MACHINE LEARNING APPLIED TO AUTOMATED TUNES CONTROL AT THE 1.5 GEV SYCHROTRON LIGHT SOURCE DELTA

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

Abstract

Machine learning (ML) driven algorithms are finding more and more use cases in the domain of accelerator physics. Apart from correlation analysis in large data volumes, low and high level controls, like beam orbit correction, also non-linear feedback systems are possible application fields. This also includes monitoring the storage ring betatron tunes, as an important task for stable machine operation. For this purpose classical, shallow (non-deep), feed-forward neural networks (NNs) were investigated for automated adjusting the storage ring tunes. The NNs were trained with experimental machine data as well as with simulated data based on a lattice model of the DELTA storage ring. With both data sources comparable tune correction accuracies were achieved, both, in real machine operation and for the simulated storage ring model. In contrast to conventional proportional–integral–derivative (PID) controller methods, the trained NNs were able to approach the desired target tunes in fewer steps. The report summarizes the current status of this machine learning project and points out possible future improvements as well as other possible applications.

INTRODUCTION

DELTA is a 1.5–GeV electron storage ring facility operated by the TU Dortmund University supplying radiation ranging from THz to the hard x-ray regime [1,2]. Since 2018, machine learning (ML) methods were extensively studied for beam orbit control. For more details see [4]. Due to thermal orbit movements and magnetic current-dependent field changes, the tunes may vary during machine operation. Therefore, automatic tunes correction is equally important, especially for the DELTA storage ring, as otherwise sudden beam losses can occur.

To provide a precise, reliable and fast tune reading, the complete measurement setup was renewed in 2006 [3]. It is based on broadband beam excitation with an x,y-kicker magnet and measurement of the relaxation betatron oscillations turn-by-turn. The betatron frequency detection utilized a classical numeric approach applying the Levenberg-Marquardt algorithm for data fitting. Thus, a tune measuring accuracy of better than $2 \cdot 10^{-5}$ can be achieved [3].

A simple PID-tune feedback loop based on this measurement compensates for tune shifts. This method is regularly in use since many years as the standard tune control method at the DELTA storage ring.

* detlev.schirmer@tu-dortmund.de

MACHINE LEARNING APPROACH

Simulation results

An alternative approach is the application of machine learning methods. For this purpose, simulations were initially carried out to figure out if there exists a ML-trainable correlation between quadrupole strength variations and corresponding tune shifts. A lattice model of the DELTA storage ring served as the basis for Twiss parameter (optics) computations within the Accelerator Toolbox (AT) framework [7, 8]. The lattice contains all main accelerator components including all insertion devices (IDs) such as the superconducting asymmetrical wiggler magnet (SAW) [5]. Furthermore, non-linear magnetic saturation and cross-talk effects of the combined function magnets were taken into account [6]. Based on this detailed ring model, the AT software library calculates all parameters of the x,y-coupled linear optics and thus also the simulated tunes (Q_x, Q_y). The simulated data correspond to the measured values of the real machine in the range of a few percent [5].

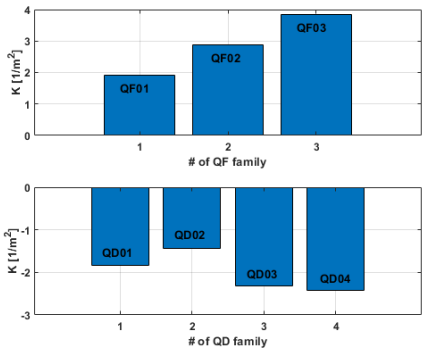

Figure 1: Initial quadrupole strength settings for the standard DELTA storage ring optics with 7 independent power supply circuits, 3 horizontal (QF01-03, top) and 4 vertical (QD01-04, button) focusing families located in the ring arcs (settings for AT optics simulations, SAW switched off).

The linear optics of the DELTA storage ring can be flexibly adjusted with a total of 30 independent quadrupole families distributed over a circumference of 115 m. It is mainly composed of a sequence of so-called triplet unit cells arranged in the arcs of the racetrack-shaped storage ring [14]. The Twiss functions in the straight sections can be set by matching sections at the respective ends of each arc. In this region, the SAW generates a large impact on the vertical betafunction due to vertical edge focusing effects [15]. In order not to change the optics in the straight sections, only 7 quadrupole families in the arcs are used for tune control.

Due to relatively small betafunctions at these quadrupole positions, significant tune shifts can be generated by moderate quadrupole strength variations. Inside the arcs a total of 3 horizontally and 4 vertically focusing quadrupole families are available. They can be operated independently by dedicated power supplies (PS). The initial quadrupole settings for the standard optics (SAW off) are shown in Fig. 1.

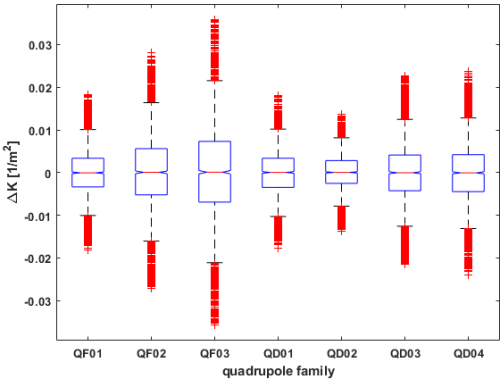

Figure 2: Distribution of 3000 random quadrupole strength variations (±1%). Settings for 3 horizontal (QF01-03) and 4 vertical (QD01-04) focusing independent magnet families. On each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The outliers are plotted individually using the '+' symbol. See also Fig. 16 in appendix.

