# Machine Learning Applications for ALS & ALS-U

Simon C. Leemann, ALS Accelerator Physics Program Deputy, for the ALS-AP ML Team Accelerator Technology & Applied Physics Division



Director's Review May 16, 2024



ACCELERATOR TECHNOLOGY & ATAPO



Office of Science

# Intro: Machine Learning (ML) at the ALS

- ALS has been kept at the forefront of soft x-ray light sources for 3 decades by continuous upgrades and R&D effort
- ML presents excellent new opportunity for accelerator R&D to extend ALS leadership
- ALS ML efforts have so far been enabled by a 3-year grant funded jointly by DOE BES ADRP & ASCR as well as by ALS operations funds
- Initial ALS ML R&D effort: use ML as a powerful "new" tool to solve "old" accelerator problems:
  - Accelerator operations: automated tuning, replace feedback approaches, virtual diagnostics → Project #1
  - Accelerator development: improve physics understanding, augment/extend lattice optimization, accelerate multi-objective optimization (e.g. MOGA) → *Project* #2

# **#1 ML for Acc Ops: Stabilizing Beam Size at ALS**

- State-of-the-art light sources achieve excellent stability in terms of beam position/angle & current (orbit feedbacks, top-off injection)
- In spite of <u>extensive</u> correction efforts, beam size is still perturbed by insertion device (ID) config changes → can affect experimental resolution
- Problem is nonlinear, complex, and non-stationary
- Previous solutions relied on approximations & required extensive dedicated machine time for frequent recalibration (feed-forward tables)
- Resulting level of performance has started to become a limitation at most demanding experiments (STXM, XPCS, ptychography, ...)
- Expected to become a serious issue in next-generation light sources (diffraction-limited storage rings, eg. APS-U, ALS-U, ...)







# **Developing a Solution Based on Machine Learning**

- Machine Learning can exploit large amounts of data that are already collected during routine operations → "training"
- Once trained, neural network (NN) provides predictions for beam size changes that result from ID config changes & skew quad corrections (V disp. wave)





# **Developing a Solution Based on Machine Learning**

- Machine Learning can exploit large amounts of data that are already collected during routine operations → "training"
- Once trained, neural network (NN) provides predictions for beam size changes that result from ID config changes & skew quad corrections (V disp. wave)
- These predictions can serve as a dynamic lookup
   → which skew quad correction required to compensate for changes resulting from currently applied ID config?
- If such a lookup is incorporated into the accelerator control system as a feed forward (FF), we can stabilize the storage ring over prolonged periods of time & online retraining can mitigate drift

Training

am Size

Outpi

Beam Size

Prediction

Input

Magnets

IDs

# **Results: NN-based FF Off vs. On During User Ops**



#### Stabilization Confirmed at STXM @ ALS Beamline 5.3.2.2



#### **Recently: Improved Model With Online Fine-tuning & Implementation**

- Split up **DNN** into 2 models
- Revised primary DNN based on broad hyper-parameter optimization
- Revised & accelerated training DAQ: ID scanning based on one prior year of user ops
- Online retraining → online fine-tuning anchored DNN using sliding window buffer → measured performance now at diagnostic BL noise floor
- Robustness & integration → new high-level interface with event logging & monitoring
- One-click operation provides for becoming standard part of everyday user ops





#### #2 ML for Acc Dev: Improving Multi-Objective Optimization

- 4th-generation storage rings (4GSRs) leverage multi-bend achromat (MBA) lattices to deliver ultra-high brightness & large coherent fraction
- But MBA lattices are very challenging: dense & exploit very strong focusing → drives large chromatic terms & higher-order corrections
- Solutions not only highly nonlinear but involve many degrees of freedom (DoF) → demanding optimization:
  - tough objectives, many often in direct competition
  - large number of parameters, many boundary constraints
- Multi-objective genetic algorithms (MOGA) are highly successful at such optimization & have become tool of choice among community



