CLEANING TURN-BY-TURN DATA FROM THE LHC WITH AUTOENCODERS

J. Gray*, F. Carlier, K. Skoufaris, European Organization for Nuclear Research, Geneva, Switzerland

Abstract

Turn-by-turn (TbT) BPM data in the LHC is often affected by noise, limiting the extraction of resonant driving terms (RDTs) and reducing the precision of nonlinear optics studies. We developed a denoising autoencoder trained on simulated tracking data to reconstruct clean transverse oscillations and suppress noise directly in the time domain. The method produces cleaner frequency spectra and significantly improves RDT visibility compared to established methods such as singular value decomposition, even when trained on fewer turns. In its current form, the autoencoder performs well on data that resemble the training set. However, when applied to new conditions such as different noise levels, excitation amplitudes, tunes, or beam configurations, its ability to generalise decreases. These results demonstrate that autoencoders can substantially improve TbT data quality in specific conditions. Establishing broader and more diverse training datasets is a promising next step toward applying this technique to real LHC measurements.

INTRODUCTION

During LHC commissioning, linear and nonlinear corrections are required to maximise beam lifetime and luminosity [1]. Resonances from horizontal and vertical tune combinations, $(j-k)Q_x + (l-m)Q_y$, are characterised by resonant driving terms (RDTs) f_{jklm}, making RDTs useful diagnostics for nonlinear errors [2–4]. They are inferred from AC-dipole-excited turn-by-turn (TbT) beam position monitor (BPM) data, so reconstruction quality depends directly on the TbT signal quality.

The standard cleaning strategy is singular value decomposition (SVD) [5]. SVD is attractive because it is computationally efficient and effective at removing uncorrelated noise, as demonstrated in [6]. However, it can also suppress important spectral content across all BPMs, which limits RDT reconstruction precision [7].

Autoencoders are widely used for dimensionality reduction and denoising [8], similar to SVD. By learning a compact latent representation and reconstructing the input, they can capture underlying structure while filtering noise. For TbT data, this could improve weak spectral-line visibility for RDT analysis.

We first demonstrate how SVD can remove spatial RDT information in simulated LHC TbT data, then present the convolutional autoencoder and compare both methods in terms of spectral visibility, RDT reconstruction, and generalisation.

* joshua.mark.gray@cern.ch

SVD EFFECT ON RDT RECONSTRUCTION

LHC simulation data are generated by tracking with Xsuite [9] for 1000 turns, then artificial Gaussian noise is added. For frequency analysis, the TbT matrix is decomposed with SVD and the first 12 singular values are retained, a common choice for LHC TbT data [10]. An RFFT is then applied to the cleaned data [10]. From the resulting spectral amplitudes and phases, RDTs are reconstructed [3].

To reduce complexity, the simulation used fixed initial coordinates rather than an AC dipole, giving a maximum physical peak-to-peak amplitude of 2 mm.

The chosen scenario is the LHC ring at 6800 GeV, with tunes $Q_x = 0.28$ and $Q_y = 0.31$, and $\beta^* = 0.3$ m. Although typical LHC BPM noise is around 100 μm, we use 500 μm for an oscillation spanning about -1 mm to 1 mm to highlight the negative impacts of SVD cleaning. The analysis is therefore framed by spectral signal-to-noise ratio (SNR), the usual figure of merit for RDT reconstruction quality [11].

The RDTs f_{3000} and f_{0111} were selected because they are the dominant RDTs in this scenario for their respective planes. Figure 1 shows that these frequencies remain clearly visible after SVD cleaning in the spectrum. However, the reconstructed RDTs along the ring in Fig. 2 show that local RDT information is almost entirely lost. Instead, the signal is driven toward the ring-average RDT value. This removes local diagnostic information, although the average RDT value is still recovered. Consequently, analyses that rely only on ring-averaged RDT values are much less affected by SVD cleaning [4, 11].

