ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-THPL019

MACHINE LEARNING FOR COMBINED SCALAR AND SPECTRAL LONGITUDINAL PHASE SPACE RECONSTRUCTION*

J. Kaiser†, A. Eichler, S. Tomin, Z. Zhu
Deutsches Elektronen-Synchrotron DESY, Germany

Abstract

Longitudinal beam diagnostics are a useful aid during tuning of particle accelerators, but acquiring them usually requires destructive and time intensive measurements. In order to provide such diagnostics non-destructively, computational methods allow for the development of virtual diagnostics. Existing Fourier-based reconstruction methods for longitudinal current reconstruction, tend to be slow and struggle to reliably reconstruct phase information. We propose using an artificial neural network trained on data from a start-to-end beam dynamics simulation to combine scalar and spectral information in order to infer the longitudinal phase space of the electron beam. We demonstrate that our method can reconstruct longitudinal beam diagnostics accurately and provide the reconstructed data with adaptive resolution. Deployed to control rooms today, our method can help human operators reduce tuning times, improve repeatability and achieve pioneering working points. In the future, ML-based virtual diagnostics will help the deployment of feedbacks and autonomous tuning methods, working toward the ultimate goal of autonomous particle accelerators.

INTRODUCTION

Knowledge and control of the shot-to-shot longitudinal phase space (LPS) of an electron beam is critical to achieve the performance requirements of modern particle accelerators. However, conventional methods of measuring the LPS, such as using a transverse deflecting cavity (TDS), are destructive measurements, meaning that the LPS cannot be measured during beam delivery to experiments. Furthermore, these measurements can take a long time to set up, often requiring extensive calibration before accurate data can be taken. All of this means, that the LPS is either not taken advantage of during accelerator operation, leading to inferior beam quality, or that a lot of time is spent measuring the LPS during accelerator setup and tuning, significantly reducing the time available for useful beam operations.

Methods have been developed to reconstruct the LPS from other non-invasive measurements. At European XFEL, a Fourier-based method [1] is deployed. This method uses spectral information from the THz spectrometer *CRISP* to reconstruct the current profile of each bunch. While this method works and has shown to have good accuracy in operation, it lacks phase information, meaning that it must take intelligent guesses regarding the phase. This can lead to various faults in the reconstructed current profile such as a flip in the longitudinal dimension. Furthermore, this method is expensive to compute. With up to 2700 bunches per train and 10 Hz operation, this makes real-time bunch-by-bunch reconstruction of the entire bunch train infeasible. Unlike other methods mentioned in the following, this method can also not reconstruct the full 2-dimensional LPS image, but only the 1-dimensional current profile.

Artificial neural networks (ANNs) are general purpose function approximators that have, in recent years, been shown to be capable of achieving incredible results in numerous domains [2–5]. As a result, the particle accelerator community has also started to adopt ANNs in a number of different capacities, ranging from surrogate modelling [6] to autonomous control and optimisation [7] of accelerators. Among these use cases for ANNs in the particle accelerator domain are *virtual diagnostics*, where learned surrogate models provide information about the state of an accelerator that would otherwise be difficult or even impossible to acquire using a physical diagnostic. The provided information can give useful data about experiments, help operators tune the accelerator as well as jump-start the development of intelligent controllers for autonomous tuning [7, 8].

In previous work, ANNs have been used to learn both so-called *scalar* [9–11] and *spectral*[12, 13] virtual diagnostics for 1-dimensional and 2-dimensional LPS reconstruction, using scalar inputs, such as radio frequency (RF) settings, or spectral inputs, such as THz radiation spectra, to infer the LPS. Both neural scalar and spectral virtual diagnostics improve on the Fourier-based method in that they are more efficient to compute, especially on the right hardware, than the Fourier-based method. Furthermore, they allow for the reconstruction of the full 2-dimensional phase space image instead of just the 1-dimensional current profile. Neural scalar and spectral virtual diagnostics, however, each have disadvantages. Spectral virtual diagnostics, just like the Fourier-based reconstruction, does not have access to phase information. Instead, being a machine learning method, the ANN must learn statistics over the training dataset to make more educated guesses on the phase information. As a result, neural network-based spectral virtual diagnostics are likely to fail in cases, where the phase is one that occurs less often. Scalar virtual diagnostics, on the other hand, function more like a surrogate model and therefore have access to phase information. They are however susceptible to false RF setting readbacks [9, 10, 12], such that small calibration errors can lead to very wrong LPS reconstructions.

