

Application of deep learning methods for online beam optics and coupling feedback in synchrotron light sources

Liyuan Tan[✉], Shouzhi Xuan[✉], Yihao Gong[✉], Xinzhong Liu[✉], and Xu Wu[✉]

Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201204, China

Shunqiang Tian^{✉*} and Wenzhi Zhang

*Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201204, China
and Shanghai Institute of Applied Physics, Chinese Academy of Sciences, Shanghai 201800, China*



(Received 22 November 2024; accepted 16 July 2025; published 25 August 2025)

This paper presents a beam optics and coupling feedback system for storage rings during operation, integrating deep learning methods with turn-by-turn (TBT) beam position monitor data. A deep learning network has been applied to extract phase advance and betatron coupling information implicit in TBT data. Subsequently, these data were fed into a multilayer neural network for rapid modeling and correction. The simulation data and sampled spectrum data from the Shanghai Synchrotron Radiation Facility (SSRF) were used to create training datasets for the network. The efficacy of this feedback system has been successfully demonstrated at the SSRF, enabling continuous monitoring and correction during each injection cycle. This method reduces optical distortion to less than 1% after correction, effectively suppresses optical and coupling distortion caused by insertion device movements during storage ring operation, and ensures robust performance under various operational conditions.

DOI: [10.1103/1gf-558m](https://doi.org/10.1103/1gf-558m)

I. INTRODUCTION

Ensuring the stability of beam optics in storage ring light sources is essential for optimal performance. A variety of factors can affect the optics during routine operations, including magnetic field drifts in dipole, quadrupole, and insertion device (ID) movements. Drifts in dipole and quadrupole power supplies typically happen over a certain period. Additionally, although most of the IDs with substantial optical disturbances have been compensated through an independent optics feedforward system [1], some residual disturbances remain.

Optical distortion can degrade the machine performance, reduce injection efficiency, shorten beam lifetime, and increase radiation dose. In some fourth-generation storage rings, the tunes must be precisely controlled to satisfy differential resonance conditions for the generation of round beams [2,3]. The traditional tune feedback system [4] is unable to simultaneously minimize the optical distortion. Therefore, an optics correction system is essential for

maintaining high performance and stability throughout the operation.

The traditional approach for optics correction is the linear optics from closed orbits (LOCO) method [5]. The primary disadvantage of the LOCO method is that it is time consuming due to the necessity of measuring the orbit response matrix. In recent years, several improvements have been proposed to combine fast measurement with high accuracy [6,7]. However, due to the necessity of fitting the orbit response matrix, these methods are not suitable for optics correction during user time. Turn-by-turn (TBT) orbits can be sampled without affecting user experiments by the beam position monitor (BPM). The model-independent analysis method [8] was used to calculate linear optics parameters from TBT data and the independent component analysis (ICA) technology [9] has greatly improved the accuracy of betatron mode extraction under coupled beam motion. Reference [10] recognized that small amplitudes are workable with the ICA methods and proposed using residual oscillations for online optics monitoring and correction. The N beam position monitor (N-BPM) method [11,12] has greatly improved the accuracy of model fitting by relying solely on phase advance. All these methods rely on extracting optics information hidden in beam orbit data and use constrained fitting techniques to achieve modelization.

In recent years, machine learning methods have been widely applied in accelerator physics. For instance,

*Contact author: tiansq@sari.ac.cn

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

clustering techniques and supervised learning models have been utilized to accelerate computationally costly simulations, enabling sample-efficient dynamic aperture estimation [13] and electron beam size optimization [14]. Physics informed neural networks, Bayesian optimization, and reinforcement learning methods have demonstrated success in optics model tuning [15], beam size correction [16], and beam trajectory steering [17]. Additionally, virtual diagnostics techniques have been implemented for 6D phase space reconstruction [18]. Regarding storage ring optics correction, Fol *et al.* achieved faulty BPM detection through unsupervised learning [19] and predicted quadrupole errors using multilayer neural networks [20]. Subsequent studies have shown that image-based analysis employing convolutional neural network (CNN) can effectively extract beam information from turn-by-turn (TBT) data [21]. In this research, we extend the CNN-based image analysis to optics correction in synchrotron radiation light sources. The proposed methodology extracts transverse betatron information of the beam and subsequently employs a multilayer neural network for quadrupole and skew-quadrupole error prediction.

In order to realize the online optics measurement and correction, in this study, the betatron oscillation is caused by the residual disturbances of the injection kicker. To minimize the impact on the beamline users, the perturbations are usually small, which make the measurements more sensitive to BPM noise. At the same time, the amplitude decreases rapidly due to the decoherence and recoherence of the betatron oscillations [22] and the transverse feedback system. In addition, potential beam instability under conditions of high beam current and low chromaticity makes the ICA method challenging for accurate extraction of betatron modes. The faulty BPMs and BPM timing errors also need manual identification and calibration in traditional methods. These factors present significant challenges to the measurements and correction of linear optics and coupling in daily operation.

To address these issues, a deep learning method was used to extract the betatron motion information from the TBT data. Only phase advance information was used for optics correction. The amplitude ratios of the projected betatron motion onto the horizontal and vertical planes for the two normal modes [23,24] were used for coupling correction. The sampled TBT data must be preprocessed at first by removing faulty BPMs and reducing BPM noise. A CNN was applied to extract linear optics and betatron coupling parameters from the TBT data. Then, a multilayer neural network referred to as a constrained fitting neural network (CFNN) in this research was studied to replace the traditional constrained fitting method. The workflow is illustrated in a flowchart, as shown in Fig. 1. Compared to the traditional TBT method, the application of the deep learning method realizes the automatic calibration of BPM timing errors and significantly mitigates the impact of

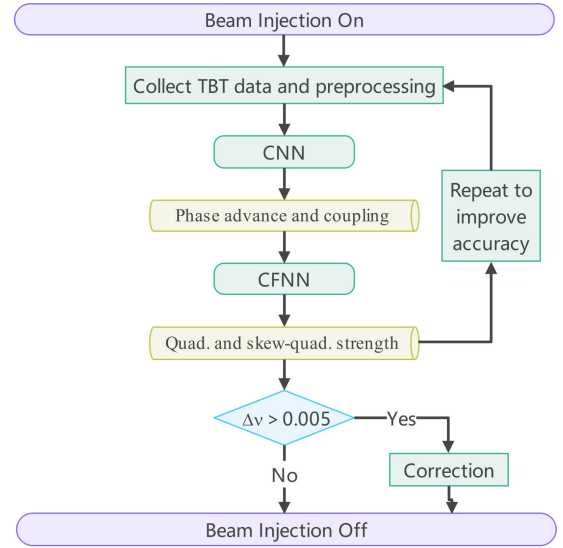


FIG. 1. Operational flowchart of the deep learning-based beam optics and coupling feedback system, showing real-time TBT data acquisition from the residual disturbance during beam injection, feature extraction from CNN, and fitting quadrupole and skew-quadrupole errors for optics correction from CFNN. Corrective action is only taken when the tune shifts are greater than 0.005.

longitudinal beam instability under high beam current, thereby improving the robustness and achieving higher accuracy. The CFNN streamlines and accelerates this process of modelization, providing results that are comparable to those obtained using the LOCO method. The optics feedback system has been successfully tested in the Shanghai Synchrotron Radiation Facility (SSRF) storage ring.

In this study, we used ELEGANT software [25] to track particles and combined this with the sampled spectrum to build our training dataset. The neural networks were designed and trained using the PyTorch library. The model training process was accelerated by two NVIDIA A6000 GPUs.

The organization of this paper is as follows. Section II describes the experimental setup and the preprocessing of the TBT data. Section III presents the deep learning network and the training method. Section IV discusses the online optics and coupling feedback experiments at SSRF. Section V concludes the research presented in this paper.

