ISBN: 978-3-95450-232-5 ISSN: 2673-7000 doi:10.18429/JACoW-NAPAC2022-THXD2

6D PHASE SPACE DIAGNOSTICS BASED ON ADAPTIVE TUNING OF THE LATENT SPACE OF ENCODER-DECODER CONVOLUTIONAL NEURAL NETWORKS

A. Scheinker*, Los Alamos National Laboratory, Los Alamos, NM, USA

Abstract

We present a general approach to 6D phase space diagnostics for charged particle beams based on adaptively tuning the low-dimensional latent space of generative encoder-decoder Convolutional Neural Networks (CNN). Our approach first trains the CNN based on supervised learning to learn the correlations and physics constrains within a given accelerator system. The input of the CNN is a high dimensional collection of 2D phase space projections of the beam at the accelerator entrance together with a vector of accelerator parameters such as magnet and RF settings. The inputs are squeezed down to a low-dimensional latent space from which we generate the output in the form of projections of the beam's 6D phase space at various accelerator locations. After training the CNN is applied in an unsupervised adaptive manner by comparing a subset of the output predictions to available measurements with the error guiding feedback directly in the low-dimensional latent space. We show that our approach is robust to unseen time-variation of the input beam and accelerator parameters and a study of the robustness of the method to go beyond the span of the training data.

INTRODUCTION

Particle accelerators are large complex systems whose beams evolve according to dynamics governed by nonlinear collective effects such as space charge forces and coherent synchrotron radiation. Because of their complexity, the control of charged particle beams in accelerators and diagnostics of these beams can greatly benefit from the application of machine learning (ML) [1] methods and advanced control theory techniques [2].

The development of ML-based tools for particle accelerator applications is an active area of research. At CERN, supervised learning techniques are being applied for the reconstruction of magnet errors in the incredibly large (thousands of magnets) LHC lattice [3]. At the LCLS, Bayesian methods have been developed for online accelerator tuning [4], Bayesian methods with safety constraints are being developed at the SwissFEL and the High-Intensity Proton Accelerator at PSI [5], and at SLAC Bayesian methods are being developed for the challenging problem of hysteresis [6] and surrogate models are being developed for the beam at the injector [7].

Convolutional Neural Networks (CNN) have been used to generate incredibly high resolution virtual diagnostics of the longitudinal phase space (LPS) of the electron beam in the EuXFEL [8]. A laser plasma wakefield accelerator has also been optimized by utilizing Gaussian processes at the Central Laser Facility [9].

Although ML tools such as deep neural networks can learn complex relationships in large systems directly from data, a major challenge faced by standard ML methods is that of time-varying systems or systems with distribution shift, which require extensive re-training whenever a system significantly changes. Accelerators continuously change and detailed beam measurements either interrupt operations or are only available for a few limited projections of the beam's 6D phase space. Therefore repetitive re-training is not a feasible solution except for very simple problems.

Recently, powerful model-independent feedback control methods, known as Extremum Seeking (ES), have been developed which can handle unknown and quickly time-varying nonlinear systems in which the direction of the controller's input is unknown and quickly time-varying [2, 10, 11]. For example, it is possible to use ES for RF cavity resonance control based only on ambiguous reflected power measurements [12]. While model-independent feedback such as ES is incredibly robust to un-modeled disturbances, noisy measurements, and can automatically track time-varying systems, a major limitation of local model-independent feedback is the possibility of getting stuck in a local minimum when operating in a complex high-dimensional parameter space.

Adaptive ML (AML) attempts to combine the complimentary strengths of ML and model-independent feedback, to provide the best of both worlds: an ability to learn directly from large complex data, while maintaining robustness to time variation and distribution shift. The first demonstration of the AML approach was the use of neural networks together with ES for automatic femtosecond-level control of the time-varying longitudinal phase space distribution of the electron beam in the LCLS [13].

