domain. If the vector was cut or downsampled too much, the time domain signal would show significant variations from the true time domain signal. Fig. 3d–f show both the true time domain field and a version that is the direct inverse Fourier transform of the downsampled and cut frequency domain signal (labeled as direct). However, this direct inverse Fourier transform does not yield a reasonable signal because the initial vector lengths and domain spacings are chosen carefully for the Fourier transforms to yield physical results. Thus, we first resample and smooth the downsampled and cut field back onto the original domain prior to taking the inverse Fourier transform (labeled as resampled). This comparison of resampled and true time domain vectors provides a means of evaluating our data reduction (see supplementary material).

The second round of preprocessing concatenates the fields such that they are compatible with the LSTM model. We show two ways this can be done: 1) three field intensities normalized to energy followed by three field phases normalized by π along with the 3 corresponding energy values from the fields at the end (see fig. 4a) and 2) real portion of three fields followed by imaginary portion of three fields (see fig. 4c). Version 1 is a total of $2 \cdot 1892$ (SHG 1) + $2 \cdot 1892$ (SHG 2) + $2 \cdot 348$ (SFG) + 3 (energies) = 8267 elements long, and version 2 is $2 \cdot 1892$ (SHG 1) + $2 \cdot 1892$ (SHG 2) + $2 \cdot 348$ (SFG) = 8264 elements long. Due to the fluctuations present in the phase of version 1, we proceed in the main manuscript with version 2 processing (see supplementary for more information).