Courtesy: Dave Robin



# **Improving MOGA: ML to the Rescue**

- But MOGA's stochastic nature is inherent weakness → need to evaluate vast number of lattice candidates, most ultimately rejected
- Do <u>not</u> want to artificially limit DoF, search ranges, or make many initial assumptions about attractive solutions → so what can we do?
- ML can be employed to render deep neural networks (DNNs)

   → surrogate models used in place of computationally expensive evaluation (many-turn nonlinear tracking)
- Lattice candidate evaluation becomes near instantaneous
- And ideally, we'll want to target:
  - speed up MOGA <u>without</u> modifying MOGA/tracking tools or existing workflows & without sacrificing physics fidelity
  - direct optimization of relevant physics quantities (ε<sub>0</sub>, DA, MA)
  - combined linear/nonlinear optimization employing all free parameters (quadrupoles & sextupoles & ...)





### **ALS-U** Optimization as a Test Case for ML

ALS-U storage ring calls for challenging 9BA lattice to achieve ≈75 pm rad (round beam) @ 2 GeV in <200 m</li>
 → dense, strong focusing, very strained optics





### **ALS-U** Optimization as a Test Case for ML

- ALS-U storage ring calls for challenging 9BA lattice to achieve ≈75 pm rad (round beam) @ 2 GeV in <200 m</li>
   → dense, strong focusing, very strained optics
- Initial optimization: 9 quadrupoles, 4 sextupoles → 11
   free knobs (later: include reverse bending & superbends)
  - Roughly a dozen magnet/lattice **constraints** on top of pre-determined quadrupole ranges
  - **Objectives**: ε<sub>0</sub>, MA, and on-momentum DA (modeled as integrated diffusion rate)



Beta in central arc bends (B3) Fractional tune difference Chromatic sextupole strength (SF, SD)	$\begin{aligned} \beta_{x,y}^{B3} &< 4 \mathrm{m} \\  \nu_x - \nu_y  &< 0.01 \\ b_2 &< 900 \mathrm{m}^{-3} \end{aligned}$	
$\sum_{i=1}^{n}$	1	



12

#### **ALS-U** Optimization as a Test Case for ML

- ALS-U storage ring calls for challenging 9BA lattice to achieve ≈75 pm rad (round beam) @ 2 GeV in <200 m</li>
   → dense, strong focusing, very strained optics
- Initial optimization: 9 quadrupoles, 4 sextupoles → 11
   free knobs (later: include reverse bending & superbends)
  - Roughly a dozen magnet/lattice **constraints** on top of pre-determined quadrupole ranges
  - **Objectives**: ε<sub>0</sub>, MA, and on-momentum DA (modeled as integrated diffusion rate)
- Ultimately, a highly staged MOGA approach resulted in
  - ±1 mm DA (compatible with on-axis swap-out & AR)
  - ≈1 hr overall lifetime (including x4 boost from 3HCs)
- ...but required *months* of CPU time on large clusters



Courtesy: Changchun Sun

# **ML for Full Linear & Nonlinear ALS-U Optimization**

- **Training data** for 11D problem cannot be acquired through systematic sampling of input space
- But do not want to make too many assumptions or "wise choices" → retain generality of approach
- Instead: use first few generations of conventional MOGA as training data for deep neural networks (DNNs)
- **Two 8-layer DNNs** used in MOGA instead of calls to Tracy for DA and MA (via many-turn tracking)
  - Traditional MOGA requires about 640 gen (5000 children/gen) → ≈8 days on 1000-core cluster
  - Training 2 DNNs to get DA/MA predictions ≈1% rms requires data from about 10 gen
- Once DNNs trained (≈30 min on desktop CPU) → quasiinstantaneous lookup (16 ms) vs. DA/MA tracking (88 sec)