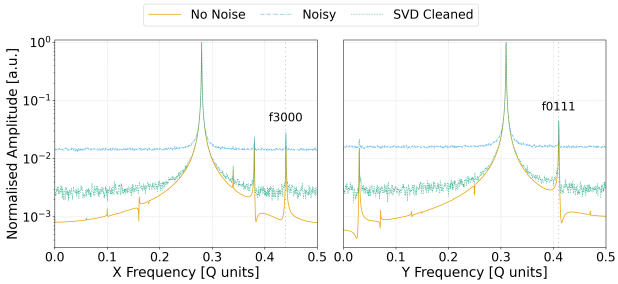

Figure 1: Effect of SVD cleaning on spectral content. Average frequency spectra: raw noisy signal, SVD-cleaned signal, and the no-noise reference.

This loss of local RDT information limits nonlinear-error localisation and motivates denoising methods that preserve local spectral structure.

AUTOENCODER METHOD

Model Architecture and Training

To address this limitation, we use a convolutional autoencoder family implemented in PyTorch and PyTorch Light-

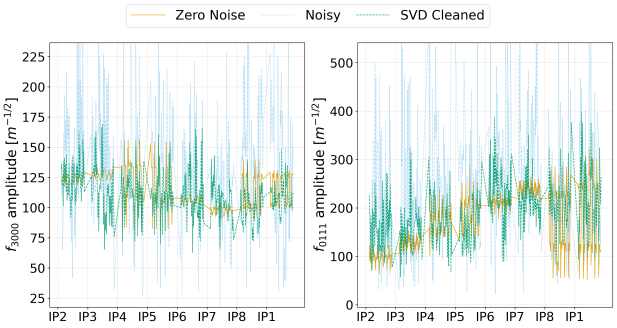

Figure 2: Effect of SVD cleaning on RDT reconstruction.

ning [12, 13]. The current configuration is a fixed-depth U-Net variant [14], chosen for its encoder–decoder structure with skip connections, which captures multi-scale dynamics while preserving local turn-by-turn oscillation information.

Clean targets are generated with Xsuite tracking [9] (without AC dipole) from an LHC beam 1 model at 6.8 TeV. Two tune pairs were chosen, $(Q_x, Q_y) = (0.28, 0.31)$ and $(0.31, 0.32)$, corresponding to typical injection and collision tunes, each tracked for 6600 turns. The x and y planes are normalised by $\sqrt{\beta_{x,y}}$ from the Twiss functions [15], min–max scaled to $[-1, 1]$, and split into 1000-turn windows with seeded random offsets.

Noisy inputs are created on the fly by adding independent Gaussian noise per BPM and turn, with amplitudes from 900 μm to 10 μm equivalent, to help the model distinguish features from noise. A total of 200 samples are split 80/20 into train/validation sets. Horizontal and vertical samples are processed separately by the same U-Net with shared weights. The loss combines time-domain reconstruction and spectral similarity (MSE+SSP), and optimisation uses AdamW [16] with a learning-rate scheduler.

Training used a single NVIDIA T4 GPU for approximately seven days with batch size 20. The model had not fully converged, so the results represent a still-improving model rather than a final optimum. Model size was kept modest to reduce overfitting risk, but larger models may be required once more working points are included.

Loss Definition

Let y denote the clean target and $\hat{y}$ the model output.

The baseline objective is the mean-squared error (MSE),

$$\mathscr{L}_{\text{MSE}} = \frac{1}{N} \sum_i (\hat{y}_i - y_i)^2. \tag{1}$$

To better preserve spectral content, the Surface Similarity Parameter [17] (SSP) is used, a frequency based loss function,

$$\mathscr{L}_{\text{SSP}} = \frac{\|\mathscr{F}(\hat{y}) - \mathscr{F}(y)\|_2}{\|\mathscr{F}(\hat{y})\|_2 + \|\mathscr{F}(y)\|_2 + \epsilon}, \tag{2}$$

where $\mathscr{F}(\cdot)$ denotes the Fourier transform, ϵ is a small non-zero constant and $\|\mathscr{F}\|_2$ describes the Euclidean norm on $\mathscr{F}$.

