ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-TUPOPT058

A MACHINE LEARNING APPROACH TO ELECTRON ORBIT CONTROL AT THE 1.5 GeV SYNCHROTRON LIGHT SOURCE DELTA

D. Schirmer*
Center for Synchrotron Radiation (DELTA), TU Dortmund University, Germany

Abstract

Machine learning (ML) methods have found their application in a wide range of particle accelerator control tasks. Among other possible use cases, neural networks (NNs) can also be utilized for automated beam position control (orbit correction). ML studies on this topic, which were initially based on simulations, were successfully transferred to real accelerator operation at the 1.5–GeV electron storage ring of the DELTA accelerator facility. For this purpose, classical fully connected multi-layer feed-forward NNs were trained by supervised learning on measured orbit data to apply local and global beam position corrections. The supervised NN training was carried out with various conjugate gradient backpropagation learning algorithms. Afterwards, the ML-based orbit correction performance was compared with a conventional, numerical-based computing method. Here, the ML-based approach showed a competitive orbit correction quality in a fewer number of correction steps.

INTRODUCTION

Stable electron orbit control is an important task especially for modern synchrotron light sources. For this purpose, singular value decomposition (SVD) of the orbit response matrix is a standard numerical tool at storage rings worldwide. An alternative concept applies machine learning techniques as an heuristic method, inspired by the pioneering work done at NSLS/BNL [1]. Since 2018, machine learning (ML) based orbit correction (OC) methods have extensively been studied and applied at DELTA [2–5], a 1.5–GeV electron storage ring operated as a synchrotron light source [6] and a new facility for ultrashort pulses in the VUV and THz regime [7, 8].

MACHINE OPERATION

The OC hardware setup of the DELTA storage ring consists of 54 beam position monitors (BPMs) which determine the position in both orbit planes simultaneously as well as 30 horizontal and 26 vertical corrector magnets (steerers) [9]. The ML methods, which were successfully tested by means of simulations on a DELTA storage ring OC model, were applied to the real storage ring.

ML-Data Acquisition

Compared to the uncoupled case in previous real machine studies [3], now a dedicated data acquisition script randomly varies all steerer strengths in both planes at once within intervals from typically ±200 mA up to ±500 mA. The interval limits are a compromise between risk of beam losses and minimizing relative measurement errors due to the limited steerer strength resolution of 2.4 mA. After each perturbation, the steerer strength changes and the emerging closed orbit differences in both planes are measured, error-cleaned (e.g., deleting of hardware and software related readout errors) and recorded. The data pool noise level was estimated to approximately 2% caused by the combined error of BPM and steerer strength readback accuracies, mainly dominated by the limited steerer strength granularity.

Weighted Beam Position Monitors

To increase the impact of orbit deviations at more important storage ring positions (e.g., synchrotron radiation source points or the injection region) each BPM can be assigned with an individual weight factor. With $\tilde{w}^{\mathrm{BPM}}_{x,z}$ as a diagonal matrix of BPM weight factors, the weighted closed-orbit error $\chi^{w}_{x,z}$ can be evaluated as a scalar quantity for both planes (x, z) by

$$\chi^{w}_{x,z} = \left\| \tilde{w}^{\mathrm{BPM}}_{x,z} \cdot (\overrightarrow{\Delta d}_{x,z} + \tilde{R}_{x,z} \cdot \overrightarrow{\Delta I}_{x,z}) \right\|_2 . \qquad (1)$$

The goal for an orbit correction algorithm is to minimize the residual closed-orbit error $\chi^{w}_{x,z}$ for arbitrary orbit deviations $\overrightarrow{\Delta d}_{x,z}$ with respect to any desired reference orbit. The product of the response matrix $\tilde{R}_{x,z}$ and the steerer strength changes $\overrightarrow{\Delta I}_{x,z}$ in Eq. 1 can be determined by means of a reverse NN (see Fig. 1). It can also be trained with the experimental data patterns, but now each squared network error $e^2_{pj} = (o^{\mathrm{BPM}}_{pj} - t^{\mathrm{BPM}}_{pj})^2$ must be weighted by an individual BPM weight factor w^{BPM}_{pj} as follows:

$$E^{R}_{\mathrm{mse}} = \frac{1}{P} \sum_{p=1}^{P} \frac{1}{N} \sum_{j=1}^{N} w^{\mathrm{BPM}}_{pj} (o^{\mathrm{BPM}}_{pj} - t^{\mathrm{BPM}}_{pj})^2 . \qquad (2)$$

