Fermilab

September 14, 2026

AI for Accelerators: Current Challenges and Future Directions

Katsuya Yonehara
CLARIPHY Collaboration Meeting

U.S. DEPARTMENT of ENERGY
Fermi National Accelerator Laboratory is managed by FermiForward for the U.S. Department of Energy Office of Science

FERMI NATIONAL ACCELERATOR LABORATORY

Where AI can contribute to accelerator systems

- AI can extract useful information from large, heterogeneous datasets and support complex accelerator decisions
- Applications span accelerator operations, beam diagnostics, and component design

Operations and reliability

- Beam tuning and operating-point optimization
- Anomaly detection and equipment protection
- Predictive maintenance and fault diagnosis

Measurement and design

- Beam diagnostics and virtual instruments
- Phase-space reconstruction from distributed measurements
- Accelerator component and target design

Most accelerator AI applications use local measurements to optimize an individual machine or subsystem.

Unique focus of this talk: Integrate accelerator instrumentation with measurements from physics detectors to optimize beam delivery and experimental performance

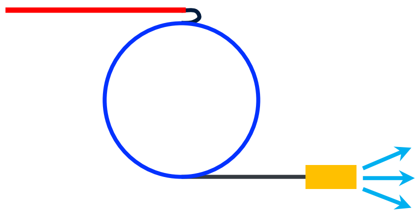
FERMI NATIONAL ACCELERATOR LABORATORY

Experimental Configuration Defines Accelerator AI Priorities

Optimal accelerator performance depends on how the experiment uses the beam

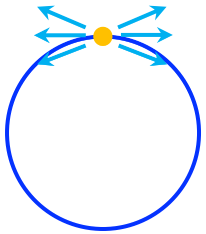
Fixed-target experiments: extracted single-pass beam

- The experiment receives each extracted pulse only once
- Beam conditions must be reconstructed rapidly from distributed measurements
- Tuning decisions must occur between pulses or spills
- Experimental performance depends on the target and downstream detector response



Circular colliders: circulating beams

- Beams circulate and collide over many turns
- Repeated measurements enable continuous feedback
- Phase-space errors affect stability, luminosity, and collision-energy resolution
- AI supports orbit correction, phase-space control, and long-term optimization



Accelerator AI should optimize experimental performance by connecting accelerator, target, and detector measurements

FERMI NATIONAL ACCELERATOR LABORATORY

Current Challenges and Future Directions

Current challenges

- Sparse fault data and incomplete labels
- Different data rates, timing systems, formats, and calibrations
- Synchronizing accelerator, target, and physics-detector measurements
- Changing machine configurations and operating conditions
- Requirements for reliability, interpretability, and response time
- Secure access across controls and machine-protection networks

Future directions

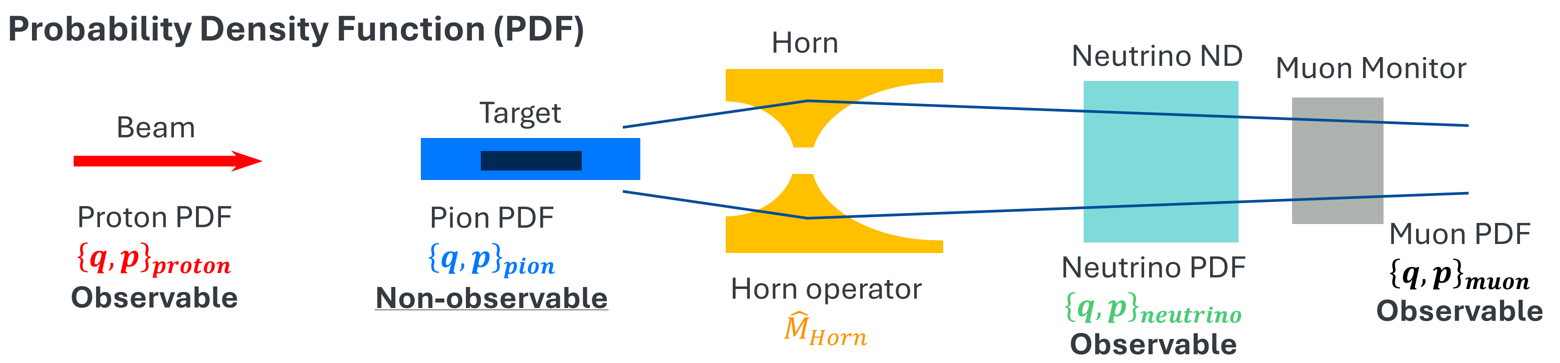
- Physics-informed models and accelerator emulators
- Common timing and metadata across accelerator instrumentation and physics-detectors
- Uncertainty-aware operator decision support
- Common data standards and reusable benchmark datasets
- Staged deployment with continuous validation and human oversight