II. TBT DATA PREPROCESSING

A. Experimental setup

In SSRF (Fig. 2), a total of 138 BPMs are evenly distributed among four superperiods consisting of 20 standard cells [26]. There are three high-precision BPMs located at the first, fifth, and end of each standard cell, which are supported by independent structures, thus ensuring enhanced measurement accuracy. During our experiment, there were five faulty BPMs, including BPM10-5x,y, BPM11-1y, BPM11-5x,y, BPM17-1x, and

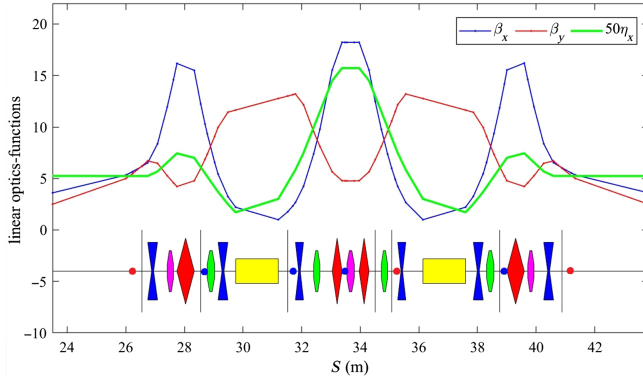


FIG. 2. BPM layout in the standard cell of SSRF, including three high-precision BPMs (red dots) with independent support structure, while the remaining four BPMs (blue dots) are installed directly on the vacuum chamber.

BPM18-1x. Consequently, a total of 133 BPMs are available for obtaining TBT data in the following study.

In daily operation, the injection kicker is carefully optimized to ensure that the injection process causes minimal disruption to the user's experiments. Typically, the initial oscillation amplitude during injection is kept within 100 μm in the horizontal plane and the vertical betatron motion induced by coupling can be ignored. The top-up beam current is 220 mA and the chromaticity is set to 1.5 in both the horizontal and vertical planes. The transverse feedback system [27] is turned on to suppress transverse beam instabilities.

The TBT data were sampled for online optics and coupling correction during the injection process. The BPM switch must be set to “off” to ensure that the closed-orbit measurements do not interfere with the TBT measurements. The slow orbit feedback system is suspended during beam injection, so this does not affect the closed orbit. In general, a larger kicked oscillation amplitude is beneficial to the accuracy of the measured TBT data (smaller proportion of BPM noise). In third-generation light sources, coupling is typically kept to a minimum to improve brilliance and injection efficiency. The coupling provided by the storage ring itself is insufficient to generate enough betatron oscillation amplitude in the vertical plane. To solve this problem, the injection kicker was rotated by 3 mrad to induce additional betatron oscillation with an amplitude of approximately 75 μm in the vertical plane.

Due to the transverse feedback, the beam damping time is much smaller than the design value. The disturbance of the injection phase is in the order of microseconds. In third-generation light sources, users are more concerned with the average brightness. Such a short perturbation will not have much influence on the user experiments. In addition, under the influence of the short damping time, the ratio of BPM noise increases rapidly with increasing turns. To ensure the reliability of the results, only TBT data from turns 1 to 512 are used in the following study.

B. BPM data correction

To simplify the subsequent neural network training process, we precorrected the BPM gains and coupling coefficients in the first step. This pre-correction is beneficial for the neural network to better learn the actual relationships between TBT data, linear optics parameters, and the lattice model more effectively.

The impact of BPM gain and coupling coefficients on the beam orbit is expressed as follows:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} g_x \bar{x} + c_x \bar{y} \\ c_y \bar{x} + g_y \bar{y} \end{pmatrix}, \quad (1)$$

where x, y are the measured orbit coordinates and $\bar{x}, \bar{y}$ are the actual orbit coordinates, $g_{x,y}$ are BPM gain coefficients and $c_{x,y}$ are BPM coupling coefficients.

In the preliminary correction, we assume that the gain and coupling coefficients for each BPM have nearly constant values, with only small variations over the course of long-term operation. The BPM coefficients for each measurement were treated as distributed random errors centered around these long-term stable values. Consequently, we used these long-term stable values to perform a preliminary correction of the BPM gain and coupling coefficients.

Data from the weekly routine maintenance implemented in SSRF storage ring from the spring start of 2024 over a four-month period have been sampled, as shown in Fig. 3. During this period, the root mean square (rms) values of the variation in the gain and coupling coefficients were approximately 0.03 and 0.04 in both the horizontal and vertical planes. The median values of these long-term fitting results were used to precorrect the TBT data.

During TBT measurements, external disturbances or BPM jumps can cause sudden changes in the measurement results, which are occasional failure of the betatron modes extraction. Hardware or communication failures can also result in BPM errors, with the read-back values being

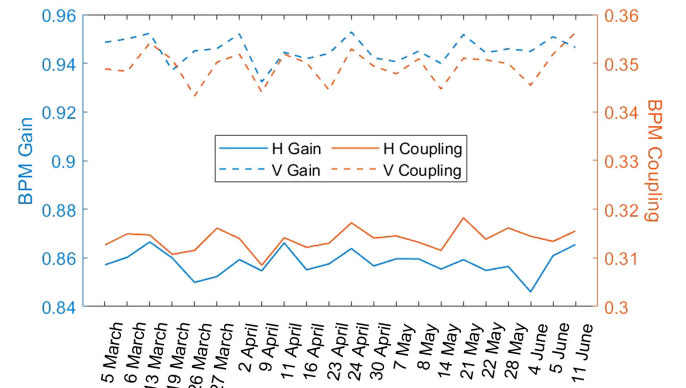


FIG. 3. Variations in BPM gain and coupling coefficients for BPM01-1 in the SSRF storage ring sampled during weekly maintenance in 2024.

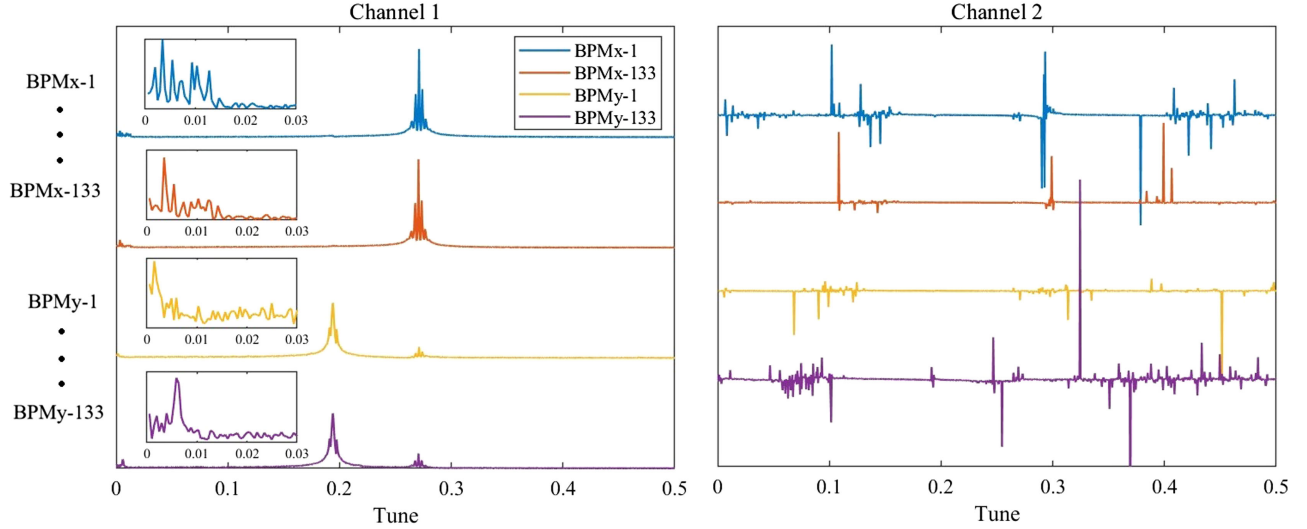


FIG. 4. Example of the input data from TBT FFT results. In the figure, only the horizontal and vertical corresponding results of 2 BPMs are plotted, and the other 131 BPMs are omitted. Subplots show details near longitudinal tunes. Channel 1 is the absolute value of FFT results, and channel 2 is the arctangent component of the angle (the imaginary part divided by the real part of FFT results). The dimension of the input data is $2 \times 266 \times 255$.

modulated noise, fixed values, or NaN (not a number) values. These outliers can make the measured beta beating inaccurate. For the deep learning method, matrices containing NaN values cannot be computed, so it is crucial to fill in or remove the affected BPM data.

