



OPEN

Real-time rebright
energy, ultrafast
deep learning

Matt S¹ d¹ w¹ n¹ f¹ e¹ J² d² r² o² t² a² n² C¹ h¹ r¹ i¹ s¹ G¹ a¹ p¹ h¹ l¹ e¹ N¹ e¹ i¹ r¹ c¹ h¹ F¹ o¹ b¹ a¹ i¹ s¹ e¹ r¹
Dea¹ M¹ f¹ a¹ r¹ i¹ n¹ C¹ e¹ h¹ l¹ r¹ i¹ a¹ s¹ A¹ m¹ a¹ m¹ e¹ c¹ u¹ s¹ i¹ o¹ b¹ a¹ f¹ e¹ d¹ i¹ F¹ r¹ a¹ n¹ d¹ o¹ l¹ f¹ h¹ a¹ r¹

We report a method for the phase reconstruction of ultrashort laser pulses by deep learning of the nonlinear response. The model was trained on simulated pulses with random initial phase and 65 fs. The reconstruction is valid with moderate pulse energy. The proposed method was validated on experimental data produced by a femtosecond laser. Time phase reconstructions were performed. This method can be used for phase and nonlinear responses, even when the pulse energy is low. The method can be used to measure beams such as the high energy, large aperture, and long duration beams.

Short pulse duration lasers (< 100 fs) focused to relativistic intensities ($> 10^{18} \text{W cm}^{-2}$) are used for a wide variety of applications, such as x-ray generation^{1,2}, electron acceleration³⁻⁶, and ion acceleration⁷⁻⁹. Creating the short pulse durations used in these interactions requires the pulse to have a large spectral bandwidth along with having minimal phase shifts between the frequencies. If spectral phase differences are introduced to the pulse, significant changes to the temporal structure of the pulse can occur. The change in temporal structure could affect physical interactions, either due to a decrease in peak intensity or through complex temporal profiles introducing non-trivial interactions when used in an experiment. While knowing the temporal profile is essential for many applications, measuring the temporal profile of a high energy beam in situ is non-trivial.

Numerous techniques exist to measure the temporal profile of an ultrafast laser pulse¹⁰⁻¹⁸. This is a non-trivial task, as the fundamental theorem of Algebra prevents single spectral measurements from solving for the spectral phase relationship commonly used in tandem with the power spectrum to measure the temporal profile of an ultrafast laser pulse. This is due to non-trivial ambiguities that are introduced when converting a single power spectrum to a complex electric field, meaning no unique solutions can be found. To solve for the phase of a ultrafast laser pulse, additional information about the pulse is require to constrain the system in a way that the non-trivial ambiguities get resolved and unique solutions can be found.

A common method to gain this additional information is to measure the spectrum after a nonlinear interaction inside of some media. The nonlinear interaction is typically used to encode the phase information into a signal measurable by a square-integrable detector. Techniques such as intensity auto-correlation and frequency resolved optical gating (FROG) rely on measuring a nonlinearity induced from the interaction between two or more pulses¹³. Other techniques, such as dispersion scan (D-Scan), rely on changing the phase of the initial pulse by a known amount and monitoring how that affects the nonlinear interaction¹⁴. Commonly, a second harmonic mechanism is used for the nonlinear effect but other nonlinearities, such as the effects originating from the third-order term of the nonlinear electric susceptibility, χ_3 , have been used^{10,19}. While these techniques have been highly successful, in situ measurements of high energy beams are practically difficult.

Another nonlinearity that can be used for pulse measurement is self-phase modulation (SPM). SPM is a nonlinear optical effect that occurs due to an intensity dependent index of refraction called the optical Kerr effect, which is a χ_3 effect²⁰. The nonlinear change in index takes the form of $n = n_0 + \gamma P(\tau)$, where n_0 is the linear index of refraction, γ is the nonlinear coefficient, and $P(\tau)$ is the temporal power profile of the pulse. SPM can be modeled by the generalized nonlinear Schrödinger equation (GNLSE), which takes into account the effects of material dispersion, delayed Raman effect, and self-steepening²¹. If these effects are negligible, then the GNLSE is able to be solved analytically, taking the form of a nonlinear temporal phase shift, $E(z) = E(0)e^{i\gamma P(z)z}$, where

¹STROBE, NSF Science and Technology Center for Data-Driven and Data-Intensive Computer Science, University of California, San Diego, Computer Science and Engineering, Michigan State University, and UC Berkeley, equally: Matthew Stearns, mstearns@cs.ucsd.edu; Jieming Dai, jiemingdai@ucsb.edu; and Jieming Dai, jiemingdai@ucsb.edu.

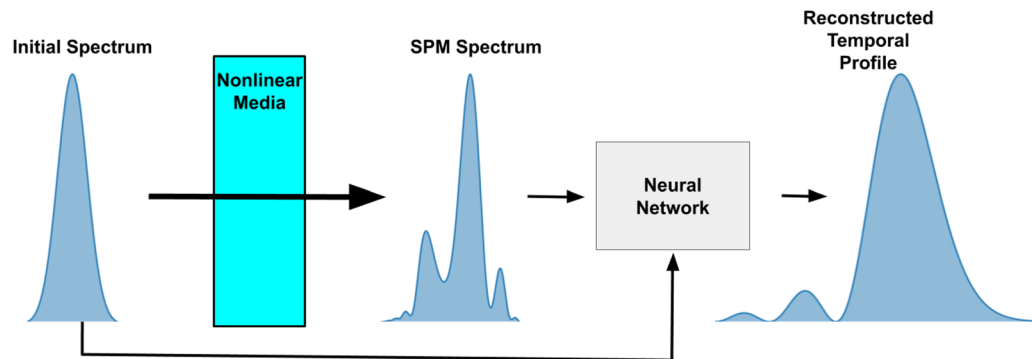


Figure 1. Example setup for using self-phase modulation to measure the initial temporal profile of a pulse. The initial spectrum passes through a material with a Kerr nonlinearity, causing a change in the spectrum due to self-phase modulation. A neural network takes the initial and final spectrum as inputs and extracts the initial phase and fluence of the initial pulse, allowing the initial temporal profile to be reconstructed.

$E(z)$ is the electric field after propagating through a material of thickness z . The total amount of nonlinearity of the system is described by the B-integral, which is the integral of the nonlinear phase shift accumulated through self-phase modulation. Under the approximation that the temporal profile is constant during propagation, the B-integral simplifies to $B = \gamma P_{max} z$ with P_{max} being the peak power of the pulse. In general, the inclusion of effects such as material dispersion, delayed Raman effect, and self-steepening requires the GNLSE to be solved numerically.

Since the spectral change between the initial and SPM spectrum are directly dependent on the temporal intensity profile. This nonlinear relationship between the initial and SPM spectra can be utilized to obtain the additional information required to reconstruct both the initial spectral phase and peak power. While SPM in dispersionless media has not been explicitly proven to uniquely find the phase for all laser pulses, it has been shown it is able to accurately and robustly solve for the phase in a wide variety of laser pulses^{15,16}. Specifically, the phase reconstruction does not require information about the peak power of the pulse and does not have a direction of time ambiguity (i.e. which side of the pulse is the leading edge of the pulse).

