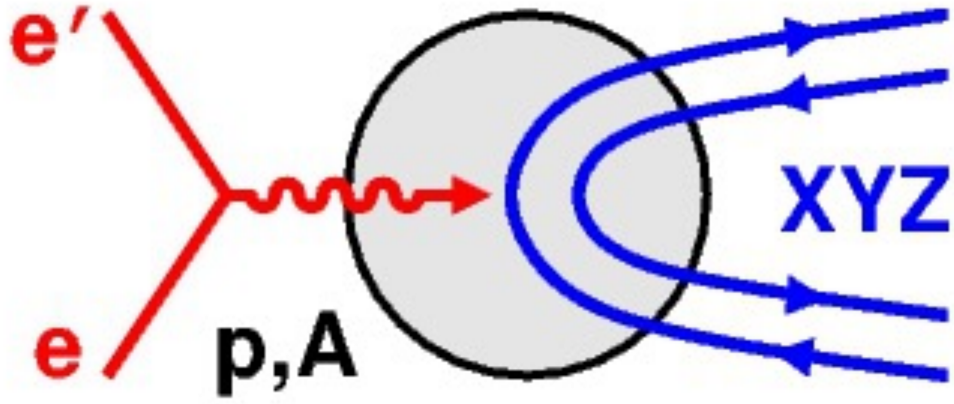


Fast Machine Learning Trigger for Heavy Flavor Measurements at the EIC

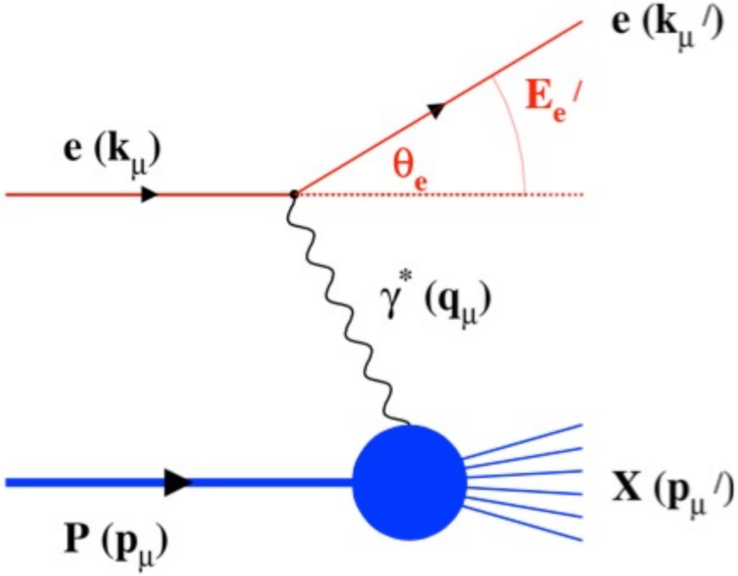
Zhaozhong Shi

Los Alamos National Laboratory
Email: zhaozhongshi@lanl.gov

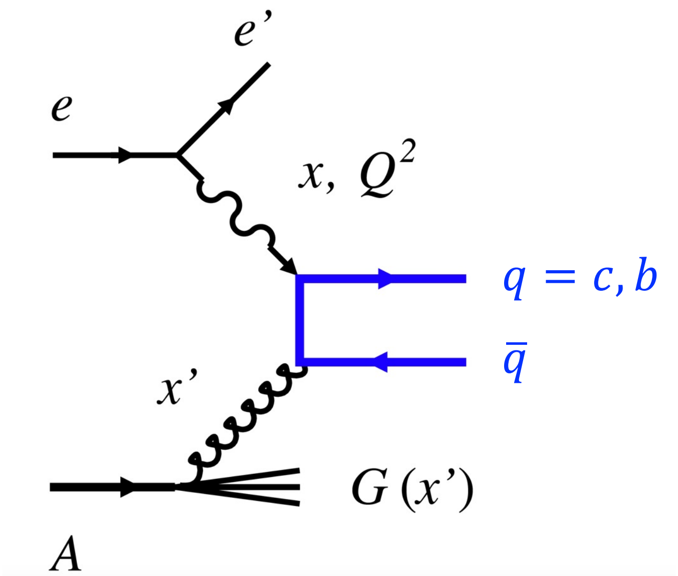
08/18/2022

Heavy Quark Production at EIC

Inclusive Deep Inelastic Scattering



Photon–Gluon Fusion



Probe the General Structure of the Proton

$$\frac{d^2\sigma}{dxdy} = \frac{4\pi\alpha^2 s}{Q^4}\left[xy^2F_1(x,Q^2) + \left(1 - y - xy\frac{M^2}{s}\right)F_2(x,Q^2)\right]$$

