MACHINE LEARNING-BASED LONGITUDINAL PHASE SPACE PREDICTION OF TWO-BUNCH OPERATION AT FACET-II[*]

C. Emma[†], A. Edelen, M. Alverson, A. Hanuka, M. J. Hogan, B. O'Shea,
D. Storey, G. White, V. Yakimenko
SLAC National Accelerator Laboratory, Menlo Park, USA

Abstract

Machine learning (ML) based virtual diagnostics predict what the output of a measurement would look like when that diagnostic is unavailable [1]. This is especially useful for cases when a particular measurement is destructive. In this paper, we report on the application of ML methods for predicting the longitudinal phase space (LPS) distribution of the FACET-II linac operating in two-bunch mode. Our approach consists of training a ML-based virtual diagnostic to predict the LPS using only nondestructive linac and e-beam measurements as inputs. We validate this approach with a simulation study including the longitudinal smearing of the bunch profile which occurs as a result of measuring the LPS using a Transverse Deflecting Cavity (TCAV). We find good agreement between the simulated LPS as measured by the TCAV and the prediction from the ML model. We discuss how the predicted LPS profile compares to the actual beam LPS distribution extracted from simulation and how the resolution limits of the TCAV measurement are reflected in the ML prediction. We discuss important challenges that need to be addressed, such as quantifying prediction uncertainty, for this diagnostics to be implemented in routine accelerator operation. Finally, we report on the use of the ML-based prediction in conjunction with a standard optimizer for tuning the accelerator settings to generate a desired two-bunch LPS profile at the exit of the FACET-II linac.

INTRODUCTION

The main running configuration for PWFA experiments at FACET-II will involve accelerating two bunches from the photocathode to the interaction point (IP) at the plasma entrance with specific longitudinal profile properties and drive-witness bunch separation. For a full description of PWFA experiments at FACET-II see Ref. [2]. The major goals for the PWFA experiments will be to demonstrate pump depletion of the 10-GeV drive beam and acceleration of the witness beam to approximately 20-GeVwhile preserving good beam quality. The figures of merit for the beam quality will be preservation of energy spread and emittance of the witness bunch, and these will need to be measured on a shot-to-shot basis for both the incoming distribution and the accelerated witness beam. To this end, accurate measurements of the bunch profile entering the plasma are essential for the success of the experimental campaign. Previous work has demonstrated the feasibility of using Machine Learning

(ML) models as virtual diagnostics to predict the LPS distribution of FACET-II single bunch operation (in simulation) and at LCLS (in experiment) [3].

At FACET-II we plan to measure the LPS distribution of the electron bunch at the entrance of the plasma with an X-band TCAV operating at a peak voltage of 20 MV. This introduces a challenge for accurately characterizing the longitudinal bunch profile, as the accelerator is expected to produce very short bunches ($\sigma_z \sim 1\,\mu m$) beyond the TCAV resolution. In this work we examine the effect of the TCAV measurement on the performance of the ML-based virtual diagnostic and discuss its application in the FACET-II two-bunch operation mode.

In the following section we describe the TCAV measurement of the two-bunch configuration at FACET-II and compare the measured LPS distributions with the actual LPS which we extract directly from particle tracking simulations. The results show very good agreement between the LPS distribution measured by the TCAV and the LPS distribution predicted by the ML model. Due to TCAV resolution limits there is some discrepancy when we use the projection of the measured LPS distribution to infer the current profile at the entrance of the plasma. This discrepancy affects the accuracy of the ML-based virtual diagnostic, which predicts the LPS using the output of the TCAV measurement as training data. We present results from 3,125 simulations of the FACET-II linac from the exit of the injector to the end of the linac with induced jitter of key accelerator and beam parameters described in Table 1. The simulations include longitudinal space charge, incoherent and coherent synchrotron radiation and wakefields and are performed using the Lucretia particle tracking code [4].