AML methods have also combined CNN and ES to track time-varying input beam distributions at the HiRES UED [14], and preliminary results have shown an ability to adaptively tune the low-dimensional latent space of encoder-decoder CNN to track all 15 unique 2D projections of beam's 6D phase space despite unknown and time-varying input beam distributions and accelerator and beam parametes [15]. Such AML methods are general tools applicable to a wide range of complex time-varying systems and have also been demonstrated for 3D reconstructions of the electron density of crystals for coherent diffraction imaging [16].

* ascheink@lanl.gov

ISBN: 978-3-95450-232-5 ISSN: 2673-7000 doi:10.18429/JACoW-NAPAC2022-THXD2

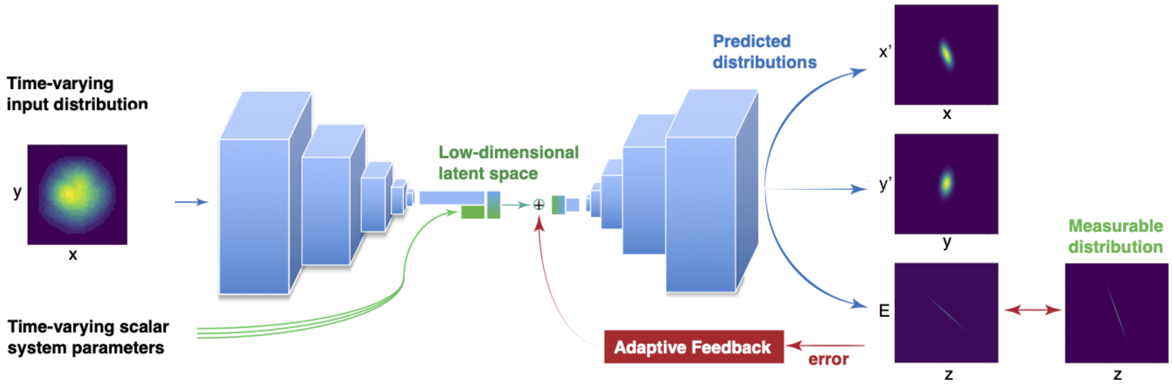

Figure 1: An input beam $\rho(x, y)$ distribution as well as beam and accelerator scalar parameters are used as input and squeezed down to a 2D latent space from which all 15 2D projections of the beam's 6D phase space are generated (only three projections shown here). Out of all of the generated projections, only the LPS (z, E) projection is used to compare to a measurement of the LPS, as is available in most modern electron accelerators such as the LCLS, FACET-II, and the EuXFEL. The LPS-based error then adaptively tunes the latent space representation of the beam to achieve a match.

AML FOR 6D DIAGNOSTICS

Time-varying systems, or systems with distribution shift, are an open problem and an active area of research in the ML community [17–21].

In this work we present simulation-based AML studies at the HiRES UED [22], for predicting all 15 unique 2D projections of a charged particle beam with unknown and time-varying input beam conditions at the photocathode, unknown beam charge and injector solenoid magnet strength, and demonstrate that this method has the capability to accurately predict beyond the span of the training set data.

We tackle the problem of distribution shift by incorporating model-independent adaptive feedback directly within the architecture of an encoder-decoder CNN which takes beam distributions and parameters (charge and solenoid current) as inputs and generates 15 256×256 pixel 2D projections ($\sim 10^6$ dimensions) of the beam's 6D phase space downstream from the injector.

Our AML setup is shown in Fig. 1 in which a CNN-based encoder-decoder takes as input the initial $\rho(x, y)$ transverse beam density as well as bunch charge and solenoid strength in the HiRES UED. These inputs are squeezed down to a 2D latent space embedding from which the generative half of the encoder-decoder then generates all 15 unique 2D projections of the beam's 6D phase space.

We demonstrate the ability of the encoder-decoder to accurately generate phase space distributions from a 2D latent space. Figure 2 shows one case of the CNN's predictions as compared to the ground truth for a beam with known input distribution and known charge passing through an accelerator with a known solenoid strength.