For simulations, the strengths of these quadrupole families were randomized, uniformly distributed within various maximum interval limits (±0.1%, ±0.5%, ±1%, ±2%), whereby these limits have also been uniformly randomized. This leads to Gaussian-like distributions without outliers beyond the interval limits[1]. Fig. 2 depicts an example for 3000 random quadrupole set variations within a maximum interval limit of ±1%. For each set of limits 3000 related tunes were computed via the AT-framework. Fig. 3 presents the corresponding results for 3000 tune calculations in two cases, SAW switched on and off, respectively (random limits ±1%).

With these overall 12000 data pairs (NN input: tune shifts and NN output: quadrupole strength changes) classical 3-layer feed-forward neural networks were trained (see Fig. 4). The neural network consists of 2 input neurons (ΔQ_x, ΔQ_y), between 7 and 14 neurons in one hidden layer (optimal quantity determined by trial and error) and 7 output neurons (quadrupole strength variations ΔK). The non-linear hyperbolic tangent serves as the transfer function between the input and hidden layer and a linear transfer function was applied between the hidden and output layer. The NNs were trained on a trial basis with a variety of training methods [11]. Best results were achieved with a conjugate gradient backpropagation with Polak-Ribiére updates [12]. Corresponding learning curves are shown in Fig. 5 as an example. The learning performance is defined by the mean squared normalized error performance function (mse) of the NN output[2].

[1] This avoids beam losses in the later application on the real machine operation.

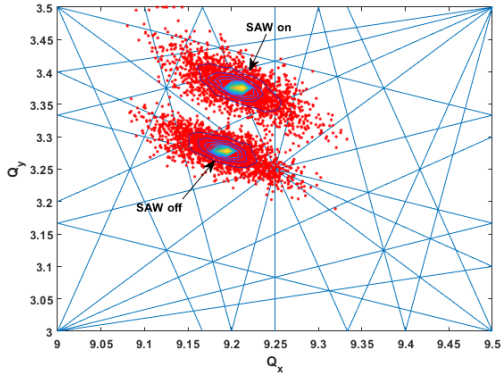

Figure 3: Distribution of 3000 tunes for quadrupole settings randomly varied by ±1%. Optics simulations with SAW switched on and off. The tune is vertically shifted due to edge focusing effects of the SAW.

$$mse = \frac{1}{P}\sum_{p=1}^{P}\frac{1}{N}\sum_{n=1}^{N}(o_{pn} - t_{pn})^2$$

It summed up the square difference between all target (t) and NN output (o) neurons (N) for all training patterns (P). With increasing number of full batchsize (80% of all data pairs P) backpropagation iterations (also referred to as epochs), the mean squared error of the network was reduced by more than one order of magnitude and reaches a minimum of $3.9 \cdot 10^{-5}$ [1/m^4] after 166 epochs. The validation learning curve (green line) demonstrates that the neural networks can be trained by simulated data and thus there exists an 'expected' trainable correlation between tune shifts and quadrupole strength variations. The regression coefficient[3] is calculated to $R = 0.65$.

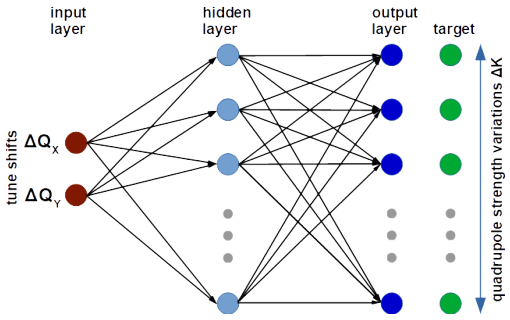

Figure 4: Three-layered neural network topology to be trained for the automated tune control.

Subsequently, the NNs were verified with the accelerator model. Figure 6 shows the tune matching for an arbitrary

[2] Mean squared error is defined as the average squared difference between NN outputs and targets. Lower values are better. Zero means no error.

[3] The regression R value measures the correlation between NN outputs and targets. An R value of 1 means a close relationship, 0 a random relationship. For details see [11, 13].

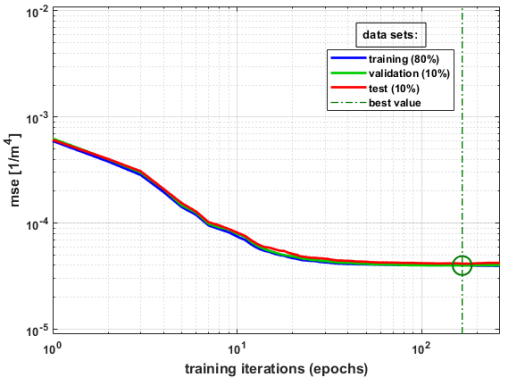

Figure 5: Training performance of different data sets determined by the mean squared error (mse) of the conjugate gradient backpropagation learning algorithm. The training was performed with a full batch size of 12000 simulated data (see Fig. 2, 3). Best validation performance is $3.9 \cdot 10^{-5}$ [1/m^4] after 166 iterations (green curve).