SIMULATED TCAV MEASUREMENTS OF LONGITUDINAL PHASE SPACE AT FACET-II

Two-bunch Simulations

Three examples of the simulated LPS profiles as measured by the TCAV are shown in in Fig. 1 with corresponding current profiles and prediction from the ML model. The three distributions shown represent an under-compressed, over-compressed and nearly fully-compressed (nominal) beam respectively. Note that the head of the bunch is on the left of the images. The ML model we used was a three-layer fully-connected neural network with (500,200,100) neurons in each successive hidden layer and a rectified linear unit activation function for each neuron. The network was trained using the open source ML library Tensorflow, and two sepa-
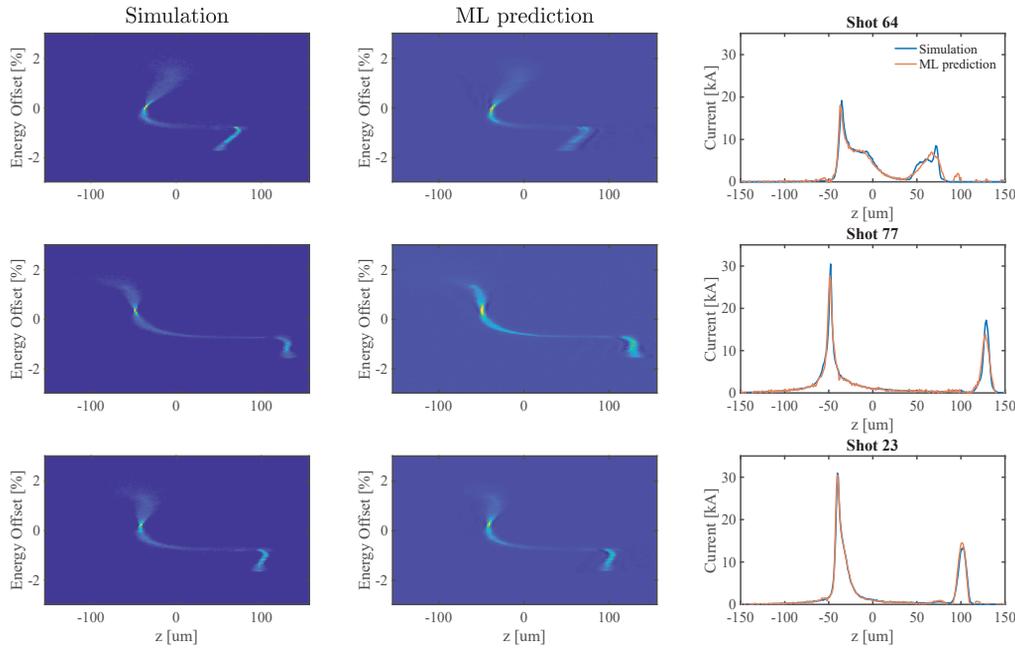

Figure 1: Example shots from numerical simulation of the FACET-II two-bunch operation mode with nominal jitter values given in Table 1. The ML model accurately predicts the LPS distribution including chirp, time separation and bunch charge ratio. The current profile matches well with what is measured on the TCAV. As shown in Fig. 2 this may deviate from the true current profile at the IP due to resolution limits of the TCAV for some high current shots.

rate models with the same architecture were trained for the 2D LPS prediction and the 1d current profile prediction. As evidenced Fig. 1, we see very good agreement between the LPS profiles measured by the TCAV and those predicted by the ML model. There is also good agreement between the ML-predicted current profiles and those extracted from the TCAV image. The variety of LPS images input to the ML model for training result from the expected shot-to-shot-jitter of linac and e-beam parameters outlined in the FACET-II Technical Design Report (TDR, see Table 1) [5]. The nominal settings produce a ~150 μm bunch spacing, a 2:1 ratio between the peak currents and a 3:1 ratio in the bunch charge between the drive and witness beam. The variation in bunch profile from shot-to-shot jitter results in a 9 % rms variation in the drive-witness charge ratio, a 30 μm rms variation in

Table 1: Linac and e-beam parameters scanned in the 5^5 simulations of the FACET-II accelerator. The ranges are chosen closely based on the jitter parameters from the FACET-II TDR [5]. The diagnostics fed to the ML model include random errors introduced artificially to approximate the measurement accuracy present in the accelerator.

Simulation Parameter Scanned	Range
L1 & L2 phase [deg]	± 0.25
L1 & L2 voltage [%]	± 0.1
Bunch Charge [%]	± 1
Input to ML model	**Accuracy**
L1 & L2 phase [deg]	± 0.25
L1 & L2 voltage [%]	± 0.05
I_{pk} at BC (11,14,20) [kA]	± (0.25,1,5)
Beam centroid BC (11,14) [m]	N/A

the bunch separation and a 36 % rms variation in the ratio of the peak current from the nominal settings. These parameter variations are well predicted by the ML model.

