ISBN: 978-3-95450-217-2 doi:10.18429/JACoW-ERL2019-WECOYBS04

# COMMISSIONING OF THE bERLinPro DIAGNOSTICS LINE USING MACHINE LEARNING TECHNIQUES*

B. Kuske†, Helmholtz-Zentrum für Materialien und Energie, Berlin, Germany,
A. Adelmann, Paul Scherrer Institut, Villigen, Switzerland

*Abstract*

bERLinPro is an Energy Recovery Linac (ERL) project currently being set up at HZB, Berlin. It is intended as an experiment in accelerator physics, to pioneer the production of high current, low emittance beams in a fully super-conducting accelerator, including SRF gun, booster and linac. RF-Commissioning of the gun is planned in mid 2020, [1]. HZB triggered and partially supported the development of release 2.0 of the particle tracking code OPAL [2]. OPAL is set up as an open source, highly parallel tracking code for the calculation of large accelerator systems and many particles. Thus, it is destined for serving attempts of applying machine learning approaches to beam dynamic studies, as demonstrated in [3]. OPAL is used to calculate hundreds of randomized machines close to the commissioning optics of bERLinPro. This data base is used to train a polynomial chaos expansion, as well as a neural network, to establish surrogate models of bERLinPro, much faster than any physical model including particle tracking. The setup of the sampler and the sensitivity analysis of the resulting data are presented, as well as the quality of the surrogate models. The ultimate goal of this work is to use machine learning techniques during the commissioning of bERLinPro.

## INTRODUCTION

As any linear accelerator, an ERL is an initial value problem: without exact knowledge of the initial parameters of the beam, a later understanding and characterization of the beam parameters is difficult. Therefore, a thorough understanding of the gun is indispensable. Before the gun is assembled and tested, many ambiguities exist, starting from the actual energy of the beam, i.e. the gun field reachable with tolerable field emission, over the bunch parameters, to the system parameters leading to successful acceleration. As the system is heavily space charge dominated with bunch charges of 77 pC, tracking calculations including space charge take minutes to hours, depending on the size of the considered structure, the number of particles and the grid size, even with a highly parallized code like OPAL. It is tempting to try to replace the tracking calculations by surrogate models, that deliver answers very close to tracking results in much shorter time, in the order of milliseconds. That would enable 'online modeling' in the control room, where the surrogate model is fed by machine parameter read-backs and would deliver expectation results for beam measurements. The paper presents first steps in setting up surrogate models for the diagnostics line of bERLinPro, where the gun is characterized. It makes strong use of earlier work and experience, published in [3] and [4]. It is intended to use surrogate models to ease and to speed up commissioning.

* Work supported by German Bundesministerium für Bildung und Forschung, Land Berlin and grants of the Helmholtz Association
† bettina.kuske@helmholtz-berlin.de

## DIAGNOSTICS LINE

The diagnostics line consists of the 1.3 GHz, 1.4 cell, single cavity SRF gun, providing up to 3 MeV electrons with a design bunch charge of 77 pC. The gun module also hosts two corrector coils (H/V) and a cold solenoid. The booster, hosting three two-cell cavities can boost the energy up to 6.5 MeV. The first cavity imprints a chirp on the bunch for velocity bunching, while the other two cavities are run on crest for acceleration. Further elements are 6 quadrupoles, a transverse deflecting cavity, a spectrometer followed by a 300 W Faraday cup, or, straight ahead, a 35 kW beam dump, Fig. 1. Optics were developed including the booster (6.5 MeV) and with three booster replacement quadrupoles (taken from the recirculator) and 2.7 MeV. Four beam position monitors (BPM) and two screens (FOM) are available for diagnostics.