One benefit of SPMs ability to measure both the initial spectral phase and peak power is that by combining the peak power reconstruction with knowledge of the spatial profile of the laser pulse, the peak intensity of the laser pulse can also be calculated. With the usage of a high damage threshold dielectric as a nonlinear media, self-phase modulation can be performed with high intensity beams, enabling in situ pulse reconstructions of the beam for high energy beams. Commonly, the analytical solution for the GNLSE has been used in an Gerchberg-Saxton style iterative phase reconstruction algorithm, in which the phase is reconstructed from the measured spectra of subsequent thin dielectric plates^{15–17}. Iterative phase retrieval methods have seen success for many-cycle laser pulses, however relying on the analytical solutions to the GNLSE limits the application to systems with negligible material dispersion. For broadband laser systems, even a small amount of material dispersion can substantially alter the way the spectrum changes from SPM, requiring the GNLSE to be solved numerically.

Deep learning based algorithms recently have shown great promise for ultrafast laser pulse reconstructions. Deep learning has been applied to other pulse measurement techniques, such as SHG FROG and D-Scan, where neural networks replaced iterative algorithms^{22,23}. Deep learning can directly learn nonlinear relationships between various features within data and map them to the desired target variables²⁴. Since the information is present in the data used in the iterative phase reconstruction algorithms, deep learning can be used to directly learn the transformation between the data and the reconstructions without the need for the iterative algorithms. Since deep learning methods bypass the need for an iterative algorithm during reconstruction, deep learning approaches can be significantly faster than their iterative counterparts, enabling real-time reconstructions. Problems involving self-phase modulation have already been shown to be well suited for usage of neural networks, as neural networks have been applied to both solve the NLSE²⁵ and to solve for the peak power and pulse duration of a soliton transversing an optical fiber²⁶.

In this study, we demonstrate a technique to perform a robust deep learning reconstruction of the spectral phase and peak fluence of ultrafast laser pulses, which enables phase reconstruction to occur even if material dispersion is non-negligible. The reconstruction is based on using neural networks trained to extract the initial phase and fluence from the relative shapes of the laser spectrum before and after self-phase modulation through a dielectric medium. The training for the neural networks is done by generating a large simulated dataset of various spectra and phases before and after self-phase modulation. By training the neural network on data generated using the GNLSE, which includes effects such as material dispersion, self-steepening, and the Raman response, the neural network is able to account of these effects in its reconstruction.

Results

The technique is based on predicting the phase of the initial laser pulse, as shown in Fig. 1 by using self-phase modulation in a nonlinear media and a neural network. The spectrum of the laser pulse is measured before and after sending the pulse through a known nonlinear media. The network was trained using simulated pulses to

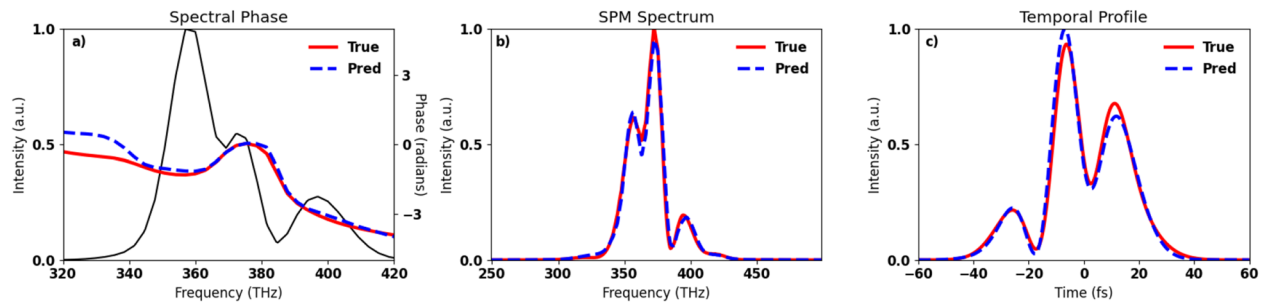


Figure 2. An example of a randomly generated simulated laser pulse used to train the neural network. (A) A randomly generated spectrum (black) centered on 374 THz along with the randomly generated phase (red) and the reconstructed initial phase (dashed blue). (B) GNLSE simulations based from the reconstructed initial pulse (dashed blue) and the true initial pulse (red). (C) Temporal profiles of the reconstructed pulse (dashed blue) and the true temporal profile (red).

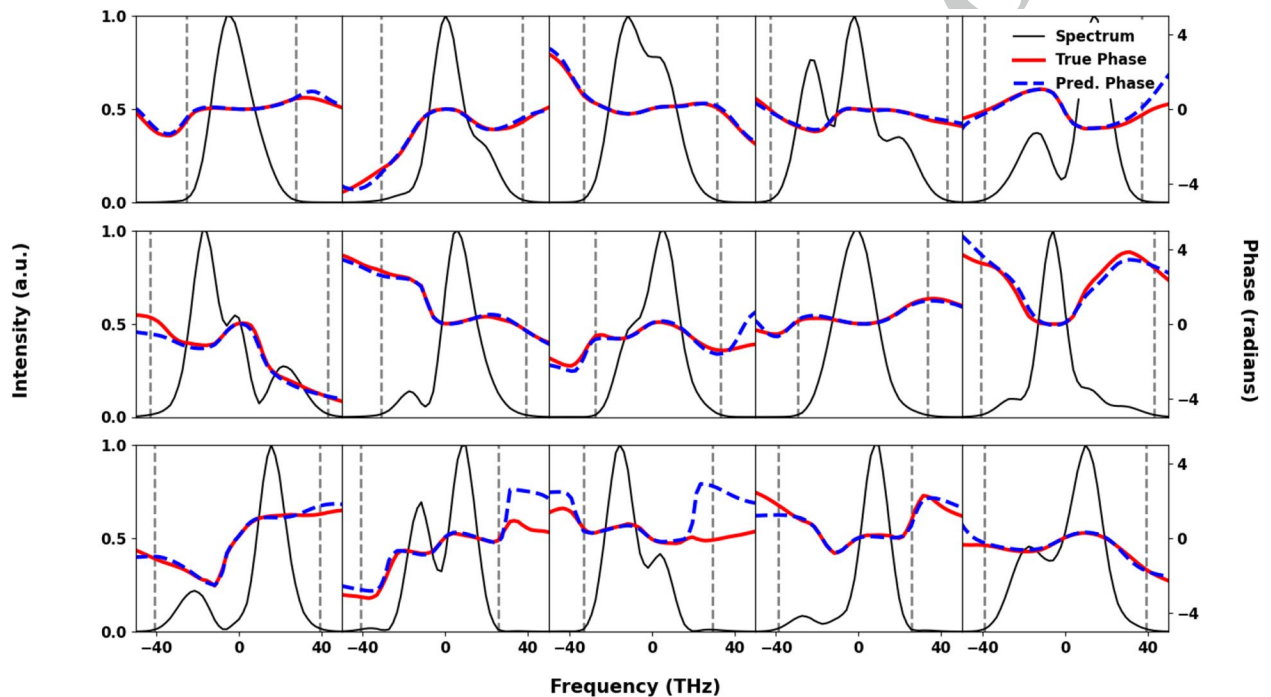


Figure 3. Examples of reconstructed pulses multiple examples of the variety of simulated initial spectra and phases predicted by the simulated few-cycle neural network (dashed blue) and the true phase (red). The predicted phase matches well in all regions of high spectral intensity, with disagreement only occurring in regions of near zero spectral intensity. Vertical dashed lines show location of the dB-20 spectral width.

predict the spectral phase of the laser pulse from the initial and final spectrum. While simple pulses such as a Gaussian spectrum with a quadratic phase may give spectral change that could be easily human interpretable, complex pulses will have a non-trivial spectral change after undergoing self-phase modulation. While non-trivial phases would not give a human-interpretable spectral change, these complex pulses can still have the phase information retrieved by a neural network (Fig. 2).

