MACHINE LEARNING APPROACH TO TEMPORAL PULSE SHAPING FOR THE PHOTOINJECTOR LASER AT CLARA

A.E. Pollard*, W. Okell, D. Dunning, E. Snedden
ASTeC, Cockcroft Institute, STFC Daresbury Laboratory, UK

Abstract

The temporal profile of the electron bunch is of critical importance in accelerator areas such as free-electron lasers and novel acceleration. In FELs, it strongly influences factors including efficiency and the profile of the photon pulse generated for user experiments, while in novel acceleration techniques it contributes to enhanced interaction of the witness beam with the driving electric field. Work is in progress at the CLARA facility at Daresbury Laboratory on temporal shaping of the ultraviolet photoinjector laser, using a fused-silica acousto-optic modulator. Generating a user-defined (programmable) time-domain target profile requires finding the corresponding spectral phase configuration of the shaper; this is a non-trivial problem for complex pulse shapes. Physically informed machine learning models have shown great promise in learning complex relationships in physical systems, and so we apply machine learning techniques here to learn the relationships between the spectral phase and the target temporal intensity profiles. Our machine learning model extends the range of available photoinjector laser pulse shapes by allowing users to achieve physically realisable configurations for arbitrary temporal pulse shapes.

INTRODUCTION

In photoinjector systems, control over the longitudinal properties of the electron bunch can be achieved through temporal shaping of the laser pulse temporal profile [1]. Following the temporal shaping concept presented in [2], we have developed an apparatus for temporally shaping the photoinjector laser pulses at CLARA, shown schematically in Fig. 1. The input laser pulse is spectrally dispersed by a transmission grating. A concave mirror one focal length away from the grating collimates the spectrum and focuses the laser pulse to a line focus, along which the laser wavelength varies approximately linearly. A fused-silica AOM is placed at the position of the focus, and a transducer driven with an RF waveform at 200 MHz central frequency generates an acoustic wave in the AOM, which propagates along the line focus of the laser. The laser pulses are diffracted from the induced refractive index modulation, and the spectral components are recombined using a second concave mirror and transmission grating.

To shape the laser pulse temporally, the spectral phase can be adjusted by varying the temporal phase of the acoustic wave via the temporal phase of the RF drive wave. The laser pulses can also be shaped temporally by varying the temporal amplitude of the acoustic wave; however, as this approach is lossy, it is necessary to carry out all shaping using

* amelia.pollard@stfc.ac.uk

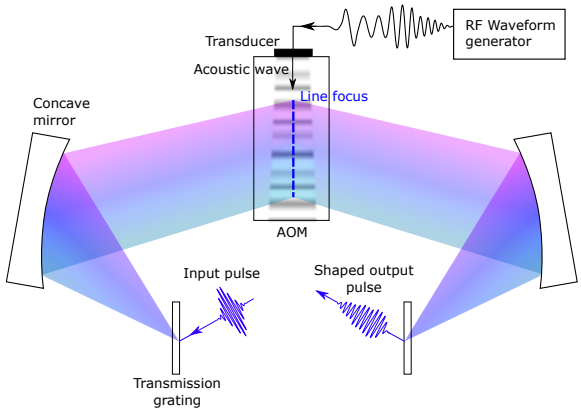

Figure 1: Schematic of the temporal pulse shaper at CLARA.

only the phase. In order to produce a particular target pulse temporal intensity profile, we need to find a suitable spectral phase mask to apply to the shaper. This is non-trivial for arbitrary shapes, as we require both the phase and amplitude in either the spectral or temporal domain to fully define the pulse. However we know only the temporal intensity and the spectral intensity, leaving the temporal and spectral phase as unknowns with many potential solutions. The complexity of real experimental systems poses additional challenges, for example, there are limitations imposed by the physical characteristics of the AOM. Modulating the spectral phase by modulating the temporal phase of the RF wave broadens the RF spectrum. The AOM has a finite acoustic bandwidth, and the RF spectrum must remain within this bandwidth for spectral phase modulations to be physically realisable.

Machine learning approaches excel for complex non-linear problems such as this. In particular, deep neural networks are known to be capable of approximating any function[3], and recent work has demonstrated that such networks can be used to learn and manipulate spectral, temporal, and shape properties of laser pulses[4][5]. Recent research has explored encoding physical laws into machine learning models with partial differential equations as priors[6] to reduce the data requirements of these otherwise data-intensive approaches. This approach has come to be known as Physically Informed Neural Networks (PINNs) and can be used to constrain the outputs of deep neural networks within physical reality, by encoding properties such as conservation of energy *a priori*.

