REAL-TIME TOMOGRAPHY OF SYNCHROTRON LONGITUDINAL PHASE SPACE BASED ON SPATIO-TEMPORAL NEURAL NETWORKS

Yixuan Luo, Peizhi Fang, Yihao Huang, Hongjuan Yao, Shuxin Zheng*
Department of Engineering Physics, Tsinghua University, Beijing, China

Abstract

Accurate acquisition of the longitudinal phase space distribution is crucial for synchrotron optimization, but traditional tomography is time-consuming and difficult to meet the needs of real-time diagnostics. This paper proposes a novel spatio-temporal network model integrating Convolutional Neural Networks (CNN) and Transformers to achieve end-to-end continuous reconstruction from one-dimensional beam projections to two-dimensional phase space dynamic evolution. Based on a high-fidelity simulation dataset incorporating nonlinear effects such as space charge, test results show that this model can accurately restore complex phase space topological structures. Meanwhile, the model achieves fast reconstruction at the sub-millisecond level. This non-intrusive diagnostic tool provides a new paradigm for high-intensity accelerators to achieve high-fidelity online real-time beam measurement and automated feedback.

INTRODUCTION

The longitudinal phase space of a synchrotron is a core dynamic parameter characterizing the longitudinal motion state of the beam. Accurate diagnosis of it is important for improving accelerator performance and ensuring stable beam operation. Traditional tomographic methods couple particle tracking with iterative corrections, which heavily rely on complex physical calculations. [1–3] This heavy computational overhead leads to long single-reconstruction times (typically several minutes), making it difficult to meet the demands of real-time diagnostics. When handling nonlinear dynamic effects in high-intensity scenarios, it will also face a dual challenge of efficiency and accuracy.

In recent years, end-to-end tomographic techniques based on deep learning have been introduced. Convolutional Neural Networks (CNN) successfully reduced the single-reconstruction time to second level by establishing a direct mapping from 1D projections to 2D phase spaces. [4–7] However, early network designs were mostly limited to the reconstruction of static single-frame images. The newly developed Transformer architecture, featuring a multi-head attention mechanism, is highly powerful in processing complex temporal dependencies and dynamic modeling. This paper proposes a novel spatio-temporal neural network integrating the CNN and Transformer to continuously reconstruct the dynamic evolution process of the phase space. The training dataset is optimized using the BLonD code, introducing collective effects to make the data more closely resemble real beam conditions and enhance the network's generalization capability.

* zhengsx@mail.tsinghua.edu.cn

MODELING AND DATASET GENERATION

To enable the model to learn the real physical evolution laws and possess the capability to handle high-intensity beams, this study constructs a high-fidelity spatio-temporal evolution dataset using Beam Longitudinal Dynamics (`BLonD`), developed by the European Organization for Nuclear Research (CERN). This simulation is based on the machine parameters of the Xi'an 200MeV Proton Application Facility (XiPAF) project. This study focuses on introducing the space charge effect into the simulation, which is slowly applied via an adiabatic process to ensure the stability of the simulated beam.

To enhance the generalization capability of the spatio-temporal neural network, the simulation conducts extensive random sampling on multiple physical parameters. For example, parameters such as beam intensity, bunch length, RF voltage amplitude stabilization error, and the initial phase space distribution profile are randomly varied within specific ranges. More importantly, mismatch factors are applied to excite longitudinal dipole and quadrupole oscillations, thereby enriching the network's learning features during dynamic evolution. During tracking, 1280 synchronous turns are discretely sampled to obtain the current projection data, while the corresponding 2D phase space distribution is mapped into a 128×128 pixel image.

SPATIO-TEMPORAL NETWORK

To achieve reconstruction of the multi-turn continuously evolving longitudinal phase space, this study designs a hybrid model combining CNN and Transformer. This network possesses the advantages of both CNN in high-precision spatial image processing and Transformer in highly continuous global temporal perception. The input to the network is 1D projection sequences, and output is the corresponding 2D phase space dynamic evolution video. As shown in Fig. 1, the network architecture is divided into three core modules:

1D CNN Tokenizer: To extract spatial features and compress dimensions, the front end of network employs a tokenizer composed of multi-layer 1D convolutions paired with BatchNorm and GELU activation functions. This module maps the single-frame 1D projection data into feature vectors (Tokens) with a dimension of 256, retaining the local features of the beam line density.