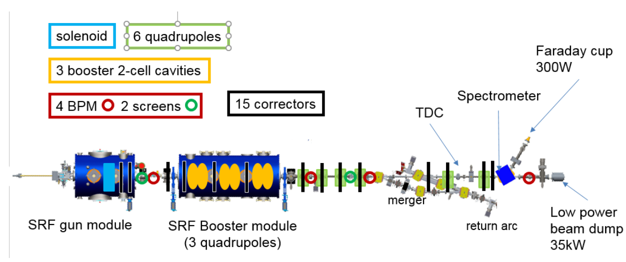


Figure 1: The diagnostics line is intended for the characterization of gun and booster and initial beam parameters.

### *Independent Modeling Parameters*

The ambiguities in the gun parameters, prior to commissioning, are caused by:

- production uncertainties like cavity geometry and field flatness, i.e. the gun field
- changes during cool down of SRF structures (f.e. cathode retraction position)
- ambiguities before first use (f.e. achievable max. gun field amplitude).

In addition, the cathode laser pulse length, spot size and intensity, the solenoid strength, the phases and amplitudes of three booster cavities and the field of the six quadrupoles in the beam line will determine the bunch properties. The transverse beam size can be measured on the two screens and four BPMs. The transverse deflecting cavity (TDC) and the spectrometer in combination enable the measurement of

ISBN: 978-3-95450-217-2 doi:10.18429/JACoW-ERL2019-WECOYBS04



the full, 6D phase space. The emittance will be determined by quadrupole scans using the first three merger quadrupoles. See Fig. 2, where the 19 independent parameters are indicated. In order to predict the expected bunch properties for all possible combinations of set points of the independent parameters, huge scans would be necessary. Even using only three set points for each parameter, $3^{19} = 1.16 \times 10^9$ possible combinations arise.

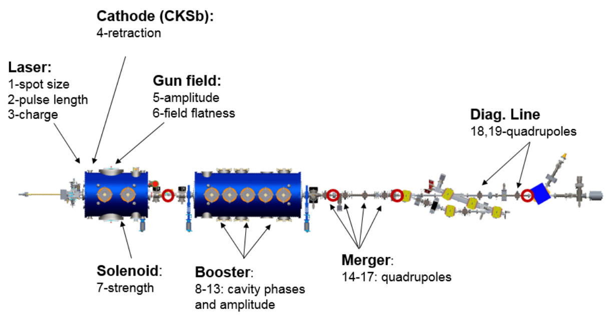


Figure 2: The most relevant 19 parameters contributing to the physics model of the diagnostics line of bERLinPro.

Therefore, there is a strong need to minimize the number of independent parameters to ease commissioning and enable modeling. In order to thread the beam successfully through the diagnostics line and enable first bunch measurements, the following restrictions can be taken:

- Bunch charge, current: During commissioning, the current will be limited to ≈ 1 µA defined by the tolerable load on the screens. Measurements can be taken by much lower currents, down to nA. Calculations show, that the limit for the bunch charge to eliminate space charge effects in the reference optics is about 0.1 pC corresponding to 1 µA in single bunch mode at 10 Hz. Low bunch charge also eliminates the dependance on the laser intensity.
- Quadrupole setting: The quadrupoles in front of the screen in the merger can be set in a way to produce a round beam on the screen. This is easily controllable. Without space charge, the settings can be linearly scaled to the achieved energy.
- Booster amplitudes and phases: We assume, that these parameters can be well measured and keep them fixed in the simulations.
- The two last quadrupoles serve to adjust the transverse beam size into the dump. They are located behind the diagnostics elements and need not be part of the modeling.

This leaves six independent variables for the first commissioning phase. The expected variation range and set points used in the sampling run, are listed in Table **??**. The five cathode retreat positions and three field flatnesses are encoded in 15 CST field files.

### *OPAL*

OPAL is one of the few tracking codes, that includes a 3D space charge routine, coherent synchrotron radiation effects and is open source. OPAL was planned from the beginning as a highly parallelized tracking code for the calculation of large accelerator systems and many particles, and its underlying structure has been developed by computer scientists and mathematicians. HZB triggered and partially supported the development of release 2.0 of OPAL, as no other code was available at the time, that was suitable for the modeling of a high current, low energy ERL, such as bERLinPro. OPAL can now cope with with specifics like double passage through elements, and thus is a very adequate tool to study this type of machines.