In this study, we first flagged the NaN values as outliers. Subsequently, an unsupervised learning method based on isolation forest was employed to identify faulty BPMs and local outliers [19], with the contamination parameter set to 0.01 to ensure detection reliability. Faulty BPMs were completely removed, while the identified local outliers were set to random noise levels close to the mean value. Such noise information tends to dissipate effectively during propagation through convolutional neural networks. Following outlier processing, singular value decomposition (SVD) was applied to the input TBT matrix, retaining only the first ten eigenvalues for data denoising. Additionally, BPM timing errors induce data shifts in measured TBT matrices, causing significant deviations in phase advance calculations. Traditional manual calibration of BPM timing errors method is not suitable for online correction. In our neural network, BPM timing errors were automatically corrected by incorporating such errors into training samples (see Sec. III B), eliminating tedious manual intervention. In the following study, fast Fourier transform (FFT) was applied to the input TBT data, concentrating betatron oscillation-related information around corresponding tunes and further mitigating the impact of residual outliers.

C. FFT data transformation

Subsequently, we performed an FFT on the 512 turn TBT results from each BPM, retaining the spectral results from

2 to 256 (removing the DC component at zero frequency and the mirror symmetric part). Different BPMs were used as rows and the frequencies of the FFT results were used as columns to form the input data. The absolute value and arctangent component of the angle (the imaginary part divided by the real part of FFT results) constitute the two channels of the input feature graph, as shown in Fig. 4. Finally, the preprocessed TBT data forms a two-channel feature graph, with the dimension of the input data being $2 \times 266 \times 255$.

III. NEURAL NETWORK CONSTRUCTION AND MODEL TRAINING

This section introduces the CNN used to extract linear optics and betatron coupling parameters from TBT data, and the CFNN used for modelization. The training dataset was divided into two parts, including simulated spectrum and sampled spectrum. The simulated spectrum contains the ideal single-particle motion and the beam central motion. The sampled spectrum includes oscillations caused by beam instability, the residual disturbance of the transverse feedback, and BPM noise.

A. Convolutional neural network and constrained fitting neural network

The TBT data contain both the temporal and spatial characteristics of the beam. The CNN is adept at processing these dimensions simultaneously, making it well suited for extracting the spatiotemporal correlations present in TBT data. By focusing on small areas within the data through local perception, the CNN can identify local patterns and spatial distribution characteristics within the frequency

spectrum, which enhances its performance when dealing with complex signals.

Extracting features from the FFT spectrum is similar to extracting the betatron modes using ICA method. The difference is that traditional SVD and ICA algorithms usually assume that the signal is linear, while the deep learning network is able to model complex nonlinear relationships. For many practical applications, especially in the daily operation, these nonlinear features are crucial to be filtered out.

The whole convolutional neural network is composed of a convolutional part and a fully connected part, as shown in Table I. The convolution part consists of six convolution structures for feature extraction and the fully connected part consists of three fully-connected layers for feature analysis and result computation. Each convolution structure consists of 2D convolution (Conv2D) layers, activation layers, and a 2D max pooling (MaxPool2d) layer. The kernel size of all Conv2d layers is [3 3] with the stride of [1 1]. In the first four convolution structures, two Conv2d layers with the kernel size of [3 3] were cascaded to achieve the effect of kernel size of [5 5] while increasing the computational speed. The kernel size of the MaxPool2d layer is [2 2]. Pooling operation reduces the feature dimensions by half while each convolution structure doubles the number of features. All padding steps were performed in Conv2d layers, and the padding of MaxPool2d layers was set to [0 0]. The activation function used throughout is the Leaky ReLU, with the negative input coefficient set to 0.01. At the end of the convolution part, there are 512 feature maps with the size of 5×4 . These output feature maps were then fed into the fully connected part to compute the phase advance and betatron coupling parameters corresponding to each BPM in both the horizontal and vertical planes, with the final output size being 1×532 .

The traditional constrained fitting method requires calculating the Jacobian matrix at first for model fitting. If the skew quadrupole has a large offset or the coupling is large, the nonlinear effect of the skew quadrupole will be obvious, the Jacobian matrix needs to be recalculated at each iteration, which is extremely time consuming. In this study, we have replaced the traditional constrained fitting method with CFNN. The difference is that the trained neural network was used directly to fit the optical and coupling parameters and achieve modelization. This means that the BPM weights cannot be changed in the network and must be determined when generating training samples. Additionally, unlike the N-BPM method, the neural network can automatically identify the relationship between BPMs and generate BPM weights according to the training samples instead of manual selecting the BPM combination.

The phase advance and betatron coupling values derived from the CNN were then fed into a three layer fully connected CFNN to fit the K value errors of the quadrupoles and skew quadrupoles. These results were then

TABLE I. CNN for extracting betatron oscillation information from TBT data consists of six convolutional layers and two fully-connected layers. In the first three convolutional layers, two [3 3] convolutional kernels are employed to replace the [5 5] kernels. The network takes processed FFT spectra as input and outputs phase advance and betatron coupling parameters for both horizontal and vertical planes across 133 BPMs.

Layer	Kernel size	Padding	Output shape
Input layer			[2,266,255]
Conv2d	[3 3]	[1 1]	[16,266,255]
LeakyReLU			[16,266,255]
Conv2d	[3 3]	[2 2]	[16,268,257]
LeakyReLU			[16,268,257]
MaxPool2d	[2 2]	[0 0]	[16,134,128]
Conv2d	[3 3]	[1 1]	[32,134,128]
LeakyReLU			[32,134,128]
Conv2d	[3 3]	[2 1]	[32,136,128]
LeakyReLU			[32,136,128]
MaxPool2d	[2 2]	[0 0]	[32,68,64]
Conv2d	[3 3]	[1 1]	[64,68,64]
LeakyReLU			[64,68,64]
Conv2d	[3 3]	[1 1]	[64,68,64]
LeakyReLU			[64,68,64]
MaxPool2d	[2 2]	[0 0]	[64,34,32]
Conv2d	[3 3]	[1 1]	[128,34,32]
LeakyReLU			[128,34,32]
Conv2d	[3 3]	[2 1]	[128,36,32]
LeakyReLU			[128,36,32]
MaxPool2d	[2 2]	[0 0]	[128,18,16]
Conv2d	[3 3]	[2 1]	[256,20,16]
LeakyReLU			[256,20,16]
MaxPool2d	[2 2]	[0 0]	[256,10,8]
Conv2d	[3 3]	[1 1]	[512,10,8]
LeakyReLU			[512,10,8]
MaxPool2d	[2 2]	[0 0]	[512,5,4]
Linear			[4096]
LeakyReLU			[4096]
Linear			[2048]
LeakyReLU			[2048]
Linear			[532]

transformed into currents of power supplies for optics and coupling correction with the ideal current-field curve. The summary of the CFNN is shown in Table II. The input size is 532, includes 133 BPM pairs of horizontal and vertical phase advance and betatron coupling. The output size is 226, including the K value error of 206 quadrupoles and 20 skew-quadrupole correctors.

Essentially, CNN and CFNN are the front and back parts of one network. However, by introducing a physics-informed intermediate layer, we partitioned them into two distinct networks. Compared to treating it as a single neural network, this division endows both parts with

TABLE II. CFNN for quadrupole and skew-quadrupole error prediction consists of two fully-connected layers. The network takes phase advance and betatron coupling parameters from 133 BPMs in both horizontal and vertical planes as input and outputs K value errors for 206 quadrupoles and 20 skew quadrupoles.

Layer	Output shape
Input layer	[532]
Linear	[1064]
LeakyReLU	[1064]
Linear	[532]
LeakyReLU	[532]
Linear	[226]

distinct physical meanings while simultaneously significantly reducing the sample size required for neural network training.