As discussed above, the FACET-II two-bunch configuration operates at near full compression and will generate very short bunches with rms sizes of a few μm putting them at the limit of the TCAV resolution. This means the values measured for the peak current on the TCAV sometimes differ from the values at the IP and therefore so will the prediction from the ML model, which is trained using TCAV measurements as inputs. We examine this discrepancy in detail in Fig. 2 where we show the same current profiles from the three example shots in Fig. 1 as measured on the TCAV and compare this with the current profile we calculate from the distribution at the IP binned at 0.25 μm per pixel. There are a few observation we can make by looking at Fig. 2 (a)-(c). The first is that the peak current values measured by the TCAV under-estimate the true value for shots with peak current greater than ~ 35 kA. We note that these high peak currents are greater than those which we plan to deliver for the two-bunch pump depletion experiments outlined in Ref. [2]. Nonetheless, close to the nominal settings (as shown in Fig. 2(c)) the correct value of the ratio of the peak currents may be under-estimated if the witness bunch current profile is poorly resolved by the TCAV measurement. To quantitatively understand the limits imposed by the TCAV measurement we can estimate the longitudinal resolution as follows:

$$\sigma_z = \frac{E_e}{eV_{rf}k_{rf}|\sin \Delta\psi|} \frac{\sqrt{\sigma_S^2 + \beta_S \epsilon}}{\sqrt{\beta_T \beta_S}} \quad (1)$$
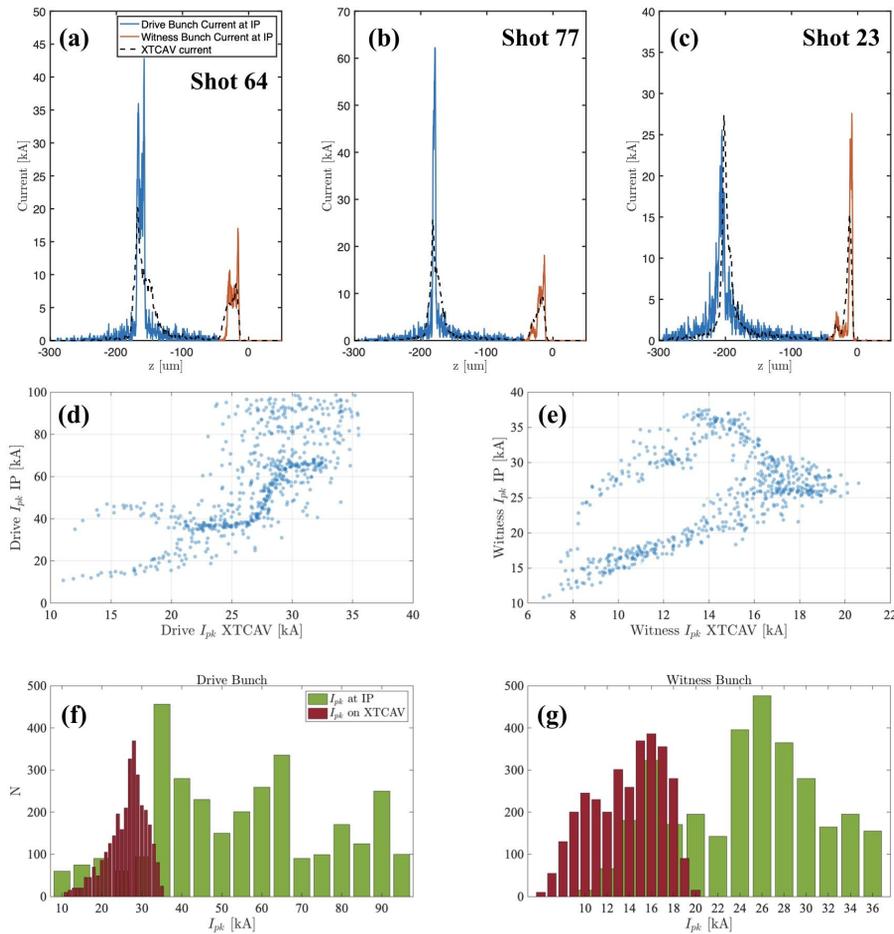
Figure 2: (a)-(e) Single shot comparisons of the current profiles measured by the TCAV and calculated from the macroparticle distribution in simulation dumped at the IP. The shots match those displayed in Fig. 1. (d)-(f) Comparison of the peak current measured by the TCAV vs. at the IP for the drive and witness beam. The plots show that the TCAV measurement underestimates the peak current value and smears out some details of the current distribution for the very short bunches which will be produced at FACET-II.

