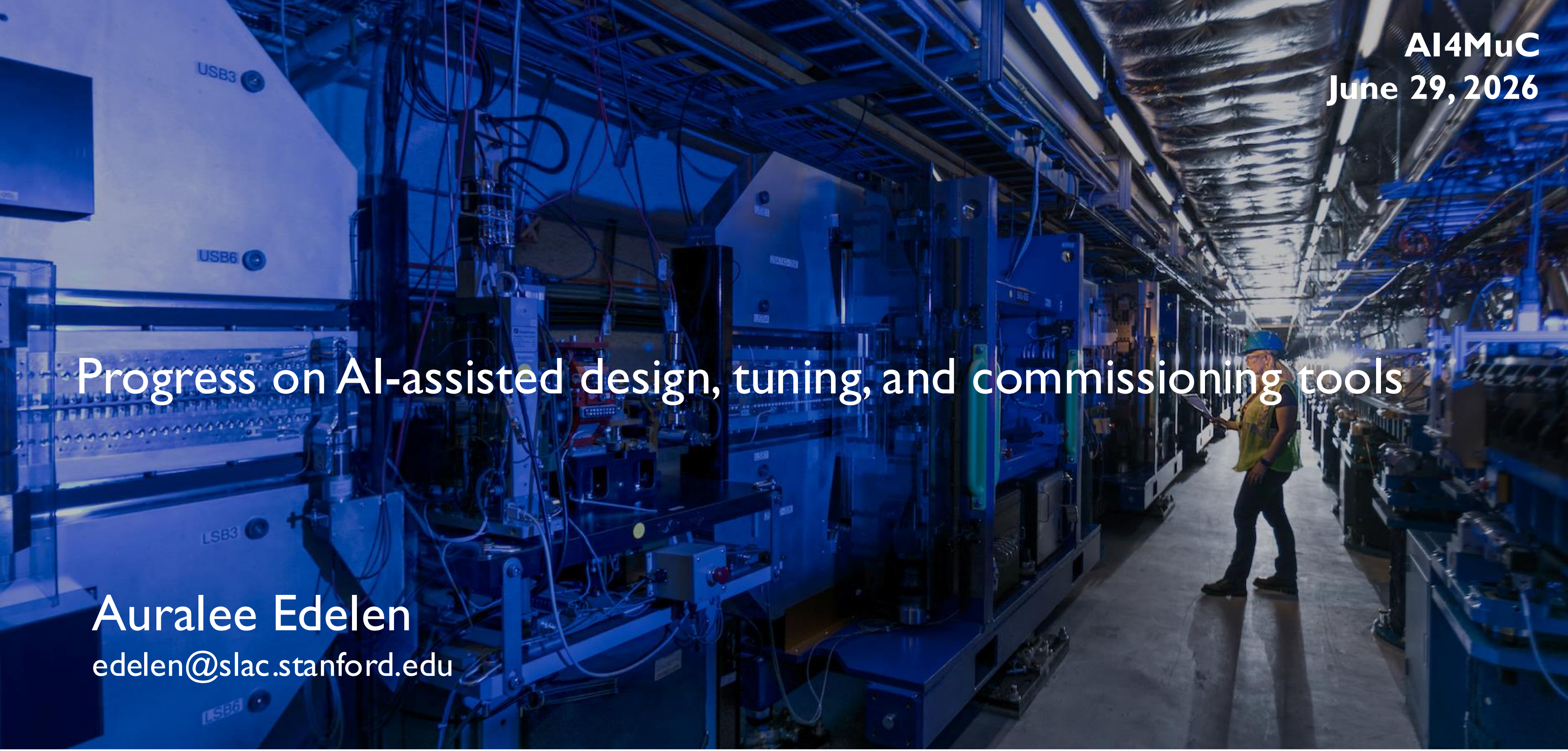


AI4MuC
June 29, 2026

Progress on AI-assisted design, tuning, and commissioning tools

Auralee Edelen
edelen@slac.stanford.edu

SLAC NATIONAL ACCELERATOR LABORATORY · Argonne NATIONAL LABORATORY · Brookhaven National Laboratory · The Center for BRIGHT BEAMS A National Science Foundation Science & Technology Center · THE UNIVERSITY OF CHICAGO · Jefferson Lab · BERKELEY LAB · UKRI Science and Technology Facilities Council ISIS Neutron and Muon Source · DESY.

Control approaches leverage different amounts of data / previous knowledge

→ suitable under different circumstances and need a combined strategy

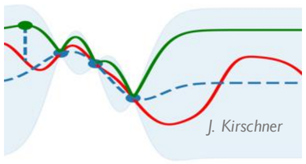
less ← assumed knowledge of machine → more

Model-Free Optimization

No previous data/model required

gradient descent
simplex
ES

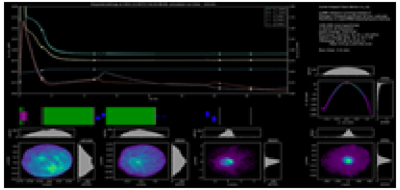
Model-guided Optimization



Minimal data / on-the-fly learning

Bayesian optimization + advanced variants
Model Predictive Control with simple models

Global Learned Representations



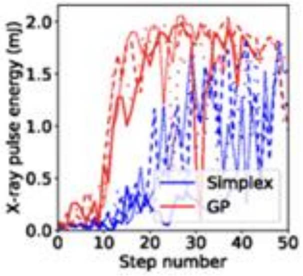
→ *provide initial guess (i.e. warm start)*
→ *model-based control*

System models, inverse models
Reinforcement Learning,
Model Predictive Control
Feed-forward corrections

General strategy: start with sample-efficient methods that do well on new systems, then build up to more data-intensive and heavily model-informed approaches.

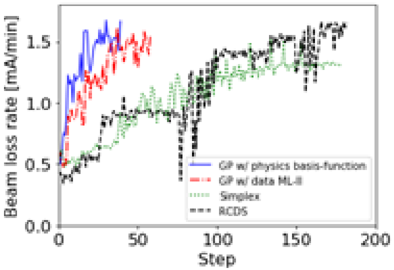
Bayesian Optimization with Advanced Variants is Mature and Operations-Ready

FEL pulse energy at LCLS



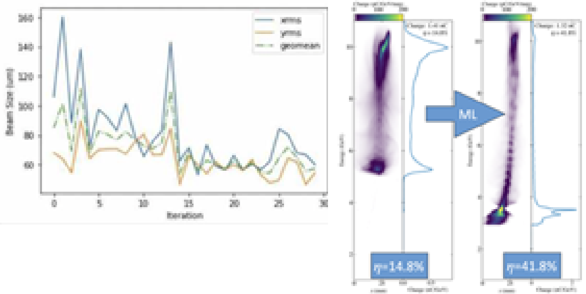
Duris et. al. PRL, **2020**

Loss rate at SPEAR3

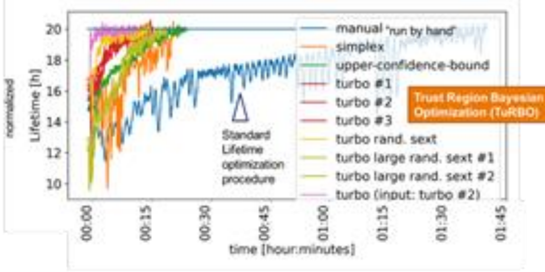


Hanuka et. al. PRAB, **2021**

Sextupole tuning at FACET-II
>2x efficiency of acceleration in plasma

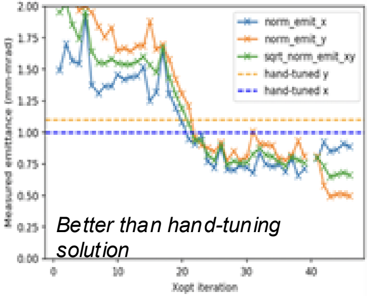


50x faster tuning and best observed lifetime at ESRF

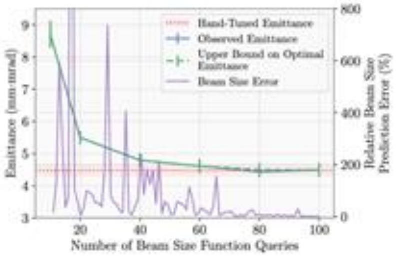


Liuzzo, ICALEPCS 2023 doi:10.18429/JACoW-ICALEPCS2023-MO3AO01

Emittance @ LCLS-II injector

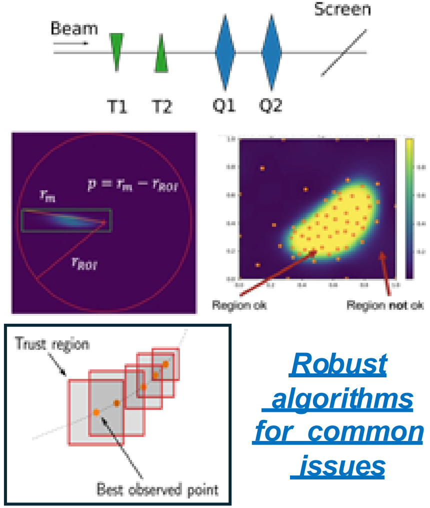


20x faster emittance tuning (BAX)