We trained our CNN in a supervised learning approach by collecting input beam distribution images at HiRES over several days and then using principal component analysis to extract a set of basis functions from those measurements in order to generate additional synthetic input beam distributions [23, 24]. For each input distribution we then scanned a grid of points in bunch charge and solenoid strength at the HiRES injector and simulated the beam transport before saving the 2D projections of the beam's 6D phase space further down the accelerator.

After training, we assume that we will lose access to the time-varying accelerator parameters and to the time-varying input beam distribution and our CNN is applied in an unsupervised adaptive way by squeezing down to a 2D latent space between the encoder and decoder sections which is adaptively tuned using ES with time-varying cost

$$C(t) = \iint \left|\rho_{z,E}(t) - \hat{\rho}_{z,E}(t)\right| dE dz, \quad (1)$$

which is a comparison between the CNN's longitudinal phase space (LPS) prediction $\hat{\rho}_{z,E}$ and the measurement of the LPS as provided by a TCAV $\rho_{z,E}$. No other projections of the beam's phase space are assumed to be available for measurement. However, by forcing the CNN to simultaneously generate all 15 projections of the 6D phase space we introduced observational biases directly through data embodying the underlying physics, allowing the CNN to learn functions that reflect the physical structure of the data [25].

The ES-based update for parameters $\mathbf{p} = (p_1, p_2)$ which perturb the latent space representation of the beam takes place according to the ES dynamics:

$$\frac{dp_1}{dt} = \sqrt{\alpha\omega}\cos\left(\omega t + kC(\mathbf{p}, t)\right),$$
$$\frac{dp_2}{dt} = \sqrt{\alpha\omega}\sin\left(\omega t + kC(\mathbf{p}, t)\right), \quad (2)$$

where $C(t)$ is as in Eq. (1), ω is a dithering frequency, α controls the perturbation size, k is a learning rate controlling convergence speed. For large ω the parameter dynamics on-average can be described by

$$\frac{d\mathbf{p}}{dt} = -\frac{k\alpha}{2}\nabla_{\mathbf{p}} C(\mathbf{p}, t), \quad (3)$$

which tracks the time-varying minimum of the analytically unknown cost function. Our encoder-decoder setup