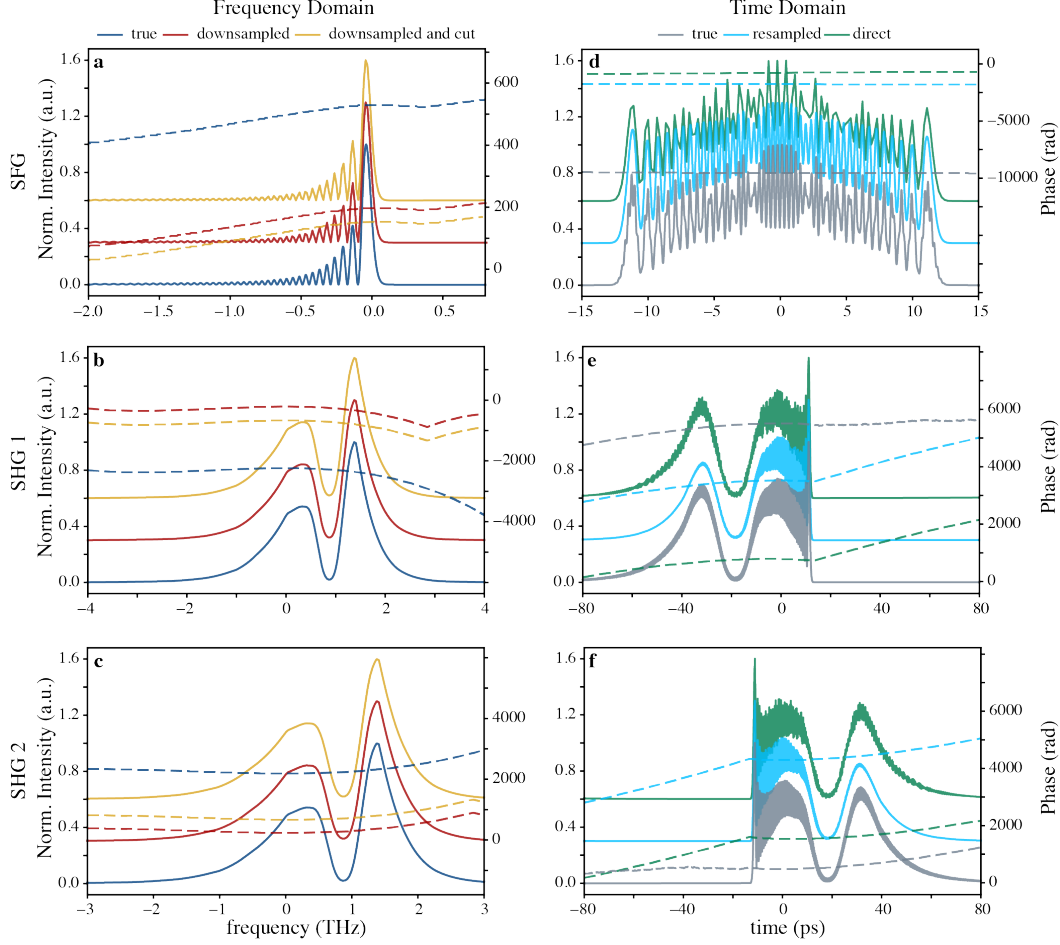


Figure 3: Shows the first step in the data preprocessing in the frequency domain (left column) for all fields (rows 1–3) as well as the time domain counterpart (right column). For all plots, the solid line is the normalized field intensity and the dashed line is the field phase.

Before feeding these vectors into the LSTM, we scale each input element to between 0 and 1 across all instances (see fig. 4b and fig. 4d for data arrangements and supplementary for details on rescaling).

4 LSTM Model & Results

While our primary goals are to present a data generation framework designed for machine learning in ultrafast optics and photonics and to provide a large dataset that explores pulse shaping effects on upconversion, our secondary goals are to demonstrate a targeted ML application using this dataset and provide an evaluation metric useful for the optics community. In particular, we want to explore an ML application that improves the ability to generate future data sets. Here, we use an LSTM network to replace the bottleneck of the simulation and evaluate the network based on accuracy and latency.

The LSTM model mimics a single slice in the nonlinear crystal where the behavior depends on previous steps. We use an LSTM with access to the 10 last spatial steps, where 10 was chosen from a balance of providing enough history—capturing information in the dynamics without over-emphasizing old behavior—while limiting total latency. The crystal has 100 spatial steps, so the simulation employs the split-step Fourier method iteratively 100 times and thus outputs a total of 101 slices. Therefore, each simulation result can be used to generate 100 instances of training data or validation data. Each instance is composed of an input array X consisting of the last 10 slices and an output array y representing the proceeding slice. In the initial 9 instances, fewer than 10 previous slices are available, so the input array is prepended by identical copies of the initial input to the crystal

