13th Int. Particle Acc. Conf. IPAC2022, Bangkok, Thailand JACoW Publishing
ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-MOPOPT069



# A DATA-DRIVEN BEAM TRAJECTORY MONITORING AT THE EUROPEAN XFEL

A. Sulc*, R. Kammering, T. Wilksen, DESY, Hamburg, Germany

*Abstract*

Interpretation of data from beam position monitors is a crucial part of the reliable operation of European XFEL. The interpretation of beam positions is often handled by a physical model, which can be prone to modeling errors or can lead to the high complexity of the computational model. In this paper, we show two data-driven approaches that provide insights into the operation of the SASE beamlines at European XFEL. We handle the analysis as a data-driven problem, separate it from physical peculiarities and experiment with available data based only on our empirical evidence and the data.

## INTRODUCTION

The European Free Electron Laser (EuXFEL) has been running with very high availability for several years. This high reliability put a lot of attention on the analysis of the operations on various levels. EuXFEL is a pulsed machine with a repetition rate of 10Hz and the properties of each shoot may change. Therefore, anomaly detection on linacs is still very limited. For beam trajectories, this is given by a very limited ability to explain how beam positions are affected by an ongoing anomaly since the beam trajectory changes for each injection. What makes the analysis of the beam even more challenging is that its trajectory can further vary from pulse to pulse due to various circumstances.

At the EuXFEL we are currently operating 103 beam position monitors (BPMs) at three SASE beamlines to measure the position of the beam passage through the undulator lines. All BPMs measure position and charge of up to 2700 bunches in a single bunch train. The absolute beam position is, unlike many other predictive maintenance tasks, a rather more approximate and global indicator, since the contribution of an issue on the beam position is often unknown.

The beam optics in the undulator lines is controlled by the use of a so-called FODO lattice. These alternating magnetic fields can introduce a periodic variation of the trajectory named betatron oscillation [1]. We can observe a specific periodic pattern of the electron bunches passing through the FODO lattice as shown in Fig. 1. This evidence imposes an assumption about the beam irrespective of its trajectory since the $\beta$-function of electron bunches will always follow the symmetry of the FODO lattice and should therefore preserve its period.

An ongoing problem might be indicated in various ways. For instance, if there is an anomaly on a magnet, the trajectory might be noticeably affected by an increased jitter.

One of the common approaches is modeling the beam trajectory and its comparison with a physical model [2].

* antonin.sulc@desy.de

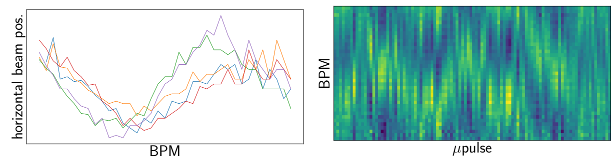


Figure 1: An example input to our methods. The left figure shows a series of the first five bunch trajectories at the SASE1 beamline after the mean of 600 bunches is subtracted. The right figure shows a series of bunches after substraction of mean. Each column is one $\mu$pulse.

A promising data-driven anomaly detection approach on synchrotons at LHC on BPM data was shown by Fol *et al.* [3, 4] to identify faulty BPMs.

We show two data-driven approaches based on our empirical evidence of the beam data at EuXFEL. The first use our assumption of the periodicity imposed by the FODO lattice and fits trajectories using a simplified assumption about the beam dynamics - mainly fitting a periodic - sine - function. The latter is a purely data-driven machine learning approach that trains to map a set of beam positions in arbitrarily long sequences to a common mode and any deviation from the mode is treated as an anomaly. This allows more flexibility in handling the input and can eventually reveal relations between bunch trains.

Our contributions can be summarized as follows: We show two data-driven approaches for the analysis of beam trajectories at the EuXFEL. The first is a method that takes into consideration a simplified empirical model of the lattice and measures the residual of this model. The second is a completely model-free approach based purely on data that models inputs from a set of multiple bunches.

In the following section, we introduce the notation and explain both proposed models. After, we show some experiments on the real data we experience at EuXFEL at SASE beamlines, and in the last section, we conclude our evaluation of the available real data.