ISBN: 978-3-95450-232-5 ISSN: 2673-7000 doi:10.18429/JACoW-NAPAC2022-THXD2

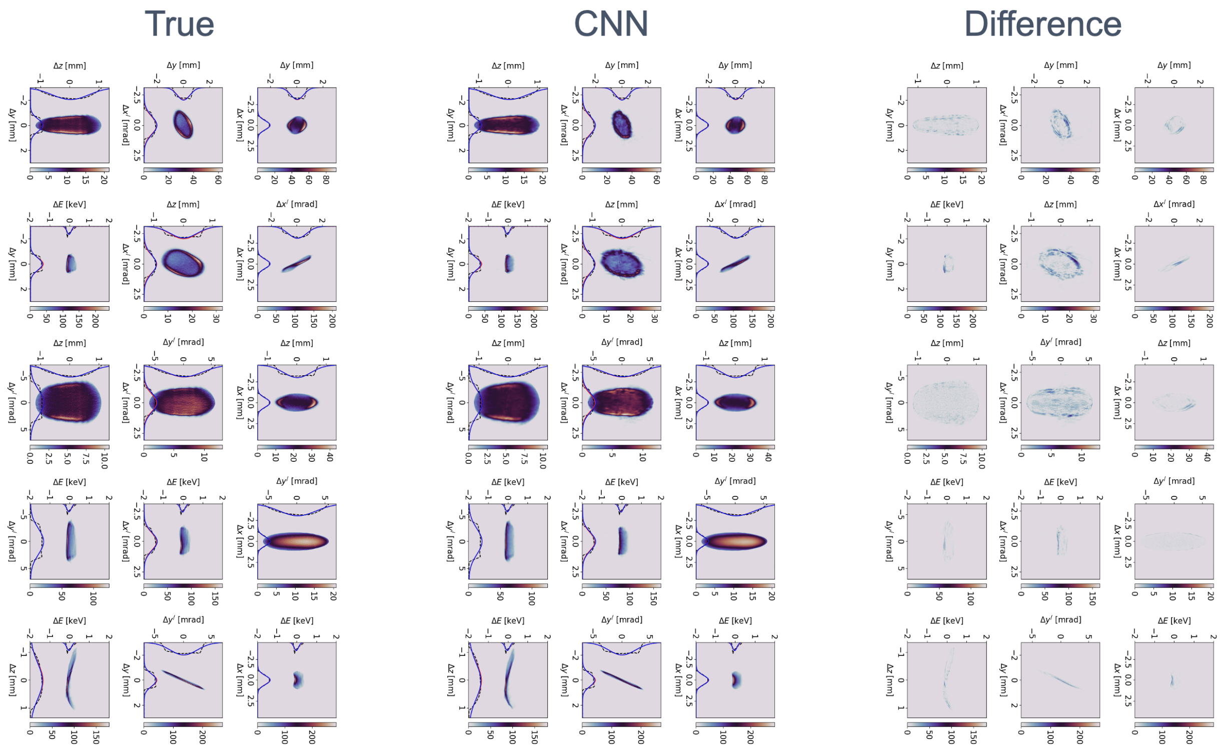

Figure 2: Projections of the beam's true 6D phase space are shown on the left, CNN predictions generated from the 2D latent space are shown in the middle, and the difference is shown on the right. Note that for each of the 15 projections the color scale of True, CNN, and Difference is the same.

translates this incredibly high dimensional problem into a 2D adaptive parameter feedback resulting in very fast convergence which can track quickly time-varying systems. A more detailed overview of the ES method including analytic proofs of convergence for noisy nonlinear and time-varying systems can be found in the references.

TRACKING PHASE SPACE

To demonstrate the robustness of this AML method for time-varying systems and for beam and accelerator parameters outside of the span of the training set we measured one additional input beam distribution at the HiRES injector 6 months after the initial training data was collected. We then generated a series of input beams in which we perform linear interpolation from one input beam distribution $\rho_0(x, y)$ which was seen during training to the new unseen distribution $\rho_u(x, y)$ over 25 steps ($n = 1, \ldots, 25$):

$$\rho(x, y, n) = \rho_0(x, y)\frac{25 - n}{24} + \rho_u(x, y)\frac{n - 1}{24}. \quad (4)$$

During this interpolation we also chose a new unseen bunch charge Q_u and solenoid strength S_u far outside of the span of the training data and interpolated their values as well starting with initial values Q_0, S_0 within the span of the training data:

$$Q(n) = Q_0\frac{25 - n}{24} + Q_u\frac{n - 1}{24}, \quad (5)$$

$$S(n) = S_0\frac{25 - n}{24} + S_u\frac{n - 1}{24}. \quad (6)$$

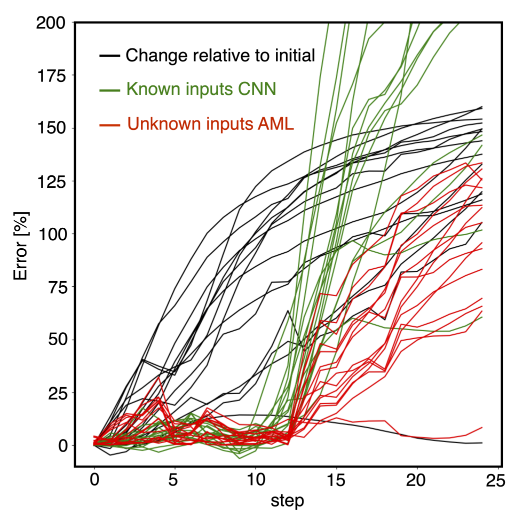

Figure 3: Error is shown in percent for each of the 15 projections of the beam's 6D phase space as the input beam distribution, charge, and solenoid strength are moved beyond the span of the training set.