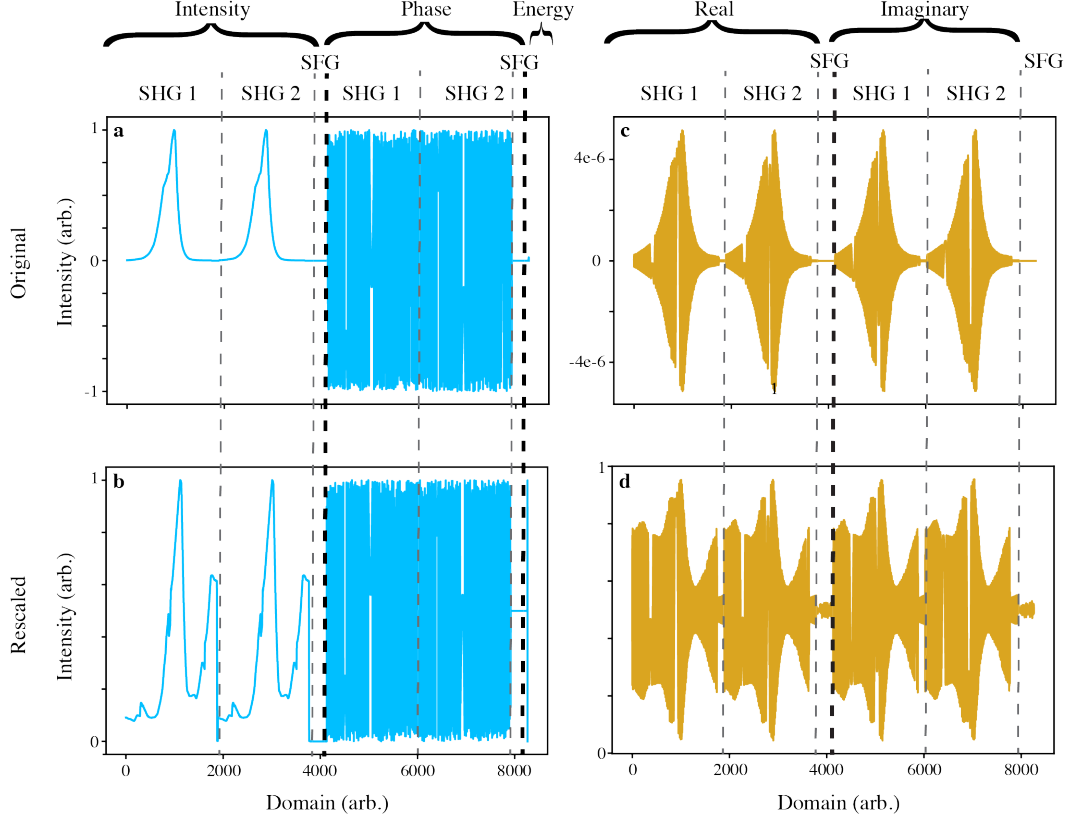


Figure 4: Shows one example demonstrating the second stage data preprocessing options. Version 1 (first column) arranges data into spectral intensity, phase, and energies. Version 2 (second column) splits data into real and imaginary portions of the field. The second row shows min-max normalized version of the first. All fields are in the frequency domain, but the exact domain points are arbitrary.

until the full input reaches a length of 10, with the subsequent slice as the output. The final shapes are (batch size, 10, 8264) for X and (batch size, 8264) for y , depending on the version of data used.

We then explored LSTM architectures of varying depth and size of LSTM and fully connected layers as well as activation functions. For the loss function, we use weighted mean squared error (wMSE)

$$\text{wMSE} = c_{\text{SFG}} \times \text{MSE}_{\text{SFG}} + c_{\text{SHG}} \times (\text{MSE}_{\text{SHG 1}} + \text{MSE}_{\text{SHG 2}}) \quad (2)$$

where c_{SFG} and c_{SHG} control the significance for the SFG and SHG outputs, respectively, and we use Adam as the optimizer with the default hyperparameters in PyTorch.

All models are trained on 890,000 instances, validated on 10,000 instances and tested on 90,000 instances. During the training process and hyperparameter search on models for small epochs, we found the combination of $c_{\text{SFG}} = 0.7$ and $c_{\text{SHG}} = 0.3$ works best for minimizing the training loss.

The best-performing model consists of 1 LSTM layer with hidden size 2048 and three linear layers with shapes (2048, 4096), (4096, 4096), and (4096, 8624) and activation ReLu, Tanh, and Sigmoid, respectively. We trained this model for 180 epochs for 160 hours on one NVIDIA A10G GPU hosted by Amazon Web Services resulting in a final training error of $2.05\text{E-}5$ and validation error of $2.03\text{E-}5$.

For inference, the LSTM loop is initiated with a “cold start” where the input sequence contains only the profile of pulses injected into the crystal, repeated 10 times. The LSTM model then runs 100 iterations for full propagation through the crystal using a sliding window mechanism, where, for each new prediction, the oldest slice in the input is discarded, and the latest model prediction is appended.