where E_e is the electron beam energy V_{rf}, k_{rf} is the TCAV voltage and wavenumber, $\Delta\psi$ is the phase advance between the TCAV and the measurement screen, σ_S is the resolution of the screen (we assume 4 μm for a transition radiation target), β_S is the beta function at the screen ϵ is the beam emittance and β_T is the beta function at the TCAV. The ~35 kA max resolvable peak current come from the constrained optimization of the beta function at the screen and at the TCAV while meeting the beam stay-clear constraints in the experimental area and mitigating loss of resolution from chromatic errors and emittance growth in the transport. For a 10 μm normalized emittance at 10 GeV with a phase advance of $3\pi/2$ between TCAV and screen, the optimized values of β_T and β_S are 107 and 6.5 m give a resolution of $\sigma_{z,min} = 4.58$ μm. Given a Gaussian drive bunch at 1.5 nC charge this corresponds to $I_{max} = 39.2$ kA which is in reasonable agreement with the trend shown in the scatter plot in Fig. 2(d).

For shots that are not beyond the TCAV resolution we can see from Fig. 2(d)-(e) that we can correlate the TCAV measured peak current with the peak current at the IP. These shots are mostly in the region defined by $I_{pk,drive} < 30$ kA

and $I_{pk,wit} < 16$ kA as measured on the TCAV. Some shots in this region still show large discrepancy between the TCAV current profile and that measured at the IP and these represent the spiky 'double-horn' type distributions in the drive and witness beam exemplified in Fig. 2(a).

One of the challenges this particular virtual diagnostic faces is to flag wether or not a single shot falls within the 'high-current' region beyond the TCAV resolution. Accurately determining this on a shot-to-shot basis will provide added assurance that the current profiles predicted by the ML model map to the electron beam current profile at the IP. One potential method to address this challenge would be to use a secondary non-destructive diagnostic in tandem with the ML prediction that is sensitive to changes in the peak current beyond the TCAV resolution. This would help identify the region in which a given shot falls. The secondary diagnostic may be a mid-IR and/or Thz spectrometer similar to those described in Ref. [6, 7] and could use diffraction or bend radiation as a non-destructive source. It may also be possible to implement a simple upgrade (adding an appropriate set of spectral filters) to the existing radiation-based bunch length monitor at the exit of the final bunch compressor (see

THBO01

Ref. [8]) to mimic a more complicated spectroscopic measurement. This would allow us to measure the integrated radiation signal over a given frequency band proportional to the bunch length for the high peak current shots. These options are currently under study and details on progress will be reported in future work.

OPTIMIZATION OF LPS PROFILE USING ML PREDICTION

One application of an accurate ML based virtual diagnostic for predicting the output LPS distribution is tuning of accelerator parameters to tailor the LPS to a desired distribution. This can be done in two different ways:

- Using a ML-based inverse model to predict the machine settings required to generate a distribution (and potentially use a conventional optimizer starting from the initial ML guess), or

- Using a conventional optimizer with the ML-prediction of the LPS as its input and the accelerator settings as output.

One advantage of both approaches for online optimization of the accelerator compared to using an online simulation of the accelerator (e.g. using a tracking code [9]) is the speed with which ML models can produce predictions ($\sim$ ms timescale). This can be orders-of-magnitudes faster than running start-to-end tracking simulations, which can take seconds or minutes to run depending on the physics effects which are included in the code [10].
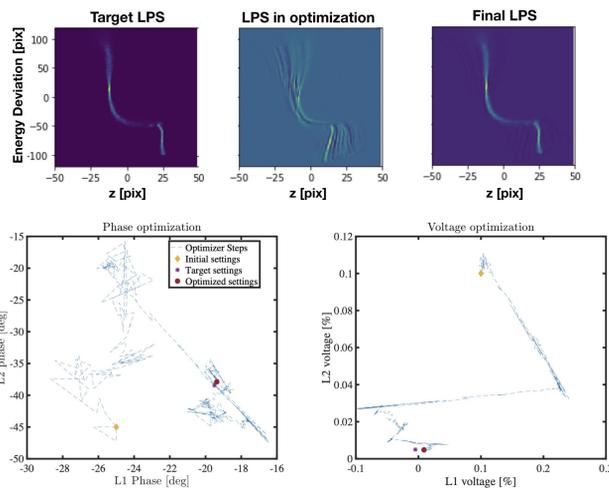

Figure 3: (Top Row) Target LPS profile, the LPS prediction at one iteration during the optimization of settings and the LPS prediction with the final optimized settings. (Bottom row) Trajectory of settings suggested by the Nelder-Mead optimizer with a ML-based prediction of the LPS as inputs.