Table 1: Range and Number of Equidistant Set Points of the Six Independent Parameters

| Parameter | Range | Set point |
|---|---|---|
| Laser pulse length | ± 10% | 3 |
| Laser spot size | ± 10% | 3 |
| Gun field flatness | 100, 113, 130% | 3 |
| Cathode retreat | 0.5-1.5mm | 5 |
| Gun voltage | 15-17MV/m | 5 |
| Solenoid field | 0.035-0.065T/m | 7 |
| combinations | | 4725 |

OPAL also offers a sampler option: in a single run it allows for random, raster or mixed sampling of parameters with different distribution functions. The results are stored in a separate directory. This is a big advantage compered to other optimization procedures, like swarm calculations or generic optimizers, where most of the calculated data is deleted and not accessible for later studies. Using the sampler, a set of 4725 different realizations of the diagnostics line has been calculated over the complete length of 14 m. The calculations included space charge (to cover cases with unintended beam waists) and 100.000 particles. On the PSI cluster the run took 5 h, using 320 cores. It should be mentioned, that this computational effort has to be invested only once. The derived surrogate model, though, can be used for further optimization and as an on-line model. Once the data is available, it can be used in versatile ways:

- Data mining
- Polynomial chaos expansion
- Neural networks
- Other boost or regression algorithm

The latter three all create a surrogate model, that approximates the parameters of interest, but is MUCH faster than the under-lying tracking calculations.

## RESULTS

### *Data Mining*

Data mining is a rich source for understanding the basic mechanisms of the system, especially, if expectations based on accelerator physics knowledge can be retrieved in the data. As an example, one can look at the cases where particles get lost in the diagnostics line. From optics development it is known, that the rf focusing of the booster is strong, and

ISBN: 978-3-95450-217-2 doi:10.18429/JACoW-ERL2019-WECOYBS04

the transverse beam size at the booster entrance is a critical parameter. Figure 3 shows the number of stable particles at 10 m (left) and at 14 m (right) as a function of the rms beam size at the beginning of the booster (round beam). Clearly, all beams with transverse sizes larger than 1.5 mm will loose particles before reaching the dump.

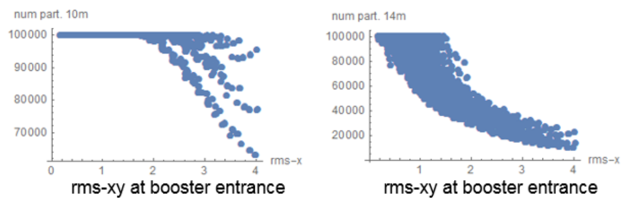


Figure 3: Number of stable particles at 10 m, left, and at 14 m, right. There is a clear dependance on the transverse beam size at the booster entry.

The correlation between four of the independent input parameters and the beam size at the beginning of the booster is displayed in Fig. 4. The distinct set values are visible. Surprisingly, the cathode position has a larger impact on the transverse beam size, than the laser spot size (within the ranges studied). In order to reach transverse beam sizes ≤1.5 mm at the entrance of the booster and secure particle transport into the dump, the setting of the solenoid is the most critical parameter. It has not been expected, that setting the solenoid field to 0.055 T/m guarantees adequate beam sizes for all other cases (red line). Selecting only cases with the optimal solenoid setting, expectation values can now be extracted for other beam parameters.