The key change for evaluating the test set is to focus on intensities of the field and allow for weighting of time and frequency domains. We calculate a normalized MSE (NMSE) error on the temporal intensity and spectral intensity profiles (see eq. 3) individually for SFG, SHG 1, and SHG 2, while all

training and training assessments are conducted in the frequency domain. NMSE is required to adjust for differences in magnitude between the time domain intensities and frequency domain intensities.

$$\text{wNMSE}_{\text{eval}} = c_t \times \text{NMSE}_{t,\text{SFG/SHG 1/SHG 2}} + c_f \times \text{NMSE}_{f,\text{SFG/SHG 1/SHG 2}} \quad (3)$$

where

$$\text{NMSE} = \frac{\sum_i (x_i - y_i)^2}{\sum_i (y_i)^2} \quad (4)$$

and where x_i and y_i are the prediction and truth intensities and c_t and c_f scale time and frequency contributions, respectively. For the generalized case, $c_t = c_f = 0.5$. This can be altered for applications that might emphasize spectrum more substantially, like spectroscopy experiments, or temporal intensity more substantially, like laser-photocathode interactions.

Table 2: Inference errors for both NMSE evaluation weighting schemes.

		$c_t, c_f = 1, 0$			$c_t, c_f = 0.5, 0.5$		
		SFG	SHG 1	SHG 2	SFG	SHG 1	SHG 2
Full Set Mean	Error	1.08E-1	2.70E-2	2.79E-2	1.03E-1	2.67E-2	2.67E-2
Minimal Hole Example	Error Percentile	8.68E-4	3.47E-4	3.00E-4	6.24E-4	4.92E-4	4.39E-4
		5th	5th	5th	5th	5th	5th
Large Hole Example	Error Percentile	3.54E-2	5.56E-2	5.41E-2	3.79E-2	7.22E-2	7.14E-2
		50th	90th	90th	55th	95th	95th

For our application of photoinjector lasers, we are primarily concerned with the temporal profile of the SFG pulse, so we select examples based on inspection of this error. We are also interested in how these errors compare with the even weighting between time and frequency. Table 2 shows the full set average errors across all test data for SFG, SHG 1, and SHG 2 for both the exclusive time domain and even weighting as well as the errors and associated percentiles for the minimal and large spectral hole examples shown in fig. 5. Specifically, fig. 5 shows the true and predicted results for SFG and SHG 1 for two examples from the test dataset—fig. 5a–d for one with only phase shaping and fig. 5e–h for one with amplitude and phase shaping. The first example prediction matches closely for both SFG and SHG 1 in both domains. The second example, which includes the amplitude shaping that produces a significant spectral hole in the SHG 1 spectral intensity, captures the SFG signal very closely and only misses portions of the frequency in the SHG 1 around large swings in amplitude. However, the error still seems large because generally the spectral amplitude shaping examples that contain a hole are more difficult for the network to learn, yet this example still yields a high quality output visually.

Moreover, the predictions have a significant speedup compared to numerical simulation. Table 3 shows timing analysis for running the trained LSTM network versus the traditional numerical method for various configurations of GPU and CPU setups, including single and multi-CPU/GPU systems. The baseline is the traditional numerical simulation, which can only be carried out on a single CPU sequentially. Running the LSTM with a batch size of 200 on 1 CPU does not yield a speedup over baseline. However, parallelizing across 8 CPUs does. Using one NVIDIA Tesla A100 with 40 GB of memory on the S3DF cluster at SLAC National Laboratory and with all data loaded into memory, we achieve a speedup of 252 over baseline numerical simulation (see supplementary information).