DOI 10.1088/2632-2153/ad169f

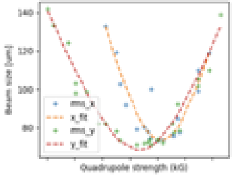
Automated beam alignment at AWA
- 20-30 minutes by hand
- 5 minutes with BAX



Robust algorithms for common issues

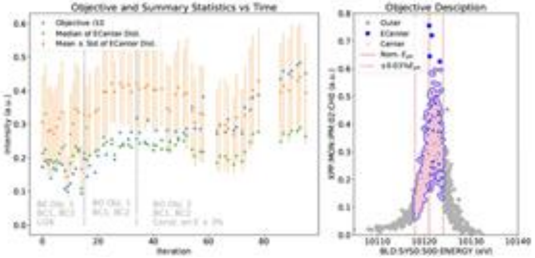
Roussel, Nat Commun **12**, 5612 (2021)

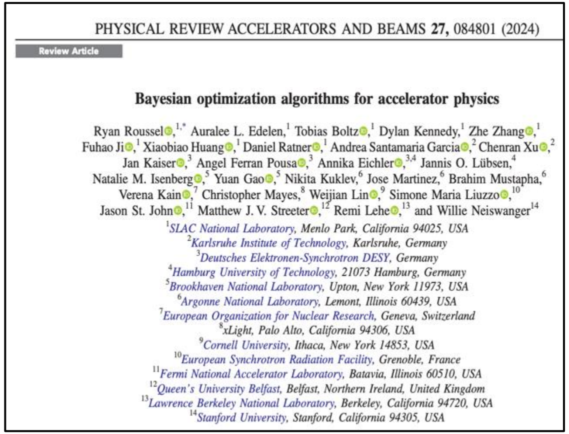
Automatic emittance measurement with BE



Roussel, Instruments **2023**, 7(3), 29

Tuning on monochrometer signal





PHYSICAL REVIEW ACCELERATORS AND BEAMS **27,** 084801 (2024)

Review Article

Bayesian optimization algorithms for accelerator physics

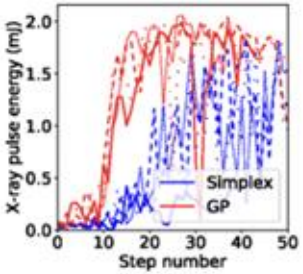
Ryan Roussel,[1,*] Auralee L. Edelen,[1] Tobias Boltz,[1] Dylan Kennedy,[1] Zhe Zhang,[1] Fuhao Ji,[1] Xiaobiao Huang,[1] Daniel Ratner,[1] Andrea Santamaria Garcia,[2] Chenran Xu,[2] Jan Kaiser,[3] Angel Ferran Pousa,[3] Annika Eichler,[3,4] Jannis O. Lübsen,[4] Natalie M. Isenberg,[5] Yuan Gao,[5] Nikita Kuklev,[6] Jose Martinez,[6] Brahim Mustapha,[6] Verena Kain,[7] Christopher Mayes,[8] Weijian Lin,[9] Simone Maria Liuzzo,[10] Jason St. John,[11] Matthew J. V. Streeter,[12] Remi Lehe,[13] and Willie Neiswanger[14]

[1]SLAC National Laboratory, Menlo Park, California 94025, USA
[2]Karlsruhe Institute of Technology, Karlsruhe, Germany
[3]Deutsches Elektronen-Synchrotron DESY, Germany
[4]Hamburg University of Technology, 21073 Hamburg, Germany
[5]Brookhaven National Laboratory, Upton, New York 11973, USA
[6]Argonne National Laboratory, Lemont, Illinois 60439, USA
[7]European Organization for Nuclear Research, Geneva, Switzerland
[8]xLight, Palo Alto, California 94306, USA
[9]Cornell University, Ithaca, New York 14853, USA
[10]European Synchrotron Radiation Facility, Grenoble, France
[11]Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA
[12]Queen's University Belfast, Belfast, Northern Ireland, United Kingdom
[13]Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA
[14]Stanford University, Stanford, California 94305, USA

includes methods that can work on > 100s of parameters (e.g. TRBO)

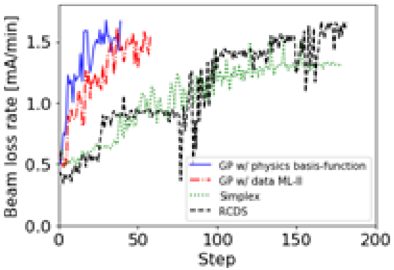
Bayesian Optimization with Advanced Variants is Mature and Operations-Ready

FEL pulse energy at LCLS



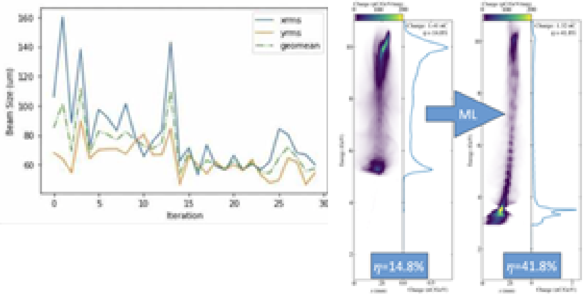
Duris et. al. PRL, **2020**

Loss rate at SPEAR3

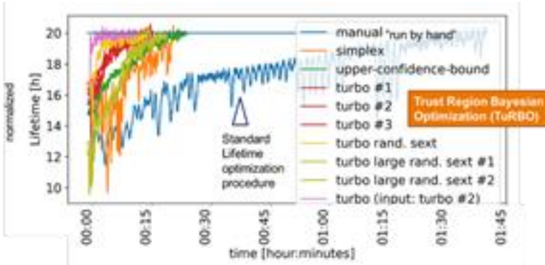


Hanuka et. al. PRAB, **2021**

Sextupole tuning at FACET-II
>2x efficiency of acceleration in plasma

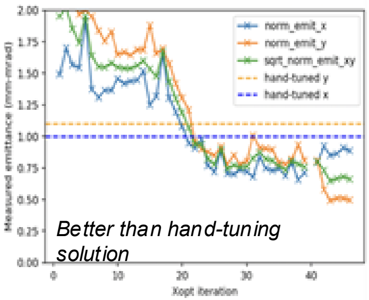


50x faster tuning and best observed lifetime at ESRF

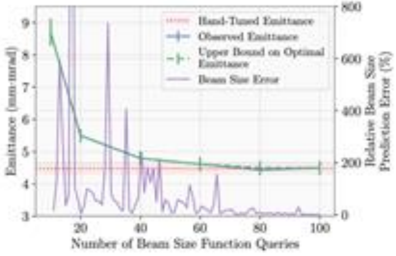


Liuzzo, ICALEPCS 2023 doi:10.18429/JACoW-ICALEPCS2023-MO3AO01

Emittance @ LCLS-II injector

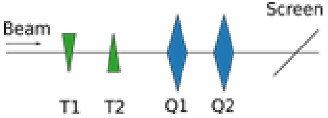


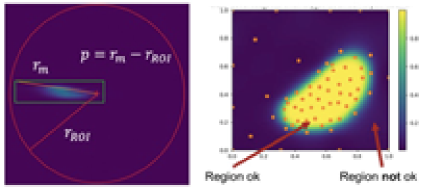
20x faster emittance tuning (BAX)