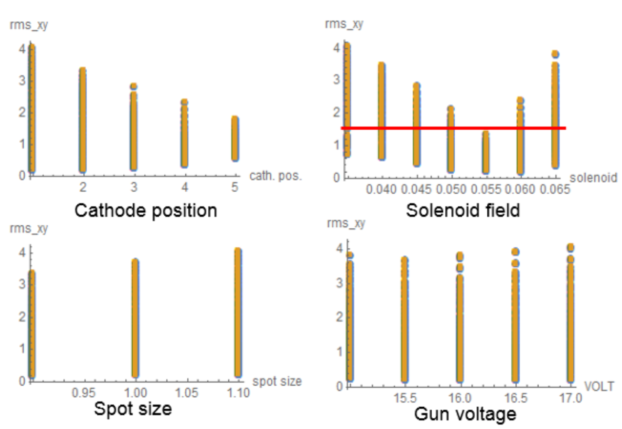


Figure 4: The transverse beam size at the entrance to the booster is displayed as a function of the cathode position, the solenoid setting, the laser spot size and the gun voltage. The red line indicates the goal of $\sigma \leq 1.5$ mm.

## *Polynomial Chaos Expansion*

Polynomial chaos expansion (PCE) and its application in accelerator physics is explained in detail in [4]. PCE is applicable, whenever system parameters with a given uncertainty or distribution are mapped by a physical model or evolution, onto quantities of interest (QoI), that depend on these system parameters. As f.e. the laser pulse length entering tracking calculations will lead to beam parameters like the bunch length, that depend on the pulse length and will have a distribution, that will depend of the distribution of the laser pulses, Fig. 5.

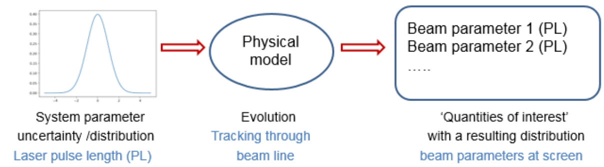


Figure 5: Polynomial Choas Expansion is a method to determine evolution in dynamical systems, when there is probabilistic uncertainty in the system parameters.

It can be shown, that:

- the distribution space of the QoI always has an orthonormal polynomial basis, i.e. every element in the QoI space can be represented as a polynomial sum
- the kind of polynomials to be used depends on distribution of the independent variables (uniform ⇔ Lagrange, Gaussian ⇔ Hermite, … )
- the determination of the sum coefficients reflects the model involved (regression technique)
- PCE is applicable in multi-dimensions and for mixed distributions

The mathematics behind PCE is conveniently wrapped up in the python package 'chaospy', [5].

The PCE will always be trained with a subset of the available samples, in order to check the results with the left over testing samples. The quality of the PCE model can be represented by plotting the results for a single QoI calculated by tracking and calculated by PCE against each other. For a perfect model all points would lie on the diagonal. Figure 6 shows the results for the horizontal emittance for different sizes of training sets. While the prediction of the model for the training samples is quite independent of the number of samples used, the results for the testing samples differ strongly. 1500 training samples were found to be sufficient, no further improvement was found for more samples.

A further optimization parameter is the order of the polynomials used in the expansion. Figure 7 compares the results for the horizontal rms beam size at the screen in the merger for third and sixth order polynomials. Blue dots represent the results for the training data and red dots that for testing samples. Sixth order polynomials show sufficient accuracy.

The accuracy of the PCE model might differ for different parameters. This reflects the sensitivity of the QoI in the system. The beam size in Fig. 7 is taken at the screen in the merger. The beam goes through a waist close to the screen for many samples. The position of the waist, and thus the beam size measured on the screen, is sensitive to small changes in all focusing parameters. For comparison, the results for the bunch length (left) and the energy (right) are shown in Fig. 8. While the energy is perfectly matched independent of the polynomial order, the bunch length profits from using sixth order polynomial (blue dots).