The mixed objective is therefore used,

$$\mathscr{L}_{\text{total}} = \alpha\, \mathscr{L}_{\text{MSE}} + (1 - \alpha)\, \mathscr{L}_{\text{SSP}}. \tag{3}$$

MSE penalises phase shifts and encourages denoising across all turns, rather than neglecting turns that are less important for frequency reconstruction. SSP better preserves spectral phase and amplitude at increased computational cost.

AUTOENCODER PERFORMANCE

We compare SVD-cleaned data with autoencoder-denoised data using spectra and reconstructed RDTs. The autoencoder is trained only on 1000-turn samples but tested on 6600-turn data, where performance remains comparable.

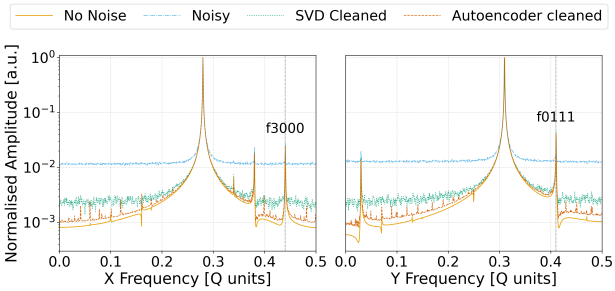

Figure 3: Frequency-domain comparison between SVD cleaning and autoencoder denoising for a model with tunes (0.28, 0.31).

In Fig. 3, the autoencoder removes more broadband noise than SVD around the RDT lines, but also introduces artefacts from attempting to predict resonances with limited capacity.

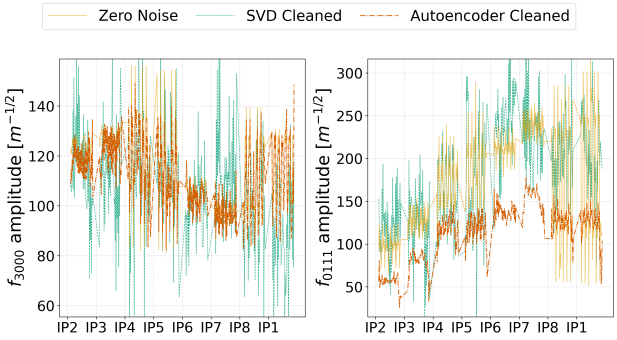

Figure 4: RDT reconstruction comparison after SVD cleaning and autoencoder denoising.

The autoencoder improves the reconstructed f_{3000} spatial structure, while the visual improvement for f_{0111} is less apparent. To quantify this, the average relative complex error is computed across all BPMs as

$$\epsilon_f = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{f_i^{\text{reconstructed}} - f_i^{\text{no noise}}}{|f_i^{\text{no noise}}|} \right|. \tag{4}$$

The resulting values are shown in Table 1, where smaller values indicate better RDT reconstruction.

Both methods reduce ϵ_f relative to the noisy case. The autoencoder gives the lowest error for f_{3000}, while SVD is

Table 1: Average Relative Complex Error ϵ_f for RDT Reconstruction.

RDT	Noisy	SVD	Autoencoder
f_{3000}	46.3	29.1	11.7
f_{0111}	165	75.5	77.3

slightly lower for f_{0111}. However, the f_{0111} phase follows the no-noise reference more closely after autoencoder denoising, even though the amplitude in Fig. 4 appears less accurate.

GENERALISATION LIMITS

The current autoencoder was trained on four spectra and can outperform SVD in these in-distribution cases. With around 190 thousand parameters and over 14 million target data points, the network appears to learn structural features rather than memorise data. However, incomplete convergence and phantom FFT peaks indicate limited capacity; larger models may be needed for robust operation. Owing to its convolutional structure, the model can be trained on shorter turn windows and applied to full 6600-turn measurements.

When very low-noise data were included in training, the autoencoder struggled to reproduce RDTs at any noise level. This suggests that it learns features of both signal and noise rather than directly reproducing spectra, so training on real data or more realistic noise simulations will be important.