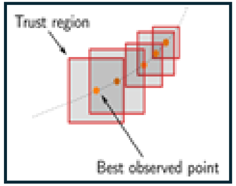


DOI 10.1088/2632-2153/ad169f

Automated beam alignment at AWA
- 20-30 minutes by hand
- 5 minutes with BAX



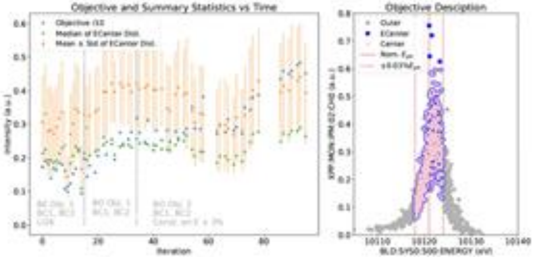




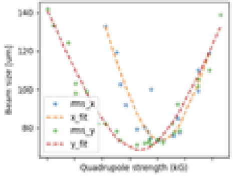
Robust algorithms for common issues

Roussel, Nat Commun **12**, 5612 (2021)

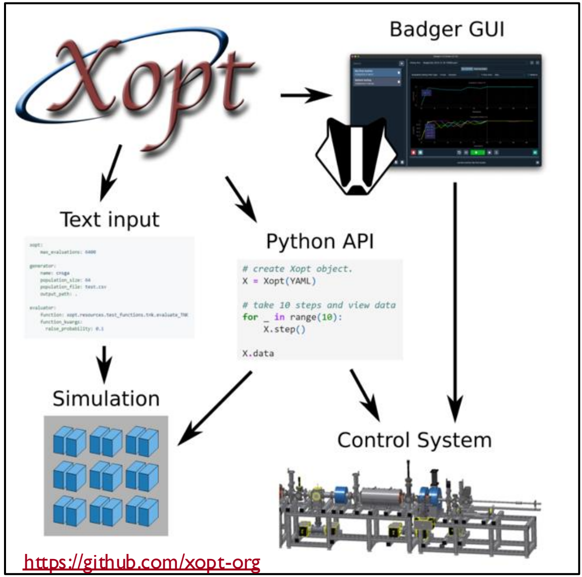
Tuning on monochrometer signal



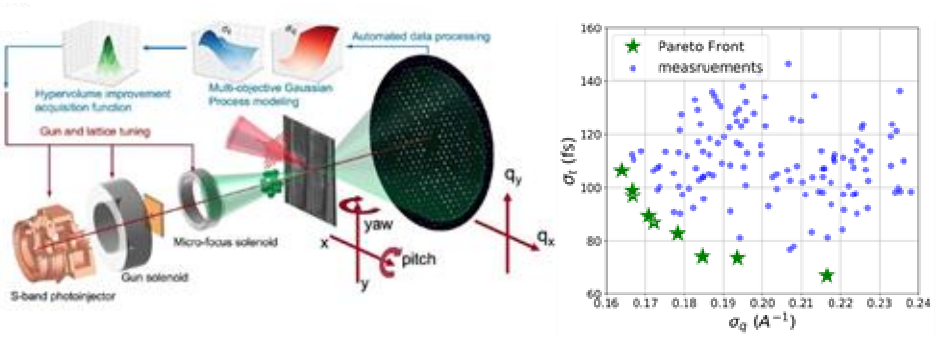
Automatic emittance measurement with BE



Roussel, Instruments **2023**, 7(3), 29

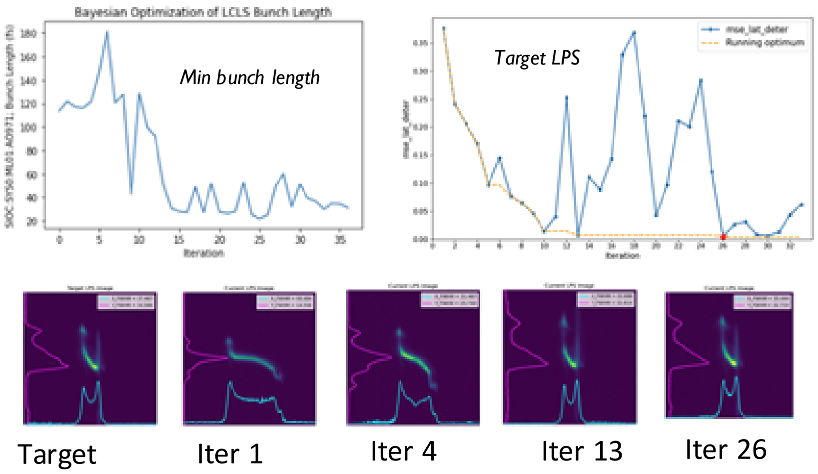


Multi-objective BO → experimental Pareto front

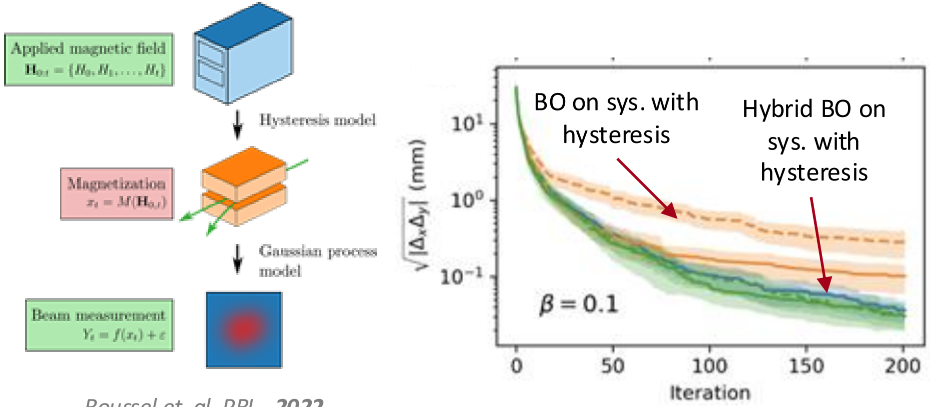


F. Ji, et al., Nat. Commun. 15, 4726 (2024)

Longitudinal phase space tuning

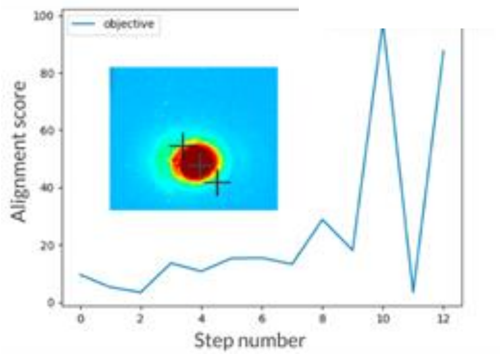


Hysteresis-aware magnet tuning



*Roussel et. al. PRL , **2022***

Photon beamline alignment at MFX

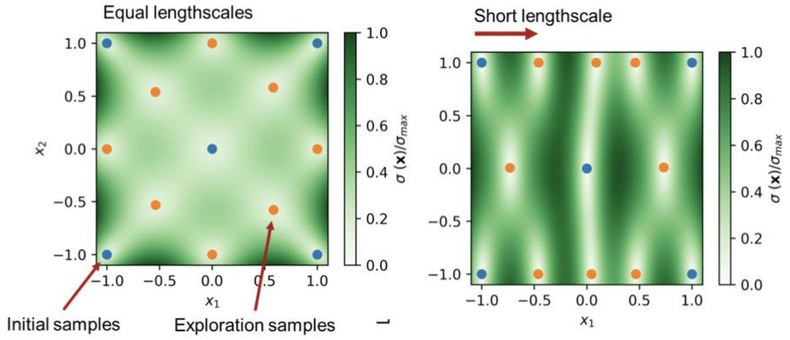


Efficient Characterization with Bayesian Exploration

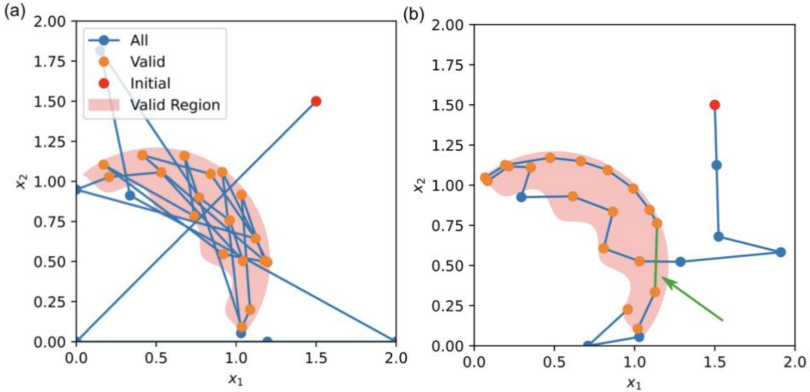
R. Roussel et. al.
Nat. Comm. **2021**

$$\alpha(\boldsymbol{x}) = \sigma(\boldsymbol{x}) \prod_{i=1}^{N} p_i(g_i(\boldsymbol{x}) \geq h_i)\, \Psi(\boldsymbol{x}, \boldsymbol{x_0})$$