First, to illustrate the limitations of traditional ML approaches, we used $\rho(x, y, n), Q(n), S(n)$ as inputs to the trained encoder-decoder CNN. As expected, the CNN's predictions were very accurate for the first ~12 steps until we hit the edge of the span of the training data, resulting in catastrophic failure as presented in Fig. 3.

ISBN: 978-3-95450-232-5 ISSN: 2673-7000 doi:10.18429/JACoW-NAPAC2022-THXD2

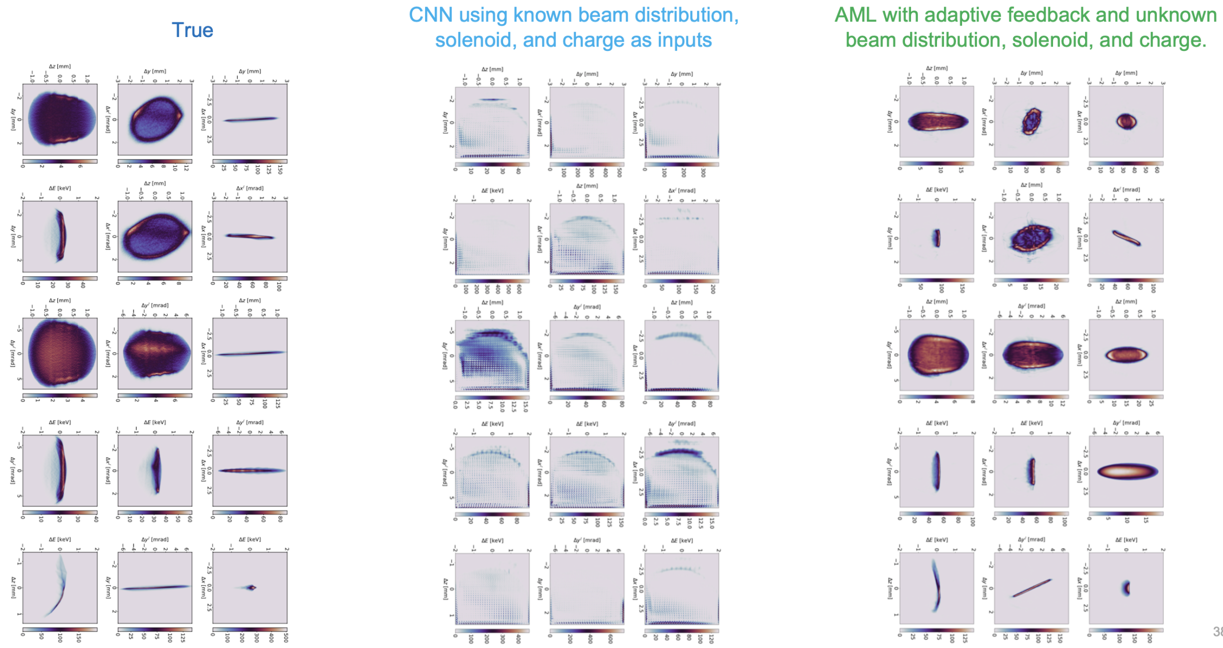

Figure 4: All 15 projections of the beam's 6D phase space are shown on the left for a previously unseen input beam distribution and charge and solenoid current values outside of the span of the training set. The middle column shows the CNN's saturated and completely wrong predictions when these new inputs are used. The right column shows the predictions made by the CNN when the input distribtuion, charge, and solenoid current are unknown, but adaptive feedback was used to try and track the LPS (z, E) projection.

Figure 3 shows the results of changing the input beam distribution, beam charge, and solenoid current far beyond the span of the training data. The black lines show the change relative to the initial starting condition. The green lines show the CNN's errors if assuming known beam distribution, charge, and solenoid strength, with catastrophic failure beyond the span of the training set where the CNN's predictions are far worse than simply doing nothing.

To demonstrate the strength of AML we fixed the CNN's input as:

$$(\rho(x, y, 1), Q(1), S(1)) = (\rho_0(x, y), Q_0, S_0), \quad (7)$$

and then as the unknown input beam, charge, and solenoid strength were changed, we compared a measurement of the 2D LPS (z, E) projection with the prediction of the encoder-decoder and adaptively tuned the latent space in order to track the LPS according to the ES approach described above with cost function Eq. (1).