In this paper, we present a PINN for finding the spectral phase mask required to produce a target temporal intensity profile in our photoinjector laser pulse shaper, subject to the physical limitation of the AOM bandwidth. Our approach both reduces the data requirements of our model and con-

strains the search space within a physically realisable range. Thus, we can be confident that predictions of temporal intensity profiles produced by our model will be experimentally achievable. Using a PINN also increases the speed with which the required phase mask can be found, compared to algorithmic and iterative methods.

In addition to the constraints of the physical system, we can also encode the inherent symmetries of the underlying physical system. By employing the principles of geometric deep learning[7], we can choose to exploit those symmetries to reduce the complexity of the underlying parameter space. In particular, we can exploit the translation-invariant nature of the temporal pulse shape, since the structure of the pulse signal is our primary concern and pulse timing can be adjusted without consequence. This guides our choice of loss function away from the mean squared error functions of other works towards alternative signal matching algorithms, such as the Pearson correlation coefficient. This significantly improves our results by expanding the potential search space.

RELATED WORK

In their paper on applying an iterative Fourier transform to this issue, Hacker et al.[8] propose an algorithm which can quickly approximate the spectral phase corresponding to a particular target waveform. This is done by iteratively performing a Fourier transform into the spectral domain, correcting for differences between the current and target spectrum, and then transforming back into the time domain and correcting for temporal amplitude differences. This process is repeated until an adequate match is found. This achieves results comparable with a genetic algorithm approach but in a significantly shorter time.

In their work on applying neural networks to predicting temporal and spectral pulse profiles in optical fibres, Boscolo et al.[4] demonstrate that supervised models are capable of learning the complex relationships between temporal pulse shape and spectral intensity. They also used a neural network to determine the non-linear propagation properties of a pulse observed at the fiber output and classify output pulses according to the initial pulse shape.

Though their work concerns spatial shaping as opposed to the temporal shaping discussed here, Xu et al.[5] demonstrate that neural networks are again capable of learning to manipulate laser profiles with an SLM to generate arbitrary output shapes.

For a more complete overview, see Genty et al.'s review[9].

METHODOLOGY

Using simulated data, we developed and tested a machine learning model to find the required phase mask to achieve a particular target pulse temporal profile. The simulated laser pulses used for training and testing the model have a spectral intensity with Gaussian shape in wavelength, central wavelength of 266 nm, and FWHM bandwidth of 1.5 nm. As pulses in the CLARA photoinjector laser system are temporally stretched in a grating stretcher before entering the shaper, our simulated unshaped pulses have 8×10^4 fs^2 of spectral phase applied. For our training set, we generate 10^5 pairs of spectral phase profiles and corresponding temporal intensity profiles, with a further 10^3 pairs generated for the test set. Each pair consists of a spectral phase profile consisting of 2642 samples over 5.78 nm and a temporal intensity profile of 294 samples over 12 ps.

So we can constrain our model to the physical limits of the AOM bandwidth, we consider the effect of AOM bandwidth on the spectral phase mask. Modulating the temporal phase of the RF drive wave broadens its spectrum; the instantaneous RF frequency at a particular point in time is given by the gradient of its temporal phase at that point. The AOM bandwidth limits the gradient of the acoustic wave temporal phase modulation, and consequently the limits on the available phase modulation per unit length along the acoustic wave propagation direction are

$$\frac{\mathrm{d}\varphi}{\mathrm{d}x} = \pm\frac{\pi\Delta f_{\mathrm{ac}}}{v_{\mathrm{ac}}}, \tag{1}$$

where φ is the phase modulation, x is the spatial coordinate across the AOM window, $v_{\mathrm{ac}} = 5968\,\mathrm{m\,s^{-1}}$ is the acoustic velocity in fused silica, and $\Delta f_{\mathrm{ac}} \approx 100\,\mathrm{MHz}$ is the AOM acoustic bandwidth. The change in laser wavelength per unit length across the AOM window is

$$\frac{\mathrm{d}\lambda}{\mathrm{d}x} \approx \frac{\Delta\lambda}{W}, \tag{2}$$

where $\Delta\lambda \approx 5\,\mathrm{nm}$ is the optical bandwidth covered by the AOM window, and $W = 20\,\mathrm{mm}$ is the width of the AOM window. From Eq. 1 and Eq. 2, the limits on the laser spectral phase gradient are therefore

$$g_\varphi = \frac{\mathrm{d}\varphi}{\mathrm{d}\lambda} \approx \pm\frac{\pi\Delta f_{\mathrm{ac}} W}{v_{\mathrm{ac}}\Delta\lambda}. \tag{3}$$

For our experimental parameters, $g_\varphi = \pm\pi/0.015$ rad/nm.