The reverse mean squared NN error E^{R}_{mse} sums up the squared differences between all numbers of neurons N at the NN target t and output o for a specific quantity of data patterns P. Thus, the reverse trained NN, as a representation of the orbit response matrix $\tilde{R}_{x,z}$, is able to determine orbit deviations $\overrightarrow{\Delta d}_{x,z}$ at all BPMs for given steerer strength changes $\overrightarrow{\Delta I}_{x,z}$. Afterwards, the weighted orbit error $\chi^{w}_{x,z}$ in Eq. 1 can be minimized using a numerical optimizer, e.g., the BFGS Quasi-Newton method [10]. In addition, the optimizer itself has also been replaced by a pre-trained NN. For this purpose, the optimizer has to pre-calculate the optimum $\chi^{w}_{x,z}$-values for all measured orbit deviations $\overrightarrow{\Delta d}_{x,z}$. These data pairs again serve as labeled input/target data to train a dedicated NN as an optimizer substitute.

* detlev.schirmer@tu-dortmund.de

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-TUPOPT058

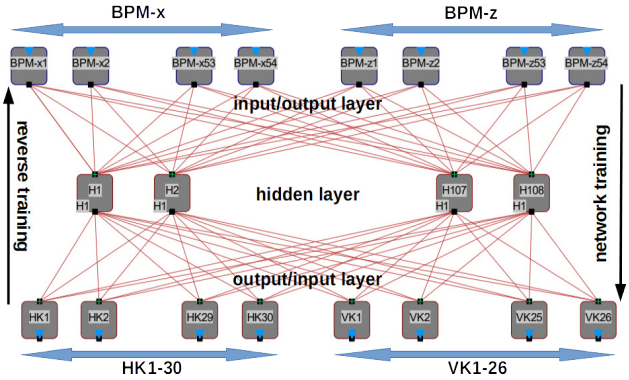

Figure 1: Schematic layout of the neural network topology for a full x,z-coupled orbit correction at the storage ring DELTA. The input layer represents 54 BPMs for each plane, fed by measured orbit deviations. It is connected via a 'hidden' layer with the output values for 56 corrector magnets (HK1-30, VK1-26). The correction considers both transverse coordinates (x,z) as well as their coupling. In total, the network consists of approx. 17700 connections (red lines, only partially shown). For reverse training, the input and output layers are swapped.

Neural Network Training

Since each corrector strength variation normally affects the beam displacement at all BPM positions in the storage ring, a fully connected feed-forward neural network (FFNN) was specified as the NN connection architecture. Thus, NNs to be trained by the experimental data are composed of three layers with a total of 272 neurons (108/108/56) and approximately 17700 weights and biases (see Fig. 1).

The supervised network training was tested with various conjugate gradient backpropagation learning algorithms. Most effective learning was achieved with scaled conjugate gradient (scg) algorithms [11]. Typical forward and reverse training curves for pure training data (blue) and the related validation data sets (green) are shown in Fig. 2. In both cases, the network output error is reduced continuously, without over- or under-fitting issues, mainly limited by the data noise level. The best forward validation performance with an E_{mse}-value of $7.3 \cdot 10^{-4}$ A^2 was reached with no significant improvement beyond approx. 1600 scg-training iterations applying a full batch training. The reverse training reduced the network output error E^R_{mse} to $1.2 \cdot 10^{-2}$ mm^2 after about 700 iterations. In both cases, the network's fitness gains approximately three orders of magnitude compared to the starting values, which indicates sufficient fitness performance. Hence, conventional NNs are able to learn and generalize the correlation between orbit deviations and the related steerer strength variations.