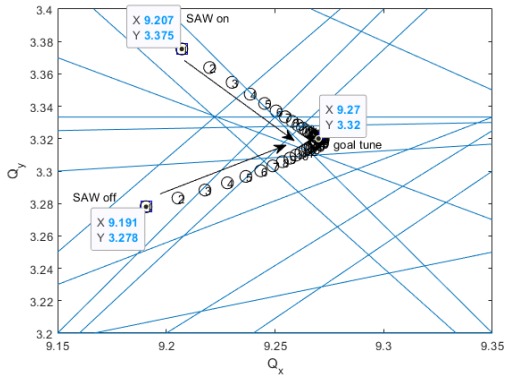

Figure 6: Example for verification of NNs trained by simulated data and applied to the simulation model. The desired goal tune was reached in iterative steps from different start values, SAW switched on and off, respectively.

goal tune as an example. The initial tunes were defined by the SAW status (switched on/off). In both cases the desired tune has been reached iteratively. The number of steps is adjustable. Small tune shifts (< 0.01) can be performed in a few steps. The resolution of the step width depends primarily on the quantity of training data for the respective value range.

Application in real machine operation

On basis of the simulations described above, corresponding experimental data were recorded during real storage ring operation. Again, only the current set values of the quadrupole families in the arcs were randomly changed and subsequently the associated tunes were measured with an accuracy of better than $1 \cdot 10^{-4}$ [3] (see Fig. 7, Fig. 8 and Fig. 9). To minimize the probability of beam losses due to e.g. large tune jumps, the variation interval of the quadrupole settings was initially limited to ±0.5% for each power supply family. During a short machine run of about 2 hours, more than 600 data pairs were recorded (see Fig. 9). Therefore, in average approximately 12 seconds was required for each tune measurement cycle.

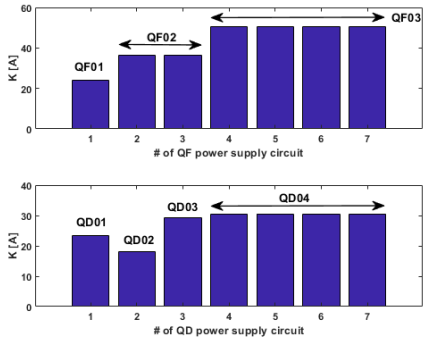

Figure 7: Initial quadrupole settings for 7 independent quadrupole families in the storage ring arcs. Circuits of equal strength K are combined into quadrupole families.

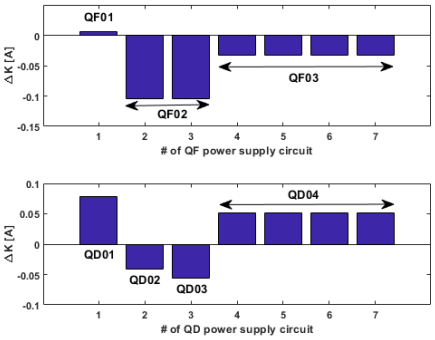

Figure 8: Random quadrupole strength variations ΔK. Settings for 3 horizontal (QFs) and 4 vertical (QDs) focusing independent magnet families (one example from 600 measurements).

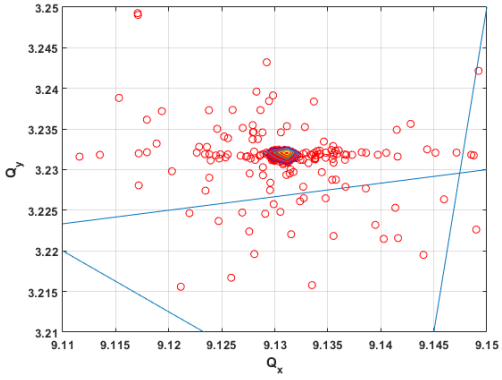

Figure 9: Distribution of 600 measured tunes generated by random quadrupole strength variations of ±0.5% with SAW switched off (see Fig. 8).

The neural networks described above (see Fig. 4) could also be used to successfully train with these measured experimental data (see Fig. 10). Because the measured data

are more noisy[4] than data from simulations, the regression coefficient is reduced to $R = 0.45$.

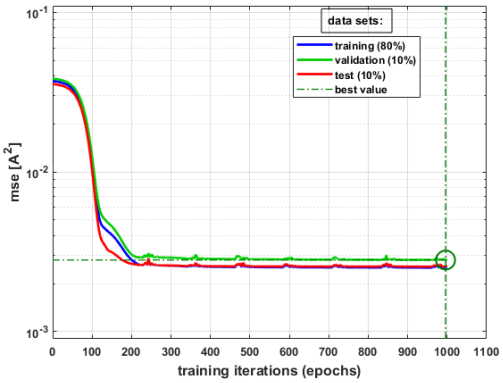

Figure 10: Training performance for different data sets determined by the mean squared error (mse). NN learning was performed by the gradient descent with momentum and adaptive learning rate backpropagation algorithm [9, 10]. NN trained by experimental data (see Fig. 7, 8, 9). A validation performance of $2.8 \cdot 10^{-3}$ [A^2] was reached after 996 iterations. No significant performance improvement was achieved after approx. 200 iterations. Training has been terminated after 1000 epochs (green curve).