5 Conclusion

Merging ultrafast optics with machine learning and machine assisted design will bring about a new era of laser system development—an era focused on efficient development and deployment, real-time tunability, and inverse design. Start-to-end modeling frameworks play a crucial role in generating data for these designs and enable ML networks to have access to large quantities of simulation data or physical experiment data modulated by simulation. In this paper, we presented a dataset generated for the photoinjector laser system at SLAC that involves complex cascaded coupled nonlinear processes. The generated dataset can be used for a number of different ML applications, from learning the pulse shaper parameters used to create downstream output fields to pulse characteristic extraction and reconstruction. We target the data generation process itself, identifying and replacing the major

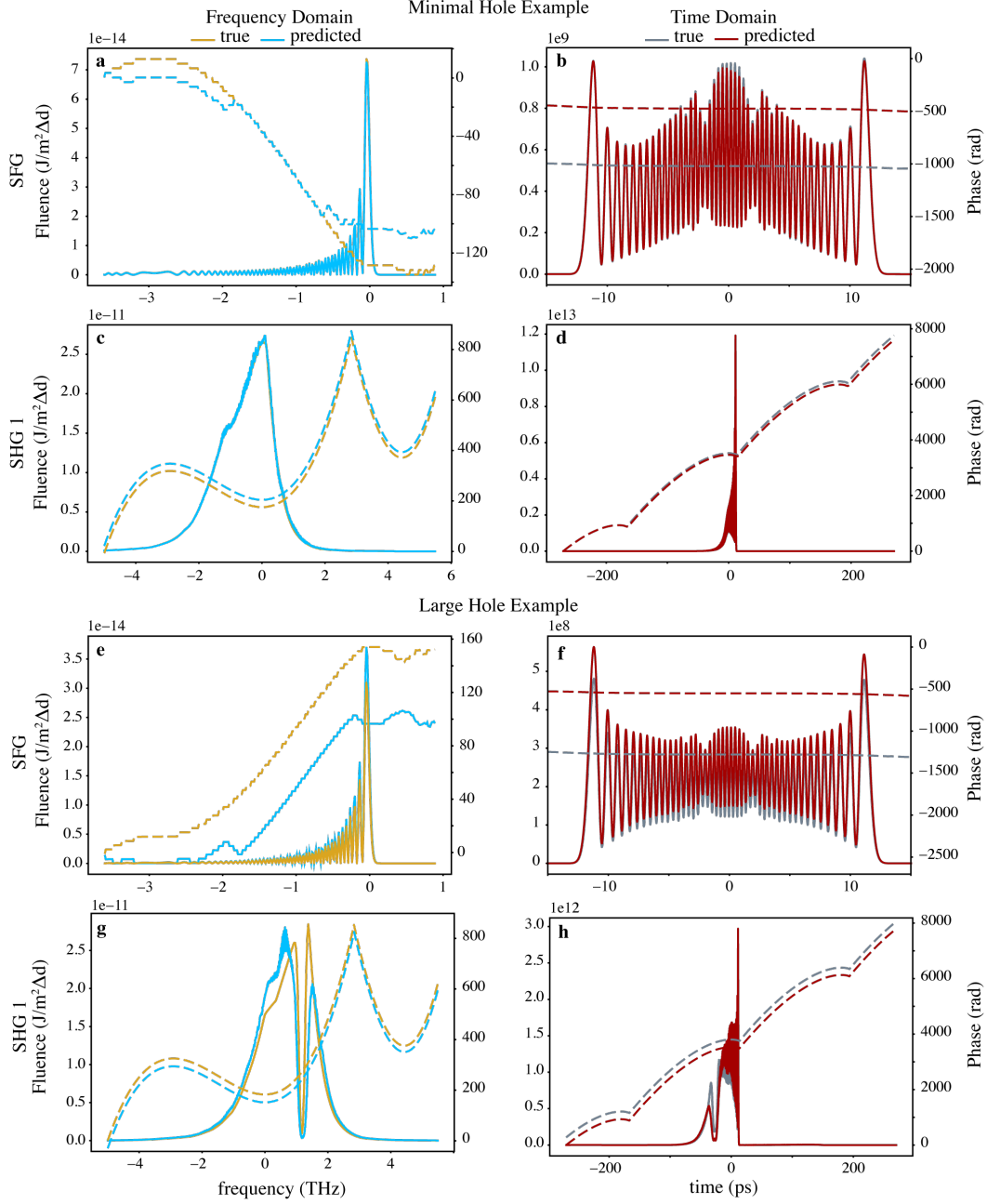


Figure 5: Shows 2 example predictions from test dataset comparing prediction versus true for SFG (output 2) and SHG 1 (parasitic output 1) in frequency (left) and time (right) domains. Solid line represents field intensity (or fluence where Δd is the domain spacing) and dashed line is the field phase (offset manually when overlapped). a-d are for minimal spectral amplitude shaping from the pulse shaper, and e-f have a significant spectral hole (seen clearly in g) carved from the pulse shaper.