Broadband simulated For the broadband simulated spectra, two separate neural networks were trained on randomly generated simulated pulses. Both networks were designed to make predictions off of the spectral measurements of the initial pulse and the pulse after SPM, with one network used to predict the initial phase of the pulses and the other network used to predict the initial fluence of the pulses. After training, 20,000 pulses withheld from the training data were run through the networks to test the accuracy of the reconstructions for previously unseen data. A set of example reconstructions are given in Fig. 3. To quantify the accuracy of the neural network on the physical qualities we are predicting, the relative reconstruction error is calculated for the fluence and the peak value of the energy normalized temporal profile. To quantify the combined accuracy of the two networks the peak value of the intensity profile is calculated.

Comparing the reconstructed fluence to its known value provides a way to measure the accuracy of the neural network's predictions on the physical values we are trying to predict. For 99% of the pulses in the test data, the

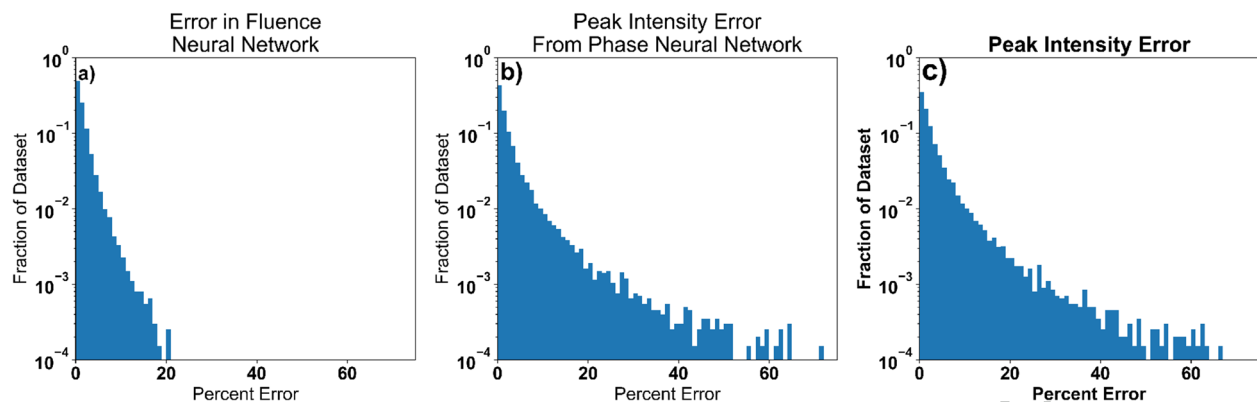


Figure 4. Reconstruction error in peak intensity prediction. To show the accuracy of the reconstructions of the simulated few-cycle neural networks the percent error is shown for the following systems. **(a)** Reconstruction error in the predictions for the fluence of the pulse. **(b)** Reconstruction error in the predictions for the maximum of the normalized temporal profile of the reconstructed pulse. **(c)** Reconstruction error in the predictions for the peak intensity of the reconstructed pulse using predicted fluence of the pulse.

neural network was able to predict the peak fluence within an error of $< 10\%$, a mean fluence reconstruction error of 1.6% and a standard deviation of the fluence reconstruction error being 2.1% , as shown in Fig. 4. When calculating the error of the phase reconstruction, we only considered regions within the pulses dB-20 spectral width, since the phase is ill-defined and not physically meaningful outside of areas with significant power spectrum. The mean standard deviation of the predicted phase was 0.13 radians. To examine how the accuracy of the phase neural network translates into the temporal domain, the reconstruction error in the predicted maximum of the normalized temporal profile was calculated. For this calculation, 93% of the pulses had a reconstruction error below 10% , a mean reconstruction error of 3.3% and a standard deviation of the error of 6.7% . An example of a reconstructed pulse is shown in Fig. 2.

By combining the results from the fluence neural network with the results of the phase neural network, the temporal intensity profile can be reconstructed, including the direction of time. Using the peak temporal intensity as an estimate of the intensity reconstruction error, which has a mean intensity reconstruction error of 3.7% and a standard deviation of the intensity reconstruction error of 7.1% with over 90% of the data set has less than 10% error.

Even in the presence of noise, accurate reconstructions can be obtained. In Fig. 5, three pulses are shown with their phase reconstructions in the presence of 0% , 20% and 40% Gaussian noise. When applied to the entire test data set, the 20% Gaussian noise caused an increase of the mean intensity reconstruction error to 5.9% and a standard deviation of the intensity reconstruction error of 8.3% with over 85% of the data set has less than 10% error. For 40% Gaussian noise the mean intensity reconstruction error to 50.1% and a standard deviation of the intensity reconstruction error of 61.2% with 25% of the data set has less than 10% error.

Experimental pulse reconstruction. In addition to validating the technique on broadband simulated data, we also utilized this method to reconstruct the phase of a experimental Ti:Sapphire laser pulse. This experiment was performed on a commercially available, 1 kHz repetition-rate laser system (Spectra-Physics Solstice ACE) with an energy of 6.6 mJ, beam diameter of 12 mm, central wavelength of 800 nm, and FTL pulse duration of 34 fs. The output of the laser was characterized using an SHG FROG. The nonlinear media, 8 mm of fused silica, is oriented at Brewster's angle such that the effective propagation length after taking into account refraction is 9.6 mm. The collimated laser beam has a peak fluence of 11.7 mJ cm $^{-2}$. Due to the large amount of material that the laser is propagating through, material dispersion will significantly impact the SPM spectrum. The spectra were taken by isolating the center of the beam with a hard aperture and sent to a fiber spectrometer utilizing an optical diffuser to minimize spectral interference from occurring inside the optical fiber.

To predict the initial phase of the laser, a new neural network was trained. While the network trained on simulated broadband pulse performed well in the previous section it has limited application to the experimental system due to the training data for that network having a bandwidth $\sim 3\times$ larger than the experimental laser system.

Due to the laser having $< 0.5\%$ root mean square energy fluctuations and a standard deviation of transform limited pulse duration < 0.75 fs, the initial laser spectrum and pulse energy can be assumed to be constant and is not needed to be included in the features given to the neural network.