Transformer Encoder: The extracted feature sequence is first injected with sine- and cosine-driven absolute positional encoding, and subsequently fed into the Transformer encoder. Transformer encoder contains 4 encoding layers, each utilizing a multi-head attention mechanism with

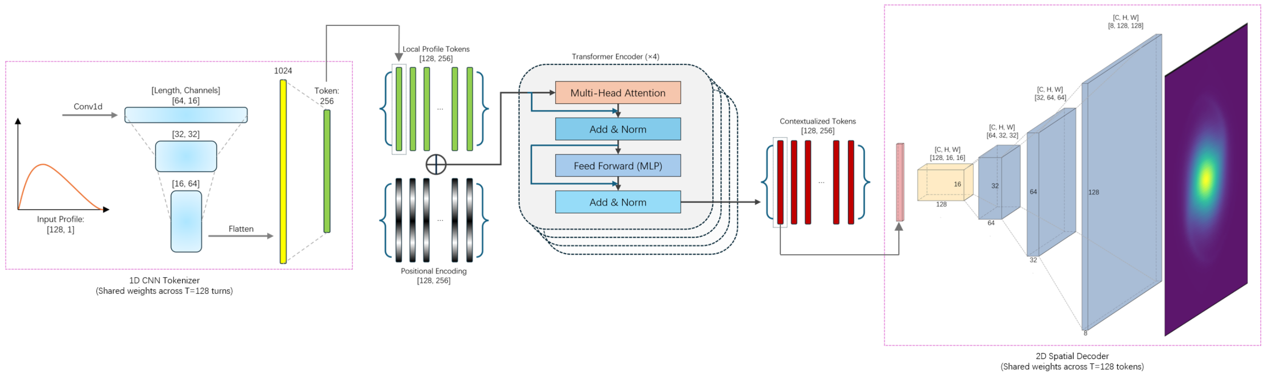

Figure 1: Schematic diagram of the proposed spatio-temporal neural network architecture. The model consists of a 1D CNN tokenizer for spatial feature extraction, a Transformer encoder for capturing global temporal correlations, and a 2D spatial decoder for end-to-end reconstruction of the longitudinal phase space.

8 heads. This global attention mechanism enables the network to span a long time window and capture the nonlinear dynamic correlations among projection sequences over continuous turns.

Spatial Decoder: The feature vectors carrying global temporal information are independently fed into the decoder to restore the 2D distribution. The decoder maps the latent variables to low-resolution feature maps via fully connected layers, and then employs a progressive upsampling module based on bilinear interpolation to restore the spatial resolution to 128×128 layer by layer. The end of the network utilizes a Softplus activation function to ensure that the reconstructed phase space density is non-negative everywhere from a physical mechanism perspective.

TRAINING STRATEGY

Considering both image-domain comparison metrics and physical observations of phase space evolution, we designed a multi-term composite loss function to strictly anchor the network's output to the true physical dynamic evolution. The total loss function is defined as Eq. (1).

$$\begin{aligned}\mathcal{L}_{total} = 0.5\mathcal{L}_{MSE} + 0.15\mathcal{L}_{SSIM} + 0.2\mathcal{L}_{Laplace} \\ + 0.2\mathcal{L}_{Proj} + 0.05\mathcal{L}_{Temp}\end{aligned} \tag{1}$$

where $\mathcal{L}_{MSE}$ is the mean squared error loss for pixel-wise differences; $\mathcal{L}_{SSIM}$ is the structural similarity; $\mathcal{L}_{Laplace}$ is the image gradient computed via the Laplace operator to control the sharpness of beam halo edges; $\mathcal{L}_{Proj}$ is the residual loss between the reconstructed distribution projection and the input profile to ensure the network satisfies the Radon transform; and $\mathcal{L}_{Temp}$ evaluates the difference between adjacent frames to characterize temporal smoothness, aiming to suppress non-physical abrupt mutations during continuous multi-turn evolution predictions.

We employ the AdamW optimizer with an initial learning rate set to 2×10^{-4} in training configuration. Meanwhile, gradient clipping with a maximum norm of 1.0 is applied to prevent Transformer from experiencing gradient explosion.