bottleneck with an LSTM to drastically reduce latency. We present an LSTM that is able to learn this complex process and predict the output shape of three full fields from the three-input, three-output system. Moreover, scanning through a large input parameter space and using this LSTM on 1 GPU for the full S2E simulation presented here reduces the total per instance simulation time by an impressive 93%, from 1.98s to ~ 0.137 s.

Table 3: Inference performance comparison across various configurations of CPU and GPU setups.

	Resources	Batch size	Time for 1,000 instances (s)	Time per instance (s)	Speedup
Baseline Simulation	1 CPU	1	1,875	1.875	-
LSTM	1 CPU	200	2015.6	2.0156	0.9
	8 CPU	200	640.3	0.6403	2.9
	1 GPU	200	7.43	0.00743	252.4

These initial results demonstrate the efficacy of the LSTM approach and associated error functions. In future work, we will explore alternative data formats and ML models. For example, reformatting the fields as spectrograms in each slice of the crystal could lend itself to diffusion models to predict the field changes through the nonlinear medium. Furthermore, the complexity of the data can be expanded. We only altered the pulse shaper parameters in the simulation chain. In future studies, we will focus on adjusting the parameters controlling the DCNS process itself, making any learned model much more robust. Finally, we want to take advantage of the recent advances in physics-informed neural networks (PINNs) [26], which use the known governing dynamics of the system to both improve output predictions and increase model robustness. Such PINNs have been used to replace a split-step Fourier method for solving nonlinear dynamics in fiber optics [27]. More recently, LSTMs and PINNs have been combined together for various applications of parameter estimation in nonlinear systems [28, 29]; however, to the best of our knowledge, have not been jointly used to solve optical pulse propagation in fibers or in free-space. Thus this combination of LSTM and PINN is the next major area we want to explore for accelerating data generation in ultrafast optics and laser system design.

All of these studies are applicable in other areas of ultrafast optics outside of just this DCNS method as well. For instance, another popular upconversion scheme is four-wave mixing in fibers, which use similar principles as the SFG process but with an additional field and with gas-filled or hollow-core photonic crystal fibers as the nonlinear medium [30, 31, 32]. Achieving robust models that can handle the multitude of different optical systems will be increasingly important as photonics transitions into the new age of machine learning. Our work on robust modeling frameworks for large data generation will play a crucial role in bringing these advances to fruition.

Acknowledgments and Disclosure of Funding

Use of the Linac Coherent Light Source (LCLS), SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. Additionally, this research is supported by the U.S. Department of Energy under Contract No. DE-SC0022559 and DE-SC0022464, the National Science Foundation under Contract No. 2231334, and by the U.S. Department of Defense under Air Force Office of Scientific Research Contract No. FA9550-23-1-0409 and a National Defense Science and Engineering Fellowship.

References

- [1] Donna Strickland and Gerard Mourou. Compression of amplified chirped optical pulses. *Optics communications*, 55(6):447–449, 1985.
- [2] Mangirdas Malinauskas, Albertas Žukauskas, Satoshi Hasegawa, Yoshio Hayasaki, Vyantas Mizeikis, Ričardas Buividas, and Saulius Juodkazis. Ultrafast laser processing of materials: from science to industry. *Light: Science & Applications*, 5(8):e16133–e16133, 2016.
- [3] Jinshi Wang, Fengzhou Fang, Haojie An, Shan Wu, Huimin Qi, Yuexuan Cai, and Guanyu Guo. Laser machining fundamentals: micro, nano, atomic and close-to-atomic scales. *International Journal of Extreme Manufacturing*, 2023.