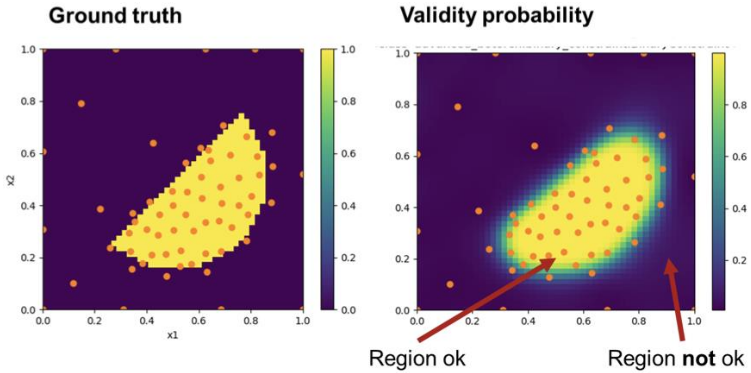
adaptive sampling



proximal biasing



learning constraints



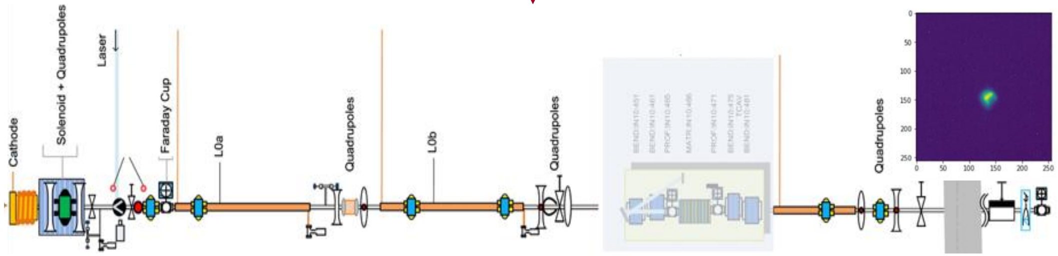
Enables sample-efficient characterization of high-dimensional spaces, while respecting both input and output constraints

Bayesian Exploration for Efficient Characterization

Automatic Exploration
(constrained to useful values of emittance and match)

Comprehensive ML Models of Injector

Setting changes on 10 variables (solenoid, bucking coil, corrector quads and matching quads)

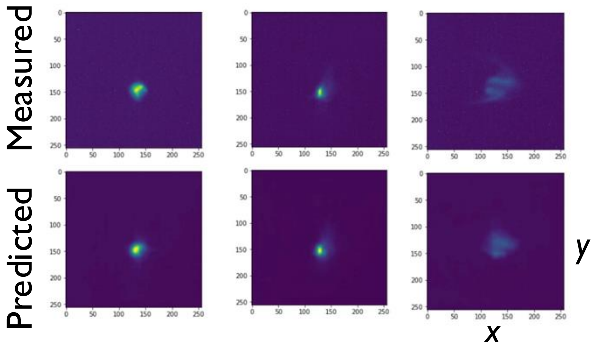


FACET-II Injector

x-y emit, match, and beam images

- Used Bayesian Exploration for efficient high-dimensional characterization (10 variables) of emittance and match at 700pC: **2 hrs for 10 variables compared to 5 hrs for 4 variables with N-D parameter scan (~8x faster)**
- Data was used to train neural network model of injector response predicting x-y beam images. GP ML model from exploration predicts emittance and match.
- **Example of integrated cycle between characterization, modeling, and optimization → now want to extend to larger system sections and new setups**

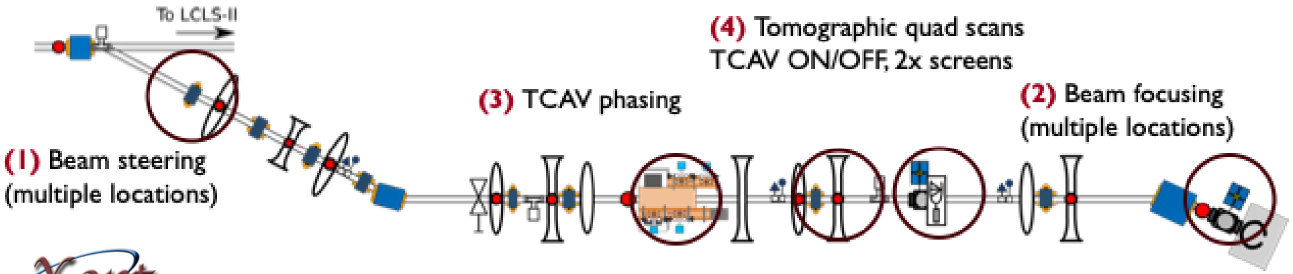
transverse phase space



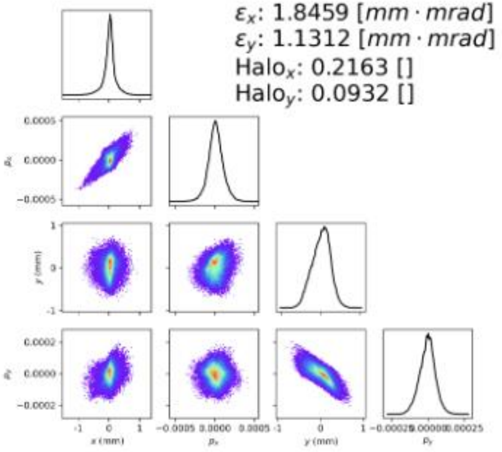
https://www.nature.com/articles/s41467-021-25757-3

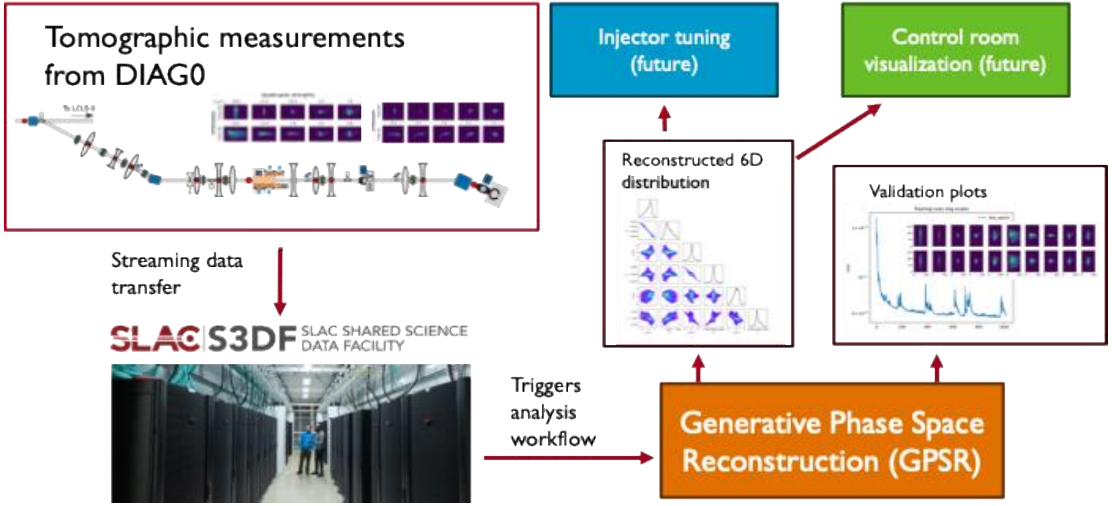
Use of Bayesian exploration to generate training data was sample-efficient, reduced burden of data cleaning, and resulted in a well-balanced distribution for the training data set over the input space. ML models were immediately useful for optimization.

