ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-MOPOTK038

BPM ANALYSIS WITH VARIATIONAL AUTOENCODERS

C. C. Hall*, J. P. Edelen, J. Einstein-Curtis, M. Kilpatrick, RadiaSoft LLC, Boulder, CO, USA

Abstract

In particles accelerators, beam position monitors (BPMs) are used extensively as a non-intercepting diagnostic. Significant information about beam dynamics can often be extracted from BPM measurements, and used to tune the accelerator. Common measurement tools such as measurements of kicked beams may become more difficult when very strong nonlinearities are present or generally when data is very noisy.

In this work we examine the use of variational autoencoders (VAEs) as a technique for extracting measurements of the beam from simulated turn-by-turn BPM data. In particular we show that VAEs may have the possibility to outperform other dimensionality reduction techniques that have historically been used to analyze such data. When the data collection period is limited, or the data is noisy, VAEs may offer significant advantages.

INTRODUCTION

The beam position monitor (BPM) is a ubiquitous diagnostic tool in particle accelerators for monitoring the transverse position of a passing charged particle beam. Source separation techniques such as independent component analysis (ICA) are commonly used [1] on data from sets of BPMs to extract measurements of an accelerator's operating point, beyond what might be available from raw BPM signals.

Autoencoder (AE) neural networks seek to create a reduced-dimensionality representation of their input by training to reproduce that input after data is compressed to a latent space at lower dimension than the input. The variational autoencoder (VAE) can be seen as an adaptation of the vanilla AE structure that instead seeks to represent a parameterization of the input data as distributions over the latent space. This representation by distribution allows for smooth interpolation over the latent space, and makes the VAE a useful tools for performing inference to make measurements of the accelerator.

In this work we demonstrate the ability of VAEs to measure tune from an analytic, continuous focusing model that is representative of the idealized transverse dynamics of a charge particle beam in an accelerator. In particular we explore the use of variational autoencoders for this task due to their ability to smoothly interpolate between latent space values. The capabilities of the VAE are compared against a vanilla autoencoder, and against a typical ICA analysis. Even when given data with a very short measurement period, or noisy measurements VAE models are shown to give good tune measurements.

* chall@radiasoft.net

METHODOLOGY

Analytic Accelerator Model

For exploration of this technique we use data generated from a simplified model of a circular (periodic) accelerator with analytic solutions [1]. Rather than composing the accelerator of discrete focusing magnets we consider a uniform focusing channel with coupled optics. This reduces the problem to that of a coupled harmonic oscillator. We consider only motion in the transverse plane so that we have coupled differential equations:

$$\begin{aligned} \frac{d^2x}{d\theta} + \nu_x x + Cy = 0 \\ \frac{d^2y}{d\theta} + \nu_y y + Cx = 0. \end{aligned} \tag{1}$$

Where C is the coupling strength and $\theta = 2\pi f t$ is the fractional revolution period. The solutions to the coupled equations of motion will then be:

$$\begin{aligned} x(\theta) = A_x cos(\nu_+\theta) + B_x cos(\nu_-\theta) \\ y(\theta) = A_y cos(\nu_+\theta) + B_y cos(\nu_-\theta), \end{aligned} \tag{2}$$

with the coupled oscillation frequency given by:

$$\nu_\pm^2 = \frac{1}{2}\left(\nu_x^2 + \nu_y^2 \pm \sqrt{\left(\nu_x^2 + \nu_y^2\right)^2 + 4C^2}\right). \tag{3}$$

The principal goal of the analysis tools developed in this paper will be to extract the correct independent frequencies from noisy, periodic measurements of x and y. These are referred to as the tunes, ν_x and ν_y in Equation 1. The amplitude coefficients may be uniquely determined from the initial x and y positions, but are not included here as we do not use them in analyzing performance of the methods presented.

To create test and training data the continuous x and y position data generated from the model is sampled as if from M BPMs placed around a ring so that BPM m will have a phase offset of $\varphi = 2\pi m/M$. Noise in the measurements at each turn N is sampled from a normal distribution $\mathcal{N}(0, \sigma_M)$. Where the variance σ_M, representing the noisiness of each BPM, has been set by sampling from a normal distribution $\mathcal{N}(0, \sigma_{noise})$. An example of data produced from the model is shown in Fig. 1.