B. Generation of training datasets

For a new network, a large number of samples containing essential features are required for initial training. At low beam current and low chromaticity, the ideal single-particle results differ little from the TBT of a real machine. The ideal single-particle tracking can generate a large number of training samples through parallel computation, which is very suitable for initial training of the model. Therefore, single-particle tracking was used to generate training datasets at first.

Random variations were introduced into all 206 quadrupoles and 140 sextupoles (skew-quadrupole errors) to induce changes in phase advance and betatron coupling. In general, approaching the problem from specific cases to more general ones tends to yield higher network convergence efficiency than the reverse. Therefore, for most samples, only a few magnets were adjusted at a time. To facilitate network convergence, we created a specific sample mixing ratio. The detailed parameters, including probability, number, and rms value, are presented in Table III.

In reality, while linear optics correction demonstrates sufficiently high linearity so that complex training sample mixtures are not required for model training, practical accelerator facilities present additional challenges. Specifically, skew-quadrupole corrector coils in actual machines are often intertwined with sextupoles, and their potential offsets may introduce nonlinear effects. This causes the response matrix of the skew-quadrupole correctors to vary under different coupling conditions, necessitating continuous updates of the Jacobian matrix in ICA methods. In this study, these factors make training neural networks for determining skew-quadrupole corrector strengths particularly challenging, as simplistic training samples fail to capture the full features of nonlinear effects. This fundamental limitation compels us to ensure that

TABLE III. Statistical distribution of error magnet quantities, rms magnitudes, and sample proportions associated with quadrupole and skew-quadrupole errors. TBT samples were generated by introducing these errors into the ideal lattice through particle tracking simulations.

Probability %	Number skew quadrupole	Quadrupole/ rms $\times (\Delta K/K)$	Magnitude rms $\times (\Delta K/K)$
10	1/1		0.0210/0.018
10	2/2		0.0180/0.014
10	4/4		0.0140/0.010
10	8/8		0.0130/0.007
10	16/16		0.0115/0.005
10	32/32		0.0070/0.004
10	64/64		0.0065/0.003
10	128/80		0.0050/0.002
10	160/100		0.0040/0.001
10	206/140		0.0030/0.0005

sample randomness comprehensively covers potential operational scenarios, the primary rationale for implementing this specific sample mixing ratio. Furthermore, enforcing physically reasonable constraints in varying conditions helps mitigate the risks of failures in linear optics calculations or particle loss during the tracking of TBT data.

For single-particle tracking, the initial 6D coordinates are specified directly. In training datasets generation, the initial amplitude was set as a random number in the range of 50–300 μm , and the energy spread was set to zero. The longitudinal oscillation was completely eliminated at this stage and will be added to the training datasets in future studies. Figure 5 shows an example of the TBT results and corresponding FFT spectrum. It can be seen that the ideal sample contains only transverse betatron oscillations.

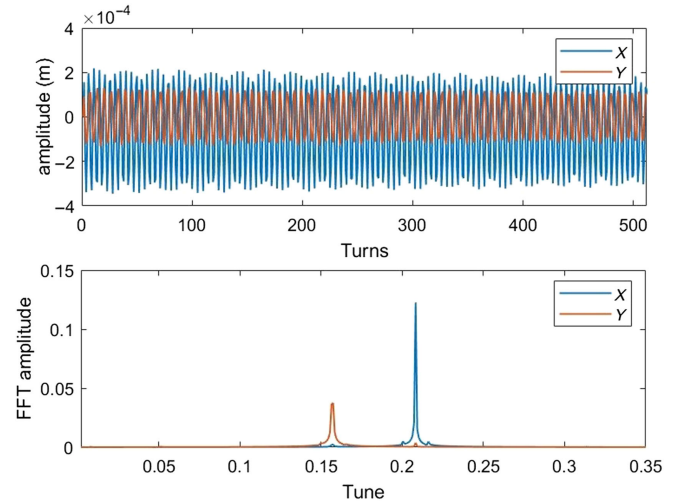


FIG. 5. Example of TBT data and corresponding FFT spectrum at BPM01-1 generated from 512 turns single-particle tracking, while the initial 6D coordinates were specified as [180 μm 0 100 μm 0 0 0].

These samples are helpful for training the extraction of the most basic features and relationships.

These simulated FFT results were used as inputs. The phase advance error (difference between the actual phase advance and the ideal lattice phase advance) and the betatron coupling were used as outputs. Using the phase advance error instead of exact values can increase the sensitivity of the outputs to variations in different inputs and help the network converge.

At high beam current, due to the bunch lengthening effect of the third harmonic cavity, the decoherence and recoherence motions are significant. In order to enable the network to identify these cases, a part of training samples focused on the decoherence and recoherence motions was incorporated into the training datasets. A total of 1024 particles were generated to form the initial bunch based on the equilibrium parameters of the storage ring. The quadrupoles and skew quadrupoles were set to follow the same rules as in Table III. The horizontal and vertical chromaticities of the beam were randomly set between 1 and 4. Figure 6 shows an example of the simulated TBT result and corresponding FFT spectrum, while the bunch lengthening factor is 2 and the chromaticities $\xi_{x,y} = [3\ 3]$.

In both the single-particle and coherence motion datasets, the beam was assumed to be in equilibrium. In daily operation, low chromaticities are often selected to increase beam lifetime. In such cases, potential beam instability may decrease the accuracy of betatron mode extraction by PCA and ICA methods (discussed in Sec. IV). In general, transverse beam instability can be suppressed by transverse feedback. Therefore, the longitudinal spectrum becomes more important. The deep learning network method must

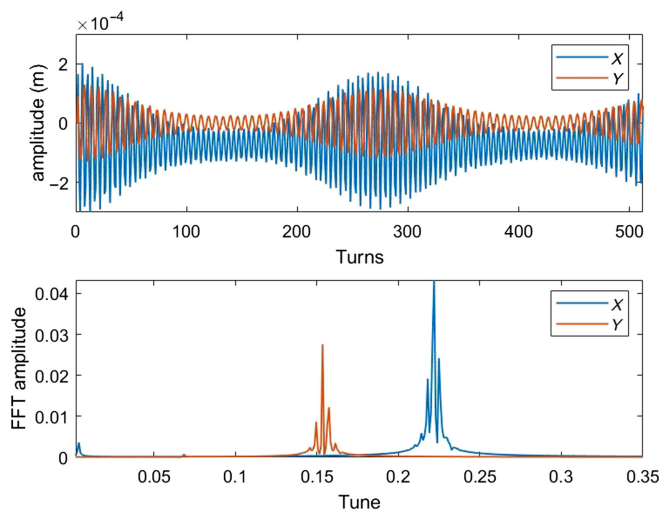


FIG. 6. Example of decoherence and recoherence motion TBT data and corresponding FFT spectrum at BPM01-1 generated from 512 turns tracking, while 1024 particles were generated to form the initial bunch based on the equilibrium parameters of the storage ring. The bunch lengthening factor is 2 and the chromaticities $\xi_{x,y} = [3\ 3]$.

be able to distinguish the transverse betatron motion from a complex environment, including betatron motion, residual disturbance from the transverse feedback kicker, beam instability, and noise. The core of this process is the extraction of spectral features.

To generate training dataset that includes the actual operation conditions, we sampled the TBT FFT spectrum during operation at SSRF. In most cases, the peaks present in the FFT spectrum are almost invisible. An extreme example is presented here to clarify the approach in this study, as shown in Fig. 7. It can be seen that there are many spectral harmonic peaks near the longitudinal tune. The transverse oscillation spectrum (near the frequency corresponding to the horizontal and vertical tunes) is generated by residual disturbances from the transverse feedback kicker (which was not functioning optimally). These sampled FFT spectra were then superimposed onto the simulation results as final training datasets. Figure 4 shows an example of the input data from the first and the last BPM, the other 131 BPMs are omitted.