The first approach was prototyped at LCLS as described in Ref. [11] with successful results. Here we apply the second approach to the FACET-II two-bunch operation case. In order to use the ML model in an iterative feedback which acts only on controllable machine parameters we re-train

the virtual diagnostic using only L1 and L2 amplitudes and phases as inputs to the model. A cost function is then defined using the structural similarity index between the target and the given LPS profile [12] and is minimized using a Nelder-Mead downhill simplex algorithm, which is called iteratively until a given convergence criterion is met. We can see from the bottom row of Fig. 3 that the final solution for the parameters is very close to the target parameters which generated the phase space distribution and the achieved distribution is very similar to the target. Another important feature of this example is that the starting point for the neural network optimization lies outside the training range for both the L1-2 amplitudes and phases, whose set points were -19.2 and -38.35 degrees for the phases and zero for the amplitude. Notably, the optimization algorithm is still robust when initialized with starting points outside of the training data set. We can also conclude from the example that the ML model is able to successfully interpolate within the training set, as the settings suggested by the optimizer and the predictions returned by the ML model are not limited to the discretized points sampled by the numerical simulation inputs supplied as training data.

Robust implementation of the ML-based LPS prediction model in tandem with a conventional optimizer requires an associated estimate of the uncertainty of each prediction as well as the risk involved in making each adjustment in settings to regular accelerator operations. Model ensembling, MC dropout techniques, and Bayesian neural networks are all being considered as possible options for assessing the prediction uncertainty on a shot-to-shot basis. These questions will be addressed in future work.

CONCLUSION

We have discussed the application of using ML-based methods to predicting the LPS of the FACET-II accelerator non-destructively and on a single-shot basis for the two-bunch operation mode. We have also explored using these predictive models with optimization algorithms to tailor the LPS distribution. Compared to the feasibility study presented in Ref [3] we have examined in detail the impact of the TCAV measurement on the accuracy of the virtual diagnostic prediction, especially for shots which lie beyond the TCAV resolution. We have found that for the existing FACET-II configuration there is very good agreement between the measured TCAV LPS profiles and those predicted by the ML-based virtual diagnostic. For the current profile, the resolution limits of the TCAV result in a smeared out profile, with values of the peak current deviating from those extracted directly from the particle distribution in simulations. The deviation is most severe for very high peak current shots with $I_{pk,drive} > 30\,\mathrm{kA}$ and $I_{pk,wit} > 16\,\mathrm{kA}$. We have suggested flagging these shots with additional radiation-based diagnostics which would be sensitive to changes in the bunch profile beyond the TCAV resolution. Incorporating such diagnostics in the FACET-II beamline is planned and the results will be reported on in the future.

REFERENCES

[1] A. L. Edelen *et al.*, in *Proc of IPAC 2018*, paper THYGBE2.

[2] C. Joshi *et al.*, *Plasma Phys. and Contr. Fus.*, vol. 60, p. 034001, 2018.

[3] C. Emma, A. Edelen, *et al.*, *PRAB*, vol. 21, p. 112802, 2018.

[4] Tennenbaun, P., in *Proceedings of PAC 2005*, Knoxville, May, 2005, pp. 4197-4199.

[5] Technical design report for the FACET-II project at SLAC National Accelerator Laboratory, `http://www.slac.stanford.edu/pubs/slacreports/reports19/slac-r-1072.pdf`

[6] S. Wesch, B. Schmidt, C. Behrens, H. Delsin-Hashemi, and P. Schmuser, *NIMA*, vol. 665, p. 40-47, 2011.

[7] T. Maxwell *et al.*, *PRL* vol. 111, p. 184801, 2013.

[8] H. Loos *et al.*, SLAC-PUB 12619, 2007.

[9] A. Scheinker and S. Gessner, *PRAB* vol. 18, p. 102801, 2015.

[10] A. Edelen, A. Adelmann, N. Neveu, Y. Huber, and M. Frey, 2019. `arXiv:1903.07759`

[11] A. Scheinker, A. Edelen, D. Bohler, C. Emma, and A. Lutman, *PRL*, vol. 121, p. 044801, 2018.

[12] Z. Wang, E.P. Simoncelli, and A.C. Bovik, "Conference Record of the Thirty-Seventh Asilomar Conference on Signals", *Systems and Computers*, vol. 2, pp. 1398-1402, 2004.