- [4] Ursula Keller. Recent developments in compact ultrafast lasers. *Nature*, 424(6950):831–838, 2003.
- [5] Tzu-Ming Liu. *Ultrafast Imaging and Spectroscopy for Biomedicine*. AIP Publishing LLC, 2022.
- [6] R Betti. A milestone in fusion research is reached. *Nature Reviews Physics*, 5(1):6–8, 2023.
- [7] Sergio Carbajo, Emilio A Nanni, Liang Jie Wong, Gustavo Moriena, Phillip D Keathley, Guillaume Laurent, RJ Dwayne Miller, and Franz X Kärtner. Direct longitudinal laser acceleration of electrons in free space. *Physical Review Accelerators and Beams*, 19(2):021303, 2016.
- [8] Neil V Sapra, Ki Youl Yang, Dries Vercruysse, Kenneth J Leedle, Dylan S Black, R Joel England, Logan Su, Rahul Trivedi, Yu Miao, Olav Solgaard, et al. On-chip integrated laser-driven particle accelerator. *Science*, 367(6473):79–83, 2020.
- [9] Anthony Valenzuela and Daniel Matyas. Review of recent capability improvements in ultrashort pulse laser sources: Closing the relevancy gap for directed energy applications. Technical report, US Army Combat Capabilities Development Command Army Research Laboratory, 2021.
- [10] Goëry Genty, Lauri Salmela, John M Dudley, Daniel Brunner, Alexey Kokhanovskiy, Sergei Kobtsev, and Sergei K Turitsyn. Machine learning and applications in ultrafast photonics. *Nature Photonics*, 15(2):91–101, 2021.
- [11] Jungtaek Kim, Mingxuan Li, Oliver Hinder, and Paul Leu. Datasets and benchmarks for nanophotonic structure and parametric design simulations. *Advances in Neural Information Processing Systems*, 36, 2024.
- [12] Jack Hirschman, Randy Lemons, Minyang Wang, Peter Kroetz, and Sergio Carbajo. Design, tuning, and blackbox optimization of laser systems. *Optics Express*, 32(9):15610–15622, 2024.
- [13] Hao Zhang, Sasha Gilevich, Alan Miahnahri, Shawn Christopher Alverson, Axel Brachmann, Joseph Duris, Paris Franz, Alan Fry, Jack Hirschman, Kirk Larsen, Randy Lemons, Siqi Li, Brittany Lu, Agostino Marinelli, Mikael Martinez, Justin May, Erel Milshtein, Krishna Murari, Nicole Neveu, Joseph Robinson, John Schmerge, Linshan Sun, Theodore Vecchione, Chengcheng Xu, Feng Zhou, and Sergio Carbajo. The lcls-ii photoinjector laser infrastructure. *High Power Laser Science and Engineering*, page 1–33, 2024.
- [14] Peter Kroetz, Axel Ruehl, Krishna Murari, Huseyin Cankaya, Franz X Kärtner, Ingmar Hartl, and RJ Dwayne Miller. Numerical study of spectral shaping in high energy hofly amplifiers. *Optics express*, 24(9):9905–9921, 2016.
- [15] Randy Lemons, Nicole Neveu, Joseph Duris, Agostino Marinelli, Charles Durfee, and Sergio Carbajo. Temporal shaping of narrow-band picosecond pulses via noncolinear sum-frequency mixing of dispersion-controlled pulses. *Physical Review Accelerators and Beams*, 25(1):013401, 2022.
- [16] Robert W Boyd. *Nonlinear optics*. Academic press, 2020.
- [17] A Couairon, E Brambilla, T Corti, D Majus, O de J. Ramírez-Góngora, and M Kolesik. Practitioner’s guide to laser pulse propagation models and simulation: Numerical implementation and practical usage of modern pulse propagation models. *The European Physical Journal Special Topics*, 199(1):5–76, 2011.
- [18] Kane Yee. Numerical solution of initial boundary value problems involving maxwell’s equations in isotropic media. *IEEE Transactions on antennas and propagation*, 14(3):302–307, 1966.
- [19] John Charles Butcher. *The numerical analysis of ordinary differential equations: Runge-Kutta and general linear methods*. Wiley-Interscience, 1987.
- [20] Lauri Salmela, Nikolaos Tsipinakis, Alessandro Foi, Cyril Billet, John M Dudley, and Goëry Genty. Predicting ultrafast nonlinear dynamics in fibre optics with a recurrent neural network. *Nature machine intelligence*, 3(4):344–354, 2021.