Table IV lists the primary data components included in the mixed training dataset and their corresponding sources of errors. In addition to the quadrupole and skew-quadrupole field errors already considered in simulations, we incorporated BPM timing errors with a maximum range of 10 turns, randomly added to the simulated results of each BPM. This approach allows the neural network to autonomously correct deviations in phase advance caused by BPM timing errors. The final training dataset contains real BPM noise, necessitating preprocessing through the cleaning and denoising methods described in Sec. II B before training. By replicating the FFT spectral characteristics of the actual machine in the synthetic training data generation process, the deep learning network learns to extract betatron motion features from complex operational environments.

Figure 8 illustrates the complete workflow from generating mixed training samples to K value prediction, along with the architecture of the neural network. In this framework, the phase advance and betatron coupling parameters output by the CNN are treated as a physics-informed intermediate layer. The loss associated with this intermediate layer is combined with the loss of the final prediction results during model training to enhance convergence.

C. Model training

A total of 200,000 single-particle samples and 50,000 coherent motion samples were generated, with 70% allocated to the training set, 15% to the test set, and 15% to the validation set. Before model training, each feature of these samples was normalized and standardized. The mean squared error (MSE) loss function and the root mean square propagation (RMSprop) algorithm were selected for model training. The initial learning rate was set to 0.001, which is adaptive and decreases dynamically during the training process. The batch size for the training, test, and

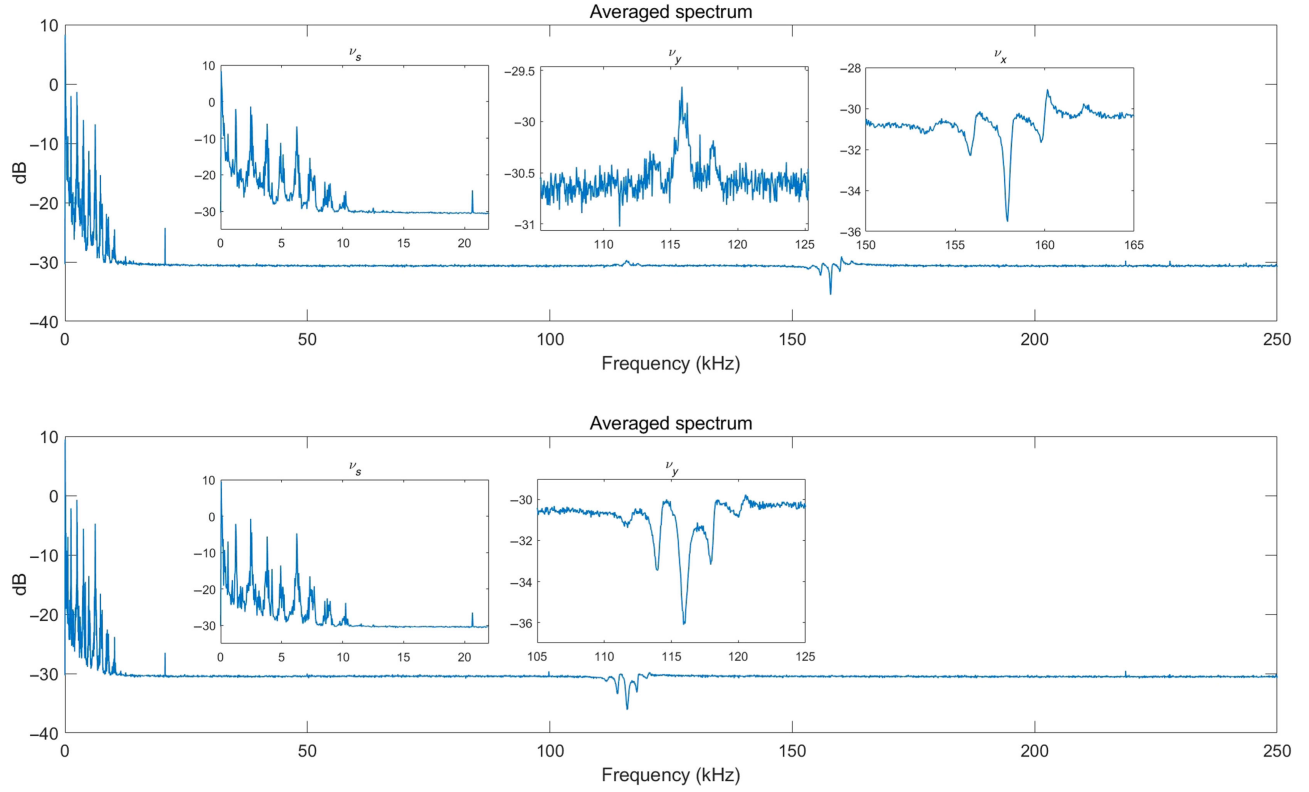


FIG. 7. Example of sampled FFT spectra in the horizontal (above) and vertical (below) planes is shown with the injection kicker “off” during routine 220 mA beam current operation at SSRF. Subplots magnify spectral features near horizontal, vertical, and longitudinal tunes, where perturbations adjacent to transverse tunes arise from residual disturbance of the transverse feedback kicker, while harmonics near the longitudinal tune originate from beam longitudinal instabilities.

validation datasets was set to 1024. Figure 9 shows the MSE loss as a function of epochs. There is strong consistency between samples. The loss values of the training, test, and validation datasets decrease simultaneously. After about 800 epochs, the learning rate of the network dropped to 0.0001 and the MSE loss converged to the minimum value of about 0.022. After the loss has fully converged, the training loss remained at a comparable level to that of both the validation and test sets, indicating that the model did not exhibit overfitting. The total training time of the model is about 3 h. Figure 10 shows an example

of the validation set for quadrupoles after training and Fig. 11 shows an example of the validation set for skew quadrupoles. It can be seen that the model performs well in predicting the K value error for quadrupoles and skew quadrupoles.

IV. EXPERIMENTS AT SSRF

This section discusses the experimental results of the deep learning network for optics and coupling measurement and correction. In our experiment, the transverse betatron motion information inside the mixed spectrum was obtained from the trained deep learning network. The output of the intermediate layer, including phase advance and betatron coupling, was only used for online monitoring, while the final output quadrupole and skew quadrupole K value errors were used directly for optics and coupling correction.

Figures 12 and 13 present comparisons of the phase advance errors and beta beating obtained from LOCO and the deep learning network. The error bars represent the standard deviations calculated from 12 consecutive TBT measurements during the injection cycle. The mean values of the 12 measurements were used for the optical model fitting to mitigate the impact of phase advance errors

TABLE IV. Primary components comprising the mixed training dataset and their corresponding sources of errors.

Components	Sources
Quadrupole field errors	Table III
Skew-quadrupole field errors	Table III
BPM timing errors	Simulated
BPM noise	Sampled
Residual disturbance from the transverse feedback kicker	Sampled
Longitudinal beam instability	Sampled
Faulty BPM data	Sampled

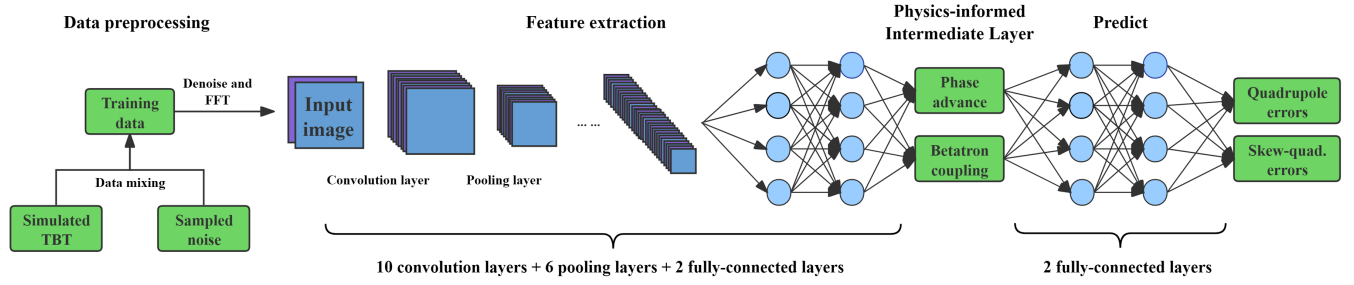


FIG. 8. Schematic diagram of the complete network architecture, including mixed training dataset generation, data denoise and preprocessing, a CNN for feature extraction, and fully-connected layers for quadrupole and skew-quadrupole error prediction. The phase advance and betatron coupling parameters output by the CNN are treated as a physics-informed intermediate layer and integrated into the model training process.