Dipole Nucleus Cross Section

$$\sigma_A^{q\bar{q}}(x,r^2) \sim \frac{\pi^2}{3}\alpha_S r^2 x G_A(x, 10/r^2)$$

- At the EIC, for ep at $\sqrt{s} = 100$ GeV: $\sigma_{c\bar{c}}^{Inc} \sim 100\ nb$ and $\sigma_{b\bar{b}}^{Inc} \sim 10\ nb$
- The luminosity at the EIC is $\sim 10\ fb^{-1}$/year
- Multiplying to the fragmentation function for exclusive heavy flavor spectra

Exotic Quarkonia Program at the EIC

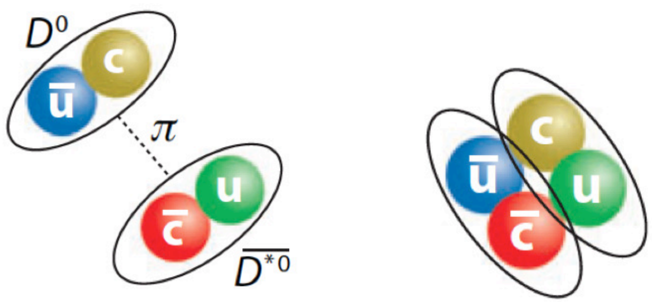
Recent Study of X(3872) by LHCb in pp, pPb and by CMS in PbPb

Citation: R.L. Workman *et al.* (Particle Data Group), Prog.Theor.Exp.Phys. **2022**, 083C01 (2022)