The reconstruction of the experimental neural network for two separate pulses is shown in Fig. 6 and is compared to a reconstruction from a second harmonic generation FROG. The pulse duration for the positive chirped pulse predicted by the neural network was 73 fs compared to the FROGs 76 fs. The pulse duration predicted by the neural network for the near transform-limited pulse was 36 fs compared to the FROGs 35 fs pulse duration.

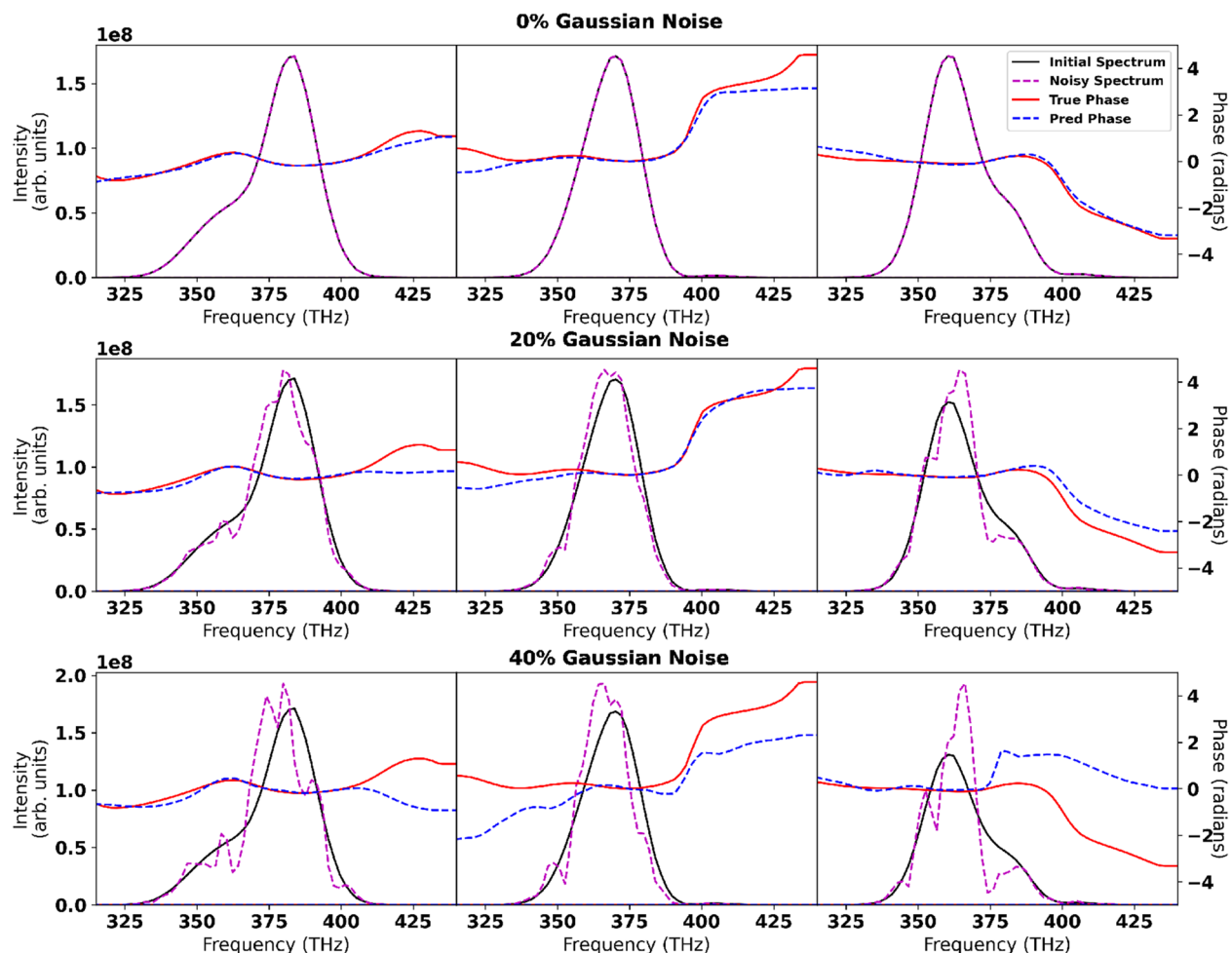


Figure 5. Reconstructions with Gaussian noise. To show the robustness of the phase predictions from the simulated few-cycle neural network above is three random initial spectra (black) and the noisy spectra (magenta) after adding a 0%, 20%, and 40% Gaussian noise. The true phase (solid red) is shown in comparison to the noisy phase (blue dashed) predicted from the noisy spectrum. The vertical dashed lines denote the location where the spectral intensity falls below 1% of the maximum value, outside this region the phase is ill-defined.

Discussion

We have shown an inexpensive and easy to experimentally implement method for measuring the temporal intensity profile of the laser pulse by utilizing self-phase modulation. While two separate neural networks were trained to predict the phase for the initial pulse, this was only done to simplify the data generation process. Since a neural network is only able to predict pulses of similar structure to what it was trained on, a specifically tailored network was designed for the broadband simulated pulses and many-cycle experimental pulses. Since the majority of experimental systems output laser spectrum that are relatively consistent in spectral width this constraint does not cause significant limitations when the method is applied to experimental systems.

Due to relying only on the localized intensity of the pulse and requiring no scanning, single-shot intensity profile characterization can be performed in-situ for a collimated beam of with energies ranging from the milli-Joule to Joule level without requiring attenuation of the beam. With the usage of a imaging spectrometer and sampling the beam at various spatial locations, a 3-dimensional intensity mapping of the beam could be created. These capabilities make this technique ideal for large aperture multi-petawatt laser systems such as ZEUS²⁷, which often measure the beam using a low energy sampling of the beam. When combined with the knowledge of the focal spot of an experimental system, this information would enable peak intensity of the focal spot to be calculated.

When applied to experimental data, as shown in Fig. 6, we see that both the FROG and the neural network method are in good agreement with the SPM spectrum, and the reconstructed temporal profiles are nearly identical. Since the spectral modulations tend to not contain fast varying features, high spectral resolution is not needed. For example, the neural network trained on simulated broadband pulses was trained on data with a wavelength resolution of > 4 nm per pixel, compared to the resolution < 2 nm resolution for the spectrometer used. The phase was able to be reconstructed using the neural network in under 10 ms on a commercially available desktop computer, implying that real-time display of the spectral phase and reconstructed temporal profile is possible.