Variational Autoencoders Model

For the VAE model, since we are analyzing time series data, we use a recurrent network architecture for the encoder and decoder. The implementation of the VAE architecture is based on [2] and uses Long Short-Term Memory (LSTM) units for both the encoder and decoder. A schematic of the VAE is shown in Fig. 2 together with the regressor for

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-MOPOTK038

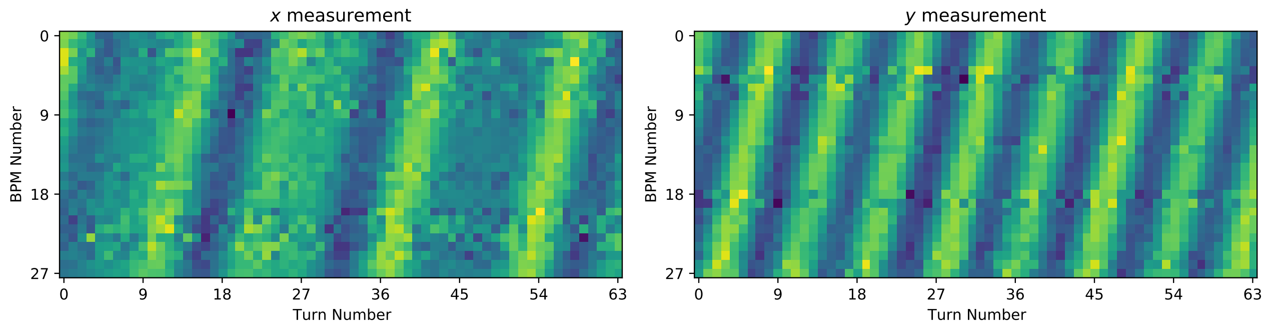

Figure 1: Example of data produced from discrete sampling of Equation 2 to simulate BPM placement around a ring. Data was created with noise of $\sigma_{noise} = 0.25$ applied.

extracting measurement of the tune. Both encoder and decoder contain 1 hidden layer of size 90. Linear layers are used to transform from the LSTM output to the distribution parameterization for sampling the latent space. The loss function is composed of two terms: the Kullback–Leibler divergence [3] — which acts as a regularization term — and the reconstruction loss. For the reconstruction loss mean squared error between the encoder input and decoder output is used.

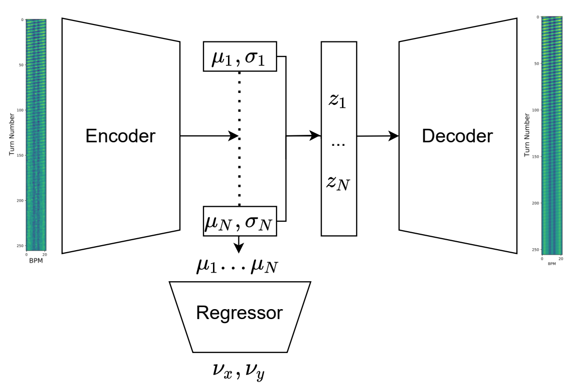

Figure 2: Structure of the VAE together with regressor which determines tunes from the latent space distribution parameterization.

The VAE was trained on generated data with a sequence length of 28 turns and with 56 features (28 BPM measurements of x and 28 measurements of y). Initially training data amplitude was normalized, however, this seemed to result in over-fitting and the model would show poor performance when applied to datasets with noise levels or period that differed from the training set. Training with a dataset of random initial amplitudes significantly improved overall performance.

The latent space of the trained VAE is shown in Fig. 3, reduced to a 2 dimensional representation using t-SNE, generated from a portion of the validation data. The coloring of points in the latent space by tune suggests very good correlation between tune of the data and the latent space vector. It is expected that the VAE will show significantly better performance in use for predicting tune, over a standard autoencoder, given its ability to smoothly interpolate across the latent space.

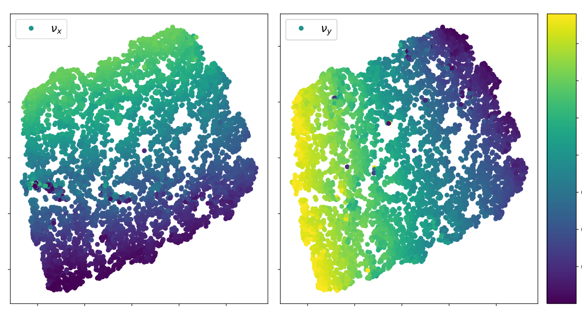

Figure 3: Latent space of the VAE visualized in 2 dimensions using t-SNE. Left: coloring of points corresponds to horizontal tune ν_x. Right: coloring of points corresponds to vertical tune ν_x.