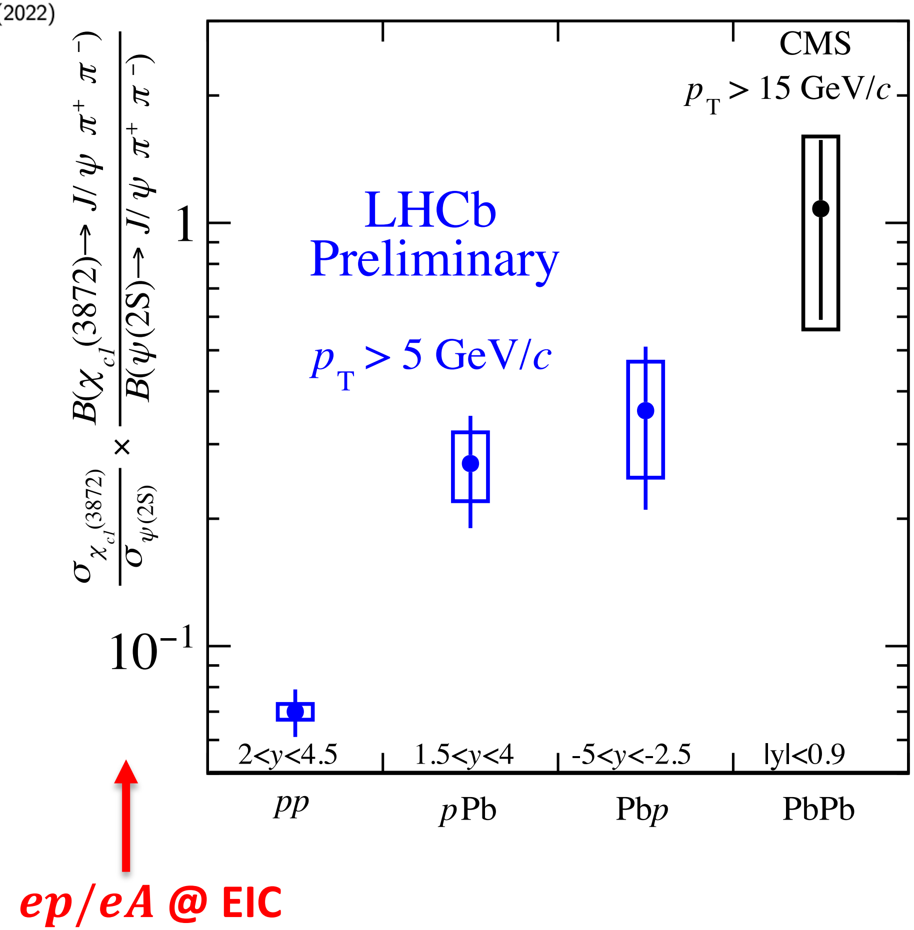
$\chi_{c1}(3872)$

$I^G(J^{PC}) = 0^+(1^{++})$

also known as $X(3872)$



D^0–$\overline{D^{*0}}$ "molecule" **Diquark–diantiquark**



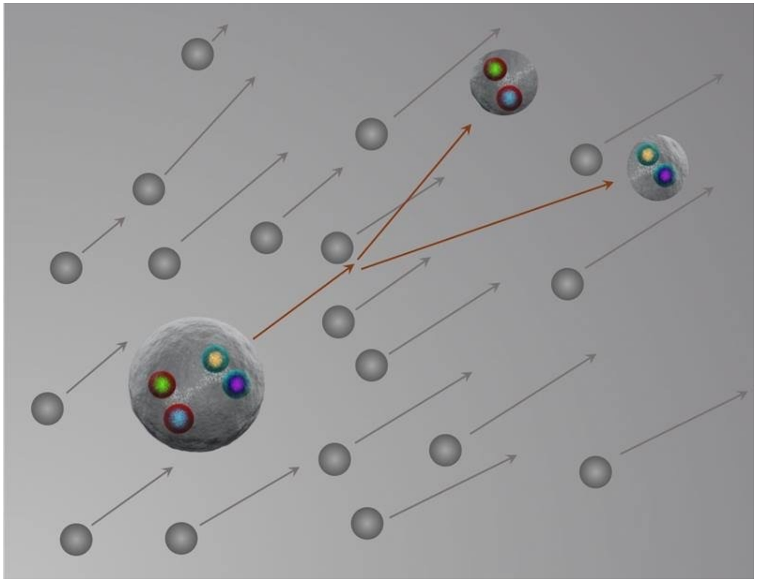
<u>Opportunities at the EIC</u>

- Perform exotic hadron spectroscopy
- Decipher the structure of X(3872)
- Understand heavy quark hadronization to exotic quarknonia

Example: X(3872) Measurement

Analysis Challenges

- 4-prong decay with an intermediate J/ψ state: $X(3872) \rightarrow J/\psi\pi^+\pi^- \rightarrow l^+l^-\pi^+\pi^-$
- Small production cross section
  - Prompt X(3872): 2.6 × 12% pb at ep for $\sqrt{s}$ =100 GeV (Xiaojun Yao talk in 2018)
  - Non-Prompt X(3872) from B decay: $BR[B \rightarrow X(3872)] \sim 10^{-5}$
- Short decay lifetime: $c\tau < 1$ pm
- Small acceptance
- Track reconstruction efficiency
- Combinatorial background