While the assumption of a constant laser spectrum and pulse energy was able to be assumed for the experimental results discussed, environmental effects could cause a long timescale spectral drift, invalidating this

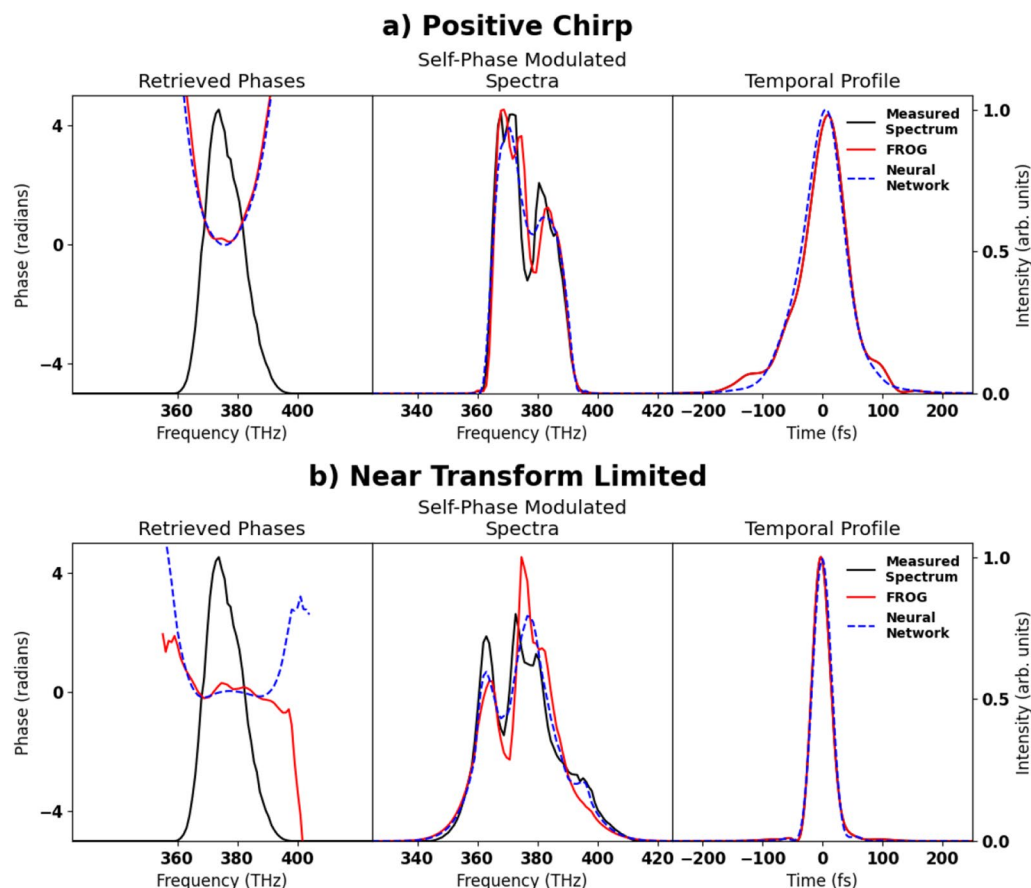


Figure 6. Experimental phase reconstruction phases measured from experimental data. Spectra measured after propagating through 8 mm of fused silica at Brewster's angle. Spectra and temporal profiles are normalized to the area under the curve. **(a)** FROG reconstruction of a positively chirped laser pulse in comparison to the reconstructed results from the neural network. Self-Phase modulated spectra and temporal profile are normalized to area under the curve for ease of comparison. **(b)** FROG reconstruction of a near transform-limited pulse in comparison to the reconstructed results from the neural network.

assumption. For a laser system where this is a concern, a network architecture where the initial spectrum is given as an input to the neural network can be used instead of including it as an assumption of the neural network. An example of such a network is the neural networks used for the simulated few-cycle sections of this manuscript, which has been shown to be able to account for a wide range of spectral profiles and pulse energies.

OPCPA laser systems have enabled high power large bandwidth laser systems in the mid-infrared (MIR) wavelength regions. In these regions traditional silicon-based detectors no longer work, meaning forcing a reliance on more expensive InGaAs detectors. While this technique was demonstrated for wavelengths from a Ti:Sapphire laser system, the technique could be applied to the wavelengths in the MIR region. Scaling to other wavelengths would only require knowledge of material properties of the nonlinear media used along with retraining of the neural network. Being able to reconstruct the phase from only two spectral measurements enables the phase information to be readily obtainable from a field auto-correlator, meaning phase information could be reconstructed from a single power diode. Field auto-correlators are already commonly used for techniques such as Fourier transform infrared spectroscopy, making this technique simple to implement into such systems.

The ability for self-phase modulation to spectrally broaden a pulse is utilized in many pulse compression techniques to generate a few-cycle laser pulses. Modifying the neural network to predict the spectral phase after self-phase modulation would enable the reconstruction of the temporal profile after pulse compression. With this modification, the same system could be used to generate and characterize a few-cycle laser pulse.

Conclusion

The presented pulse measurement technique shows a general technique of measuring the intensity profile of a laser pulse in single-shot applications using inexpensive and readily available components, only requiring a piece of glass and a spectrometer. By using a fully connected neural network phase reconstruction based on the generalized nonlinear Schrödinger equation is able to be done, which includes material dispersion, delayed Raman effect, and self-steepening. Since material dispersion is included in the modeling this technique is able to be used to characterize broadband laser pulses in real-time. With minor modifications, this technique enables

measuring the fluence and spectral phase of the pulse across the wavefront, enabling measuring variance in the temporal profile across the beam for large aperture beams that are common at facilities such as ZEUS.

Method

Training of the neural network occurs by running many examples of the problem through the neural network and updating the parameters of the network to find the optimum weights for predicting the desired target values. Once trained, neural networks can be used to predict the phase of pulses similar to the training data. While neural networks can excel in extract information from data similar to what it was trained on, prediction on data vastly different from the training data can drastically hurt the performance of the network. Due to the quasi-infinite range of laser pulses possible, generating a single generalized network is non-trivial, as one would be required to ensure all possible laser spectrum and phases are present in the training data. While a single neural network may not be generalize to a complete range of spectral bandwidths, the method of generating a neural network still holds.

In this work, we focus on two systems of interest and train neural networks separately for each system. One system is purely simulation based and simulates a wide range of broadband spectral shapes and initial phases. The second system uses simulations to model an experimental Ti:Sapphire laser. For the experimental system, due to the stability of the laser spectra the initial spectrum is able to be assumed to reduce the number of input features into the neural network.

Broadband simulated Data generation. To generate sufficiently large data sets for training neural networks, numerical simulations of a wide range of ultrafast laser pulses with varying phases were produced. The nonlinear propagation was modeled with PyNLO, a python based 1-dimensional GNLS solver using the split-step Fourier method^{21,28}. PyNLO numerically models material dispersion, self-steepening, and the delayed Raman response. The central frequency of the simulations was set to 374.0 THz, which is the central frequency of Ti:Sapphire lasers. The material properties were based on the values for fused silica, which is a common optical glass that is able to be obtained with high optical quality and is well characterized. The material dispersion was modeled by using the second, third, and fourth order expansion curves of the Sellmeier equation for fused silica, which are $36.1 \text{ fs}^2 \text{ mm}^{-1}$, $27.49 \text{ fs}^3 \text{ mm}^{-1}$, and $-11.4335 \text{ fs}^4 \text{ mm}^{-1}$.

To ensure a representative set of phases and spectrum were present, the training data was generated from a randomly generated vector. The vector has a Gaussian envelope applied in the temporal and then spectral domains, generating a pulse with a random spectrum and spectral phase. The temporal envelope used to generate the data is 30 fs. The spectral envelope used has a width of 40 THz centered on 374 THz.