These needs also align with the Genesis Mission

Operational value requires trustworthy predictions under changing machine conditions

FERMI NATIONAL ACCELERATOR LABORATORY

Experiment-Integrated Digital Ecosystem for LBNF



$$\binom{q}{p}_{muon} \oplus \binom{q}{p}_{neutrino} = \hat{M}_{Horn} \cdot \binom{q}{p}_{pion}$$
$$= \hat{M}_{Horn} \cdot \hat{M}_{Hadron\ prod.} \cdot \binom{q}{p}_{proton}$$

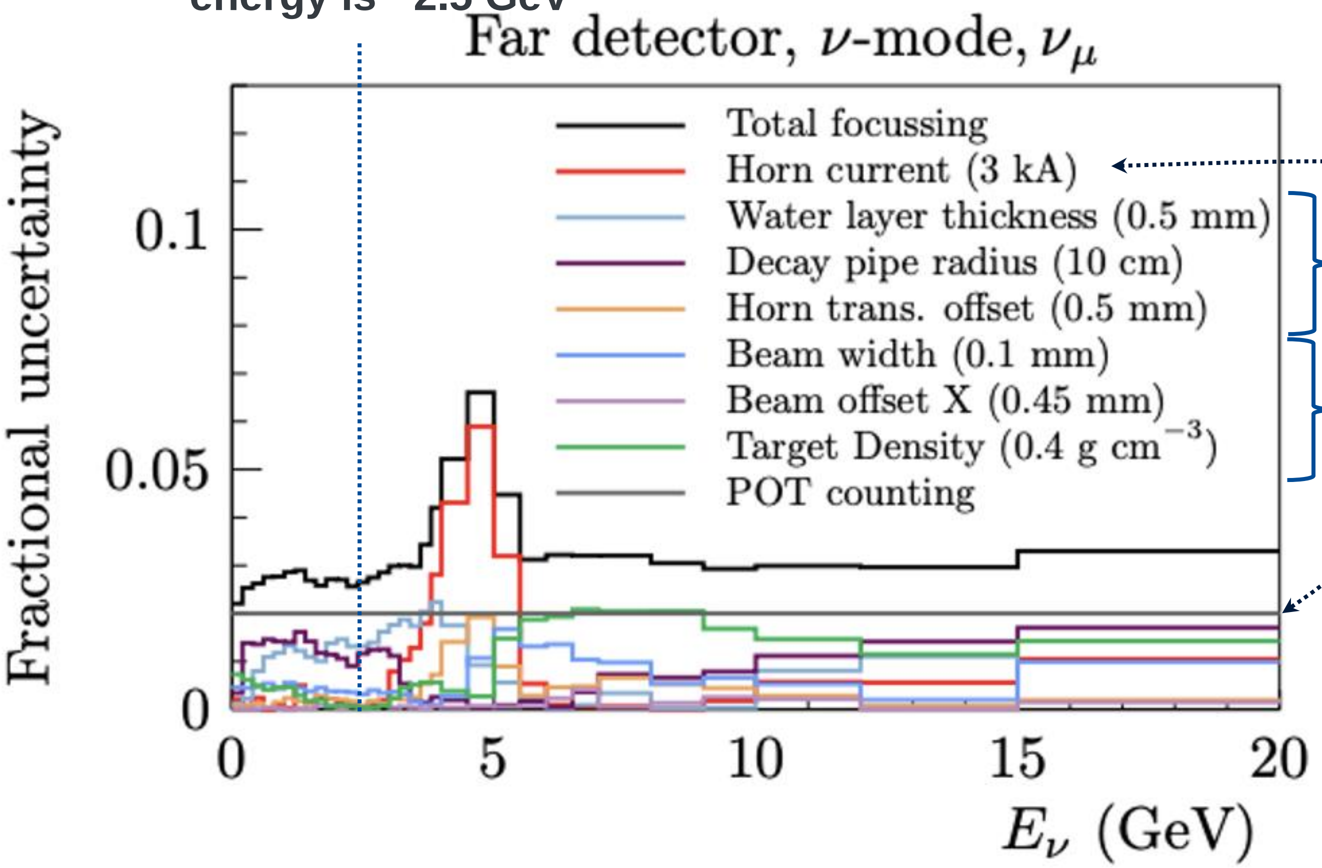
- Surrogate model emulates magnetic-horn focusing
- Acquire bunch-resolved accelerator data
  - Synchronize fast beam instrumentation, such as Wall Current Transformer (WCT), Beam Position Monitor (BPM), and Beam Loss Monitor (BLM) with beam and machine settings
- Use AI-assisted muon monitors (e.g. MuMS) to infer the unobserved pion phase space
- Integrate measurements from the SAND ND, including its liquid-argon volume, into the accelerator data ecosystem