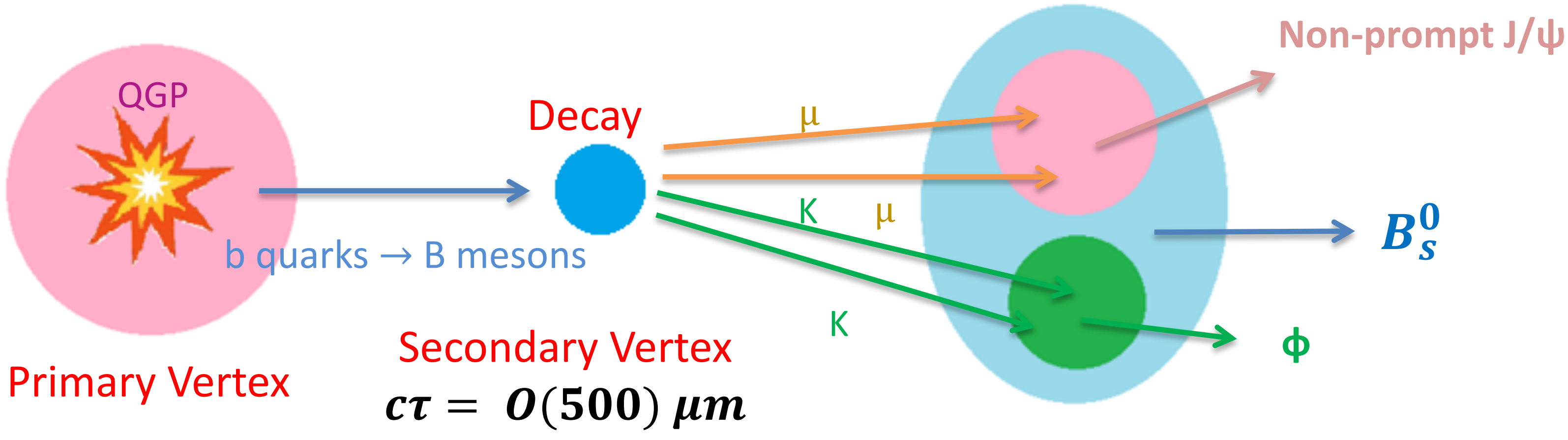


Experimental Setup and Analysis

- DAQ: continuous streaming readout
- Online triggers: heavy flavor events, particularly with displaced muon vertex
- Tracking: state-of-the-art tracking detector
  - High granularity forward silicon vertex/tracking detector with excellent spatial and timing resolution (LANL LDRD project by Dr. Xuan Li)
- PID: broad momentum range
- Machine learning techniques: online triggering and offline analysis

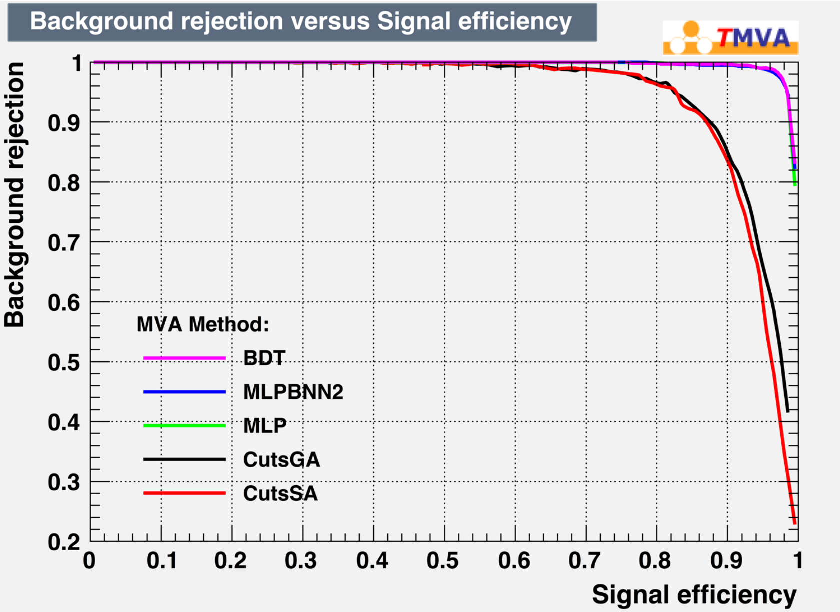
CMS B_s^0 Offline Cut Optimization

$B_s^0 \to J/\psi\phi \to \mu^+\mu^-K^+K^-$ Decay Topology



Application of Machine Learning

- Framework: TMVA package (standard C++ on ROOT with different algorithms)
- Tune the training parameters to optimize the performance of the algorithm
- Best performance among the algorithms: **Boosted Decision Tree (BDT)** (large area under the ROC curve), much better than traditional rectangular cuts (**CutsSA**/**CutsGA**)