RECONSTRUCTION RESULTS

We randomly selected 400 independent samples from the validation dataset for large-sample statistical analysis, evaluating a total of 51,200 frames of evolving phase space images, which comprehensively evaluate the performance of the network in complex dynamic environments.

In terms of visual presentation, Fig. 2 shows the phase space tomographic results over an evolution cycle. Four representative turns are extracted for comparison, with the ground truth on the left and the reconstructed images on the right. As observed, driven by the RF field and space charge forces, the mismatched beam undergoes dipole and quadrupole oscillations, as well as phase space filamentation. The network precisely reconstructs the dynamic topological structure of the phase space, capturing both the high-density core trajectory and the low-density edge halo simultaneously, free from the over-smoothing or artifacts common in traditional algorithms.

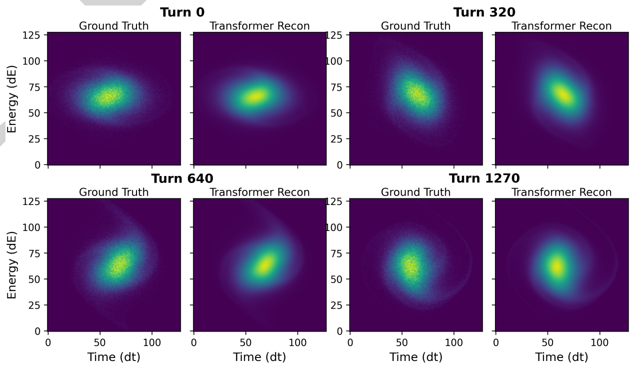

Figure 2: Snapshots of the longitudinal phase space evolution. For each specified turn, the ground truth (left) is compared with the Transformer reconstruction (right).

In terms of quantitative evaluation, similar to the loss functions, we selected two image-domain metrics: Mean Squared Error (MSE) and Structural Similarity Index (SSIM). Alongside the Kullback-Leibler (KL) divergence characterizes the agreement of statistical probability distributions. Furthermore, to verify the consistency between the 1D projection of the reconstructed distribution and the input profile, we utilized the Normalized Total Variation Distance (NTVD, $NTVD = \frac{1}{2}\sum |p_i - q_i|$, , where p and q are the 1D profiles normalized to a unit area) to calculate the absolute reconstruction error of the longitudinal line density projection. The evaluation results and ideal values for all metrics are summarized in Table 1.

Table 1: Evaluation Metrics of the Reconstruction Results

Metric	Model Output	Target
MSE	0.001053	0
SSIM	0.8627	1
KL Divergence	0.01108	0
NTVD	0.0089	0

The evaluation results presented in the table demonstrate the exceptional precision of the proposed model, with the absolute error of the longitudinal line density projection strictly controlled within 1%. This series of accuracy metrics outperforms traditional algebraic reconstruction techniques by more than an order of magnitude. These results fully prove the advantage of the global temporal attention mechanism in solving the continuous modeling problem of high-intensity nonlinear dynamics.

To further validate the model's applicability in real-world scenarios, longitudinal phase space reconstruction was performed using actual Fast Current Transformer (FCT) beam measurement data from the XiPAF facility. The tomographic results are illustrated in Fig. 3. Given that the ground-truth 2D phase space distribution is physically inaccessible during non-destructive operations, the reconstruction fidelity is quantitatively assessed by comparing the 1D projected current with the measured profile. In this real-beam experiment, the average NTVD between the two 1D waveforms is successfully constrained to less than 2%.

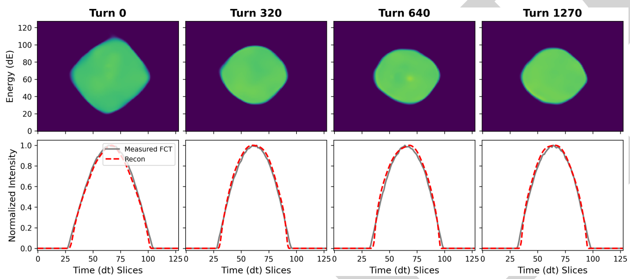

Figure 3: Longitudinal phase space tomographic reconstruction using real FCT measurement data from the XiPAF facility. The top row displays the reconstructed 2D phase space evolution, while the bottom row compares the 1D projected line density (dashed red) with the actual FCT measurements (solid grey).