To encode this physical limit associated with the AOM bandwidth into the network, we developed a regulariser which acts to limit the gradient of the spectral phase profile to a physical limit of $\pi/0.015$ rad/nm, corresponding to a maximum phase change per wavelength step of $\delta\varphi \approx 0.153\pi$ rad/step. For the purposes of limiting the gradient over a discrete sampling, we define the discrete gradient as Eq. 4. To account for the cyclic nature of angular frequency in a differentiable manner, we calculate the gradient by projecting into the complex plane. We then take the absolute value of the discrete gradient of the network's output vector. We multiply the resulting function by a high-gradient sigmoid function with an offset of $\delta\varphi$ to provide a differentiable approximation of a step function. Since $Re(e^{i\theta})$ bounds between -1 and 1, and the input is between $-\pi$ and π, we divide $\delta\varphi$ by π to arrive at the final definition for the regulariser, shown in Eq. 5. Note that the first and last elements are masked out from the difference calculation, since we are not concerned

with forcing the spectral phase profile to begin and end at 0 rads.

The loss function encodes the translation invariant nature of temporal pulse shaping by calculating the Pearson correlation coefficient of the target temporal pulse profile against the temporal pulse profile simulated from the spectral phase profile output by the network. The simulation code is made differentiable by the Keras[10] framework, allowing the training of the network to be guided by the gradient of the underlying function space. This further encodes physical laws into the network.

$$\Delta_+(f) = f_i - f_{i+1} \tag{4}$$

$$\frac{1}{N}\sum |\Delta_+(e^{i\varphi(\omega)})| * \eta\sigma(|\Delta_+(e^{i\varphi(\omega)})| - \delta\varphi/\pi); \eta = 100 \tag{5}$$

The architecture for the network is a simple deep neural network with three hidden layers using the ReLu[11] activation function, with batch normalisation between the layers. The final output layer uses a linear activation function. We use the Adam optimiser[12] with a learning rate schedule decaying from 0.001 at a rate of $e^{0.001}$ per epoch after the first 100 epochs.

RESULTS

We find that with the application of the principles of physically informed networks enables the proposed system to learn to extrapolate appropriate spectral phase profiles for arbitrary temporal pulse shapes in linear time, which are both accurate and contain no non-physical phase transitions. This allows users to specify arbitrary temporal pulse profiles and receive an input for the SLM which will provide that profile within milliseconds, a significant advantage over algorithmic and iterative methods.

As in other works, we calculate the mean squared error (MSE) of the output temporal intensity profile against those in the test set, and find strong agreement (6.4e−3 ± 3.7e−5 MSE over 10,000 samples). Indeed, our results match simulation extremely well, as can be seen in Fig. 2. We are also able to specify a wide variety of target pulse shapes which are well outside of those described in the test set, and receive matching physically realisable spectral phase profiles which generate them well, as in Fig. 3. Without the limitations of the SLM it is possible to achieve very high quality matches to the target patterns, however these are not physically achievable since they require spectral phase transitions well beyond what is physically possible. However, with the physical limitations imposed by the PINN, we achieve high quality matches to arbitrary temporal phase profiles which are physically realisable.

CONCLUSION

By using physically informed networks we can build better machine learning models which more accurately model the reality of the target system. In doing so, we develop a model for predicting spectral phase profile configurations for photocathode laser at CLARA, to enable arbitrary specifications of temporal pulse profiles for fine control over the bunch profile. In future work, we intend to deploy this system on the CLARA facility to enable fine temporal pulse shaping and expansion into bunch profile specification, with an eye toward use in FEL research.