The red lines in Fig. 3 show the error when we do not have access to the unknown beam distribution, charge, and solenoid strength, but with the use of adaptive feedback which has access to the (z, E) projection to be used as feedback within the latent space by continuously minimizing the cost function Eq. (1). Although the AML method also begins to lose accuracy as we go beyond the training set it does so in a very controlled manner and achieves results which are much more accurate than the CNN alone and also more accurate than doing nothing (black lines).

Figure 4 shows the results of the adaptive tracking procedure, in which we compare the true phase space projections to those predicted by the CNN alone as well as to the AML-predicted phase space projections. Clearly the CNN is failing and saturating for points so far beyond the span of its training set. On the other hand the AML-based approach was able to closely track the LPS and gave reasonable predictions for the other phase space projections, although they are also imperfect, as expected for unseen data, they are much more reasonable and give fairly accurate 1D projections of the various 2D beam profiles.

CONCLUSIONS

We have demonstrated preliminary studies of a physics-informed AML method for tracking all 15 projections of a charged particle beam with unknown and time-varying initial distribution and charge at the photocathode and unknwon and time-varying solenoid strength at the injector based only on TCAV measurements of the (z, E) LPS.

ACKNOWLEDGEMENTS

Work funded by US Department of Energy (DOE), Office of Science, Office of High Energy Physics under contract number 89233218CNA000001 and the Los Alamos National Laboratory LDRD Program Directed Research (DR) project 20220074DR.

ISBN: 978-3-95450-232-5 ISSN: 2673-7000 doi:10.18429/JACoW-NAPAC2022-THXD2

REFERENCES

[1] U. Gentile and L. Serio, "A machine-learning based methodology for performance analysis in particles accelerator facilities," in *Proc. EECS'17*, 2017, pp. 90–95. doi:10.1109/EECS.2017.26

[2] A. Scheinker and M. Krstic, "Minimum-seeking for CLFs: Universal semiglobally stabilizing feedback under unknown control directions," *IEEE Transactions on Automatic Control*, vol. 58, no. 5, pp. 1107–1122, 2012. doi:10.1109/TAC.2012.2225514

[3] E. Fol, R. Tomás, and G. Franchetti, "Supervised learning-based reconstruction of magnet errors in circular accelerators," *The European Physical Journal Plus*, vol. 136, no. 4, p. 365, 2021. doi:10.1140/epjp/s13360-021-01348-5

[4] J. Duris, D. Kennedy, and D. Ratner, "Bayesian optimization at LCLS using Gaussian processes," in *Proc. HB'18*, 2018.

[5] J. Kirschner, M. Mutný, A. Krause, J. Coello de Portugal, N. Hiller, and J. Snuverink, "Tuning particle accelerators with safety constraints using bayesian optimization," *Phys. Rev. Accel. Beams*, vol. 25, no. 6, p. 062 802, 2022. doi:10.1103/PhysRevAccelBeams.25.062802

[6] R. J. Roussel and A. Hanuka, "Towards Hysteresis Aware Bayesian Regression and Optimization," in *Proc. IPAC'21*, Campinas, Brazil, May 2021, pp. 2159–2162. doi:10.18429/JACoW-IPAC2021-TUPAB289

[7] L. Gupta, A. Edelen, N. Neveu, A. Mishra, C. Mayes, and Y.-K. Kim, "Improving surrogate model accuracy for the LCLS-II injector frontend using convolutional neural networks and transfer learning," *Machine Learning: Science and Technology*, vol. 2, no. 4, p. 045 025, 2021. doi:10.1088/2632-2153/ac27ff

[8] J. Zhu, Y. Chen, F. Brinker, W. Decking, S. Tomin, and H. Schlarb, "High-fidelity prediction of megapixel longitudinal phase-space images of electron beams using encoder-decoder neural networks," *Phys. Rev. Applied*, vol. 16, no. 2, p. 024 005, 2021. doi:10.1103/PhysRevApplied.16.024005