In this work, we present an ANN-based combined scalar-spectral virtual diagnostic, addressing the shortcomings of either scalar and spectral virtual diagnostics, for reconstructing both the 1-dimensional current profile and the 2-

† jan.kaiser@desy.de

ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-THPL019

dimensional LPS at the European XFEL at DESY in Hamburg, Germany, from RF settings and the THz spectrum. Furthermore, we introduce an additional feature to our method, that allows it to achieve high adaptive resolution for bunches covering both large and small areas of the LPS.

COMBINED SCALAR- AND SPECTRAL VIRTUAL DIAGNOSTICS

At the European XFEL, we can combine scalar and spectral data for creating a virtual longitudinal phase space diagnostic. RF settings influence the beam's current profile and can therefore provide information about the latter. The RF readbacks are, however, not always perfectly reliable and a purely scalar diagnostic is therefore not be very robust. At the same time, the THz spectrum measured by the THz spectrometer CRISP installed at European XFEL delivers information about the frequencies present in in the beam's current profile, but it is missing phase information. Any reconstruction using purely spectral diagnostics will therefore struggle to accurately infer the current's phase.

We propose using a neural network machine learning model that is trained on data from a start-to-end beam dynamics simulation to combine scalar and spectral information in order to infer either the current profile or the 2-dimensional LPS. To this end, we construct an ANN that receives two inputs: A 5-dimensional vector of RF settings $(\alpha, \rho, \psi, \alpha_{L1}, \alpha_{L2})$ and a 240-dimensional THz formfactor sampled from 0.7 THz to 58 THz. These are then concatenated and passed through a multilayer perceptron (MLP), outputting either 300 ordered samples of the current along the longitudinal dimensions, or a 300 by 300 image of the LPS. The ANN is trained in a supervised setup to minimise the difference between the reconstructed current profile or LPS from the ground truth current profile or LPS generated using an end-to-end simulation.

ACHIEVING HIGH RESOLUTION AT ANY SCALE

When predicting the current profile or the LPS, one has to choose the sample positions on each dimension for which the model is expected to predict the current or LPS magnitude. There is, however, a trade-off when choosing the samples to use. Bunches may be very long or very short. In order to cover a large bunch, the samples must be spread out over that wide range. In order to cover short bunches with adequate resolution, one must then have a large number of samples. Not only does this needlessly increase the complexity of the ANN model, it also means that a lot of the time, most of the ANN outputs are not predicting any useful information. This can lead to the model learning to output 0.0 on these simply because this is correct most of the time, instead of even attempting to learn the correct function for that particular sample.

We solve this problem by adding a second output to the final layer of the ANN that is either the length of the bunch along the longitudinal dimension when only reconstructing

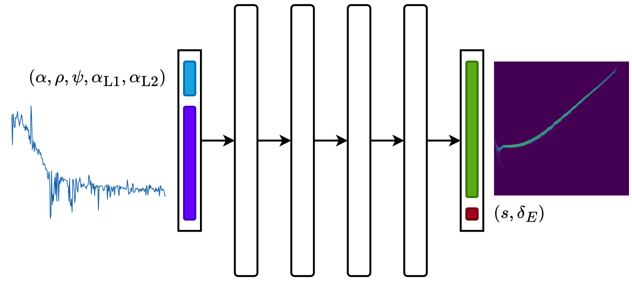

Figure 1: Flowchart of the ANN architecture used for reconstructing the LPS image. The RF settings and THz are input to the ANN on the left. The information then flows through multiple hidden layers, before the LPS image and its extents in both dimensions are output on the right.

the current profile, or the 2-dimensional range of the bunch in both dimensions of the LPS. The 300 samples are then equidistantly arranged over this range output by the same model. This way, the resolution of our model output adapts automatically to the length of the bunch, always resolving all relevant features, while ensuring that the entire bunch is captured. To our knowledge none of the previous works has done this, instead relying on fixed samples.

The resulting ANN model has the same inputs and outputs as described in the previous section, but adds either a single scalar value denoting the longitudinal length of the bunch or a tuple (s, δ_E) denoting the full 2-dimensional range of the LPS to the output. The model $\mathscr{M}$ is given as

$$\hat{y}_{\mathrm{LPS}}, \hat{y}_{\mathrm{range}} = \mathscr{M}\left(x_{\mathrm{RF}}, x_{\mathrm{THz}}\right)$$