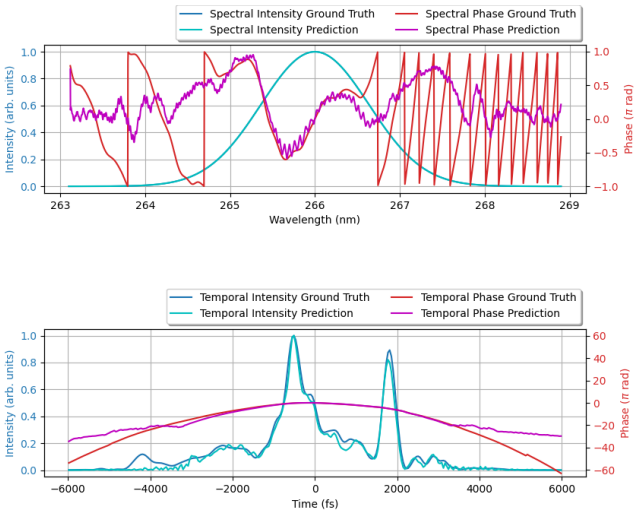

Figure 2: Randomly selected example from the test set, showing spectral phase mask and spectral intensity (top), and pulse temporal phase and temporal intensity profile (bottom). The predicted temporal pulse shape is an excellent match to the ground truth.

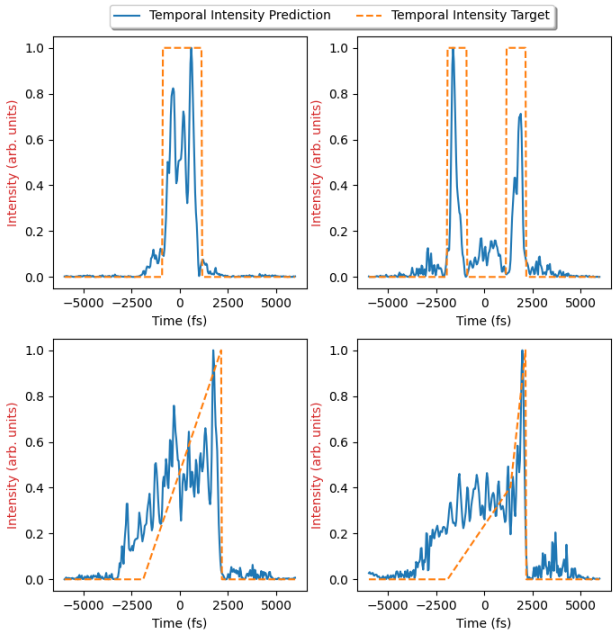

Figure 3: Demonstration of solutions found for arbitrary pulse shapes. In particular note that these are physically realisable due to the gradient constraint.

REFERENCES

[1] G. Penco *et al.*, "Experimental demonstration of electron longitudinal-phase-space linearization by shaping the photoinjector laser pulse," *Physical Review Letters*, vol. 112, p. 044 801, 2014.

[2] C. W. Hillegas, J. X. Tull, D. Goswami, D. Strickland, and W. S. Warren, "Femtosecond laser pulse shaping by use of microsecond radio-frequency pulses," *Optics Letters*, vol. 19, no. 10, pp. 737–739, 1994.

[3] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.

[4] S. Boscolo and C. Finot, "Artificial neural networks for nonlinear pulse shaping in optical fibers," *Optics & Laser Technology*, vol. 131, p. 106 439, 2020.

[5] C. Xu *et al.*, "Machine learning based spatial light modulator control for the photoinjector laser at flute," in *12th Int. Particle Accelerator Conf.(IPAC'21), Campinas, Brazil*, 2021.

[6] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics informed deep learning (part i): Data-driven solutions of nonlinear partial differential equations. arxiv 2017," *arXiv preprint arXiv:1711.10561*,

[7] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst, "Geometric deep learning: Going beyond euclidean data," *IEEE Signal Processing Magazine*, vol. 34, no. 4, pp. 18–42, 2017.

[8] M. Hacker, G. Stobrawa, and T. Feurer, "Iterative fourier transform algorithm for phase-only pulse shaping," *Optics Express*, vol. 9, no. 4, pp. 191–199, 2001.

[9] G. Genty *et al.*, "Machine learning and applications in ultrafast photonics," *Nature Photonics*, vol. 15, no. 2, pp. 91–101, 2021.

[10] F. Chollet *et al.* "Keras." (2015), https://github.com/fchollet/keras

[11] A. F. Agarap, "Deep learning using rectified linear units (relu)," *arXiv preprint arXiv:1803.08375*, 2018.

[12] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.