Neural network validations

Although the regression correlation is smaller in comparison to simulation calculations, these trained NNs were still tested to determine quadrupole changes for given tune shifts (i.e., differences to the so-called goal tunes) in real machine operation. Figures 11 and 12 demonstrate two successful examples for storage ring tune controls, SAW switched off and on, respectively. In general, NNs were able to calculate the correct quadrupole settings to reach the desired target tunes in just a few steps. But since the chopper power supplies of the DELTA quadrupole magnets cannot be controlled without delays in real time, the new quadrupole set values must be approached in several smaller current steps. After each single step (new quadrupole settings) the tune can be determined again (see numbered actual tunes in diagrams 11 and 12) and the NNs calculate the next step until the desired goal tune is reached iteratively. In the shown examples the goal tunes were reached in 10 respectively 50 steps with an absolute error of less than $4 \cdot 10^{-3}$. The step size is adjustable within narrow bounds. It depends on the desired total tune shift and the quadrupole variation limits during training. Resetting from the goal tune to the initial start tunes was also possible analogously.

Ideally, the trained NNs need only one single calculation to match the tune in one large step. However, a non-synchronous approach to the calculated power supply (PS) current values can lead to (partial) beam losses. For this reason, the software synchronous approach described above

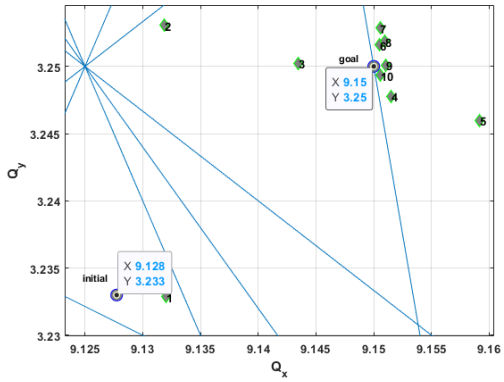

Figure 11: Usage of NNs trained with experimental data and applied to real machine operation. This initial experiment demonstrates a tune fitting in 10 steps from start to goal tune without beam losses (SAW switched off).

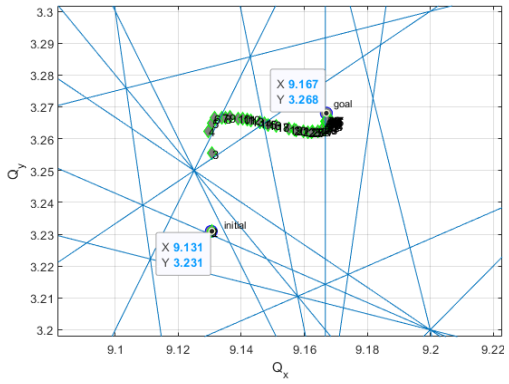

Figure 12: Application of NNs trained with experimental data and applied to real machine operation with SAW switched on. This second example demonstrates a tune matching in 50 smaller steps from start to goal tune without beam losses.

was implemented. Nevertheless, this method is not always successful either, since, as with the standard PID method, beam destroying resonances can be crossed, as indicated by the resonance lines in the $Q_{x,y}$-tune diagrams.

Finally, it was examined whether it is also possible to perform tunes control of the real storage ring with NNs only trained by simulated model data; and vice versa NNs only trained with real machine data and then applied to the simulation model. This should be possible as long as the relationship between quadrupole variation and tune change in the linear storage ring model is similar to that of the real machine[5]. Here, the correct transformation of model simulation parameters into real magnet power supply current settings and vice versa is an important task. The corresponding recalculations were performed by a Matlab version of the conversion program "i2k" [5, 6, 11]. The program also takes non-linear saturation and crosstalk effects of the combined function magnets into account [6].

[4] Mainly dominated by the limited read/set accuracy of approx. 0.025% for the quadrupole chopper power supplies.

[5] Alternatively, if required, additional quadrupole correction currents could be adjusted by an empirically determined scaling factor.

Fig. 13 illustrates a typical tune matching application as an example. It is comparable to the validation shown in Fig. 12, which has been obtained with real machine data exclusively. This example demonstrates, for the first time at the DELTA storage ring, that NNs trained by simulated model data can also be used for controlling real machine processes.

In an analogous manner, it was also possible to use NNs trained by measured data to perform tune control on the simulated storage ring model (see Fig. 14). This result corresponds to the verification example shown in Fig. 6. As previously indicated the step sizes can be adjusted within certain limits essentially only given by the value range of training data.

So far, only difference data ($\Delta Q_{x,y}$ and $\Delta K_{QF,QD}$) with small changes (approx. 1%) were used for NN training. As long as the correlation of these relative changes (gradients) are similar for a wide range of optics settings (i.e., absolute tunes), this machine learning technique is applicable in a wide tune workspace of the DELTA storage ring, without having to retrain. At least, this is the case in the domains of stable solutions for the pure triplet unit cell, which can be seen from the stability diagram of the triplet necktie island (see Fig. 15, [14]). For more details see also [16]. However, this must further be validated by future studies.

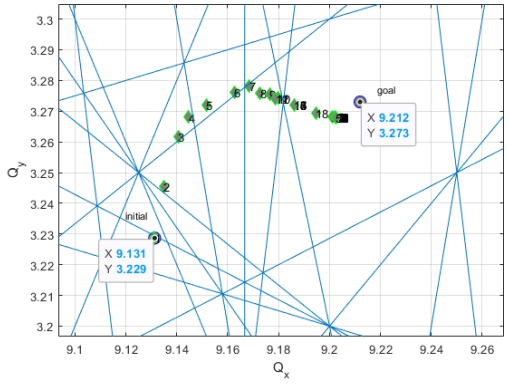

Figure 13: Example for validation of NNs trained by simulation data and applied in real machine operation with SAW switched on. In 50 iterative steps from start to desired goal tune.