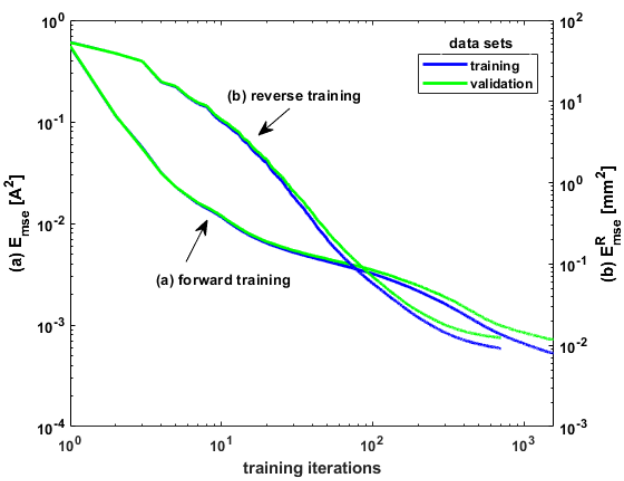

Figure 2: Supervised NN training, based on a scaled conjugate gradient (scg) backpropagation algorithm, with experimental data (blue) and verification of the NN performance with an additional 'unseen' validation test data set (green).

BENCHMARK RESULTS

The performance of the ML-based OC program was benchmarked against a recently implemented numerical approach [12, 13] at different comparison terms. The tests considered actual reference settings, which include the currently valid reference orbit and the corresponding weight factors (wf) for all K BPMs in operation. Various arbitrary steerer-induced orbit errors have been enforced (a–d) in both orbit planes, respectively (hk/vk). The benchmark results are compared in Fig. 3 (ML-based) and Fig. 4 (numerical) separately for each orbit plane.

In addition, typical but not pre-trained sources for orbit disturbances have been provoked. This includes unmatched closed orbit bumps (e.g., injection dc-bump), strong sextupole strength changes (SF/SD) or ramping of insertion devices (e.g., the undulator U250). The benchmark results

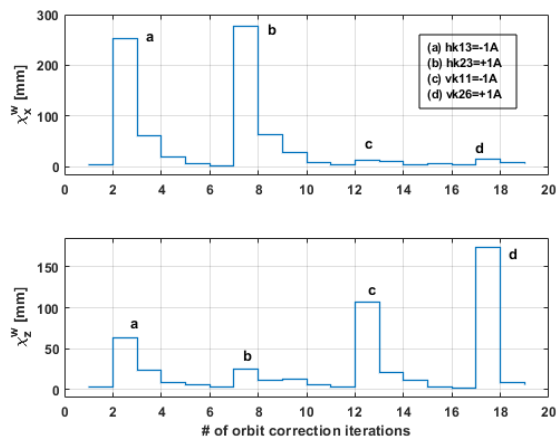

Figure 3: Weighted rms orbit error $\chi^w_{x,z}$ for different steerer-induced (hk/vk) orbit disturbances (a-d) compensated iteratively by the ML-based OC program. Due to x,z-coupling, the provoked deviations appear in both orbit planes. A total of 18 steps were required to compensate for all disturbances (a-d). In comparison to Fig. 4, similar final residual orbit errors are achieved in significantly fewer iteration steps.

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-TUPOPT058

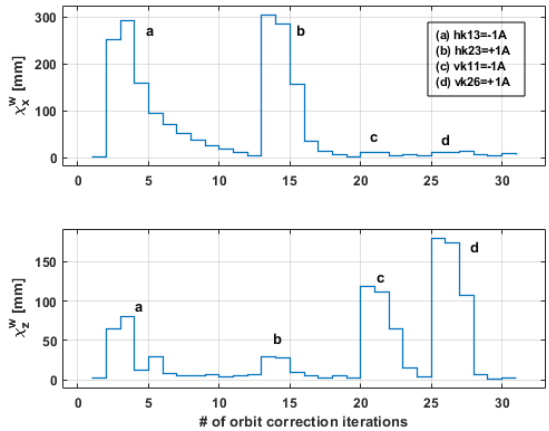

Figure 4: Weighted rms orbit error $\chi^w_{x,z}$ to benchmark the orbit correction quality and convergence for a conventional numerical method [12, 13]. The correction was performed iteratively for the same steerer-based (hk/vk) orbit distortions (a-d) as presented in Fig. 3. In total, 30 steps were required to compensate for all disturbances (a-d).