## METHOD

EuXFEL produces trains of electron bunches at a frequency up to 4.45MHz at a repetition rate of 10 Hz. These consecutive pulses vary in their individual properties and the resulting trajectory can vary from train to train. If a mean beam position is subtracted over a certain time range, we can obtain the underlying pattern formed by the magnets, as described in the introduction. Visually, it forms a characteristic periodical pattern, which can be seen after subtracting a mean trajectory which we consider as an input, see Fig. 1. Understanding these patterns provides important insights into underlying beam dynamics. The EuXFEL can

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-MOPOPT069

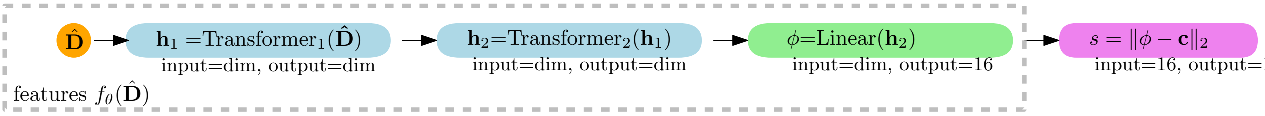


Figure 2: Proposed data-driven two-layer transformer architecture for detection of anomalies. The input $\hat{\mathbf{D}}$ is a sequence of stacked vertical and horizontal positions of a beam in selected beamlines. The dimension of input and all consecutive layers depends on the number of BPMs present at the diagnosed beamline. The inputs are passed through two transformer layers [5] to produce a vector $\mathbf{h}_2$. Its value is further passed to the final linear layer $\phi$. Per-pulse calculation of anomaly score $s$ is performed via measuring $L_2$ distance from a vector $\mathbf{c}$.

produce up to 2700 bunches in a $\mu$pulse (bunch train), but we consider only the first bunch in a bunch train. From a so-called $\mu$pulse number $\mu$ we can associate the same bunch trains at all BPMs along the trajectory and therefore monitor it. Because EuXFEL itself is not operating continuously, we consider each $\mu$pulseas one sample at a time.

We are presenting two major methods of how we tackled the analysis of anomalies on beam orbits. The first is empirical-based and assumes the periodic pattern of the beam trajectory on the $\beta$-function imposed by the FODO lattice, see Fig. 1. We fit the observed values with a sine function which should empirically approximate the data and score the anomality as deviation from this pattern, expressed by residuals of all BPM values to the fitted function.

The second approach is solely based on the data and neglects any periodicity by training a model to project input data to a standard lower-dimensional mode. Due to the absence of labels, one class loss (OCL) [6] loss is employed to train the proposed model. The OCL trains a model to learn a transformation that minimizes the volume of a data-enclosing hypersphere in feature space centered on a point. In the test phase, the anomality is measured as the distance from the hypersphere center.

It then allows modeling of more complicated unknown relations of events which are spread over multiple bunch trains (seconds timescale). It is important to highlight that we always consider only the first bunch in a bunch train.

### *Notation*

We consider that a position of the first bunch in a bunch train is observed at $n$th BPMs at a time $\mu$ at $x_n^\mu$ and $y_n^\mu$ for horizontal and vertical coordinates respectively. For each bunch train, we have a $\mu$pulse number which uniquely identifies it in different bunch trains in $N$ BPMs. We put the coordinates from $N$ BPMs into vectors with horizontal and vertical coordinates $\mathbf{x}^\mu$ and $\mathbf{y}^\mu$ respectively. We stack a series of consecutive $\mu$pulses coordinates $\mathbf{x}$ and $\mathbf{y}$ (e.g. one minute) into a data matrix $\mathbf{D}$ where each row encodes one of $N$ BPM and column one a $\mu$pulse. We subtract the column-mean of a data matrix $\mathbf{D}$ and obtain a normalized matrix $\hat{\mathbf{D}}$ shown in Fig. 1 where one can notice the periodic pattern on the $\beta$-function imposed by the FODO lattice.

### *Empirical Model*

Based on the aforementioned evidence, we assume that the $\beta$-function must follow a periodic pattern, which resembles the sine function. We built a curve fitting approach, where we fit the normalized beam positions $\hat{\mathbf{x}}^\mu$ and $\hat{\mathbf{y}}^\mu$ from the matrix $\hat{\mathbf{D}}$ to a hypothesized function $g$. We define $g$ as a sine function parameterized with amplitude, period, phase shift, and frequency. Since we distinguish between horizontal and vertical coordinates, we consider two different parameterizations $\phi_x$ and $\phi_y$ for both orientations. Under these conditions, individual beam positions $\hat{x}_n^\mu$ and $\hat{y}_n^\mu$ should mirror their fitted functions $g_{\phi_x}(n,\mu)$ and $g_{\phi_y}(n,\mu)$. The quantity which expresses the anomality can be expressed by residuals i.e.