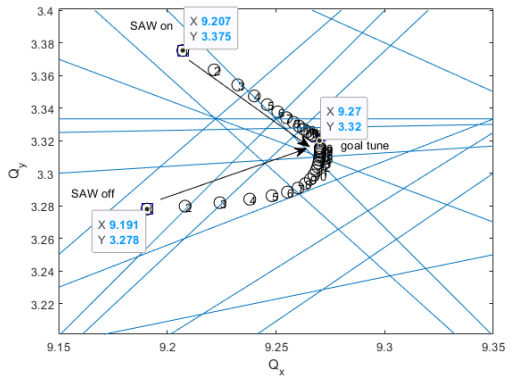

Figure 14: Example for verification of NNs trained by real machine data. In iterative steps from start to desired goal tune. NNs were applied to the simulation model with SAW switched on and off, respectively (compare to Fig. 6).

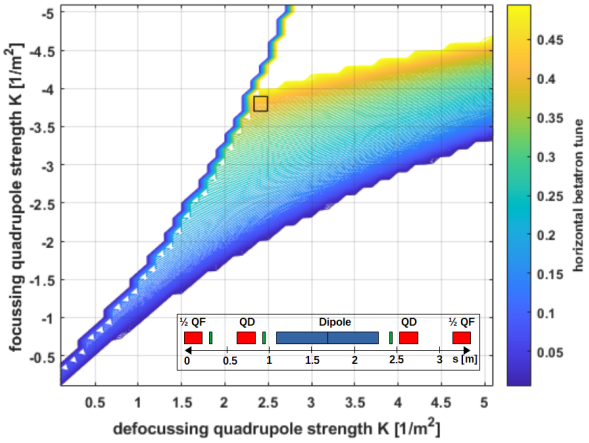

Figure 15: Necktie stability diagram for the DELTA triplet unit cell (see insert). The coloured area indicates the horizontal phase advance per unit cell in dependence of the triplet quadrupole strengths (QF, QD). White regions depict sections without periodic solutions for the transfer matrix M (trace M<2). The rectangular zone marks approximately the workspace of standard machine operation (AT simulation).

SUMMARY AND OUTLOOK

Machine learning based techniques are increasingly replacing classical, e.g. PID-based, feedback loops in the control system domains of particle accelerators. At the electron storage ring DELTA, it could be shown that trained neural networks can also be used for automated tunes adjustment. The network training was carried out with data recorded during real machine operation as well as with simulation data, based on an accelerator model. With both data sources, comparable tune correction accuracies were achieved in real machine operation. In contrast to the iterative PID methods, NNs are able to approach desired target tunes in single calculation steps, which could enable a more controlled scanning in the tune diagram. But, as long as the DELTA quadrupole power supplies lack the feature to drive synchronously, an iterative procedure for the NN-based tune control loop must necessarily be applied.

In principle, the ML-method presented in this article can also be transferred to other accelerator optimizations. For example, by varying the sextupole strengths and measuring corresponding chromaticity changes. NN trained by such data can be applied to match desired chromaticities. A similar method could also be applied to the adjustment of beam coupling and the associated beam size [17]. Therefore, one would have to vary the strengths of skew quadrupoles and determine the related coupling change.

ACKNOWLEDGEMENT

I would like to thank all colleagues of the DELTA team for many inspiring discussions, helpful suggestions and providing sufficient measuring time in accelerator shifts. I would also like to thank all my colleagues at other accelerator facilities who have provided software libraries, programming interfaces and documentations. Especially, I thank my colleague Peter Hartmann who built up the complete hard- and software setup of our reliable tune feedback system. Without this equipment this work would not have been possible in the short term.

REFERENCES

[1] M. Tolan, T. Weis, C. Westphal, and K. Wille, "DELTA: Synchrotron light in nordrhein-westfalen", *Synchrotron Radiation News*, vol. 16, pp. 9-11, Mar. 2003. doi:10.1080/08940880308603005

[2] S. Khan *et al.*, "Generation of Ultrashort and Coherent Synchrotron Radiation Pulses at DELTA", *Synchrotron Radiation News*, vol. 26, pp. 25–29, 2013. doi1:0.1080/08940886.2013.791213

[3] P. Hartmann, *et al.*, "Kicker Based tune Measurement for DELTA" , in *Proc. DIPAC'07*, Venice, Italy, pp. 277–279.

[4] D. Schirmer, "Orbit Correction with Machine Learning Techniques at the Synchrotron Light Source DELTA", presented at ICALEPCS'19, New York, USA.

[5] D. Schirmer and A. Althaus,"Integration of a Model Server into the Control System of the Synchrotron Light Source DELTA", presented at ICALEPCS'19, New York, USA.

[6] M. Grewe, "SVD-basierte Orbitkorrektur am Speicherring Delta", dissertation, Dortmund University, 2005.

[7] A. Terebilo, "Accelerator Modeling with Matlab Accelerator-Toolbox", in *Proc. PAC'01*, Chicago, USA, pp. 3203–3205.

[8] ATcollab, http://atcollab.sourceforge.net/index.html.

[9] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," Computer Journal, Vol. 7, 1964, pp. 149–154.

[10] M.F. Moller, "A scaled conjugate gradient algorithm for fast supervised learning", Neural Networks, Vol. 6, 1993, pp. 525–533.