Fully-connected (FC) NN, using ReLU as activation function, # = node depth

Courtesy: Yuping Lu



# **Results: ML-MOGA Successful & 40× Faster**

- ML predictions are not 100% accurate (training based on early data)
- Initial ML-MOGA solutions show disagreement compared to tracking validation → but can retrain DNNs with data from validation step
- Iterate cycles of *validation—retraining—ML-MOGA* using **modelindependent distance metrics** to determine **convergence**
- ML-MOGA very quickly converges (6-8 iterations) toward true Paretooptimal front → overall speedup ≈ 40× (incl. training effort)
- Once fully converged, ML-MOGA inputs & objectives match those of traditional MOGA to within "noise floor"
- **Flexible**: can be adapted to other lattice optimization problems as long as can provide reasonably accurate DNNs
- Potential to fully automate entire optimization campaign & optimize in parallel from the start for many error lattices is highly attractive → derive truly *global* optimum



# Success, Outlook & Opportunities for Collaboration

- Project success: ML stabilization now always running during user ops @ ALS & 1x PRL, 1x NIM-A, 1x submitted PRAB, 2021 Klaus Halbach Prize for Innovative Instrumentation
- ML applied to accelerators shows vast potential to enable
  - new more aggressive designs, but also
  - exploit full performance of existing & soon to be commissioned rings
- These are highly relevant issues in both present (ALS) & future 4th-generation storage rings (ALS-U)
- We have plenty of ideas for future ML applications (White Paper submitted to E. Lessner)
- ATAP & LBL continue to foster great collaboration on ML for accelerators
  - https://atap.lbl.gov/research/crosscutting-endeavors/artificial-intelligence-and-machine-learning/
  - https://ml4sci.lbl.gov/projects