FERMI NATIONAL ACCELERATOR LABORATORY

Beam-related Flux Uncertainty and Measurement Goals

An experiment-integrated digital ecosystem could reduce beam-related uncertainty for LBNF/DUNE

Peak on-axis neutrino energy is ~2.5 GeV



Protons on Target

- Improve uncertainty from 2 to 0.1%
- Introduce new current transformer

Horn current

- Improve uncertainty from 3 to 0.1 kA
- Already demonstrated at NuMI

Predict with emulator and validate with survey & beam tomography

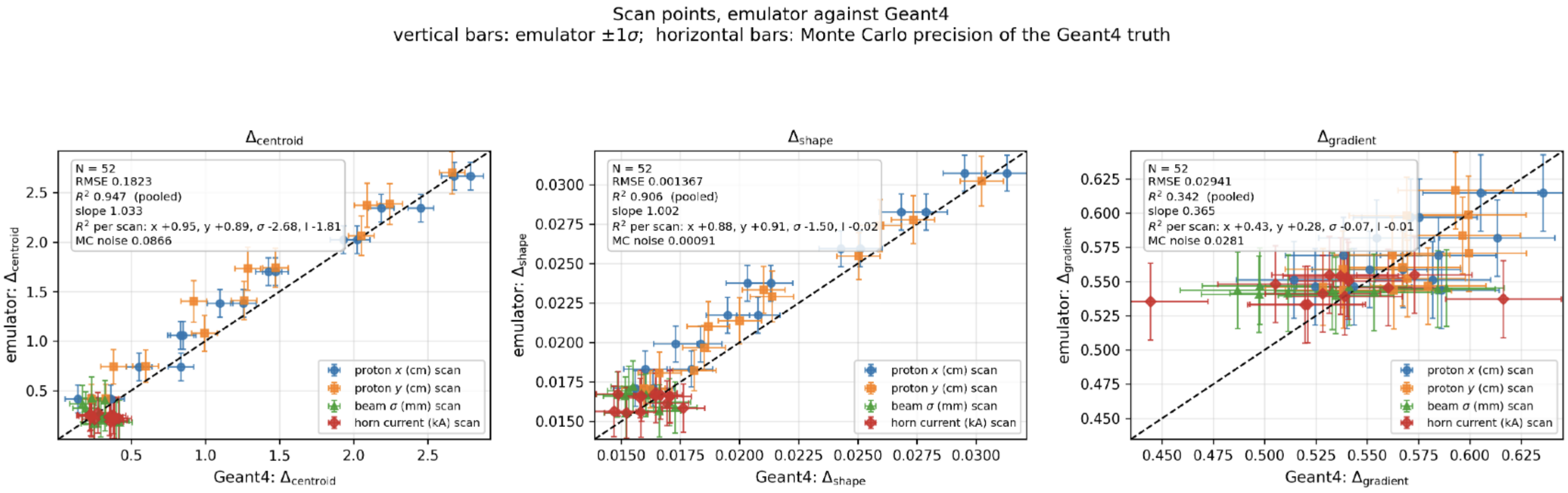
Beam and Target conditions

- Goal beam width < 0.1 mm
- Goal beam position offset ~ 0.02 mm
- Goal target Density < 5 %

FERMI NATIONAL ACCELERATOR LABORATORY

Bayesian Exploration with a Physics Informed Emulator for LBNF

Curtesy of S. Ganguly



- Plots demonstrate that the physics informed emulator approximates high-fidelity Geant4 responses across beamline parameter scans
- Fast inference enables Bayesian exploration and feedback for beam tuning

FERMI NATIONAL ACCELERATOR LABORATORY

Advancing AI for High-Power-Target Design

AI replaces routine evaluations inside a validated domain; selected and high-risk cases return to physics

DEFINE → GENERATE → PREDICT & OPTIMIZE → GATE → VALIDATE

01
Define
Geometry · material
beam · cooling

02
Generate
Parametric concepts
feasible candidates

03
AI predict
Validated surrogates
+ uncertainty

04
Screen
Optimize · constraints
UQ · novelty · risk

05
Validate
Selected high-fidelity
simulation · experiment

validated evidence retrains surrogates and expands their domain

What the model explores (Define + Generate)

Geometry, material, beam parameters, and cooling configuration

What the model predicts (AI predict)

Particle yield, temperature, stress, fatigue, radiation damage, and physics uncertainty

What makes a candidate viable (Optimize / Gate)

OBJECTIVE

Maximize physics performance and service lifetime

CONSTRAINTS

Temperature · stress · cooling capacity · manufacturability · safety