This study evaluated single-frame inference times on different computing platforms. On professional computing nodes equipped with a high-end graphics processing unit, the average time for a single phase space reconstruction is only 0.109 ms. This performance outperforms existing phase space tomography approaches, satisfying the requirements for real-time reconstruction at the conventional operating frequencies of proton synchrotrons. Furthermore, utilizing 8 threads for pure CPU inference on a standard desktop processor (i.e. Intel Core i7-8700K), the single-frame time remains 4.557 ms, demonstrating the model's friendliness to deployment hardware.

DISCUSSION

The proposed spatio-temporal neural network successfully resolves the traditional compromise between reconstruction accuracy and computational speed in phase space tomography. Comprehensive evaluations demonstrate the model's exceptional fidelity, achieving a longitudinal line density projection error of less than 1% in simulations and under 2% when processing real FCT beam measurements from the XiPAF facility. Furthermore, the framework exhibits remarkable computational efficiency, delivering single-frame inference times of 0.109 ms on a professional GPU node and 4.557 ms on a standard desktop CPU.

The sub-millisecond processing speed successfully cross the engineering threshold for online real-time diagnostics at proton synchrotron operating frequencies. By maintaining robust performance even under physical operating conditions containing complex nonlinear effects such as space charge forces , this high-fidelity, ultra-low-latency diagnostic tool establishes a reliable new continuous imaging paradigm for high-intensity accelerators. Ultimately, it lays a solid technical foundation for the future deployment of fast, automated beam real-time feedback control systems based on 2D longitudinal phase space dynamics.

FUTURE WORK

Future research will mainly focus on enhancing physical interpretability and experimental verification with real beams. Firstly, to further improve network accuracy and interpretability, we plan to embed beam dynamics constraints into the model, such as introducing the Vlasov equation into the loss function as a PDE regularization term. Secondly, beam diagnostic experiments will be conducted on a synchrotron using high-precision destructive measurements, such as transverse deflecting cavities, to acquire the true 2D longitudinal phase space distribution. These experimental measurements will test the reconstruction fidelity under real physical noise and system errors, and subsequently serve as empirical data for the network's hybrid training and fine-tuning.

REFERENCES

[1] M. Lindroos, S. Hancock, and S. Koscielniak, "Longitudinal Phase Space Tomography with Space Charge", in *Proc. EPAC'00*, Vienna, Austria, Jun. 2000, paper WEP4A03, pp. 1726–1728.

[2] H. Y. Liu *et al.*, "Reconstructing phase space of injection beam based on the measurement of accumulated beam: A novel tomography algorithm and its applications", *Phys. Rev. Accel. Beams*, vol. 27, p. 052804, May 2024. doi:10.1103/PhysRevAccelBeams.27.052804

[3] M. W. Wang *et al.*, “Application of iterative reconstruction for beam phase space measurement at Xi’an Proton Application Facility”, *Mod. Appl. Phys.*, vol. 12, no. 3, p. 30404, Sep. 2021. doi:10.12061/j.issn.2095-6223.2021.030404

[4] Y. Luo *et al.* “Longitudinal phase space reconstruction from beam current projections using deep convolutional neural networks”, presented at HB’25, Huizhou, China, Oct. 2025, paper THFA12, unpublished. doi:10.18429/JACoW-HB2025-THFA12

[5] C. Emma *et al.*, “Machine learning-based longitudinal phase space prediction of particle accelerators”, *Phys. Rev. Accel. Beams*, vol. 21, p. 112802, Nov. 2018. doi:10.1103/PhysRevAccelBeams.21.112802

[6] J. Kaiser, A. Eichler, S. Tomin, and Z. Zhu, “Machine learning for combined scalar and spectral longitudinal phase space reconstruction”, in *Proc. IPAC’23*, Venice, Italy, May 2023, pp. 4464–4467. doi:10.18429/JACOW-IPAC2023-THPL019

[7] A. Wolski, D. Botelho, D. Dunning, and A.E. Pollard, “Accelerator beam phase space tomography using machine learning to account for variations in beamline components”, *J. Instrum.*, vol. 19, p. P07013, Jul. 2024. doi:10.1088/1748-0221/19/07/P07013