In between the input and the output layer, there are from 2 to 6 hidden layers, as is illustrated in fig. 1. Each hidden layer may be followed by batch normalisation before the activation. The outputs use a different activation from hidden layers. Namely, a *Softplus* activation is chosen as the output activation function. The latter ensures that all outputs are strictly positive, as they physically should be, while also ensuring that there remains some gradient when the output should be zero according to the ground truth. If a rectified linear unit (ReLU) activation function were used on this output, gradients may be lost when 0 is output, which results in fully trained models that have single 0 samples where they stopped learning. We use an mean squared error (MSE) loss function over each output where both outputs are weighted the same as in

$$L(y, \hat{y}) = \mathrm{MSE}\left(y_{\mathrm{range}}, \hat{y}_{\mathrm{range}}\right) + \mathrm{MSE}\left(y_{\mathrm{LPS}}, \hat{y}_{\mathrm{LPS}}\right)$$

with

$$\mathrm{MSE}(y, \hat{y}) = \frac{1}{n}\sum_{i=0}^{n}\left(y_i - \hat{y}_i\right)^2$$

For training, an *Adam* [14] optimiser is used with $\beta_1 = 0.9$ and $\beta_2 = 0.999$. Hyperparameters such as the number and width of hidden layers, or the learning rate were tuned

ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-THPL019

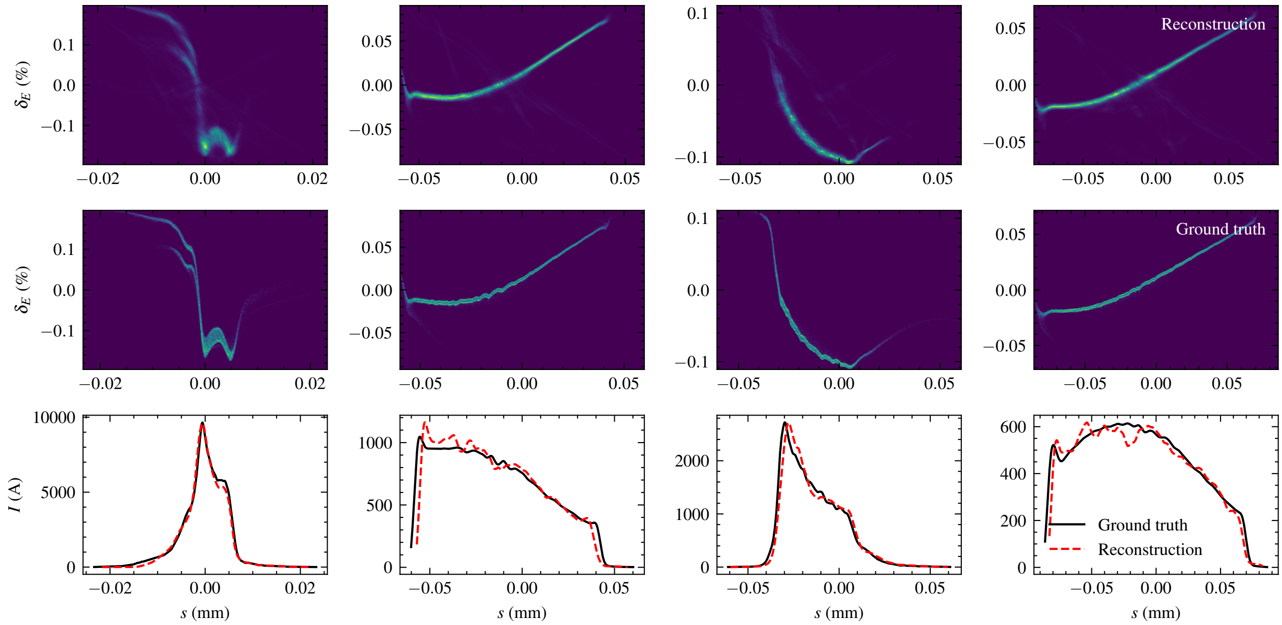

Figure 2: Comparison of ground truth and reconstructed 2-dimensional LPS images and 1-dimensional current profiles. Each column shows the samples belonging together. The 1-dimensional current profiles were not derived from the LPS images, but instead reconstructed by a purpose-built ANN.

Table 1: Hyperparameters Used During Training of the Reconstruction Models

Parameter	**Current profile**	**LPS**
# Hidden layers	3	4
Hidden layer width	267	681
Hidden activation	ReLU	ReLU
Batch normalisation	Yes	Yes
# Training epochs	76	95
Learning rate	0.003	0.0008
Batch size	66	70

using Bayesian optimisation to minimise the loss over the validation dataset. The final used hyperparameters for both 1-dimensional current reconstruction and 2-dimensional LPS reconstruction are listed in table 1. The training dataset used contained 32 000 samples and was randomly split into training, validation and test sets according to the ratios (80 %, 10 %, 10 %). The RF settings and formfactors are normalised over the training dataset follow a standard normal distribution. Current profiles, bunch lengths, LPS images and LPS image extents are normalised to a range of $(0, 1)$.