[9] R. Shalloo *et al.*, "Automation and control of laser wakefield accelerators using Bayesian optimization," *Nature Communications*, vol. 11, no. 1, pp. 1–8, 2020. doi:10.1038/s41467-020-20245-6

[10] A. Scheinker, "Simultaneous stabilization and optimization of unknown, time-varying systems," in *Proc. 2013 Amer. Contr. Conf.*, IEEE, 2013, pp. 2637–2642. doi:10.1109/ACC.2013.6580232

[11] "Bounded extremum seeking with discontinuous dithers," *Automatica*, vol. 69, pp. 250–257, 2016. doi:10.1016/j.automatica.2016.02.023

[12] A. Scheinker, "Application of extremum seeking for time-varying systems to resonance control of RF cavities," *IEEE Transactions on Control Systems Technology*, vol. 25, no. 4, pp. 1521–1528, 2016. doi:10.1109/tcst.2016.2604742.

[13] A. Scheinker, A. Edelen, D. Bohler, C. Emma, and A. Lutman, "Demonstration of model-independent control of the longitudinal phase space of electron beams in the Linac-Coherent Light Source with femtosecond resolution," *Phys. Rev. Lett.*, vol. 121, no. 4, p. 044 801, 2018. doi:10.1103/PhysRevLett.121.044801

[14] A. Scheinker, F. Cropp, S. Paiagua, and D. Filippetto, "An adaptive approach to machine learning for compact particle accelerators," *Scientific Reports*, vol. 11, no. 1, pp. 1–11, 2021. doi:10.1038/s41598-021-98785-0

[15] A. Scheinker, "Adaptive machine learning for time-varying systems: Low dimensional latent space tuning," *J. Instrum.*, vol. 16, no. 10, P10008, 2021. doi:10.1088/1748-0221/16/10/P10008

[16] A. Scheinker and R. Pokharel, "Adaptive 3d convolutional neural network-based reconstruction method for 3d coherent diffraction imaging," *J. of Appl. Phys.*, vol. 128, no. 18, p. 184 901, 2020. doi:10.1063/5.0014725

[17] J. Quiñonero-Candela, M. Sugiyama, A. Schwaighofer, and N. D. Lawrence, *Dataset shift in machine learning*. MIT Press, 2008. doi:10.7551/mitpress/7921.003.0001

[18] J. G. Moreno-Torres, T. Raeder, R. Alaiz-Rodrıguez, N. V. Chawla, and F. Herrera, "A unifying view on dataset shift in classification," *Pattern Recognition*, vol. 45, no. 1, pp. 521–530, 2012. doi:10.1016/j.patcog.2011.06.019

[19] M. Sugiyama and M. Kawanabe, *Machine learning in non-stationary environments: Introduction to covariate shift adaptation*. MIT press, 2012. doi:10.7551/mitpress/9780262017091.001.0001

[20] Y. Ovadia *et al.*, "Can you trust your model's uncertainty? evaluating predictive uncertainty under dataset shift," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[21] A. Subbaswamy and S. Saria, "From development to deployment: Dataset shift, causality, and shift-stable models in health ai," *Biostatistics*, vol. 21, no. 2, pp. 345–352, 2020. doi:10.1093/biostatistics/kxz041

[22] D. Filippetto and H. Qian, "Design of a high-flux instrument for ultrafast electron diffraction and microscopy," *J. Phys. B: At. Mol. Opt. Phys.*, vol. 49, no. 10, p. 104 003, 2016. doi:10.1088/0953-4075/49/10/104003

[23] K. Pearson, "On lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, no. 11, pp. 559–572, 1901.

[24] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: computational statistics*, vol. 2, no. 4, pp. 433–459, 2010. doi:10.1002/wics.101

[25] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," *Nature Reviews Physics*, vol. 3, no. 6, pp. 422–440, 2021. doi:10.1098/rsta.2020.0093