[11] MATLAB/SIMULINK, Release 2017b, The MathWorks, Inc., Natick, Massachusetts, United States.

[12] L. E. Scales, "Introduction to Non-Linear Optimization", New York, Springer-Verlag, 1985

[13] W. H. Press, S A. Teukolsky, W. T. Vetterling and B. P. Flannery "Numerical Recipes in C", 2nd Ed., Cambridge University Press, 1992.

[14] D. Schirmer and K. Wille, "DELTA Optics", in IEEE Particle Accelerator Conference, volume 5, pp 2859–2861, San Francisco, California, 1991 May 6-9.

[15] N. Markus, R. P. Walker, and M. Poole, "SRS behaviour with a superconducting 5-Tesla wiggler insertion", Transactions on Nuclear Science, vol. 30, pp. 3127–3129, Sep. 1983.

[16] D. Schirmer, "Entwicklung von Strahloptiken für den Testspeicherring DELTA auf Basis der Triplett-Struktur", diploma thesis, Dortmund University, 1989, unpublished.

[17] S. C. Leemann, "Applying Machine Learning to Stabilize the Source Size in the ALS Storage Ring", presented at IPAC'20, Caen, France, May 11, 2020.

APPENDIX

Machine learning validations by use of correlation analysis

The following plots show NNs inputs and the corresponding NNs outputs after training for all quadrupole families in a matrix plot representation[6]. The histograms along the diagonals display the counts of strength variations ($\Delta K_{QF,QD}$ or $\Delta I_{QF,QD}$) for each quadrupole family inside given bin edges (x,y-axis for each subplot). Line-like distributions in the scatter plots (outside the diagonals) indicate a stronger correlation between quadrupoles than blurry, cloud-like distributions. Thus, e.g., NNs recognize (learn) that QFs and QDs are more correlated in each case. This could be an indication that the same tune shift could be achieved with an alternative combination of quadrupole settings than originally trained. The correlation analysis is still in the early stages and needs to be developed by further investigations.

[6] It creates a matrix of subaxes containing scatter plots. The basic data matrix (X) consists of random quadrupole strength changes (rows of X) for each quadrupole circuit (columns of X). Along the diagonal scatter plots are replaced with histogram plots of each column of X.

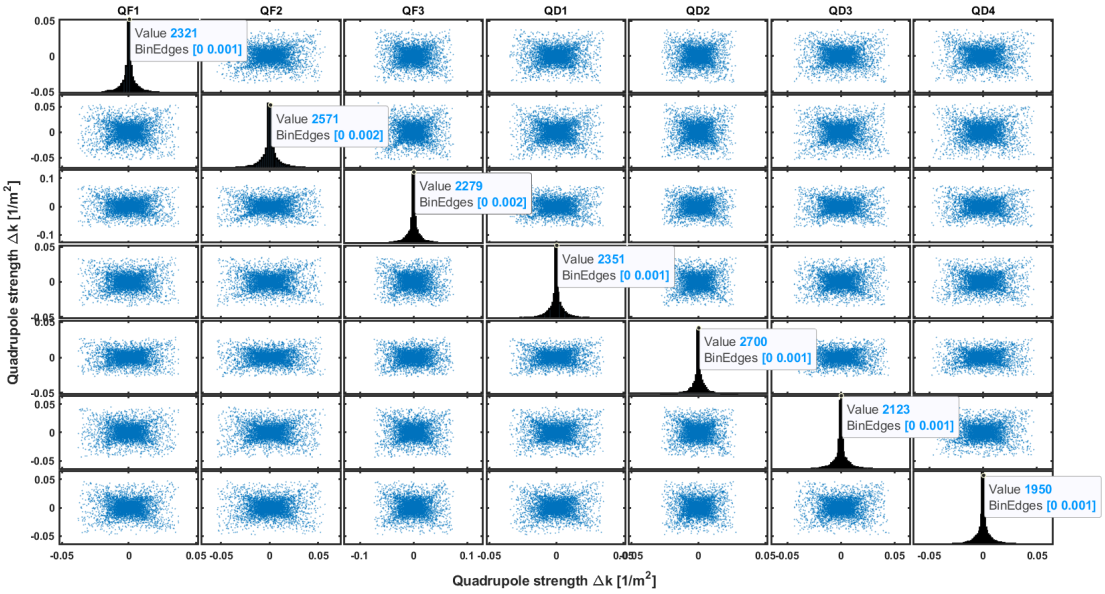

Figure 16: Distributions of 12000 random quadrupole settings as an input for NN training. Visual proof that the random settings works well (input for simulations, see Fig. 2).