After the temporal and spectral envelopes have been applied, the peak fluence is set by randomly sampling from a uniform distribution spanning the range of 16.2 mJ cm^{-2} to 43.2 mJ cm^{-2} . To remove the constant phase ambiguity, the spectral phase was defined to be zero at the central frequency. To remove the linear phase ambiguity, the temporal power's central moment was set to be centered at $t=0$. Simulations were then ran using PyNLO inside of 1 mm of fused silica assuming a nonlinear coefficient, γ , of $6 \times 10^{-8} (\text{Wm})^{-1}$. Due to the method of generating random spectra, some pulses with B-integrals > 3 are generated. These pulse are not filtered out but experimental pulses with this B-integral may run into spatial effects that break the 1-dimensional assumptions made in this work^{21,29}. The resulting simulations had a B-integral ranging from 0.65 to 4.25, with an average B-integral of 2.23.

The initial phase and initial spectrum were interpolated to 40 linearly space bins spanning a frequency range of 120 THz centered on 374 THz. The SPM spectrum was interpolated onto a linearly spaced vector with 100 bins and spanning the frequency range of 300 THz centered on 374 THz. After interpolation the area under the curve (i.e. the energy) for both the initial and SPM spectrum were normalized to unity to ensure the neural network is learning from the relative shape changes of the spectra. The initial and SPM spectra are then combined to create the feature vector that the neural network is trained on. A total of 1,830,000 samples were generated for the training and validation sets, with an additional 20,000 samples generated for the test set. An example pulse generated using this method is shown in Fig. 2.

Neural networks.

The neural network models were trained on a dataset of 1,830,000 generated samples, initially with 70% of the data in the training dataset and 30% in the validation data set. The networks were trained by gradient descent (backpropagation) using the training set. The validation dataset is then used to assess their performance and make sure the networks generalize properly to previously unseen data and do not overfit the training data. A set of 20,000 additional pulses, in neither the training nor validation data set, were then re-simulated with the phases and fluence predicted by the neural networks. These final 20,000 pulses were used to perform the analysis earlier in this paper.

The input features to the neural network are the interpolated initial and SPM spectra, with a total of 140 features. The individual features of the input tend to be right-skewed, with a majority of events taking smaller scalar values and a small minority occurring in higher regions. In order to correct this we first take the log of the input features and then normalize them, by subtracting the mean and dividing by the standard deviation. The target variables are also normalized in the same fashion. Transforming the data through this process ensures all features are on the same scale.

Separate networks were trained to reconstruct the initial phase and the fluence of the pulse. The phase neural networks were trained with the targets being the initial phase of the pulse interpolated to the same 40 length frequency grid as the initial spectrum. The fluence networks were trained with the only target being the fluence

Name	Range	Parameter type	Phase	Fluence	Baseline
Batch normalization	(Yes, no)	Choice	Yes	Yes	No
Dropout	(0, 0.25)	Continuous	0.175	0	0
Learning rate	(0.00001, 0.01)	Continuous (log)	0.008	0.003	0.001
Learning rate decay	(0.5, 1)	Continuous	0.98	1	1
Number of layers	(3, 20)	Discrete	4	8	5
Number of nodes	(128, 512)	Discrete	505	360	256
Optimizer	(Adam, SGD, RMSProp)	Choice	Adam	Adam	SGD

Table 1. Hyperparameter space. The hyperparameters from the optimized phase and fluence neural networks are shown in their respective columns, along with the baseline architecture.

of the initial pulse. All networks were implemented in Keras with a Tensorflow backend and trained on NVIDIA TITAN X GPUs.

All training samples were augmented with small amounts of Gaussian noise, $\mathcal{N}(0, 0.05)$, to mimic the imprecise fluctuations of experimental observations due to sources like laser fluctuations and thermal noise in silicon based detectors. This augmentation, added during training batches, also serves to prevent overfitting to the training set. Other models of experimental noise could be included by applying the noise model to the data, either during data generation or training. Training occurred over a maximum of 400 epochs. The performance of each network is characterized by calculating the mean square error loss of the predicted values compared to the target values. If the validation loss did not improve after fifteen epochs, training was terminated.

Building and training neural networks requires one to set many values, called hyperparameters, a priori. Hyperparameters include the number of layers, the number of nodes per layer, the kinds of activation functions, the learning rates, and the dropout rates. Dropout is a randomization procedure used during training that turns off different connections in the neural network, forcing the network to learn a more general solution which also helps avoid overfitting^{30,31}.

In the experiments, the hyperparameters were optimized using SHERPA³², a Python software library which is compatible with Keras and other modern deep learning libraries, and has been used to effectively optimize neural networks in various scientific applications (e.g.^{33,34}).

Leveraging SHERPA, a large suite of 500 models were explored using a Bayesian optimization algorithm. The Bayesian search has the advantage of learning a distribution over the hyperparameters of the network architecture, in relation to the task to be optimized. By employing this procedure we are able to evaluate a large space of possible models and test many configurations. To demonstrate the efficacy of the hyperparameter search, we compare the resulting model against an initially proposed baseline model. The baseline architecture is shown in Table 1. The optimized phase and fluence networks contained roughly 608 thousand and 970 thousand parameters, respectively. While the training time between each model varied due to the different hyperparameters, each network took approximately an hour to train.

In total, 500 network architectures were explored with differing hyperparameters for both the phase and fluence neural networks. The final architectures from the hyper-parameter search are shown in Table 1. The table displays the hyperparameters of the best performing phase and fluence network, along with the hyperparameters of the baseline network. The distribution of the validation mean squared errors (MSE) for the phase and fluence networks are shown in Fig. 7a,b respectively. These figures highlight the performance of the best optimized model compared to the initially proposed baseline network.

Following the hyperparameter search, the best performing phase and fluence networks were evaluated using tenfold cross validation. During tenfold cross validation the data is randomly partitioned into ten distinct folds. Each network is then trained on nine of the folds and tested on the remaining one, and the process is repeated ten times. The mean and the standard deviation of the performance (error bars) can then be computed over the ten experiments. The results from tenfold cross validation are presented in Fig. 8a,b. These figures demonstrate consistent performance across all ten folds. We confirm that neither the phase network nor the fluence network overfits the training data by comparing the performance on the training and validation set across all ten folds. The average difference between the training and validation loss is less than 0.01 and 0.002 for the phase and fluence networks respectively.