PHYSICAL REVIEW LET	TERS 123, 194801 (2019)	Application of deep learning methods for b	beam size control o		
Featured in Physics		the Advanced	Light Source		
Demonstration of Machine Learning-I	Based Model-Independent Stabilization	Thorsten Hellert, Tynan Ford, Simon C. Leem Lawrence Berkeley National Laborator	ann, Hiroshi Nishimura, ry, Berkeley 94720, Califo		
Source Properties in S S. C. La Lawrence Berkeley National Laborat	seman" ver, Berkeley, California 94720, USA	Andrea I Department of Electrical Engineering and Information Techno Instrumentation and Measurement for Pau (Dated: Ap	Andrea Pollastro Department of Electrical Engineering and Information Technology (DIETI), University Instrumentation and Measurement for Particle Accelerator Laborato (Date: April 1, 2024)		
Department of Chemistry, University of C A. Hessenet, M. A. Marcus, C. N. Lawrence Berleizh National Labour (2014)	200 Minosa, Berkeley, California 94720, USA Meltos, H. Nubhimzer, and C. San ory, Berkeley, California 99720, USA eived 23 August 2019; published 6 November 2019) none powerlat sooils of anodern scientific discovery, provide coders of magnitude higher beightesses and	Past research at the Advanced Light Source (A that muchine learning (ML) methods could be et icant perturbations to the transverse electron be of the instruction devices (II). However, incorporating has faced notable challenges. The complexity of significant updrops domatable how restricted theirs we introduce the developments of a more robust analysis embasisme the recovers of ML model of up	Past resurch at the Abraneck Light Source (ALS) provided a proof-of- that nather is suring (AL) methods could be discribely employed to con- form the state of the state of the state of the state of the of the insertion devices [1]. However, incorporating these methods into the has faced souther disclasses. The comparison of the system's operational we introduce the devices provided the system's discrimination into the dy-to-locy are functional and the systematic discrimination into the dy-to-locy are		
transverse coherence enabling the most demanding ex- will soon be limited by achievable levels of electron- microns, which is still 1–2 orders of magnitude larger and current. Urtil new source size stabilization has bee of static predeternianed physics models and lengthy counterset drift in the accolerator and instrumentati	erimets. In these experiments, overall source stability beam size stability presently on the outer of sloveral than alexally demonstrated stability of source position a calibration measurements, periodically repeated to calibration measurements, periodically repeated to no. We now demonstrate for the first time how the	ML-based methods over traditional feetback meth of the new algorithm during user-operation scenar I. INTRODUCTION	ods, and examines the effe tios. IDs. Since these meass quality of the boleum t		
application of machine learning allows for a physic relying only on provisionly existing instrumentation. S that can be continuously colline retrained achieves so results in overall source stability approaching the solu DOI: 10.1103/PhysRevLet.123.194801	<ul> <li>and model-independent stabilization of source size and the object of the source of the source size stability as low as 0.2 µm (0.4%) rms, which nee size stability as low as 0.2 µm (0.4%) rms, which percent noise floor of the most sensitive experiments.</li> </ul>	The performance of storage ring light sources is crit- ically reliant on the stability of the radiation output in terms of source position/angle and intensity. A major advance to constant improving the relation in- termediate the stability of the radiant of the radiant of the positive relatives 20 in the radiant	the machine short-tern process. Moreover, lon as ground motion, radii eventually compromises FF correction and vio assumption will invaria		
Anotherina – Synchroner radiation sources, point (1) the dispersion sources are respectively be a sources, how not introduced procession of the sources of the source of the source of the source of the sources of the adjustment of the source of the source of the source of the adjustment of the source of the source of the source of the adjustment of the source of the source of the source of the adjustment of the source of the source of the source of the adjustment of the source of the source of the source of the adjustment of the source of the source of the source of the adjustment of the source of the sour	well as XCS style heaving the high heast multility over encoded periods of the dist of the star of the star of the star in the star of the scores of a stern grin fight scores. The star of the Star of the star of the star of the star of the star of the Star of the star of the star of the star of the star of the Star of the star of the star of the star of the star of the Star of the star of the star of the star of the star of the Star of the star of the star of the star of the star of the Star of the star of the star of the star of the star of the star of the star	tion score position, angle stability in addicted by control of the decision bases with through a confination of the properties. Other stability at the only-an level, maintained one waves the bases, is not priority (or eq. [6, 3]). The stability of the only-an level, maintained one waves the bases, is not priority (or eq. [6, 4]). The stability is primarily affected by the decisers how more thready is primarily affected by the decisers how more reversion. While in gravest the horizontal decoses the largely independent of the central bases due to the lower is the bounding magnetic, the vertical bound after turbs to be smaller to be the central and ensemption of the lower is the bounding magnetic, the vertical bound after turbs to be smaller to be the ensemption of the lower To compression for these ensemptions of the lower of the Sance (A.C.S.). Base dust magneting light scores, or Sance (A.C.S.). Base dust magneting light scores, or Sance (A.C.S.). Base dust magneting light scores, or the stability of the score of the lower of the score of the lower of th	tion even in absence of A conceivable way to add a conventional FB size is monitored in r and the control variable ment excites a vertical combination of skew q he underlying mensur- noise, thus limiting the that an FB system fal closed-loop bandwidth that a mer disvite desired Previous investigati- that a mer effective lookup-table system w based on a neural net		
eaps with unprecedented resolut coper understanding of electroch atteries and feel cells. The mean of kinetics to study chemical as the direction of the direction immediated both the new sources.	<b>23</b> , 194801 (2019	Fr configuration [6]. Because of the difficulty of developing an accur physics model for the ID errors, the FF corrections best defined based on beam measurements. A distinct	ubmit		
n emerging technique to study this is x-ray photon	user run. While the horizontal beam size remains constant	set of measurements is conducted for each individual ID resulting in the creation of lookup tables that, in corre-	of the ML approach of NN algorithm, which is		





Acknowledgments: Shuai Liu, Nathan Melton, Yuping Lu, Hiroshi Nishimura, Changchun Sun, Matthew Marcus, David Shapiro, Alex Hexemer, Dani Ushizima, Mike Ehrlichman, Gregg Penn, Thorsten Hellert, Erik Wallen, Warren Byrne, Fernando Sannibale, Marco Venturini, Andrea Pollastro, Andreas Scholl, Rob Ryne, DOE Office of Science Contract No. DEAC02-05CH11231







Office of Science