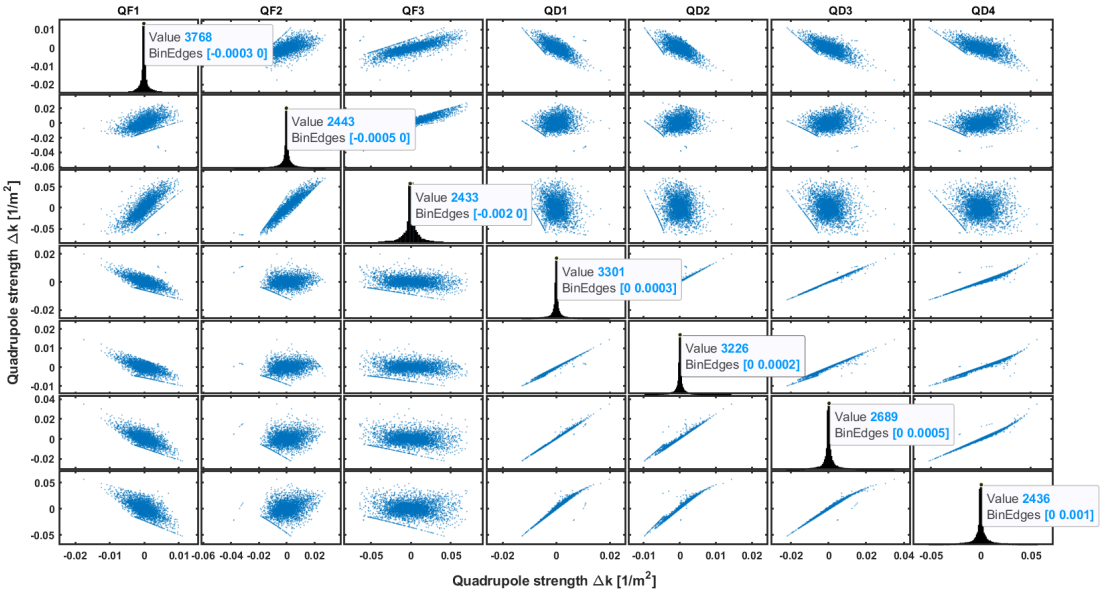

Figure 17: Outputs (quadrupole strengths distributions) of trained NN calculated for 12000 simulated tunes (compare with Fig. 16). Some QFs and QDs seems to be correlated among each other (QD01-04 and QF02-03).

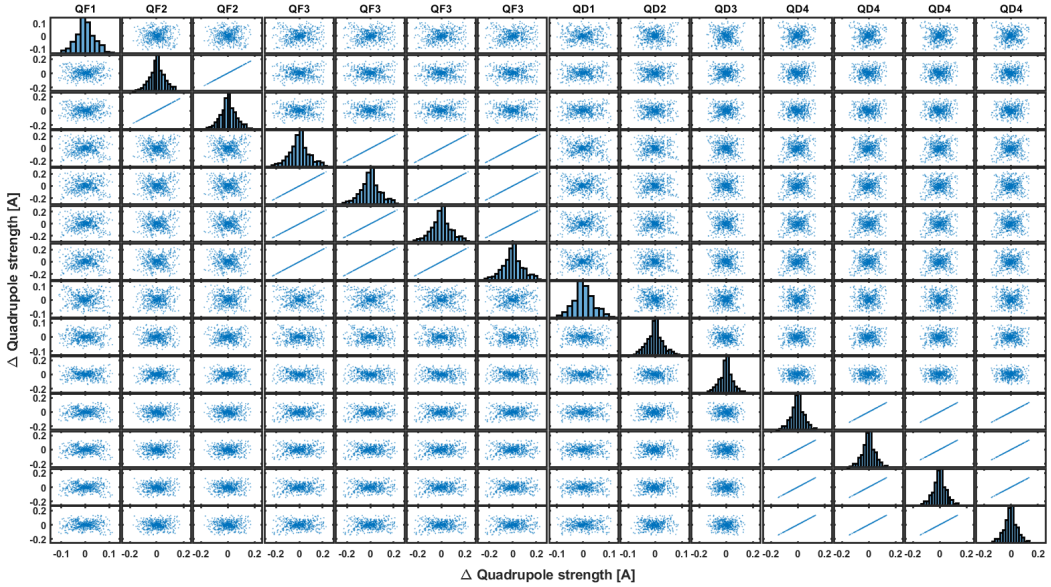

Figure 18: Distributions of 600 random real quadrupole settings as an input for NN training. Visual proof that the random settings works well. Related power supply circuits must be strongly correlated. This corresponds to lines in the scatter subplots (top left to bottom right: QF1, 2xQF2, 4xQF3, QD1, QD2, QD3, 4xQD4). Input for real machine settings.