Training for both types of autoencoders was performed on NVIDIA Tesla V100 GPUs. A dataset of 50 000 samples was generated from the analytic accelerator model with Gaussian noise of $\sigma_{noise} = 0.25$. From this dataset 90 % was used for training and 10 % for validation.

RESULTS

To assess performance of the VAE model, comparisons of tune predication — that is ν_x and nu_y in Eq. (2) — are made against ICA and a standard autoencoder (AE). The ICA identification uses the implementation in scikit-learn [4] `sklearn.decomposition.FastICA` with 4 components. The resulting source signals are analyzed using NAFF [5] as implemented by PyNAFF [1] to obtain measurements of ν_x and ν_y.

For both the AE and VAE models a multilayer perceptron regressor (MLP) model with 3 hidden layers of

[1] `https://github.com/nkarast/PyNAFF`

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-MOPOTK038

Table 1: R^2 Values for Tune Prediction

Model	28 Turns	64 Turns	128 Turns	256 Turns
	$\sigma_{noise} = 0.0$			
ICA	0.8980	**0.9951**	**0.9986**	**1.0000**
AE	0.9627	0.7222	-1.2863	-6.6645
VAE	**0.9807**	0.9869	0.9843	0.9733
	$\sigma_{noise} = 0.01$			
ICA	0.6990	0.8157	0.8996	0.9117
AE	0.9563	0.7288	-1.2566	-6.8219
VAE	**0.9814**	**0.9869**	**0.9857**	**0.9747**
	$\sigma_{noise} = 0.25$			
ICA	-1.6670	-0.8499	-0.7010	-0.5698
AE	0.9561	0.7165	-1.1897	-6.4335
VAE	**0.9807**	**0.9867**	**0.9843**	**0.9734**

100, 80 and 40 units and rectified linear unit activation functions is used to predict the tunes from the latent space. The regressor was implemented from scikit-learn `sklearn.neural_network.MLPRegressor`. Three test datasets were generated with varying levels of Gaussian noise, σ_{noise}, of 0.0, 0.01 and 0.25. In a real particle accelerator the relevant length of the turn-by-turn data to use may vary. To study how each model would handle this we also generated datasets with varying turn numbers of 28, 64, 128 and 256 turns at each noise level, each of these datasets contained 1000 samples. The ability of each model to predict the true tune values was quantified using a calculation of the R^2 score of the measured tune from each model against the true values. The results are summarized in Table 1.

While ICA performs extremely well without noise it does suffer somewhat when the number of turns available is low. At high levels of noise ICA fails to make accurate measurements for any number of turns considered. The AE works well at the training length of 28 turns at all noise levels, but fails to generalize to inputs with more turn data. In contrast, the VAE works well in all regimes for both noise level and number of turns, and is able to provide excellent predictions of the tunes even when the input data contains significant noise and very few turns.

CONCLUSION

We have shown that variational recurrent autoencoders may offer significant advantages for analysis of data from beam position monitors in particle accelerators. When the interval of collected data is short or very noisy the VAE shows extremely good performance at analyzing important information, such as the accelerator tune. Use of a model like a VAE on an actual particle accelerator does come with the requirement that sufficient data be available for the initial training. However, given the increasing adoption of machine learning for particle accelerators large volumes of data are becoming ever more accessible. In the future the use VAEs for such analysis could also offer further advantages over current methods, like ICA, due to its ability to learn nonlinear relationships in the data.

ACKNOWLEDGEMENTS

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics under Award Number DE-SC0021699.

REFERENCES

[1] X. Huang, S. Y. Lee, E. Prebys, and R. Tomlin, "Application of independent component analysis to fermilab booster," *Phys. Rev. ST Accel. Beams*, vol. 8, p. 064001, Jun 2005. doi:10.1103/PhysRevSTAB.8.064001

[2] O. Fabius and J. R. van Amersfoort, "Variational recurrent auto-encoders," 2015.

[3] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," 2014.

[4] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[5] J. Laskar, "The chaotic motion of the solar system: A numerical estimate of the size of the chaotic zones," *Icarus*, vol. 88, no. 2, pp. 266–291, 1990.doi:10.1016/0019-1035(90)90084-M