ISBN: 978-3-95450-217-2 doi:10.18429/JACoW-ERL2019-WECOYBS04

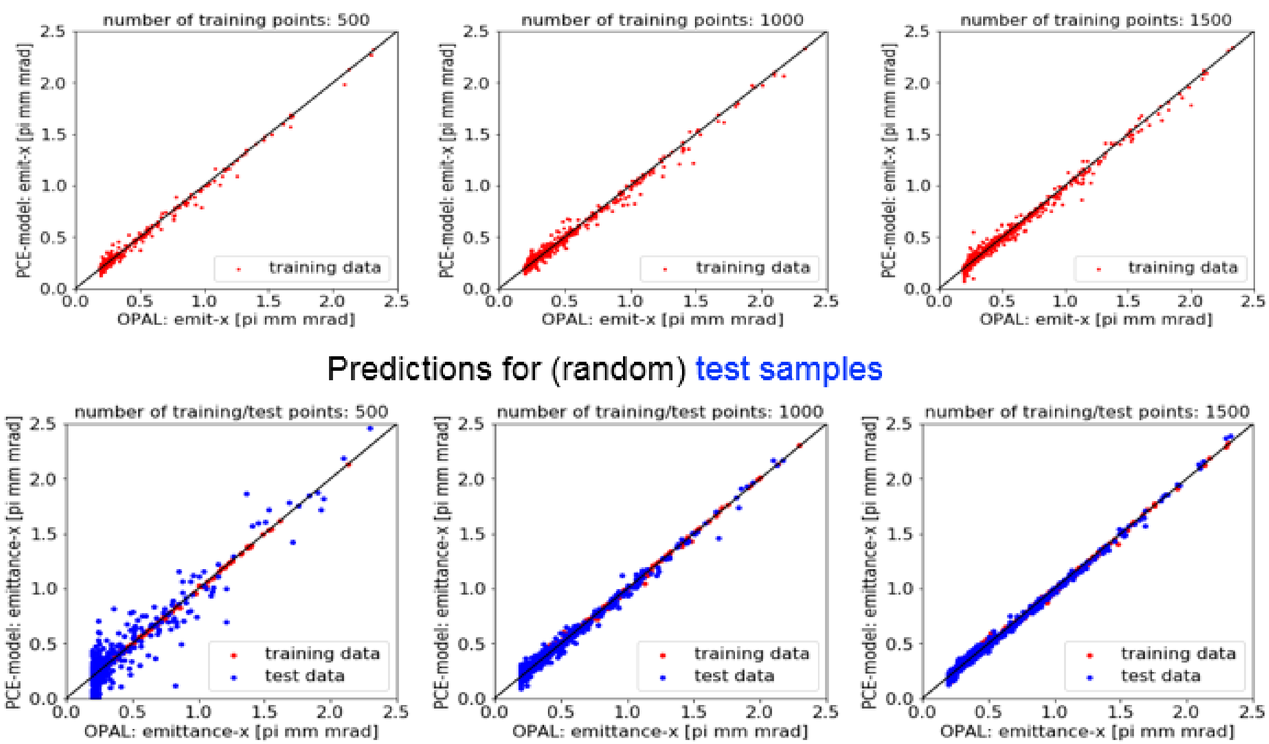


Figure 6: Results for the horizontal emittance calculated by the surrogate model and by tracking. 500 (left), 1000 (center), and 1500 (right) training samples (top) and the same number of testing samples (bottom) were used. 1500 training samples seek best results.

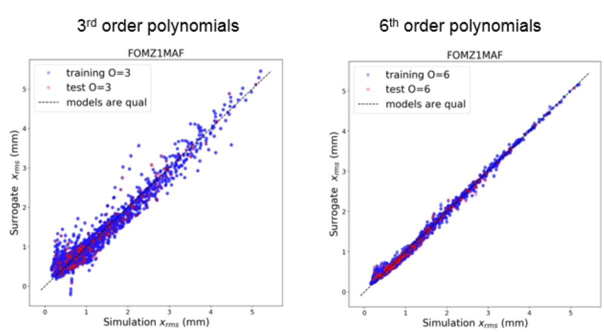


Figure 7: Results for the horizontal beam size, using third (left) and sixth order (right) polynomials. Red dots display the results for test particles, blue dots that of the training particles.