sPHENIX

Los Alamos NATIONAL LABORATORY

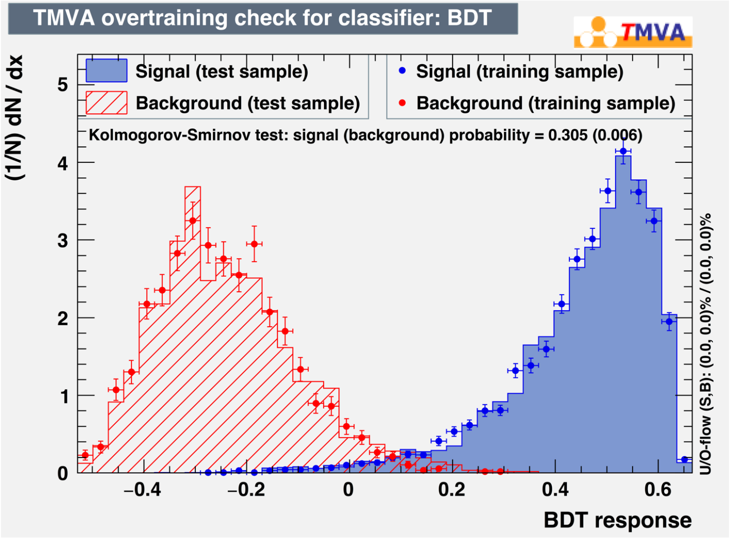
Machine Learning Performance

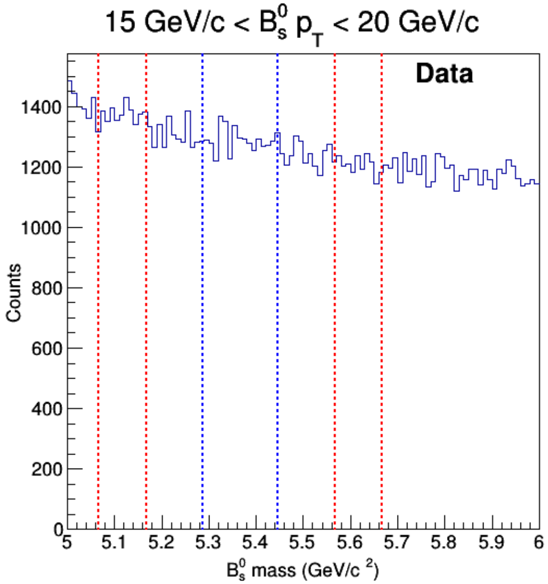
Analysis Challenges

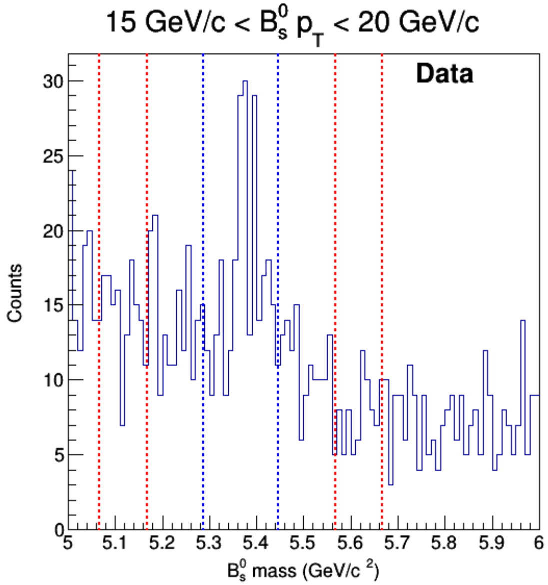
- Before selections, **B = 20672** and **S = 26**
- Require to obtain significant **B** to **S** rejection in order to observe B_s^0 signal in the PbPb data

Boosted Decision Tree Machine Learning Algorithm

- Excellent **Signal**/**Background** separation without overtraining down to 7 GeV/c in PbPb collisions
- A random BDT > 0.1 selection returns a visible signal
- Achieve better than **10^3** to **1** for the **B** to **S** rejection







Trigger Design Proposal

Intelligent experiments through real-time AI: Fast Data Processing and Autonomous Detector Control for sPHENIX and future EIC detectors