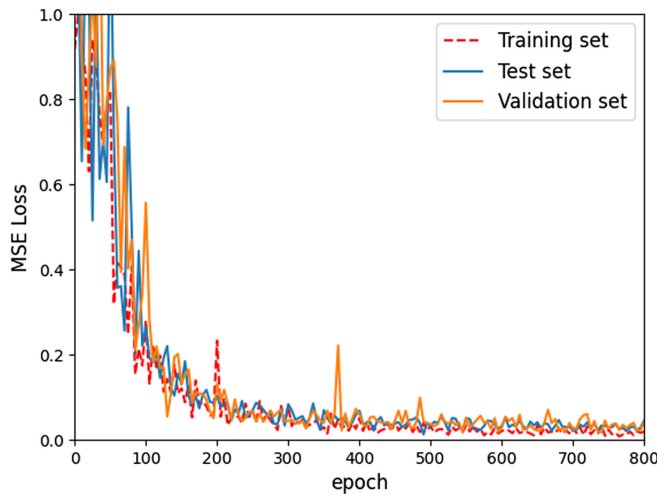


FIG. 9. MSE loss values of the training, test, and validation set as functions of epochs during model training.

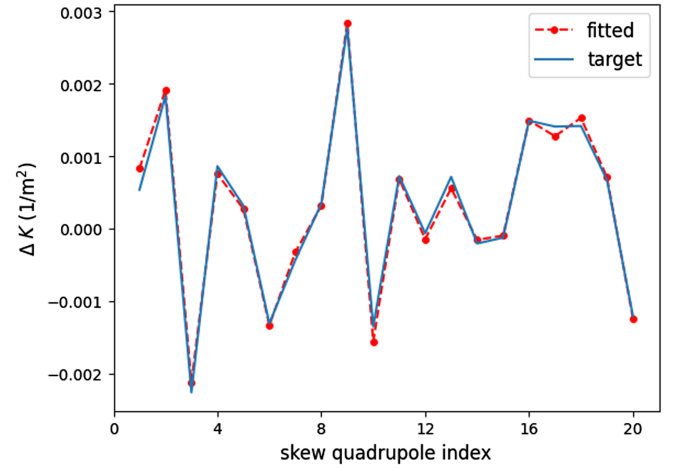


FIG. 11. Example demonstrates the comparison between the neural network's fitted values and target values for 20 skew-quadrupole errors in the validation dataset.

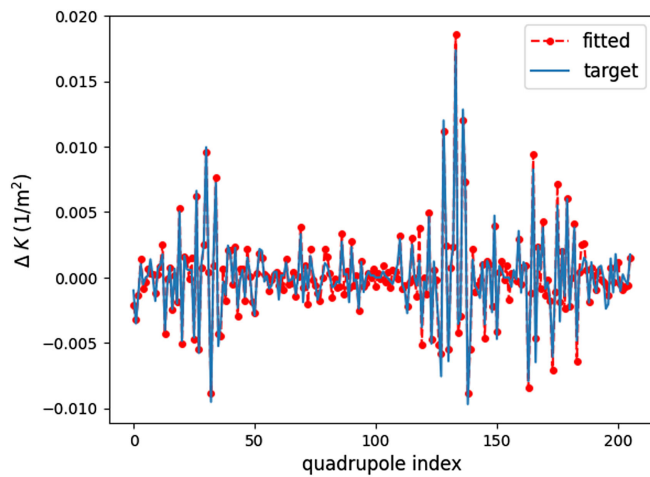


FIG. 10. Example demonstrates the comparison between the neural network's fitted values and target values for 206 quadrupole errors in the validation dataset.

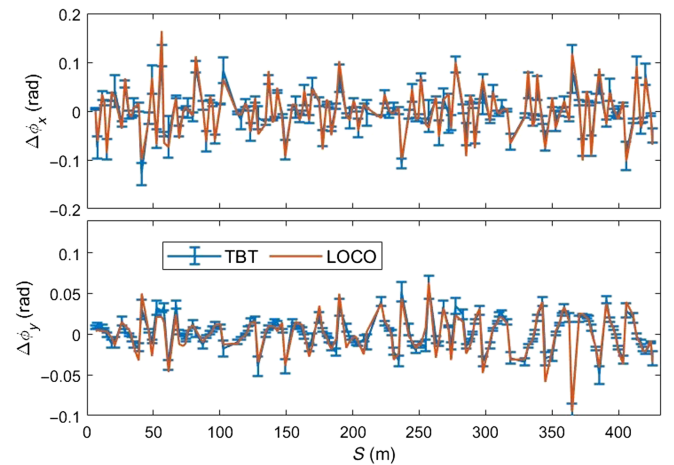


FIG. 12. Comparison of phase advance errors obtained through LOCO and TBT (deep learning network) methods in horizontal (top) and vertical (bottom) planes. The solid lines represent the mean values, and the error bars represent the standard deviations calculated from 12 continuous TBT measurements.

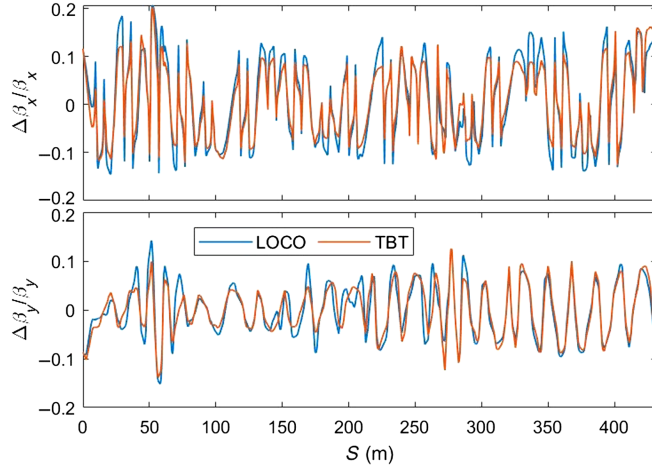


FIG. 13. Comparison of beta beating between the fitting lattice and the ideal lattice obtained by LOCO and TBT (deep learning network) methods in horizontal (top) and vertical (bottom) planes. The beta beating in TBT results is determined by averaging all 12 measurements during an injection cycle, feeding the averaged data into the neural network to predict quadrupole errors.

inherent to individual measurements. As observed, there are some local differences between the two sets of data. These differences are attributed to local under constraints caused by faulty BPMs. For LOCO methods, faulty BPMs in either the horizontal or vertical plane were accommodated by setting their gain and coupling coefficients to zero during the fitting process. However, in the deep learning method, it is necessary to remove both the horizontal and vertical components to ensure that the input matrix remains structurally consistent and that the parameter calculations maintain symmetry.

Figure 14 shows a comparison of the betatron coupling fitted by LOCO and the deep learning network and Fig. 15 illustrates the skew quadrupole result. Since the number of skew-quadrupole correctors used for coupling correction is significantly lower than the number of the BPMs (constraints), skew quadrupoles are not under constrained. It is worth mentioning that, although increasing the number of skew-quadrupole correctors may improve the fitting and correction effect, it can lead to a multisolution problem between the skew quadrupoles and betatron coupling, resulting in the network convergence failure.