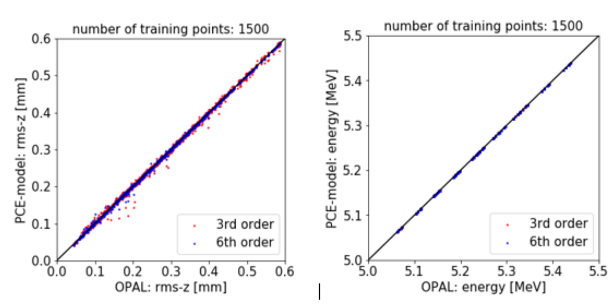


Figure 8: Results for the bunch length (left) and the energy (right). Third order polynomial expansion is shown in red, sixth order in blue.

## *Uncertainty Quantification*

Having developed a PCE from calculated or measured samples, it is straight forward to calculate Sobol's indexes, for details refer to [4]. The first order PC-based Sobol's indexes represent the individual effects of a single random input parameter on the variability of the output. Figure 9 shows the uncertainty quantification for eight quantities of interest and six independent input parameters, taken at the end of the diagnostics line. Many dependencies are expected from accelerator physics knowledge: the dependence of energy on the field flatness or the importance of the solenoid field for the beam size and emittance. The dominant importance of the gun voltage for the bunch length, is less obvious. Higher gun voltage and higher field flatness lead to increased bunch length and increased energy spread at the exit of the gun. The dependance is enhanced by velocity bunching in the booster, see Fig. 10, and diminishes slightly towards the end of the line.

The seemingly obvious impact of the laser pulse length on the bunch length turns out to be negligible. Due to the lack of space charge effects in the sampling data, the bunch length is dominated by the velocity difference due to varying fields after emission from the cathode. The change in field flatness (100 − 130 %) has a larger influence on the particle

ISBN: 978-3-95450-217-2 doi:10.18429/JACoW-ERL2019-WECOYBS04

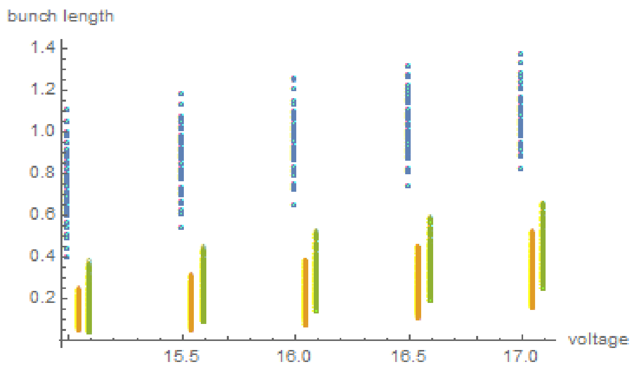


Figure 9: Sensitivity analysis of 4725 samples evaluated at the end of the diagnostics line. Blue: field flatness; light blue: cathode position; orange: laser pulse length; green: solenoid field; light green: laser spot size; red: gun voltage.

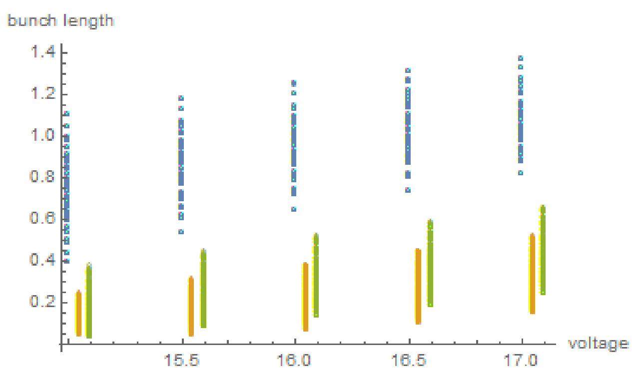


Figure 10: The dependence of the bunch length on the gun voltage. Behind the gun (blue), after compression in the booster (orange) and at the end of the line (green).

energy, than the variation of the gun voltage between 15 and 17 MV/m. After warm measurements of the field flatness, this ambiguity can be largely decreased. Sensitivity quantification helps to sort the importance of machine parameters, when trying to achieve specific bunch parameters. It is important to note, that the sensitivities can vary over the beam line.