A proposal submitted to the DOE Office of Science
April 30, 2021

DATA ANALYTICS FOR AUTONOMOUS OPTIMIZATION AND CONTROL OF ACCELERATORS AND DETECTORS

Proposing Organization:	Los Alamos National Laboratory P-3, MS H846 Physics Division Los Alamos, NM 87545
Collaborating Institutions:	Fermi National Laboratory Massachusetts Institute of Technology New Jersey Institute of Technology
Principal Investigator:	Ming Xiong Liu Phone: +1 505-667-7125 Email: mliu@lanl.gov
Administrative Point of Contact:	Srinivas Iyer Phone: +1 505-665-9402 Email: siyer@lanl.gov
FOA Number:	DE-FOA-0002490
DOE/SC Program Office:	Nuclear Physics
DOE/SC Program Office Technical Contact:	Dr. Manouchehr Farkhondeh
Requested Funding:	$1,000K/year for two years
Total Request:	$2M

Office of Science

Department of Energy Announces $5.7 Million for Research on Artificial Intelligence and Machine Learning (AI/ML) for Nuclear Physics Accelerators and Detectors

DECEMBER 2, 2021

Office of Science »
Department of Energy Announces $5.7 Million for Research on Artificial Intelligence and Machine Learning (AI/ML) for Nuclear Physics Accelerators and Detectors

Projects will advance understanding of atomic structure and the nature of matter and antimatter

WASHINGTON, D.C. - Today, the **U.S. Department of Energy (DOE)** announced $5.7 million for six projects that will implement artificial intelligence methods to accelerate scientific discovery in nuclear physics research. The projects aim to optimize the overall performance of complex accelerator and detector systems for nuclear physics using advanced computational methods.

"Artificial intelligence has the potential to shorten the timeline for experimental discovery in nuclear physics," said **Timothy Hallman, DOE Associate Director of Science for Nuclear Physics**. "Particle accelerator facilities and nuclear physics instrumentation face a variety of technical challenges in simulations, control, data acquisition, and analysis that artificial intelligence holds promise to address."

Project Participating Institutions

Los Alamos National Laboratory
MIT
Fermilab
NJIT
Oak Ridge National Laboratory
UNT University of North Texas

Online Heavy Flavor Machine Learning Trigger

Design: Suitable for Continuous Streaming Readout

- Real-time AI on tracking system to enrich **LOW p_T** hadron data taking
- Fast online processing to select interesting events
- Fit to data collection DAQ` bandwidth in with high signal efficiency
- Automated control feedback for detector operation through real-time AI