- [21] Lauri Salmela, Mathilde Hary, Mehdi Mabed, Alessandro Foi, John M Dudley, and Goëry Genty. Feed-forward neural network as nonlinear dynamics integrator for supercontinuum generation. *Optics Letters*, 47(4):802–805, 2022.
- [22] Hao Sui, Hongna Zhu, Huanyu Jia, Qi Li, Mingyu Ou, Bin Luo, Xihua Zou, and Lianshan Yan. Predicting nonlinear multi-pulse propagation in optical fibers via a lightweight convolutional neural network. *Optics Letters*, 48(18):4889–4892, 2023.
- [23] Simone Lauria and Mohammed F Saleh. Conditional recurrent neural networks for broad applications in nonlinear optics. *Optics Express*, 32(4):5582–5591, 2024.
- [24] Guoqing Pu, Runmin Liu, Hang Yang, Yongxin Xu, Weisheng Hu, Minglie Hu, and Lilin Yi. Fast predicting the complex nonlinear dynamics of mode-locked fiber laser by a recurrent neural network with prior information feeding. *Laser & Photonics Reviews*, 17(6):2200363, 2023.
- [25] Xiaotian Jiang, Min Zhang, Yuchen Song, Hongjie Chen, Dongmei Huang, and Danshi Wang. Predicting ultrafast nonlinear dynamics in fiber optics by enhanced physics-informed neural network. *Journal of Lightwave Technology*, 2023.
- [26] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.
- [27] Xiaotian Jiang, Danshi Wang, Qirui Fan, Min Zhang, Chao Lu, and Alan Pak Tao Lau. Physics-informed neural network for nonlinear dynamics in fiber optics. *Laser & Photonics Reviews*, 16(9):2100483, 2022.
- [28] Rahash Nathasarma and Binoy Krishna Roy. Physics-informed long-short-term memory neural network for parameters estimation of nonlinear systems. *IEEE Transactions on Industry Applications*, 2023.
- [29] Gyouho Cho, Di Zhu, Jeffrey Joseph Campbell, and Mengqi Wang. An lstm-pinn hybrid method to estimate lithium-ion battery pack temperature. *IEEE Access*, 10:100594–100604, 2022.
- [30] Federico Belli, Amir Abdolvand, John C Travers, and Philip St J Russell. Highly efficient deep uv generation by four-wave mixing in gas-filled hollow-core photonic crystal fiber. *Optics Letters*, 44(22):5509–5512, 2019.
- [31] SO Konorov, AB Fedotov, and AM Zheltikov. Enhanced four-wave mixing in a hollow-core photonic-crystal fiber. *Optics letters*, 28(16):1448–1450, 2003.
- [32] Hao Zhang, Linshan Sun, Jack Hirschman, Mirali Seyed Shariatdoust, Federico Belli, and Sergio Carbajo. Optimizing spectral phase transfer in four-wave mixing with gas-filled capillaries: A trade-off study. *arXiv preprint arXiv:2404.16993*, 2024.