Moving from individual tasks toward autonomous chains of them …



Xopt





Autonomous tuning and 6D phase space reconstruction → ~5 minutes online

Continuous estimate of 6D phase space

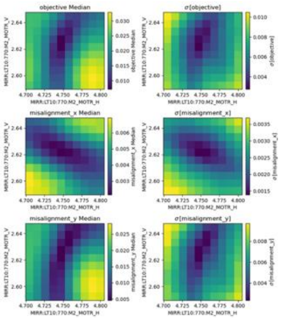
Started exploring tuning directly on reconstruction (core emittance + halo)

R. Roussel et al., https://arxiv.org/abs/2604.20125

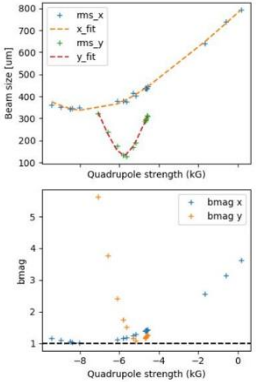
Example: FACET-II Startup Automation

Autonomous configuration of the FACET-II injector beamline w/o human intervention in < 14 min (standard is ~1hr +)

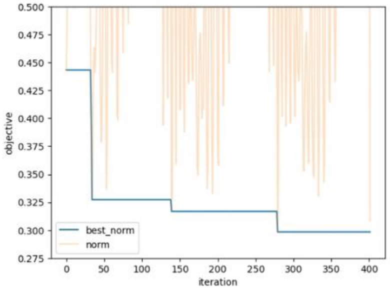
Laser alignment onto cathode (1 min)



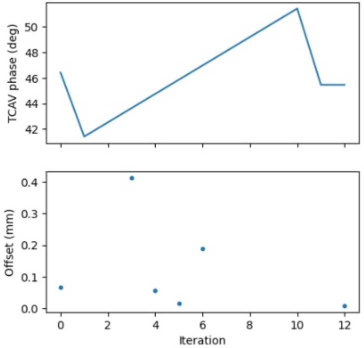
Emittance optimization + matching (7.5 min)



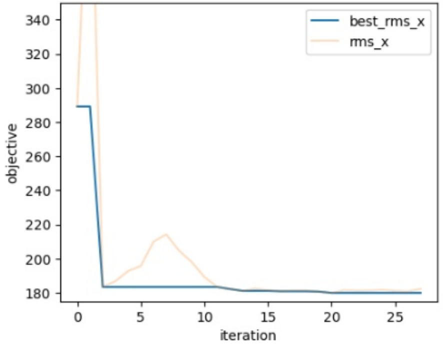
Beam steering (3 min)

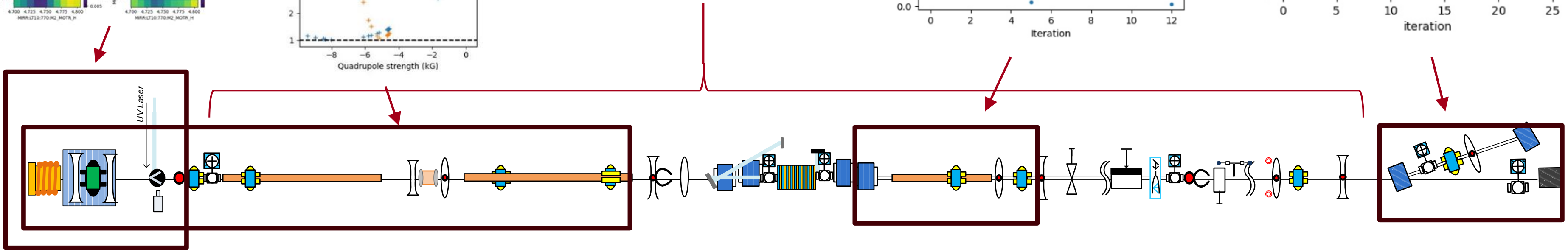


TCAV Phasing (7 sec)



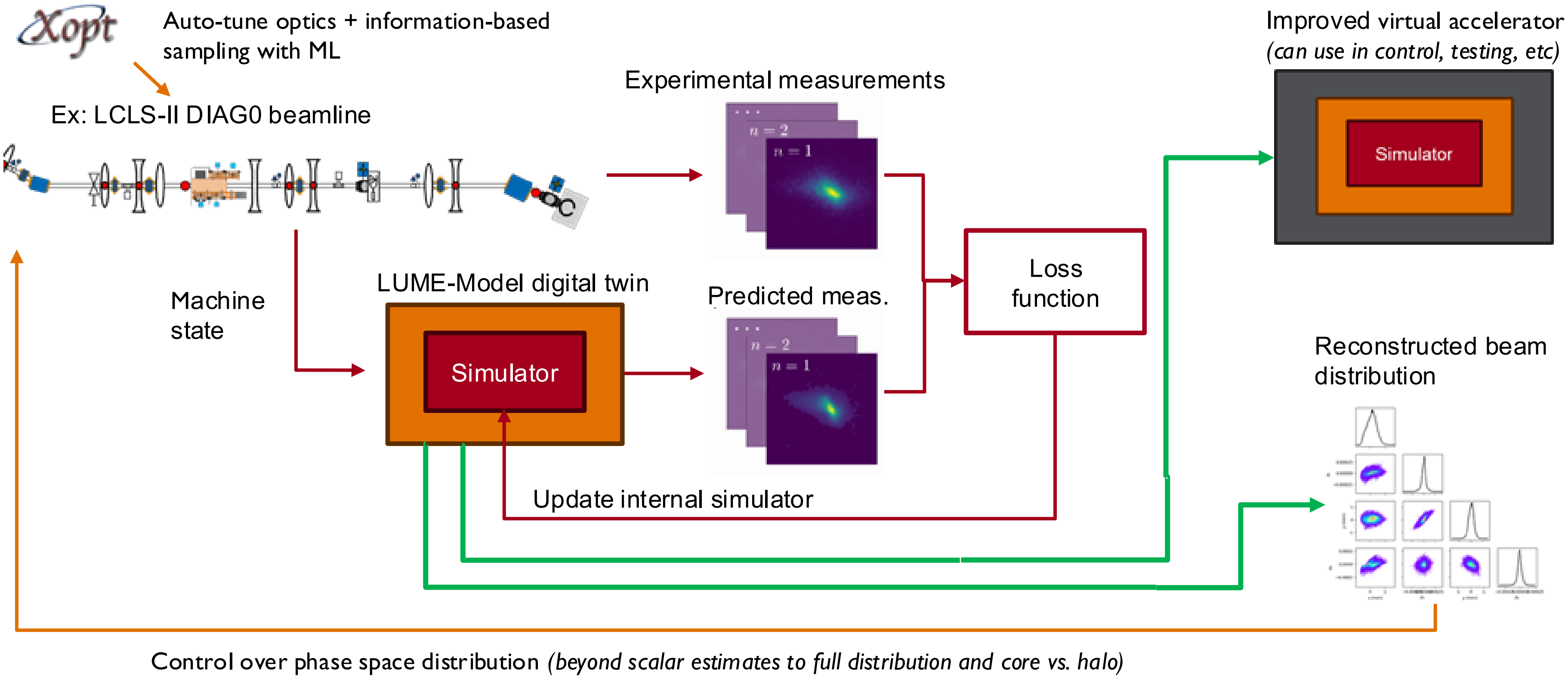
Energy spread minimization (2 min)





Example: real-time sim2real with differentiable physics → phase space reconstruction → phase space control

Physics- and ML-based **differentiable simulators** are being used inside digital twins to perform model calibration, phase space reconstruction, and in turn aid advanced phase space shaping