As can be seen in fig. 2, the output of our model matches the ground truth very well. Furthermore, unlike previous methods, our method produces no discernible artefacts.

CONCLUSION

In this work, we presented a machine learning method using neural networks, that is capable of reconstructing both current profiles and 2-dimensional LPS images with great accuracy using readily available scalar and spectral diagnostics at European XFEL. Furthermore, our method is able reconstruct current profiles and LPS images at the resolution best suited to their extent.

The reconstruction provided by our method can be provided, for example in the form of a GUI application, to human operators in order to assist them with accelerator setup and tuning. In addition, our method may also be used as an input to advanced autonomous tuning methods, reducing the engineering effort required to make such methods operational.

Future work on virtual diagnostics for LPS reconstruction may focus on improving reliability of the ANN model as well as the clarity of details in the reconstructed signals. This may, for example, be achieved using convolutional neural networks (CNNs), generative adversarial networks (GANs) or better suited loss functions. Moreover, the developed methods may also be deployed to other virtual diagnostics reconstructing different signals that may help operate future particle accelerators and push the limits of their operability.

ACKNOWLEDGEMENTS

This work has in part been funded by the IVF project InternLabs-0011 (HIR3X). The authors acknowledge support from DESY (Hamburg, Germany), a member of the Helmholtz Association HGF, as well as support through the *Maxwell* computational resources operated at DESY.

REFERENCES

[1] N. M. A. Lockmann, 'Noninvasive measurements of elec-

ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-THPL019

tron bunch current profiles with few-femtosecond resolution at mhz repetition rates', Ph.D. dissertation, Staats-und Universitätsbibliothek Hamburg Carl von Ossietzky, 2021.

[2] S. Bubeck *et al.*, *Sparks of artificial general intelligence: Early experiments with GPT-4*, 2023.

[3] J. Degrave *et al.*, 'Magnetic control of tokamak plasmas through deep reinforcement learning', *Nature*, vol. 602, pp. 414–419, 2022. doi:10.1038/s41586-021-04301-9

[4] Y. Li *et al.*, 'Competition-level code generation with AlphaCode', *Science*, vol. 378, no. 6624, pp. 1092–1097, 2022. doi:10.1126/science.abq1158

[5] A. Vaswani *et al.*, 'Attention is all you need', *Advances in neural information processing systems*, vol. 30, 2017.

[6] C. Xu, E. Bründermann, A.-S. Müller, A. S. Garcia and J. Schäfer, 'Surrogate Modelling of the FLUTE Low-Energy Section', in *Proc. IPAC'22*, Bangkok, Thailand, 2022, pp. 1182–1185. doi:10.18429/JACoW-IPAC2022-TUPOPT070

[7] J. Kaiser, O. Stein and A. Eichler, 'Learning-based optimisation of particle accelerators under partial observability without real-world training', in *Proceedings of the 39th International Conference on Machine Learning*, vol. 162, 2022, pp. 10 575–10 585. https://proceedings.mlr.press/v162/kaiser22a.html

[8] C. Xu, T. Boltz, A. Mochihashi, A. Santamaria Garcia, M. Schuh and A.-S. Müller, 'Bayesian optimization of the beam injection process into a storage ring', *Physical Review Accelerators and Beams*, vol. 26, p. 034 601, 3 2023. doi:10.1103/PhysRevAccelBeams.26.034601

[9] 'Machine learning-based longitudinal phase space prediction of particle accelerators', *Physical Review Accelerators and Beams*, vol. 21, 11 2018. doi:10.1103/PhysRevAccelBeams.21.112802

[10] C. Emma *et al.*, 'Machine learning-based longitudinal phase space prediction of two-bunch operation at FACET-II', 2019. doi:10.18429/JACoW-IBIC2019-THBO01

[11] Z. Zhu, S. Tomin and J. Kaiser, 'Application of Machine Learning in Longitudinal Phase Space Prediction at the European XFEL', Triest, Italy, Aug. 2022, presented at FEL2022, Triest, Italy, Aug. 2022, paper WEP12, unpublished.

[12] A. Hanuka *et al.*, 'Accurate and confident prediction of electron beam longitudinal properties using spectral virtual diagnostics', *Scientific Reports*, vol. 11, 1 2021. doi:10.1038/s41598-021-82473-0

[13] J. Zhu, N. M. Lockmann, M. K. Czwalinna and H. Schlarb, 'Mixed diagnostics for longitudinal properties of electron bunches in a free-electron laser', *Frontiers in Physics*, vol. 10, 2022. doi:10.3389/fphy.2022.903559

[14] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, 2017.