Finally, the deep neural network approach for online feedback correction was tested in the SSRF storage ring (July 3rd, 2024). The BPM TBT sampling is triggered by the timing signal of the injection kicker, which operates at the frequency of 2 Hz. Typically, each injection phase lasts 6 s, during which 12 sets of TBT data were sampled. Each entire measurement and fitting process was completed in half a second. The average of the multimeasurements was then used for the correction. Due to the insufficient BPM

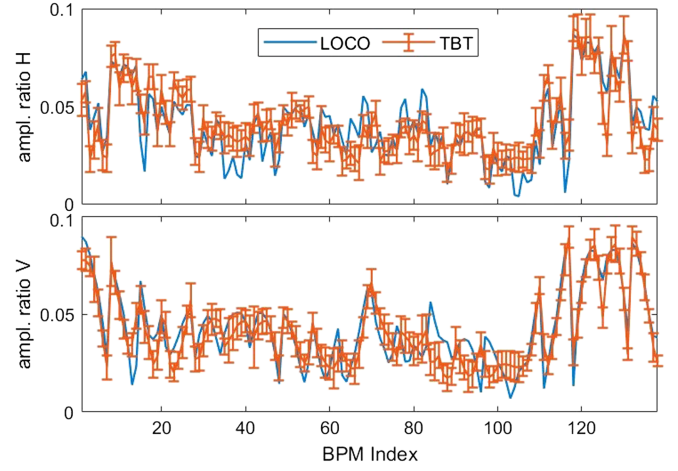


FIG. 14. Comparison of amplitude ratios of the projected betatron motion onto the horizontal (top) and vertical (bottom) plane fitted by LOCO and TBT (deep learning network) methods. The solid lines represent the mean values, and the error bars represent the standard deviations calculated from 12 continuous TBT measurements.

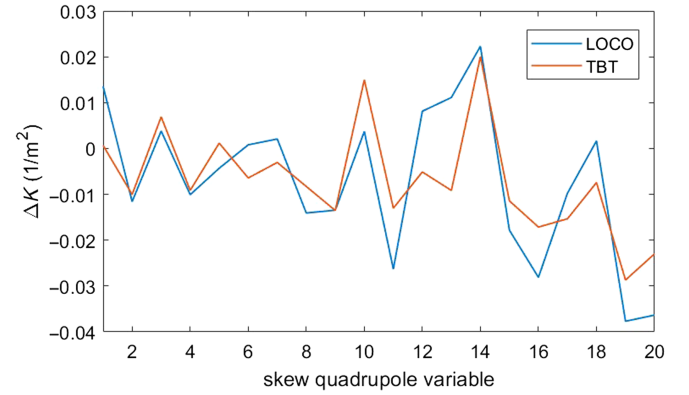


FIG. 15. Comparison of fitted 20 skew-quadrupole errors by LOCO and TBT (deep learning network) methods.

number (constraints), there is a local quadrupole overfitting phenomenon, whereby a same optics caused by multiple equivalent quadrupole combinations. This means that if the optics correction is performed at a very high frequency, the quadrupole currents could rapidly deviate from their designed values. Although this deviation does not impact the effect of optics correction, it may cause the quadrupole current to surpass their linear response range or surpass the tolerance limits of their power supply over long periods of operation. Another disadvantage of this is that the magnet with the power supply drifted cannot be located by this method. Consequently, the frequency of optics correction was limited and only performed when the difference between the measured tune and the design tune exceeded 0.005. The complete measurement and correction workflow is illustrated in Fig. 1.

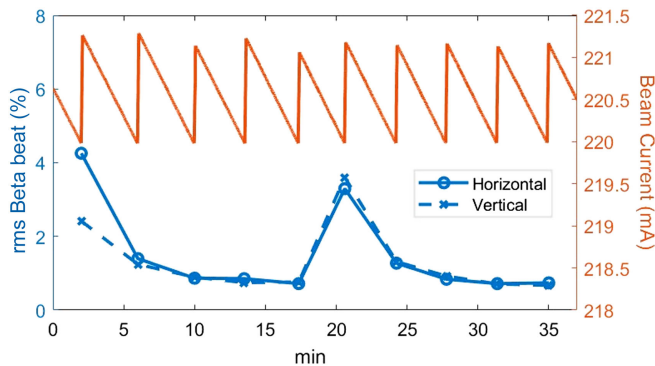


FIG. 16. Continuous optics correction during each injection period. The optical correction system performs periodic measurements aligned with injection cycles and activates optical correction when the tune shifts exceed 0.005. Prior to the 20th min, the gap of an insertion device was manually adjusted and subsequently achieved convergence.

In the experiment, the efficiency of the deep learning method was validated by randomly adjusting several ID gaps while keeping the optics feed-forward system “off,” ensuring its ability to continuously correct the optics in daily operation. Figure 16 presents the results of a continuous optics test over 35 min. The interval between each round of optics measurement (injection) is approximately 4 min. Initially, the beta beating is 4.1% in the horizontal plane and 2.2% in the vertical plane, which reduced to approximately 1.5% after the first correction and reduced to approximately 0.7% after the second correction. After two corrections, the tune limit was achieved. At the 20th min, we manually adjusted several ID gaps again. At this point, the beta beating in both the horizontal and vertical planes increased to about 3.5%, and after two iterations, it converged back to about 0.7%. Figure 17 shows the coupling correction results after the first two iterations. The rms betatron coupling was reduced from 0.039 to 0.018 in the horizontal plane and from 0.105 to 0.037 in the vertical plane. The corresponding emittance coupling (from the fitted lattice by deep learning network) reduced from 3.5% to 0.7% after the first two iterations.

Finally, we discuss the generalizability of the neural network and potential challenges. During the training process (Sec. III B), only noise data sampled at the SSRF’s routine operating beam current of 220 mA were used for sample mixing. Since differences in BPM noise across varying beam currents are negligible, this approach does not compromise neural network performance. Furthermore, as the SSRF employs a passive third-harmonic cavity whose power depends on the beam current, decoherence effects vary significantly with current. Despite these variations, our model exhibits sufficient generalizability because the mixed training samples incorporate diverse bunch lengthening factors and chromaticity-induced effects on kicked beam motion. Consequently,

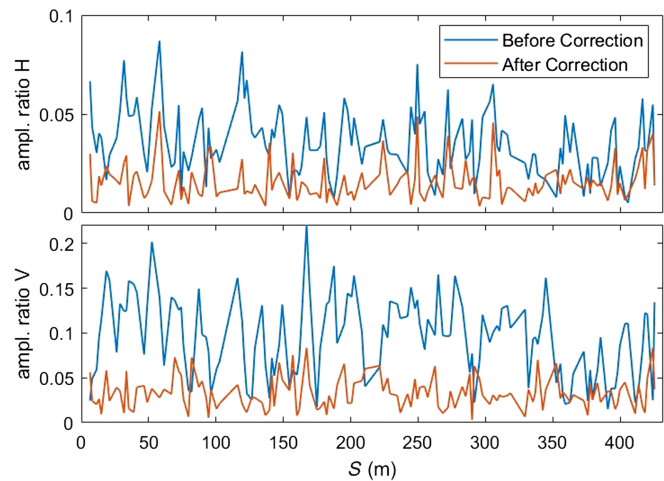


FIG. 17. Comparison of the amplitude ratios of the projected betatron motion onto the horizontal (top) and vertical (bottom) planes before correction and after the first two iterations.

the model remains robust against variations in machine operating conditions.

In this study, we have minimized the beam amplitude to avoid injection disturbances affecting user experiments. Due to limitations in noise and TBT measurement accuracy (approximately $1\ \mu\text{m}$), the precision of betatron amplitude measurements remains low, whereas phase advance measurements achieve significantly higher accuracy. Consequently, only phase advance errors were used for quadrupole error prediction. Adopting larger beam amplitudes during measurements could improve amplitude resolution. Incorporating such measurements into model fitting would not only double the number of constraints but also mitigate local under constraint issues caused by faulty BPMs.

In daily operation, the TBT measurement is highly sensitive to injection conditions. In SSRF storage ring, four injection kickers and the matching of booster to storage ring line critically shape this process. Off axis injection caused by suboptimal parameters may generate anomalous TBT data patterns that neural networks fail to recognize, causing the persistent challenge of mitigating injection system drift over prolonged operation periods. Additionally, the supervised learning framework employed in this study constrains the trained neural network from dynamically updating based on the storage ring’s real-time state. Reinforcement learning techniques typically exhibit superior adaptability when applied to online correction tasks. In future studies, replacing simulation-based supervised learning with reinforcement learning techniques appears highly promising.