### *Neural Network*

A four layer, fully connected artificial neural network has been trained on the same sample data. The results for training data (left), as well as for the test data (right), at the screen in the merger, are displayed in Fig. 11, for the horizontal emittance. Both approaches seek qualitatively good results. For the training data, the mean average error (MAE) between both models, PCE and ANN, and OPAL is 0.5 %. For the testing data, there is a small advantage for the ANN, with an error of again 0.5 %, compared to PCE with an error of 0.6 %. No sensitivity analysis is available, though, when using ANNs.

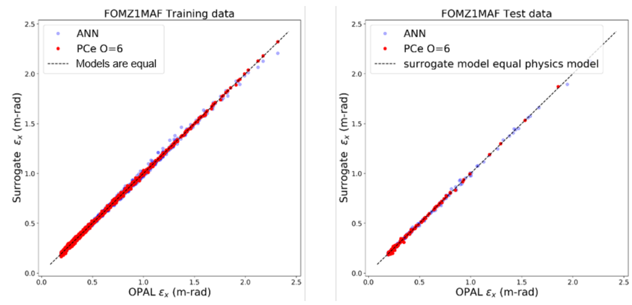


Figure 11: Comparison between PCE

## CONCLUSION

This article concentrated on showing that the application of statistical learning (PCE) or machine learning (ANN) in commissioning has manifold advantages:

- The data necessary to perform PCE/ANN in itself is a rich source of information, that can reveal helpful dependencies.
- Uncertainty quantification, a side product of PCE, gives a quick overview of dependencies of QoI on machine parameters.
- Surrogate models are fast. They can be implemented in the control room and can translate machine settings online into expectation values for beam parameters.
- Surrogate models can speed up optimizations. When approximating measured data or finding new working points, genetic optimizers can be run much faster using the surrogate model.
- The surrogate model prepared for commissioning on theoretical data can be easily improved by including measured date, when it is available.
- Ultimately, one could build a model exclusively on measured data.

Both approaches will be applied to ease the commissioning of bERLinPro. Different samples will be prepared to serve different stages and questions during commissioning, where the understanding of the gun plays the mayor role. Validation tests have to be performed for each model. The effect of increasing bunch charge is of central concern and will be studied in detail. Fortunately, bERLinPro, designed to serve accelerator physics rather than being a user facility is the ideal testbed to study this new approach.

## REFERENCES

[1] M. Abo-Bakr et al., “The Berlin Energy Recovery Linac Project bERLinPro – Status, Plans and Future Opportunities”, presented at ERL’19, Berlin, Germany, Sep. 2019, paper: MOCOXBS04, this conference.

[2] `https://gitlab.psi.ch/OPAL/Manual-2.0/wikis/home`

[3] Auralee Edelen, Nicole Neveu, Yannick Huber, Mattias Frey, Christopher Mayes, Andreas Adelmann, “Machine Learning for orders of magnitude speedup in multi-objective optimization of particle accelerator systems”, `https://arxiv.org/abs/1903.07759`

ISBN: 978-3-95450-217-2 doi:10.18429/JACoW-ERL2019-WECOYBS04

[4] Andreas Adelmann, “On non-intrusive uncertainty quantification and surrogate model construction in particle accelerator modelling”, *SIAM/ASA Journal on Uncertainty Quantification*, 7(2), (2019), doi:10.1137/16M1061928

[5] https://chaospy.readthedocs.io/en/master/