Experimental pulse reconstruction. The simulations used to generate the training data were based on the methods discussed in the Broadband Simulated Pulse Reconstruction portion of the Methods section. Due to the dominant phase terms of the pulse being group delay dispersion (GDD) and third-order dispersion (TOD), the phase was modeled primarily as a Taylor series expansion, with random GDD, TOD, and forth-order (FOD) phase terms. The dispersion coefficient were generated from a normal distribution with the standard deviation of 10^3 fs^2 , 10^4 fs^3 , 10^6 fs^4 for the GDD, TOD, and FOD phase terms. To allow for minor deviations from this expansion, a random phase was generated by taking the phase of a random spectrum generated using the Fourier technique described used to generate the broadband simulated data and was added to the Taylor series phase. This Fourier phase was generated using a temporal and frequency FWHMs used were 60 fs and 50 THz with a maximum phase deviation within 25 THz of the central frequency being sampled from a normal distribution with a standard deviation of 0.5π . The peak fluence of the pulse was set to match the fluence

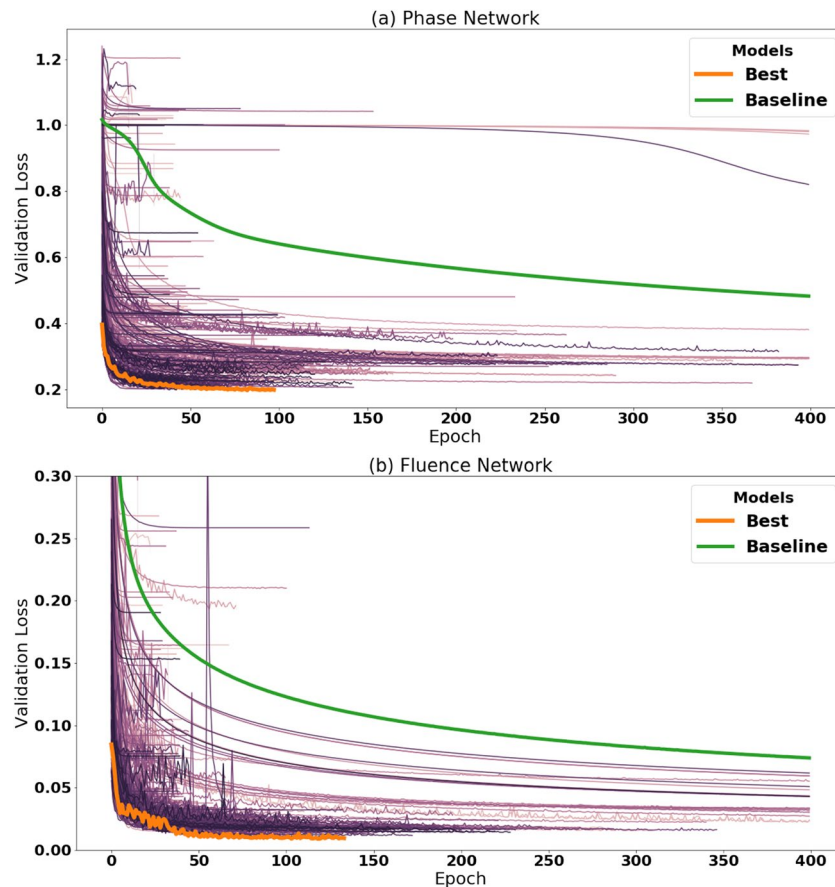


Figure 7. Validation loss of SHERPA trials, measured by the mean squared error, over time. Each line depicts the validation loss of a different SHERPA trial during the course of training of the simulated few-cycle networks. **(a)** Trials from phase networks with varying hyperparameters. **(b)** Trials from fluence networks with varying hyperparameters. Note: not all 500 trials are shown in each figure. Some trials with higher validation losses are left out for figure clarity. This discards 50 and 145 networks for a and b, respectively.

from the laser and propagated through 9.6 mm of fused silica. A total of 432 thousand simulated pulses were used for training the network.

A neural network was then trained on 432 thousand pulses total using a 80/20 split for the training and validation data sets. Due to the assumption of a single initial laser spectrum, the features used in training only needed to be based on the SPM spectrum. Both the broadened spectrum and the initial phase were interpolated to the range from 330 THz to 418 THz binned 100 linearly spaced bins. The network consisted of 8 layers with a width of 200 and was trained using a Gaussian noise of 0.1, learning rate of 0.001, drop out rate of 0.1 over 200 epochs.

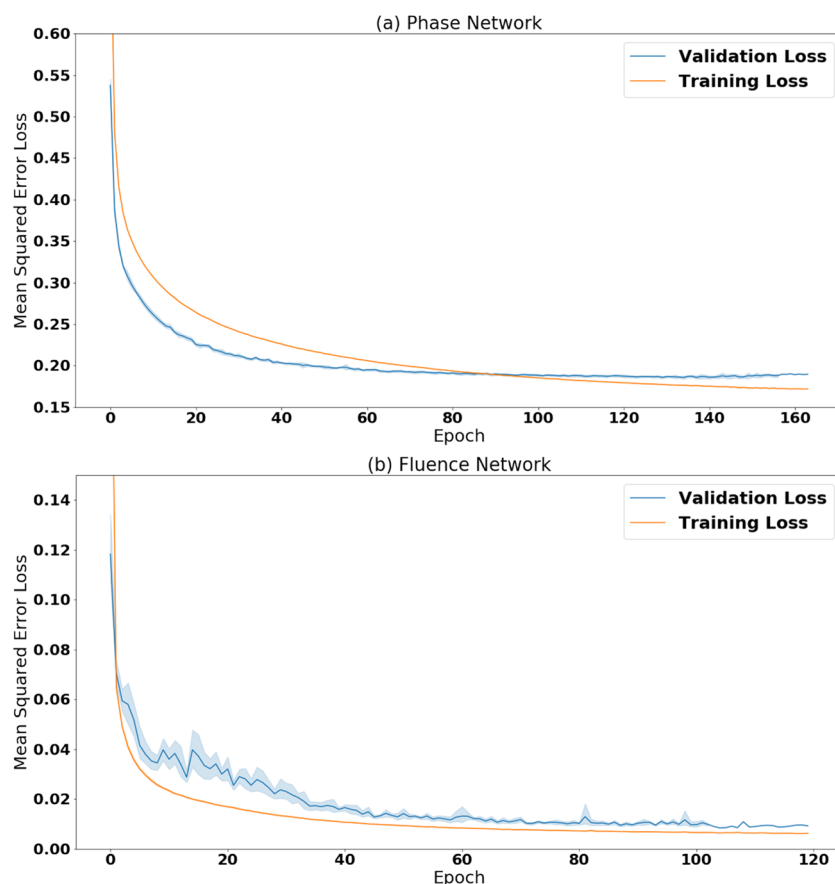


Figure 8. Training and validation loss, measured by the mean squared error, over time. Loss curves show the average (solid line) and one standard deviation (shaded region) for the ten folds of cross validation. (a) Simulated few-cycle phase network. (b) Simulated few-cycle fluence network.

Data availability

The data related to this paper is available from the corresponding author upon request.