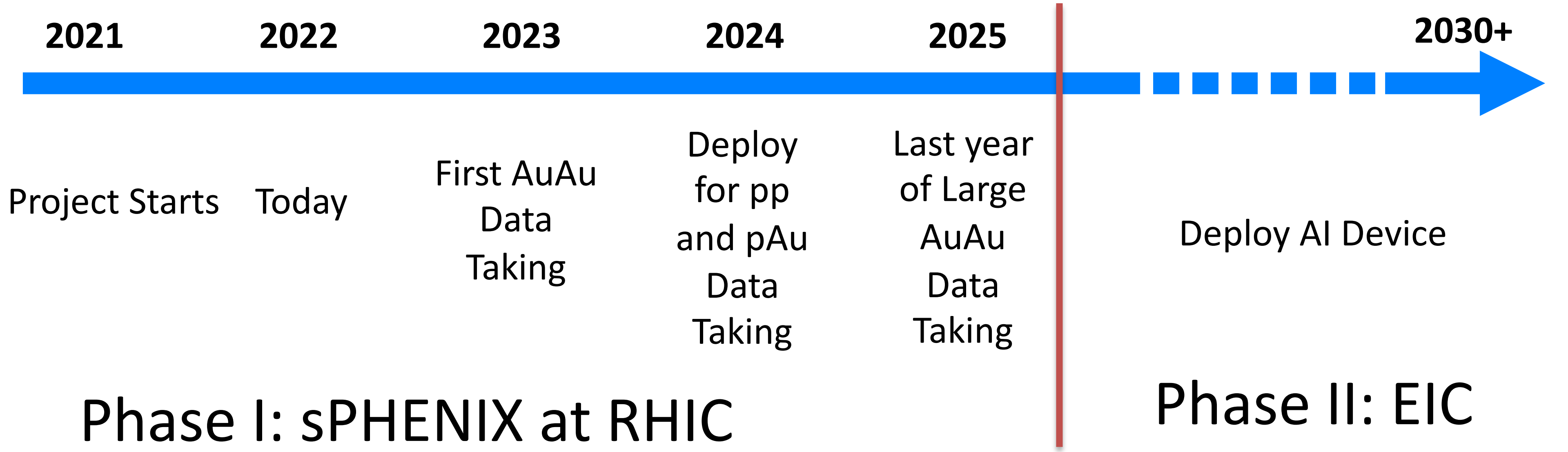
Novel Technology Applied

- Hardware and firmware interface to detect unique signal
- FPGA along with GPU parallel processing to allow large computing power
- Advanced machine learning for signal pattern recognition from the hit level

Testing Platform before EIC

- sPHENIX at RHIC: demonstration of its functionality inner tracking system streaming readout for $c\bar{c}$ and $b\bar{b}$ event trigger
- Goal: obtain 500 × heavy flavor events with only 1.5 increase of data size for fully reconstructed b hadron, both inclusive and exclusive, measurements