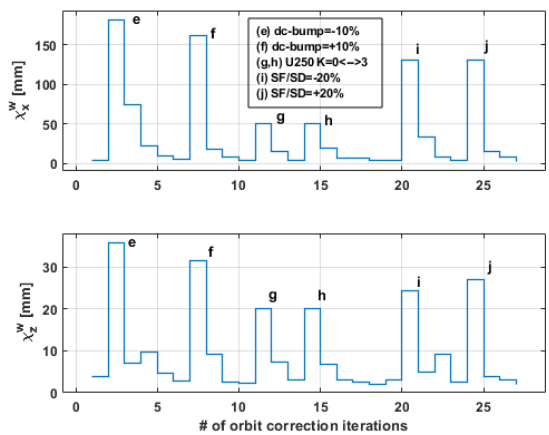

Figure 5: Individual ML-based correction steps to compensate orbit deviations caused by different error sources (e-j) which were not included in the training data sets. Although these types of error sources were not considered during NN training, all disturbances (e-j) could be compensated in 26 steps.

are shown in Fig. 5 (ML-based) and Fig. 6 (numerical) for each plane separately. In all plots, the residual orbit quality $\chi^w_{x,z} := \left(\frac{1}{K}\sum_{i=1}^{K}\left(w^{\mathrm{BPM}}_{x,z}\cdot\Delta_{x,z}\right)^2_i\right)^{1/2}$ is scored by the weighted root-mean-squared value (wrms) over all K BPMs in both orbit planes (x, z).

As can be seen from the graphs, both OC programs worked similarly stable and they were able to compensate all enforced orbit disturbances without any beam losses. The ML-based OC needs approximately 2 to 4 orbit correction iterations to equalize the individually provoked orbit deviations. Even distortions (e–j), which have not trained during

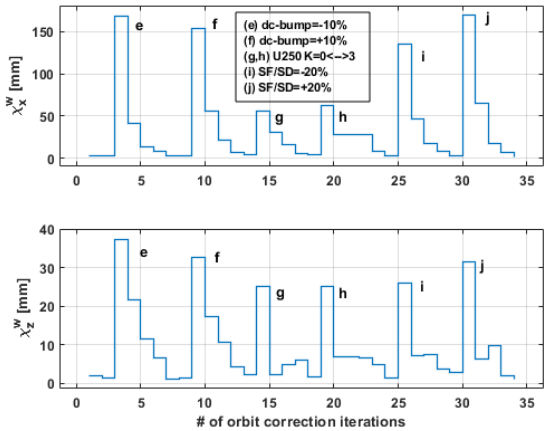

Figure 6: Correction for different scenarios of enforced orbit deviations (e-j) performed with a conventional numerical orbit correction program [12, 13]. In this case, in total 33 steps were needed to compensate the same perturbations (e-j) as depicted in Fig. 5.

supervised learning were compensated with analogous quality. After each provoked error, the residual weighted orbit error converged to less than 3 mm, which corresponds to several ten μm in real beam offset.

In direct comparison, the standard numerical OC method requires more orbit correction steps to compensate the same provoked orbit disturbances. The residual weighted orbit error converged also to less than 3 mm, but overall, in these benchmark examples, the standard OC version needs in total 63 steps (see Fig. 4 and Fig. 6) compared to 44 steps for the ML-based implementation (see Fig. 3 and Fig. 5).

SUMMARY

A machine learning-driven orbit correction algorithm has successfully been implemented at the 1.5–GeV electron storage ring of the DELTA accelerator facility. The achieved remaining orbit error is similar to the results of standard SVD-like correction approaches. In general, ML-based methods require fewer correction steps which leads to faster orbit correction convergence. Because ML techniques based on training with real machine data, this method automatically incorporates storage ring imperfections (e.g., alignment errors) and non-linearities (e.g., magnetic fringe fields). Furthermore, it has been demonstrated that NNs were also able to cope with beam disturbances that had not been trained before. In addition, even changes of BPM offsets, e.g. due to realignments of individual lattice magnets, can be taken into account, simply by editing the related orbit reference data file. Therefore, retraining of the entire neural network is not required. In general, once successfully trained, the NN-based application showed high reliability, numerical stability and robustness.