AI/ML Test Facility: Validating Predictions and Calibrating Uncertainty

Controlled beam tests establish AI performance under realistic operating conditions

Why a test facility is needed

- AI development requires frequent, short beam-access periods
- Controlled scans and off-normal conditions exceed the risk tolerance of most user operations
- Upstream beam instruments and downstream detectors require a common time base and shared metadata
- Real hardware and controls are necessary to close the simulation-to-reality gap
- AI-ready benchmark datasets and common interfaces support reproducibility and transfer to other facilities

What the test facility must evaluate

- Prediction error against independent accelerator and detector measurements
- Calibration and coverage of prediction intervals
- Robustness across operating conditions, machine drift, and faults
- Detection of extrapolation and safe fallback outside the validated domain
- Improvement in experimental performance and tuning or recovery time

Realistic beam tests measure prediction error, calibrate uncertainty, and provide operational evidence for safe AI deployment

Summary

- The experimental configuration defines the accelerator AI priority: rapid beam-state reconstruction for single-pass experiments and continuous phase-space control for circular colliders
- Our approach connects accelerator instrumentation, physics models, target systems, and HEP detectors to optimize experimental performance
- The LBNF case study combines synchronized measurements, Bayesian exploration, and an emulator to constrain beam-related neutrino-flux uncertainty
- Target AI applies validated surrogate models and uncertainty quantification to accelerator-component design
- An AI/ML Test Facility provides flexible beam access, integrated AI-ready data, and controlled experiments to validate predictions, calibrate uncertainty, and establish evidence for safe deployment

FERMI NATIONAL ACCELERATOR LABORATORY