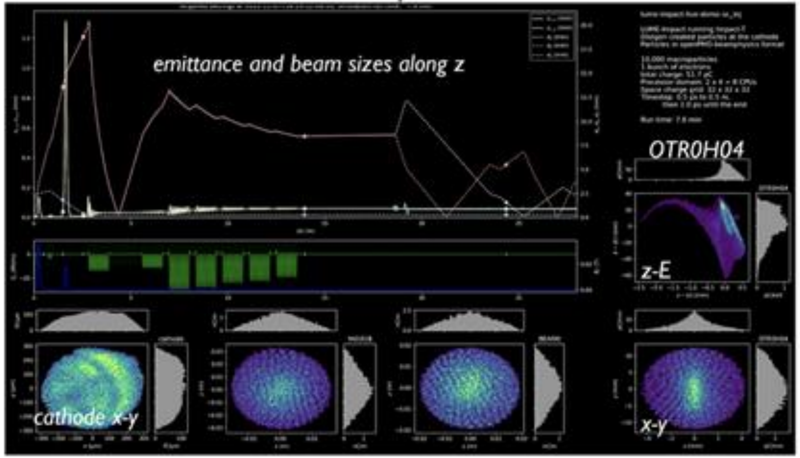


System Modeling: Warm Starts from Detailed Online Physics Models

Used combination of online physics simulation and Bayesian optimization algorithms to aid LCLS-II injector commissioning

Readings from machine via EPICS

injector settings, laser profile from VCC image

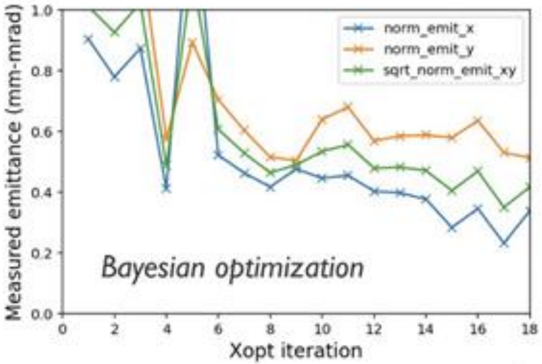


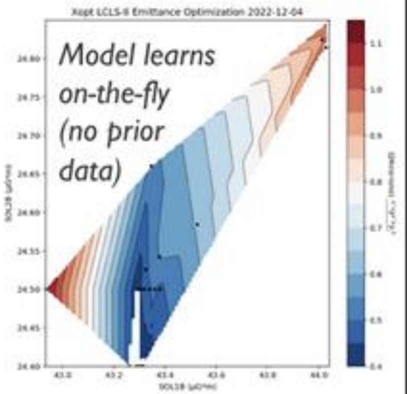
LCLS-II live sim: run on HPC and display in control room

Updates every 3-8 mins, space charge included, uses LUME-IMPACT

Adjust settings / ranges with insight from predictions

Hand over to ML-based optimization for fine tuning





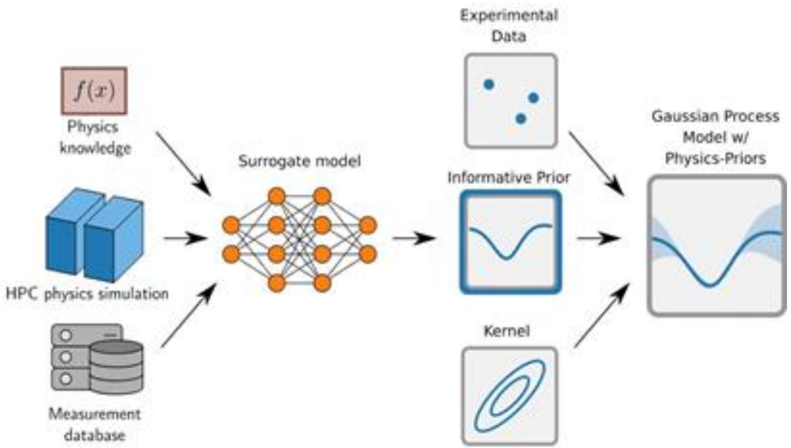
Best emittance then obtained during LCLS-II injector commissioning

High performance computing integration enables new level of detail to be predicted online → increasingly common

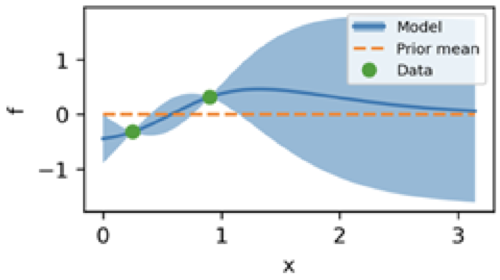
But still need improved prediction speed and accuracy!

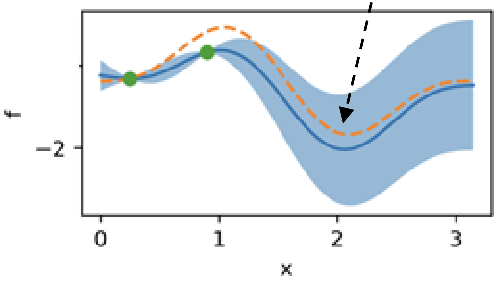
Incorporating System Models into Bayesian Optimization

Important for scaling up to higher-dimensional tuning problems



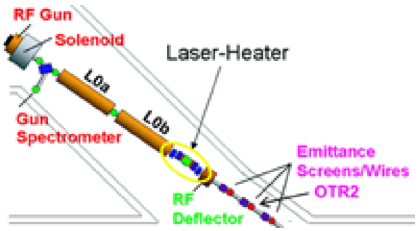
model prediction returns to prior





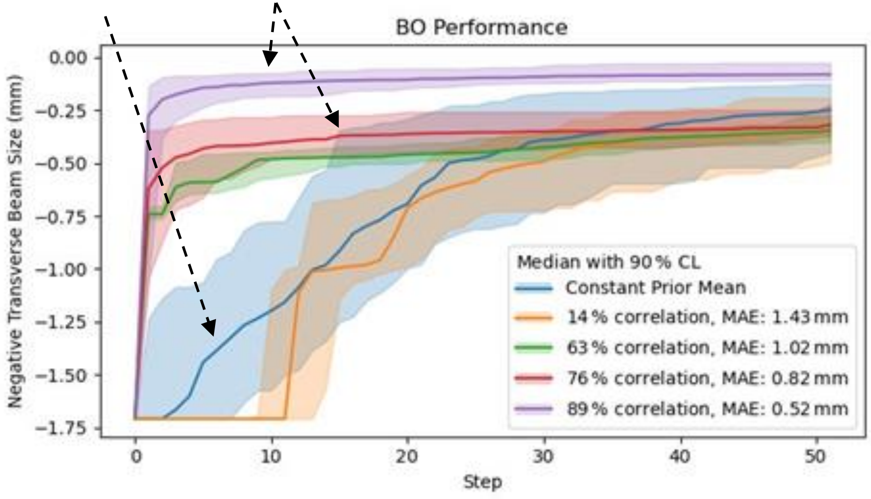
Neural network system model to provide a prior mean for a GP

Tested on LCLS injector
variables: *solenoid, 2 corrector quads, 6 matching quads*
objectives: *beam size on OTR2*



regular BO

prior means with different fidelity

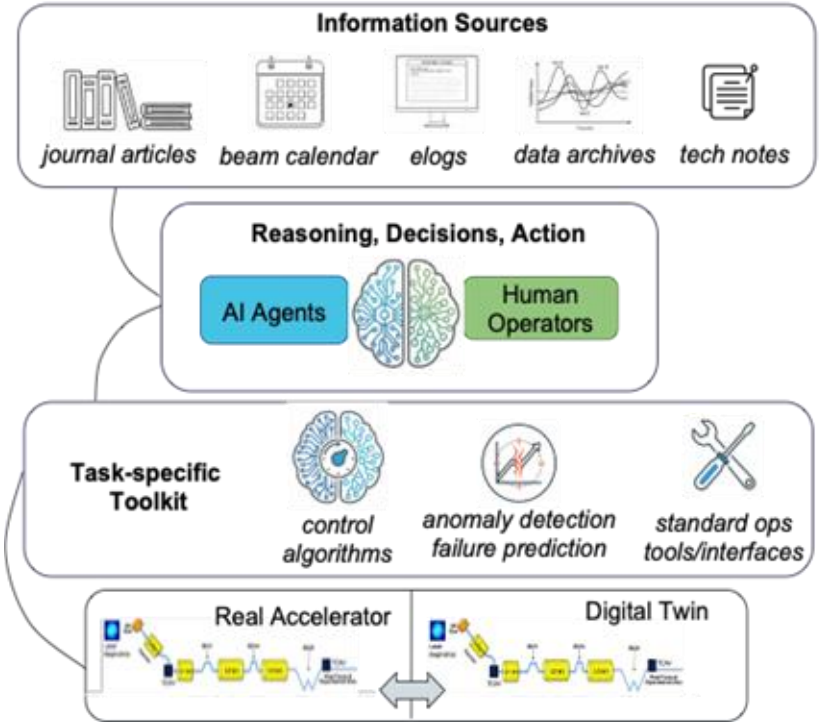


Even prior mean models with substantial inaccuracies provide a boost in initial convergence