Received: 21 December 2021; Accepted: 14 March 2022

References

- Naumova, N. M., Nees, J. A., Sokolov, I. V., Hou, B. & Mourou, G. A. Relativistic generation of isolated attosecond pulses in a λ^3 focal volume. *Phys. Rev. Lett.* **92**, 063902. <https://doi.org/10.1103/PhysRevLett.92.063902> (2004).
- Heissler, P. *et al.* Few-cycle driven relativistically oscillating plasma mirrors: A source of intense isolated attosecond pulses. *Phys. Rev. Lett.* <https://doi.org/10.1103/physrevlett.108.235003> (2012).
- Tajima, T. & Dawson, J. M. Laser electron accelerator. *Phys. Rev. Lett.* **43**, 267–270. <https://doi.org/10.1103/physrevlett.43.267> (1979).
- Faure, J. *et al.* A laser-plasma accelerator producing monoenergetic electron beams. *Nature* **431**, 541–544. <https://doi.org/10.1038/nature02963> (2004).
- Mangles, S. P. D. *et al.* Monoenergetic beams of relativistic electrons from intense laser-plasma interactions. *Nature* **431**, 535–538. <https://doi.org/10.1038/nature02939> (2004).
- Geddes, C. G. R. *et al.* High-quality electron beams from a laser wakefield accelerator using plasma-channel guiding. *Nature* **431**, 538–541. <https://doi.org/10.1038/nature02900> (2004).
- Macchi, A., Borghesi, M. & Passoni, M. Ion acceleration by superintense laser-plasma interaction. *Rev. Mod. Phys.* **85**, 751–793. <https://doi.org/10.1103/RevModPhys.85.751> (2013).
- Torrisi, L. *et al.* Protons accelerated in the target normal sheath acceleration regime by a femtosecond laser. *Phys. Rev. Acceler. Beams* **22**, 021302 (2019).
- Zhou, M. L. *et al.* Proton acceleration by single-cycle laser pulses offers a novel monoenergetic and stable operating regime. *Phys. Plasmas* **23**, 043112. <https://doi.org/10.1063/1.4947544> (2016) (publisher: American Institute of Physics).
- Trebino, R. *et al.* Highly reliable measurement of ultrashort laser pulses. *J. Appl. Phys.* **128**, 171103 (2020).
- Iaconis, C. & Walmsley, I. A. Spectral phase interferometry for direct electric-field reconstruction of ultrashort optical pulses. *Opt. Lett.* **23**, 792–794. <https://doi.org/10.1364/OL.23.000792> (1998).
- Fittinghoff, D. N. *et al.* Measurement of the intensity and phase of ultraweak, ultrashort laser pulses. *Opt. Lett.* **21**, 884–886. <https://doi.org/10.1364/OL.21.000884> (1996).
- Trebino, R. *et al.* Measuring ultrashort laser pulses in the time-frequency domain using frequency-resolved optical gating. *Rev. Sci. Instrum.* **68**, 3277–3295 (1997).

14. Miranda, M., Fordell, T., Arnold, C., L'Huillier, A. & Crespo, H. Simultaneous compression and characterization of ultrashort laser pulses using chirped mirrors and glass wedges. *Opt. Exp.* **20**, 688–697 (2012).
15. Prade, B. S., Schins, J. M., Nibbering, E. T. J., Franco, M. A. & Mysyrowicz, A. A simple method for the determination of the intensity and phase of ultrashort optical pulses. *Opt. Commun.* **113**, 79–84. [https://doi.org/10.1016/0030-4018\(94\)90595-9](https://doi.org/10.1016/0030-4018(94)90595-9) (1994).
16. Nibbering, E. T. J. *et al.* Spectral determination of the amplitude and the phase of intense ultrashort optical pulses. *JOSA B* **13**, 317–329. <https://doi.org/10.1364/JOSAB.13.000317> (1996).
17. Anashkina, E. A. *et al.* Single-shot laser pulse reconstruction based on self-phase modulated spectra measurements. *Sci. Rep.* **6**, 33749. <https://doi.org/10.1038/srep33749> (2016).
18. O'Shea, P., Kimmel, M., Gu, X. & Trebino, R. Highly simplified device for ultrashort-pulse measurement. *Opt. Lett.* **26**, 932–934. <https://doi.org/10.1364/OL.26.000932> (2001).
19. Tsang, T., Krumbügel, M. A., DeLong, K. W., Fittinghoff, D. N. & Trebino, R. Frequency-resolved optical-gating measurements of ultrashort pulses using surface third-harmonic generation. *Opt. Lett.* **21**, 1381–1383. <https://doi.org/10.1364/OL.21.001381> (1996).
20. Shimizu, F. Frequency broadening in liquids by a short light pulse. *Phys. Rev. Lett.* **19**, 1097 (1967).
21. Hult, J. A fourth-order Runge-Kutta in the interaction picture method for simulating supercontinuum generation in optical fibers. *J. Lightwave Technol.* **25**, 3770–3775 (2007).
22. Zahavy, T. *et al.* Deep learning reconstruction of ultrashort pulses. *Optica* **5**, 666–673. <https://doi.org/10.1364/OPTICA.5.000666> (2018).
23. Kleinert, S., Tajalli, A., Nagy, T. & Morgner, U. Rapid phase retrieval of ultrashort pulses from dispersion scan traces using deep neural networks. *Opt. Lett.* **44**, 979–982 (2019).
24. Baldi, P. *Deep Learning in Science: Theory, Algorithms, and Applications* (Cambridge University Press, 2021) (in press).
25. Monterola, C. & Saloma, C. Solving the nonlinear Schrödinger equation with an unsupervised neural network. *Opt. Exp.* **9**, 72–84. <https://doi.org/10.1364/OE.9.000072> (2001).
26. Herrera, R. A. Evaluating a neural network and a convolutional neural network for predicting soliton properties in a quantum noise environment. *J. Opt. Soc. Am. B* **37**, 3094–3098. <https://doi.org/10.1364/JOSAB.401936> (2020).
27. Nees, J. *et al.* Zeus: A national science foundation mid-scale facility for laser-driven science in the qed regime. in *2020 Conference on Lasers and Electro-Optics (CLEO)*. 1–2 (2020).
28. Ycas, G., Maser, D. & Hickstein, D. *Nonlinear Optics Modelling for Python*. (Pynlo, 2015).
29. Cumberbatch, E. Self-focusing in non-linear optics. *IMA J. Appl. Math.* **6**, 250–262. <https://doi.org/10.1093/imamat/6.3.250> (1970). <https://academic.oup.com/imamat/article-pdf/6/3/250/1848096/6-3-250.pdf>.
30. Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **15**, 1929–1958 (2014).
31. Baldi, P. & Sadowski, P. The dropout learning algorithm. *Art. Intell.* **210C**, 78–122 (2014).
32. Hertel, L., Collado, J., Sadowski, P., Ott, J. & Baldi, P. *Robust Hyperparameter Optimization for Machine Learning*. (SoftwareX, 2020).
33. Beucler, T. *et al.* Enforcing analytic constraints in neural networks emulating physical systems. *Phys. Rev. Lett.* **126**, 098302. <https://doi.org/10.1103/PhysRevLett.126.098302> (2021).
34. Ott, J. *et al.* A fortran-keras deep learning bridge for scientific computing. *Sci. Program.* **2020** (2020).

Acknowledgements

This work was supported by STROBE: A National Science Foundation Science & Technology Center under Grant No. DMR-1548924; This material is based upon work supported by the National Science Foundation under the CAREER program Grant No. PHY-1753165, and under grant number DGE-1633631.

Author contributions

M.S., J.O., D.F., C.M., P.B., and F.D. contributed to the initial concept of the project. M.S. generated the simulated data required for training of the networks. J.O. and P.B. designed, trained, optimized, and validated the neural networks for the broadband dataset. M.S. trained the neural network for the experimental dataset. M.S., C.G., and N.B. collected the experimental data. F.D. and P.B. supervised the project. M.S., J.O., P.B., and F.D. prepared the manuscript. All authors edited the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-09041-y>.

Correspondence and requests for materials should be addressed to M.S., P.B. or F.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022