V. COUCLUSIONS

In this study, a deep learning network was applied for online optics and coupling feedback. During data

preprocessing, BPM gains and coupling coefficients were corrected and the sampled TBT data were transformed by FFT into a feature map of size $2 \times 266 \times 255$. A total of 250,000 samples were generated and the FFT spectra measured at SSRF were incorporated into the training datasets. After 3 h of training, the MSE loss converged to 0.022. The CFNN performed well in predicting the K value error of 206 quadrupoles and 20 skew quadrupoles. This method was successfully tested in the SSRF storage ring, the rms beta beating maintained less than 0.7% and the emittance coupling after correction is about five times lower than before. Through this deep learning network, stable optics measurement and correction under high beam current and low chromaticities were realized during operation. In the future, integrating this method into real-time control systems or extending it to incorporate additional beam parameters could significantly enhance stable operation by enabling online beam diagnostics and correction. However, this approach still has inherent limitations. For instance, temporal drift in the injection system could degrade neural network performance. Furthermore, since the model was trained using supervised learning, its parameters cannot be dynamically updated based on the latest machine states. This limitation increases vulnerability to newly emerged faulty BPMs. Future research should prioritize replacing supervised learning-based surrogate models with reinforcement learning techniques, which would improve adaptability and address these challenges effectively.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

- [1] X. Liu, S. Tian, X. Wu, M. Wang, Z. Zhao, and B. Feng, Feedforward compensation of the insertion devices effects in the SSRF storage ring, *Nucl. Sci. Tech.* **33**, 70 (2022).
- [2] C. Du, J. Wang, D. Ji, and S. Tian, Studies of round beam at HEPS storage ring by driving linear difference coupling resonance, *Nucl. Instrum. Methods Phys. Res., Sect. A* **976**, 164264 (2020).
- [3] Y. Gong, S. Tian, X. Liu, S. X. L. Tan, and L. Mao, Highly coupled off-resonance lattice design in diffraction-limited light sources, *Nucl. Sci. Tech.* **35**, 163 (2024).
- [4] X. Wu, S. Tian, Q. Zhang, and W. Zhang, Operation stability improvement for synchrotron light sources by tune feedback system, *High Power Laser Part. Beams* **32**, 045107 (2020).
- [5] J. Safranek, Experimental determination of storage ring optics using orbit response measurements, *Nucl. Instrum. Methods Phys. Res., Sect. A* **388**, 27 (1997).
- [6] X. Huang, Linear optics and coupling correction with closed orbit modulation, *Phys. Rev. Accel. Beams* **24**, 072805 (2021).
- [7] X. Huang and X. Yang, Correction of storage ring optics with an improved closed-orbit modulation method, *Phys. Rev. Accel. Beams* **26**, 052802 (2023).
- [8] C.-X. Wang, V. Sajaev, and C.-Y. Yao, Phase advance and β function measurements using model-independent analysis, *Phys. Rev. ST Accel. Beams* **6**, 104001 (2003).
- [9] X. Huang, S. Y. Lee, E. Prebys, and R. Tomlin, Application of independent component analysis to Fermilab booster, *Phys. Rev. ST Accel. Beams* **8**, 064001 (2005).
- [10] X. Yang and X. Huang, A method for simultaneous linear optics and coupling correction for storage rings with turn-by-turn beam position monitor data, *Nucl. Instrum. Methods Phys. Res., Sect. A* **828**, 97 (2016).
- [11] A. Langner, G. Benedetti, M. Carlà, U. Iriso, Z. Martí, J. C. de Portugal, and R. Tomás, Utilizing the N beam position monitor method for turn-by-turn optics measurements, *Phys. Rev. Accel. Beams* **19**, 092803 (2016).
- [12] A. Wegscheider, A. Langner, R. Tomás, and A. Franchi, Analytical N beam position monitor method, *Phys. Rev. Accel. Beams* **20**, 111002 (2017).
- [13] F. V. der Veken, M. Giovannozzi, E. Maclean, C. Montanari, and G. Valentino, Using machine learning to improve dynamic aperture estimates, in *Proceedings of the IPAC-21*, International Particle Accelerator Conference No. 12 (JACoW, Geneva, Switzerland, 2021), pp. 134–137, [10.18429/JACoW-IPAC2021-MOPAB028](https://doi.org/10.18429/JACoW-IPAC2021-MOPAB028).
- [14] A. Edelen, N. Neveu, M. Frey, Y. Huber, C. Mayes, and A. Adelman, Machine learning for orders of magnitude speedup in multiobjective optimization of particle accelerator systems, *Phys. Rev. Accel. Beams* **23**, 044601 (2020).
- [15] A. Ivanov and I. Agapov, Physics-based deep neural networks for beam dynamics in charged particle accelerators, *Phys. Rev. Accel. Beams* **23**, 074601 (2020).
- [16] S. C. Leemann, S. Liu, A. Hexemer, M. A. Marcus, C. N. Melton, H. Nishimura, and C. Sun, Demonstration of machine learning-based model-independent stabilization of source properties in synchrotron light sources, *Phys. Rev. Lett.* **123**, 194801 (2019).
- [17] V. Kain, S. Hirlander, B. Goddard, F. M. Velotti, G. Z. Della Porta, N. Bruchon, and G. Valentino, Sample-efficient reinforcement learning for CERN accelerator control, *Phys. Rev. Accel. Beams* **23**, 124801 (2020).
- [18] R. Roussel, A. Edelen, C. Mayes, D. Ratner, J. P. Gonzalez-Aguilera, S. Kim, E. Wisniewski, and J. Power, Phase space reconstruction from accelerator beam measurements using neural networks and differentiable simulations, *Phys. Rev. Lett.* **130**, 145001 (2023).
- [19] E. Fol, R. Tomás, J. Coello de Portugal, and G. Franchetti, Detection of faulty beam position monitors using unsupervised learning, *Phys. Rev. Accel. Beams* **23**, 102805 (2020).
- [20] E. Fol, F. Carlier, J. C. de Portugal, A. Garcia-Tabares, and R. Tomás, Machine learning methods for optics measurements and corrections at LHC, in *Proceedings of the 9th International Particle Accelerator Conference (IPAC-18)*, Vancouver, BC, Canada, 2018, International Particle Accelerator Conference No. 9 (JACoW, Geneva,

- Switzerland, 2018), pp. 1967–1970, [10.18429/JACoW-IPAC2018-WEPAF062](#).
- [21] X. Xu, Y. Zhou, and Y. Leng, Machine learning based image processing technology application in bunch longitudinal phase information extraction, *Phys. Rev. Accel. Beams* **23**, 032805 (2020).
- [22] I. C. Hsu, The decoherence and recoherence of the betatron oscillation signal and an application, *Part. Accel.* **34**, 43 (1990), <https://cds.cern.ch/record/1108236/files/p43.pdf?version=1>.
- [23] I. M. Borchardt, E. Karantzoulis, H. Mais, and G. Ripken, Calculation of beam envelopes in storage rings and transport systems in the presence of transverse space charge effects and coupling, *Z. Phys. C* **39**, 339 (1988).
- [24] M. Vanwelde, C. Hernalsteens, S. Bogacz, S. Machida, and N. Pauly, Review of coupled betatron motion parametrizations and applications to strongly coupled lattices (2022), <https://api.semanticscholar.org/CorpusID:253080515>.
- [25] M. Borland, ELEGANT: A flexible SDDS-compliant code for accelerator simulation, Argonne Nat. Lab., Argonne, IL, Report No. LS-287, 2000, <https://cds.cern.ch/record/511027>.
- [26] X. Wu, S. Tian, X. Liu, W. Zhang, and Z. Zhao, Design and commissioning of the new SSRF storage ring lattice with asymmetric optics, *Nucl. Instrum. Methods Phys. Res., Sect. A* **1025**, 166098 (2022).
- [27] L. Zhan, L. Zhao, J. Liu, S. Liu, and Q. An, Design and testing of a bunch-by-bunch beam position transverse feedback processor, *Nucl. Instrum. Methods Phys. Res., Sect. A* **902**, 62 (2018).