$$r_x = \|g_{\phi_x}(n,\mu) - \hat{x}_n^\mu\|_2 \quad \text{and } r_y = \|g_{\phi_y}(n,\mu) - \hat{y}_n^\mu\|_2.$$

### *Purely Data-Driven Model*

Attention-based models [5] gained a lot of popularity in handling sequences. Working with sequences is particularly useful in taking into consideration anomalies that are spread over multiple bunch trains since the attention-layer can access all previous states in the same input and weight them according to the learned relevance. The location of individual elements in sequence is encoded with positional encoding. We are, however, interested mostly in the co-occurrence of relevant bunches and neglect their position in sequence at this stage.

The network architecture is shown in Fig. 2. It consists of a two-layer transformer with a linear layer in the output. For simplicity, only a single head is used in both transformers. The transformer layer associates $\mu$pulses in input data and may take into consideration various unknown properties between bunches. Since we do not have explicit labels we adopted the OCL [6] to train the model

$$L(\theta) = \|f_\theta(\hat{\mathbf{D}}) - \mathbf{c}\|_2, \tag{1}$$

where the model $f_\theta$ and hypersphere center $\mathbf{c}$ are gradually trained to transform inputs $\hat{\mathbf{D}}$ to a lower-dimensional feature space where the common inputs are transformed to be close to $\mathbf{c}$. Anomaly score $s$ is calculated identically as OCL loss by measuring the $L_2$ distance of $f_\theta$ to $\mathbf{c}$.

## EXPERIMENTS

### *Data Acquisition and Implementation Details*

Data from BPM at respective beamlines was acquired with our DAQ system [8]. For long-term analyses, the data

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-MOPOPT069

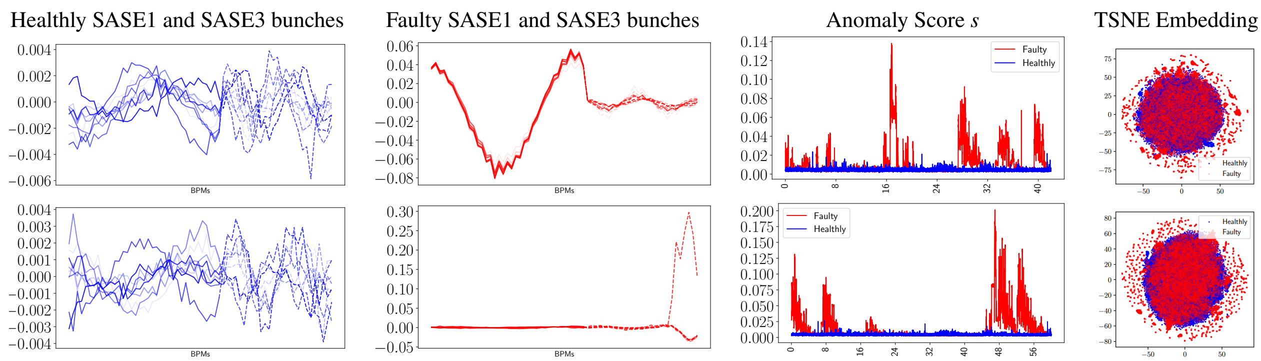


Figure 3: (Top) SASE1 undulator server crashed after an unusual selection of colours for individual cells (2022-03-07 20:45:00) (Bottom) A phase shifter at SASE3 does not move (2022-04-25 21:23:42). The data used for evaluation are taken one hour before the issue was reported. (Left) healthy bunches with the lowest scores. SASE3 is distinguished by dashed lines. (2nd) Faulty bunches with high scores, notice the increased amplitude compared to healthy bunches. (3rd) Anomaly score $s$, the horizontal axis is in minutes. (Right) TSNE Embedding [7] of $f_\theta(\hat{\mathbf{D}})$. Notice that healthy (blue) and faulty (red) points do not overlap. This should indicate a high likelihood of an issue.