C. Xu et al, NeurIPS (2022) T. Boltz, J. Martinez, et al., SciRep (2025)

BO Review: R. Roussel et al., PRAB (2024)

Agentic AI

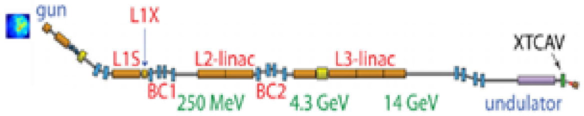


Osprey demo @LCLS

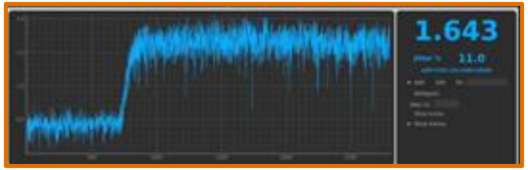
New optimization routine

Badger run history

natural language description of tuning task



Automatic execution on machine via Badger tuning tool



Rapid improvement in FEL performance

Prompt to completion: ~ 10 minutes

Digital Twins / Virtual Accelerators

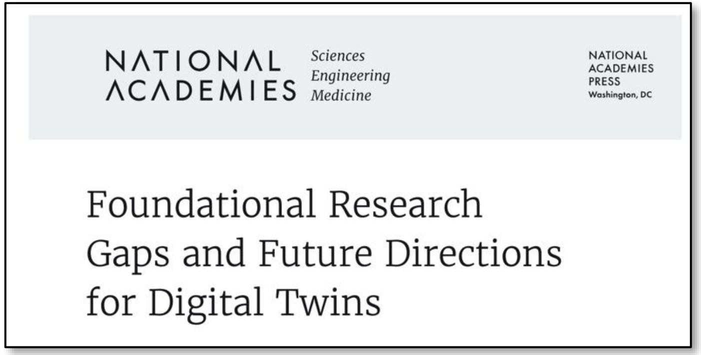
Essential characteristics of a digital twin:

- System model that tracks the physical system as it changes ("online model")
- Ideally is adaptively calibrated to match measurements
- Informs decision making in the physical system

"bi-directional information flow"

- Long history in accelerators of using simple system models online and in control → *new level of fidelity, speed, and adaptability possible with HPC and AIML*

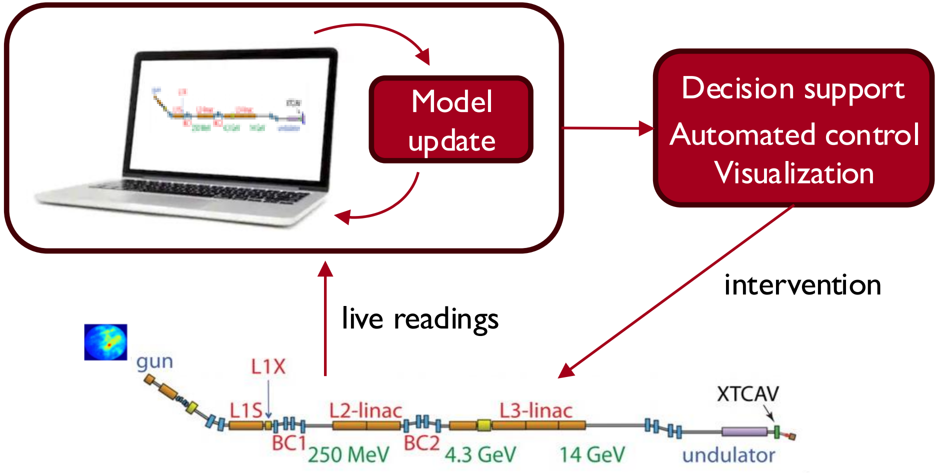
Virtual accelerator – available for interactive use by other programs and need not interact with physical system (useful in design)

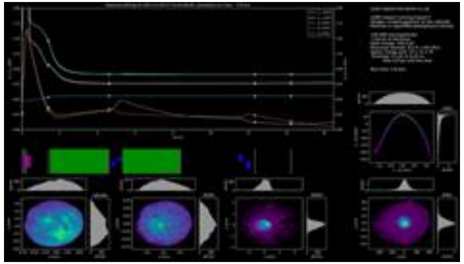

NATIONAL ACADEMIES Sciences Engineering Medicine

NATIONAL ACADEMIES PRESS Washington, DC

Foundational Research Gaps and Future Directions for Digital Twins

https://nap.nationalacademies.org/catalog/26894/foundational-research-gaps-and-future-directions-for-digital-twins





Included emulated components can span many sub-systems and processes

Typical Control Room Apps

Additional Visualization

High Level Output (e.g. EPICS PV)

Physics Process Emulators (phys. sims, ML)

beam dynamics, cooling systems, user beamlines, etc
range of fidelities and comprehensiveness

Wishlist also includes:

- Model adaptation
- Complex orchestration *(fidelity / choice of sims)*
- Interactivity of visualization
- Virtual accelerator mode and streaming toggle
- Data archiving