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-TUPOPT058

REFERENCES

[1] E. Bozoki and A Friedman, "Neural Networks and Orbit Control in Accelerators", in *Proc. 4th European Particle Accelerator Conf. (EPAC'94)*, London, UK , Jun.-Jul. 1994, pp. 1589–1591.

[2] D. Schirmer, "Intelligent Controls for the Electron Storage Ring DELTA", in *Proc. 9th Int. Particle Accelerator Conf. (IPAC'18)*, Vancouver, Canada, 2018, pp. 4855–4858. doi:10.18429/JACoW-IPAC2018-THPML085

[3] D. Schirmer, "Orbit Correction with Machine Learning Techniques at the Synchrotron Light Source DELTA", in *Proc. of 17th Int. Conf. on Accelerator and Large Experimental Physics Control Systems (ICALEPCS'19)*, New York, USA, 2019, pp. 1426–1430. doi:10.18429/JACoW-ICALEPCS2019-WEPHA138

[4] D. Schirmer, A. Althaus, S. Hüser, S. Khan, T. Schüngel, "New machine learning projects at DELTA", in *17th DELTA Annual Report 2021*, pp. 3–4, Dortmund, Germany, November 2021. https://www.delta.tu-dortmund.de/cms/Medienpool/User_Reports/DELTA_User_Report_2021.pdf

[5] D. Schirmer, A. Althaus, S. Hüser, S. Khan, T. Schüngel, "Machine learning projects at the 1.5–GeV synchroton light source DELTA", in *Proc. of 18th Int. Conf. on Accelerator and Large Experimental Physics Control Systems (ICALEPCS'21)*, Shanghai, China, 2021, pp. 631–635. doi:10.18429/JACoW-ICALEPCS2021-WEPV007

[6] M. Tolan, T. Weis, C. Westphal, and K. Wille, "DELTA: Synchrotron light in nordrhein-westfalen", *Synchrotron Radiation News 16(2)*, 9–11, (2003). doi:10.1080/08940880308603005

[7] S. Khan *et al.*, "Coherent Harmonic Generation at DELTA: A New Facility for Ultrashort Pulses in the VUV and THz Regime", in *Synchrotron Radiation News 24(5)*, 18–23, (2011). doi:10.1080/08940886.2011.618092

[8] S. Khan *et al.*, "Generation of Ultrashort and Coherent Synchrotron Radiation Pulses at DELTA", *Synchrotron Radiation News 26(3)*, 25–29, (2013). doi:10.1080/08940886.2013.791213

[9] M. Grewe, "SVD-basierte Orbitkorrektur am Speicherring Delta", dissertation, TU Dortmund, Germany, 2005.

[10] Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm, https://en.wikipedia.org/wiki/Broyden-Fletcher-Goldfarb-Shanno_algorithm

[11] Møller, M. F., "A scaled conjugate gradient algorithm for fast supervised learning", Neural Networks 6(4), 525-–533, (1993).

[12] S. Kötter, B. Riemann, and T. Weis, "Status of the Development of a BE-Model-Based Program for Orbit Correction at the Electron Storage Ring DELTA", in *Proc. 8th Int. Particle Accelerator Conf. (IPAC'17)*, Copenhagen, Denmark, 2017, pp. 673–675. doi:10.18429/JACoW-IPAC2017-MOPIK065

[13] S. Kötter, A. Glassl, B. D. Isbarn, D. Rohde, M. Sommer, and T. Weis, "Evaluation of an Interior Point Method Specialized in Solving Constrained Convex Optimization Problems for Orbit Correction at the Electron Storage Ring at DELTA", in *Proc. of 9th Int. Particle Accelerator Conf. (IPAC'18)*, Vancouver, Canada, 2018, pp. 3507–3510. doi:10.18429/JACoW-IPAC2018-THPAK114