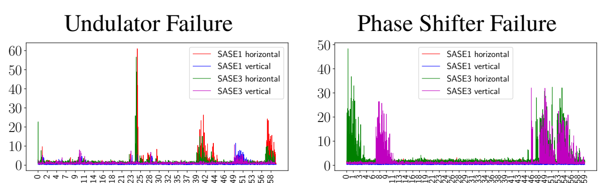


Figure 4: Residuals $r_x$ and $r_y$ of the empirical model for the same faults shown in Fig. 3 on SASE1 and SASE3.

are stored in 1-minute chunks with 8-bit float precision to minimize requirements for data storage. Each chunk usually consists of a series of 600 orbits. For training the purely data-driven model we used data recorded in the time range from Feb 25, 2022, to May 19, 2022. Alongside the records from the DAQ, we also have available a logbook with reported time-stamped issues with a detailed description of the issue. The transformer layers of the purely data-driven model are implemented with a PyTorch [9]. The inputs with no beam were neglected.

## *Results*

We selected some issues from the EuXFEL logbook and evaluated the proposed approaches on the data we have avail-able. For the purely data-driven approach, we analyzed the available beam position data from SASE1 and SASE3 beamlines, where we are currently operating 64 BPMs. The SASE2 beamline should be unaffected by the selected faults since it is in a separate branch and is therefore neglected for now. For comparison, we sampled one-minute samples ev-ery two hours in the entire dataset we have available to have a comparison with a normal setting. The first issue we selected is a crash of the undulator server after an unusual selection of colors for individual cells which took place on 7. March 2022 at 20:45. Data shown in both Fig. 3 and Fig. 4 show a notice-able increase in the anomaly score. The long-term maximum $s$ is approximately 0.03, while the highest achieved score an hour before the issue was reported is around 1.4, with numerous indications within this hour. The ridges are also visible in residuals of the empirical model. The second issue is the phase shifter at SASE3 which does not move. The issue was reported on 25.April 2022 at 21:23:42. There is a noticeable variation in SASE3 compared to healthy data. The highest score $s$ is 0.2. Analysis of the feature space with TSNE embedding [7] reveals, that issues cause variation in the features $f_\theta(\hat{\mathbf{D}})$ which does not overlap with the long-term data, see right column in Fig. 3.

## CONCLUSION

In this paper, we introduce two approaches for the analysis of beam dynamics at EuXFEL. We exploit our empirical evidence about the $\beta$-function which provides a direct interpretable indication of the beam position data by fitting a sine function. The latter is purely data-driven and allows more complex inputs of the inter-bunching relations by considering multiple bunches.

The presented approach reveals that we are already able to identify some issues taking place where beam trajectories are affected by an ongoing problem in beamline within data recorded one hour before the fault was reported. The long-term evaluation revealed a lot of variation in beam positions caused by numerous operations and therefore the scoring of purely data-driven models often yielded many false positives, which currently limits the current application for predictive maintenance. Experiments show that both approaches are similarly efficient with revealing problems on beam trajectory. One of the potential benefits of the purely data-driven approach is feature space, which allows further investigation of otherwise hidden relations between bunch trains. Additionally, there is a limited ability to correlate an issue with its effects on the position of the beam.

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-MOPOPT069

## REFERENCES

[1] B. Holzer, "Beam dynamics in synchrotrons," *arXiv preprint arXiv:1804.08873*, 2018. doi:10.48550/arXiv.1804.08873

[2] T. Hellert, "Intra-bunch-train transverse dynamics in the superconducting accelerators flash and european xfel," Ph.D. dissertation, Deutsches Elektronen-Synchrotron, 2017.

[3] E. Fol, J. M. C. de Portugal, and R. Tomás, "Unsupervised Machine Learning for Detection of Faulty Beam Position Monitors," in *Proc. IPAC'19*, Melbourne, Australia, May 2019, pp. 2668–2671. doi:10.18429/JACoW-IPAC2019-WEPGW081

[4] E. Fol, R. Tomás, J. C. de Portugal, and G. Franchetti, "Detection of faulty beam position monitors using unsupervised learning," *Phys. Rev. Accel. Beams*, vol. 23, no. 10, p. 102805, Oct.2020.doi:10.1103/PhysRevAccelBeams.23.102805

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

[6] L. Ruff *et al.*, "Deep one-class classification," in *International conference on machine learning*, PMLR, 2018, pp. 4393–4402.

[7] G. E. Hinton and S. Roweis, "Stochastic neighbor embedding," *Advances in neural information processing systems*, vol. 15, 2002.

[8] T. Wilksen *et al.*, "The Control System for the Linear Accelerator at the European XFEL: Status and First Experiences," in *Proc. ICALEPCS'17*, Barcelona, Spain, Oct. 2017, pp. 1–5. doi:10.18429/JACoW-ICALEPCS2017-MOAPL01

[9] A. P. *et. al.*, "Automatic differentiation in pytorch," 2017.