Low-level Controls

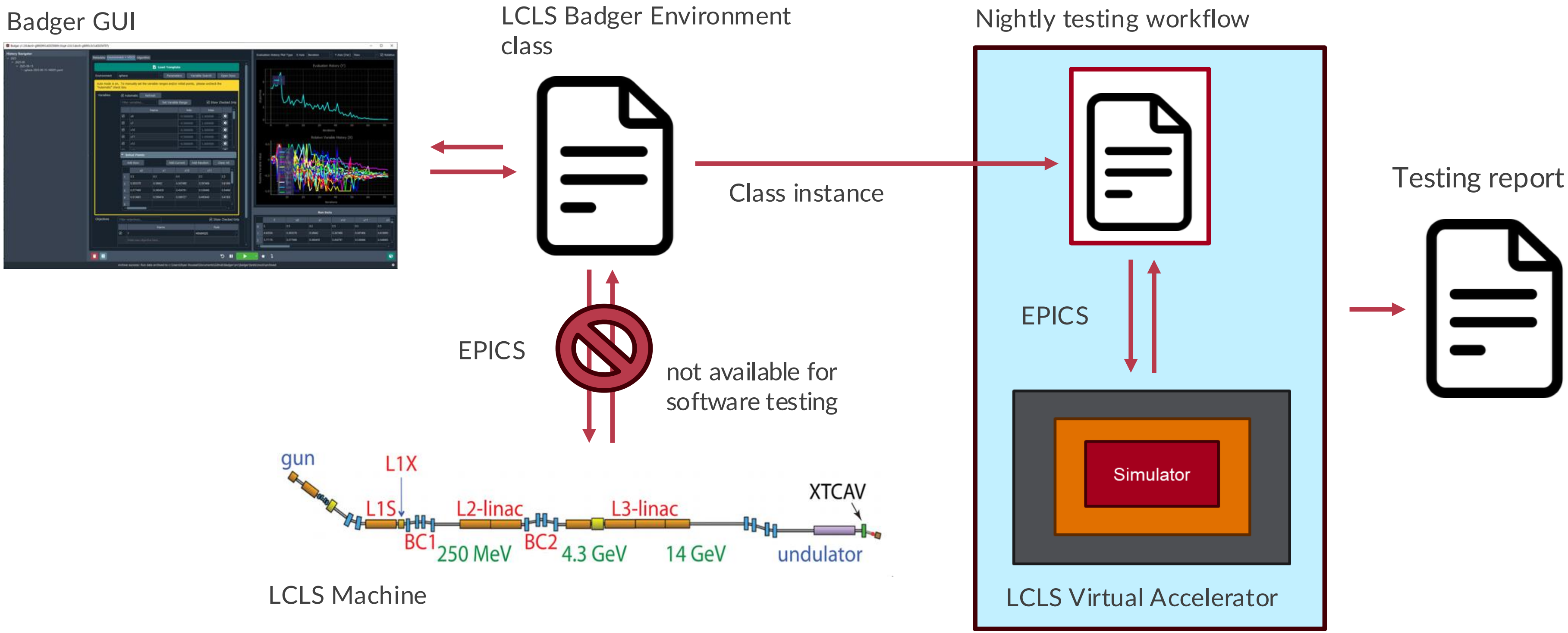
e.g. PID loops on magnet settings

Hardware I/O

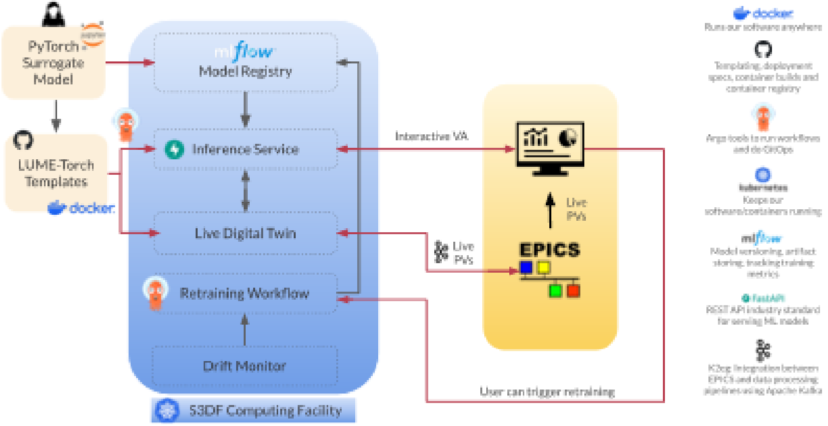
Many accelerator facilities address part but not all of this stack at present

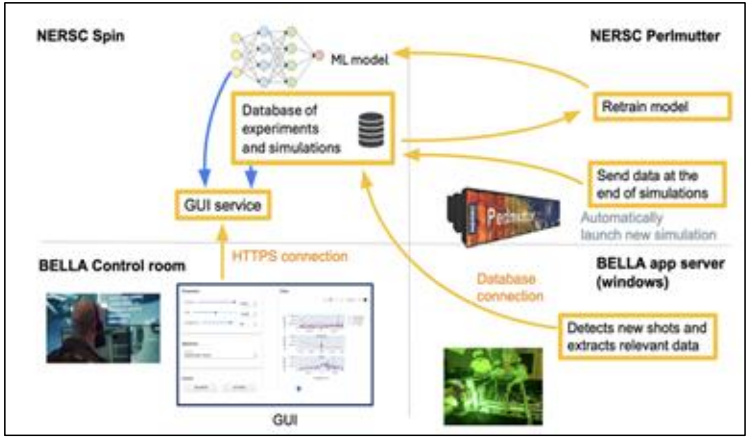
Example: Badger Environment Testing / Continuous Integration

Use VA to perform nightly testing / validation of code used to interact with the machine

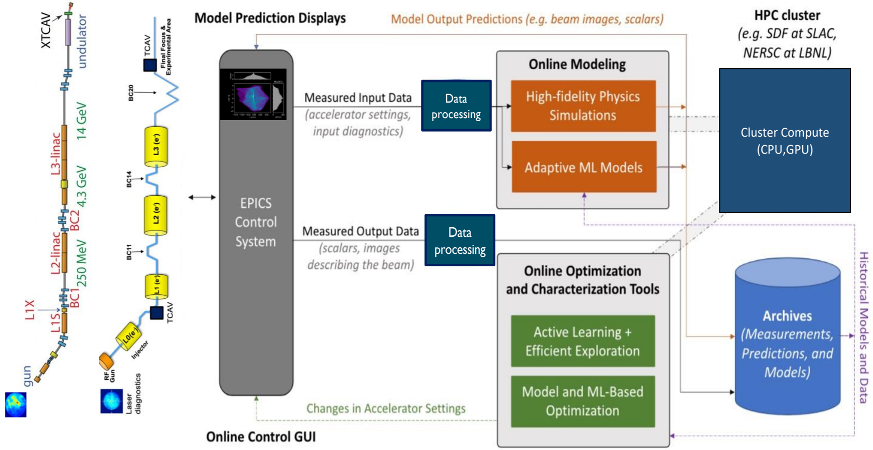


continuous integration / deployment and live model updating and retraining

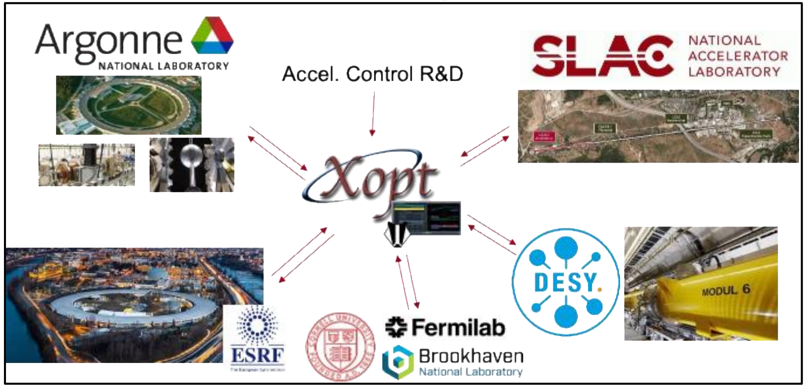




local facility infrastructure



modular/transferrable tools



Emerging Opportunities in AI/ML Assisted Design

- **Large-scale end-to-end co-design to improve achievable performance and maximize scientific output**
  - Combine subsystems *(e.g. cryogenic system, cavity design, lattice design, high level controls, detector design)*
  - Handle more sophisticated constraint tradeoffs: beam performance, construction costs / footprint, power consumption *(e.g. lower RF overhead of amplifiers through better control → lower cost)*
  - Combine with detector optimization and analysis → optimize end science + facility directly

- **Agentic workflows**
  - Streamline organization across teams (*e.g. flag design collisions based on documents in quasi-real-time)*
  - Orchestrate complex simulation / search campaigns *(e.g. co-design)*

- **AI/ML algorithms and agentic workflows for efficient design space search**
  - Numerous available, increasingly can also leverage agentic workflows
  - e.g. DAMA optimization with BAX for storage rings (*400x fewer simulations to reach equivalent solution vs. NSGA-II, 80% improvement in objective when agent-driven orchestration on previously intractable problem*)
- **AI/ML accelerated physics simulations** *(e.g. granular surrogates for complex effects)*

Design is a less-explored space than online optimization / control → potentially very large gains
Clear place where AI/ML research and MuC needs can intersect

Conclusions

- **Numerous mature tools available for commissioning**
  - Bayesian optimization and variants and ideal for commissioning / characterization (e.g. see Xopt)

- **Timing is good for emerging opportunities in AI/ML for design**
  - Mature algorithms for efficient search available (with orders of magnitude speedup)
  - Agent-driven workflows can yield performance improvement on complex problems
  - HPC + AIML enables breadth of co-design space to be addressed (emerging area)
  - Agent-assisted project management / organization of heterogeneous teams (e.g. ID design collisions) is an un-tested but interesting prospect
  - Many opportunities for novel AIML R&D coupled to MuC needs

- **Virtual accelerators are extremely useful for full stack software testing and R&D**
  - Include measurement devices/analysis and advanced control algorithms alongside physics
  - Can be converted into a digital twin once accelerator is running

- **Think about systems integration and infrastructure early**
  - Biggest barrier for AI/ML in practice is the surrounding software/hardware environment, <u>not</u> techniques

Thanks to many contributors to the work shown (not exhaustive)