No evaluation on real measurement data has been performed yet. Figure 5 shows a tune setup slightly outside the training distribution. The tune is reconstructed accurately, suggesting some adaptability, but some resonance lines are missing and extra peaks are introduced, indicating that additional training and/or capacity is required.

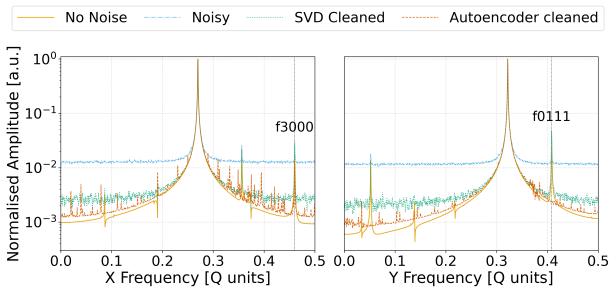

Figure 5: Frequency-domain comparison between SVD cleaning and autoencoder denoising for model with tunes (0.27, 0.322).

This behaviour reflects a known limitation of supervised denoisers trained on finite distributions [18]. SVD remains a strong operational baseline because it requires no training. To make the autoencoder viable, training must cover more conditions (tunes, resonance regimes, AC-dipole excitation), and unsupervised approaches without clean targets should be investigated for real-world performance [18]. Validation on real measurements is then needed to assess generalisation beyond simulation.

CONCLUSION

This work benchmarks autoencoder-based denoising for LHC TbT data against SVD at both signal and physics-analysis levels. The main finding is regime dependence: SVD remains robust and simple, while the autoencoder can improve RDT reconstruction when test conditions are represented in training. Generalisation beyond the training domain remains the main limitation, requiring broader training coverage and validation on real measurements.

ACKNOWLEDGEMENTS

Many thanks to Borja Rodriguez Mateos for detailed discussions on creating autoencoders and helping improve my understanding of the topic.

AUTOENCODER ARCHITECTURE

The network is a compact U-Net-like convolutional autoencoder. It maps a single-channel input of shape $(B, 1, 559, 1000)$ to an output of the same shape, using four encoder levels, a bottleneck, and four mirrored decoder levels. Skip connections pass same-resolution features from the encoder to the decoder, helping preserve local BPM and turn information while the bottleneck encourages denoising.

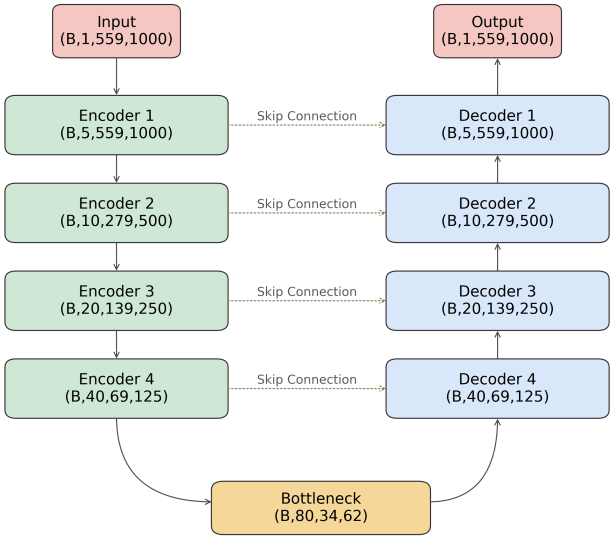

Figure 6: Slimmed layer-wise representation of the fixed-depth U-Net used for TbT denoising. The numbers underneath the labels describe the dimensions as batch (b), channel, number of BPMs, turns.