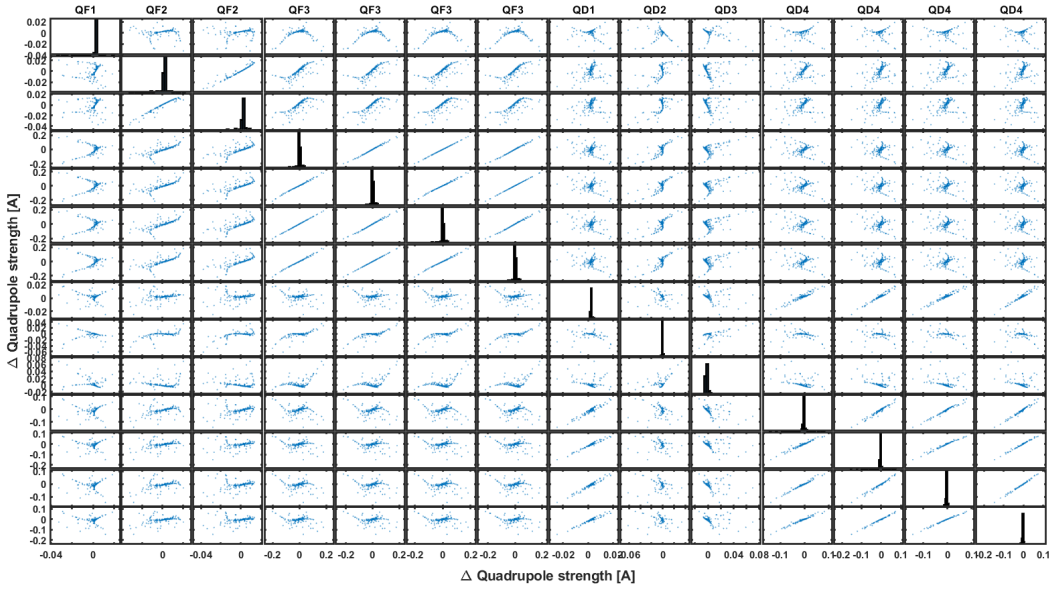

Figure 19: Trained NN outputs (quadrupole strengths distributions for individual power supply circuits) calculated for 600 measured tunes.

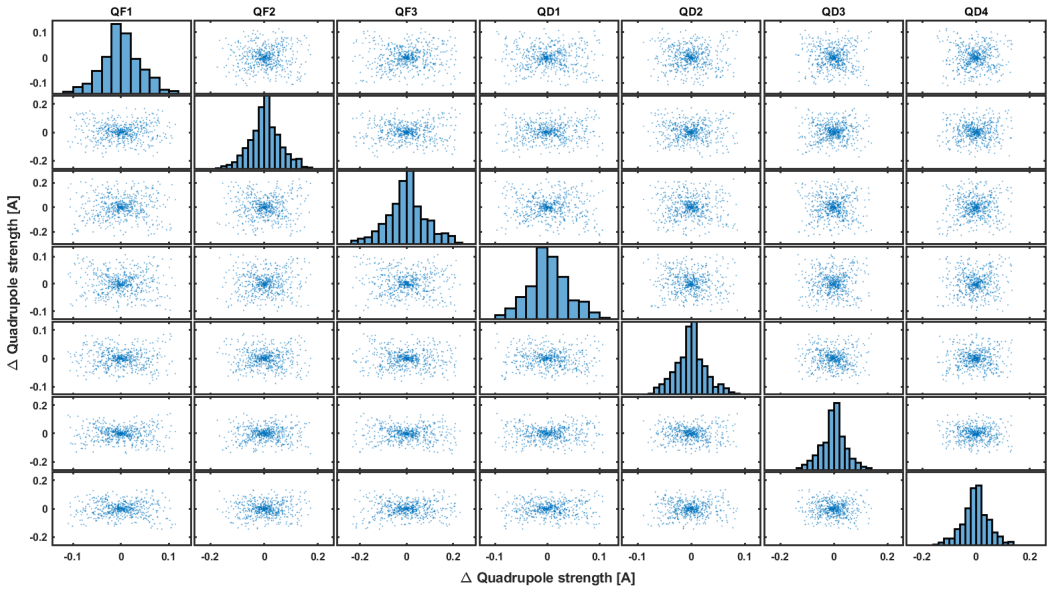

Figure 20: Distribution of 600 random real quadrupole settings as an input for NN training. Associated power supply circuits were combined into quadrupole families (compare with Fig. 18). Visual proof that the random settings works well (input for NN training with real machine data).

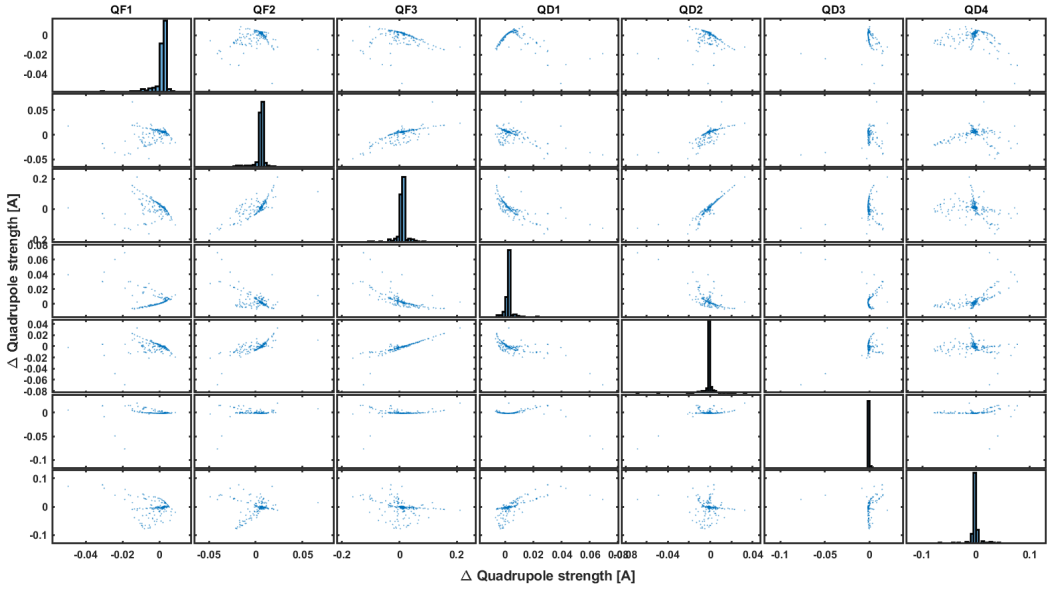

Figure 21: Trained NN outputs calculated for 600 measured tunes. Quadrupole strengths distributions for combined power supply circuits (see Fig. 20).