REFERENCES

[1] E. H. Maclean *et al.*, 'New approach to LHC optics commissioning for the nonlinear era', *Phys. Rev. Accel. Beams*, vol. 22, no. 6, p. 061004, Jun. 2019. doi:10.1103/PhysRevAccelBeams.22.061004

[2] F. S. Carlier, EH. Maclean, T. Persson and R. Tomás, 'Nonlinear Correction Strategies for the LHC Using Resonance Driving Terms', in *Proc. 9th International Particle Accelerator Conference (IPAC'18), Vancouver, BC, Canada, April 29-May 4, 2018*, Vancouver, BC, Canada, pp. 161–164, Jun. 2018. doi:10.18429/JACoW-IPAC2018-MOPMF032

[3] R. Tomás, M. Bai, R. Calaga, W. Fischer, A. Franchi and G. Rumolo, 'Measurement of global and local resonance terms', *Phys. Rev. ST Accel. Beams*, vol. 8, no. 2, p. 024001, Feb. 2005. doi:10.1103/PhysRevSTAB.8.024001

[4] S. Horney *et al.*, 'Sextupole RDTs in the LHC at injection and in the ramp', in *Proc. IPAC'24*, Nashville, TN, pp. 71–74, May 2024. doi:10.18429/JACoW-IPAC2024-MOPC13

[5] G. H. Golub and C. F. Van Loan, *Matrix computations*. Johns Hopkins University Press, 2013.

[6] R. Calaga and R. Tomás, 'Statistical analysis of RHIC beam position monitors performance', *Phys. Rev. ST Accel. Beams*, vol. 7, no. 4, p. 042801, Apr. 2004. doi:10.1103/PhysRevSTAB.7.042801

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

[8] I. D. Mienye and T. G. Swart, 'Deep autoencoder neural networks: a comprehensive review and new perspectives', *Archives of Computational Methods in Engineering*, vol. 32, no. 7, pp. 3981–4000, Oct. 2025. doi:10.1007/s11831-025-10260-5

[9] G. Iadarola *et al.*, 'Xsuite: An integrated beam physics simulation framework', in *Proc. HB'23*, Geneva, Switzerland, pp. 73–80, Mar. 2024. doi:10.18429/JACoW-HB2023-TUA2I1

[10] L. Malina, 'Harpy: A Fast, Simple and Accurate Harmonic Analysis with Error Propagation', in *Proc. IPAC'22*, Bangkok, Thailand, pp. 2326–2329, Jul. 2022. doi:10.18429/JACoW-IPAC2022-WEPOMS035

[11] E. Maclean *et al.*, '3Qy resonance correction at LHC injection', in *Proc. IPAC'25*, Taipei, Taiwan, Jun. 2025, pp. 1964–1967. doi:10.18429/JACoW-IPAC2025-WEPM008

[12] A. Paszke *et al.*, 'Pytorch: an imperative style, high-performance deep learning library', in *Advances in Neural Information Processing Systems 32*, 2019.

[13] W. Falcon and The PyTorch Lightning team, PyTorch Lightning, Mar. 2019. doi:10.5281/zenodo.3828935

[14] O. Ronneberger, P. Fischer and T. Brox, 'U-net: convolutional networks for biomedical image segmentation', in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015. doi:10.1007/978-3-319-24574-4_28

[15] ED. Courant and HS. Snyder, 'Theory of the alternating-gradient synchrotron', *Annals of Physics*, vol. 281, no. 1, pp. 360–408, 2000. doi:10.1006/aphy.2000.6012

[16] I. Loshchilov and F. Hutter, 'Decoupled weight decay regularization', in *7th International Conference on Learning Representations (ICLR)*, 2019. https://openreview.net/forum?id=Bkg6RiCqY7

[17] M. Wedler, M. Stender, M. Klein, S. Ehlers and N. Hoffmann, 'Surface similarity parameter: a new machine learning loss metric for oscillatory spatio-temporal data', *Neural Networks*, vol. 156, pp. 123–134, 2022. doi:10.1016/j.neunet.2022.09.023

[18] I. Boucherit and H. Kheddar, 'Reinforced residual encoder–decoder network for image denoising via deeper encoding and balanced skip connections', *Big Data and Cognitive Computing*, vol. 9, no. 4, 2025. doi:10.3390/bdcc9040082