Project Timeline



Two Phases for the R&D of online heavy flavor machine learning trigger

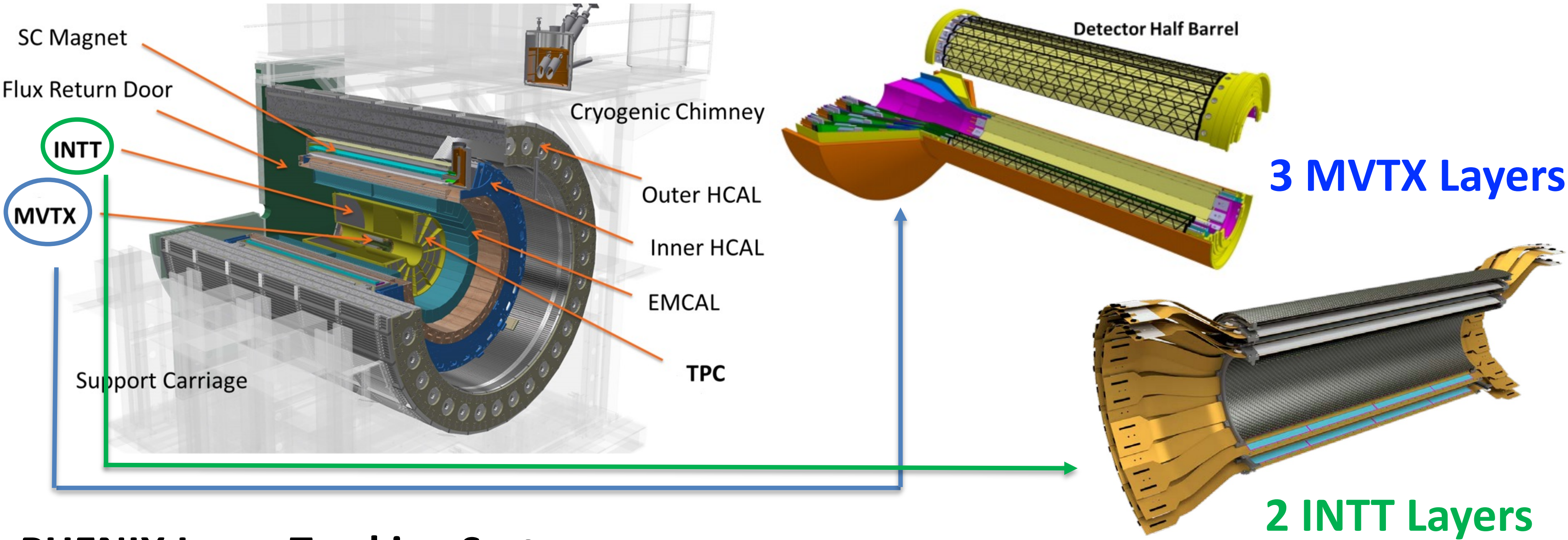
Phase I: the sPHENIX experiment at RHIC

- Research and Development
- Demonstration functionality

Phase II: Deploy device for EIC experiments

- Modify to be suitable for EIC environment and physics goals

The sPHENIX Experiment at RHIC



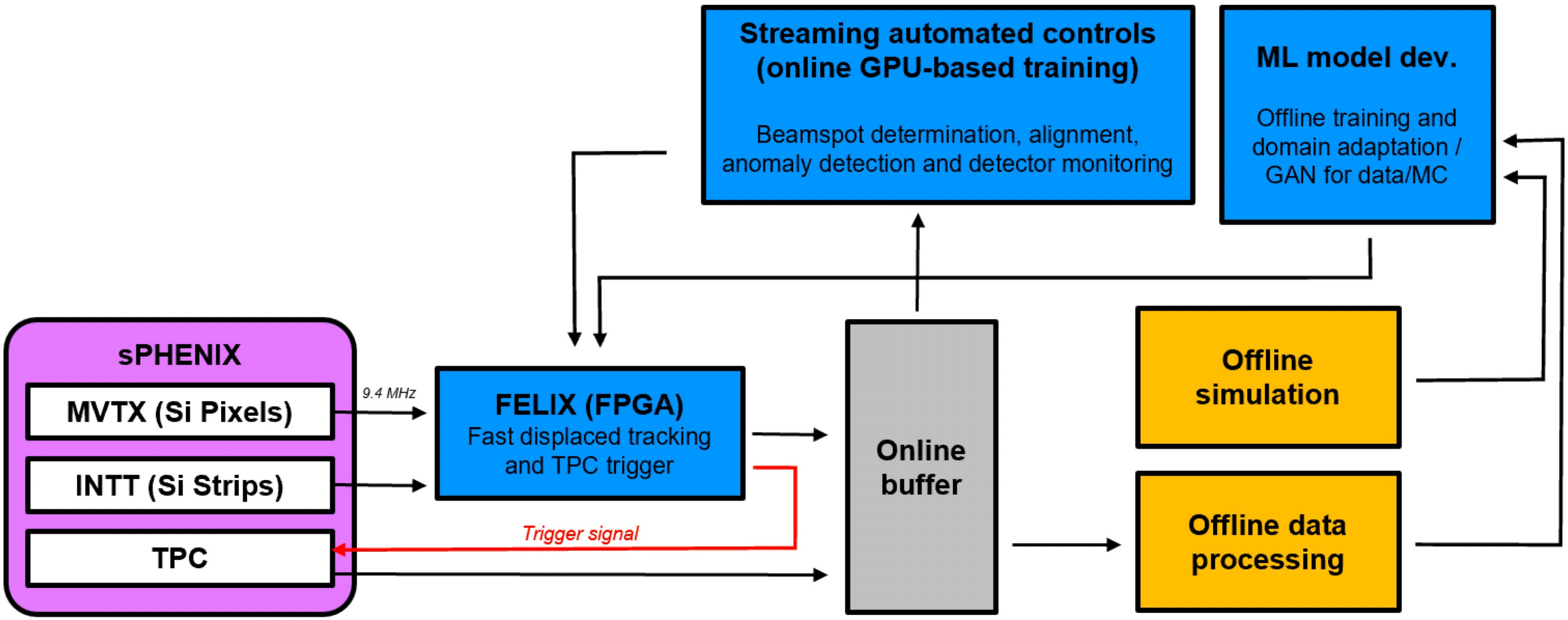
sPHENIX Inner Tracking System

- **MVTX:** Monolithic Active Pixel Sensor (MAPS)-based-vertex detector
  - High granularity of pixel pitches with excellent position resolution ($\sim 5\ \mu m$)
- **INTT:** Intermediate Tracker
  - Silicon strip with fast electronics

Readout Features

- Both MVTX and INTT are operating in the continuous streaming readout mode
- Overall DAQ rate ~15kHZ compared to RHIC 10 MHz in pp collisions

Trigger Design Schematic for sPHENIX



- Fast trigger decision based on machine learning algorithm in order of several μs using MVTX and INTT hit information to initial TPC information
- Fast track reconstruction with the entire tracking system with unsupervised machine learning to determine beam vertex (x, y)
- Autonomous close-loop feedback for the detectors monitoring and calibration as the condition evolves

AI Model for Online Trigger

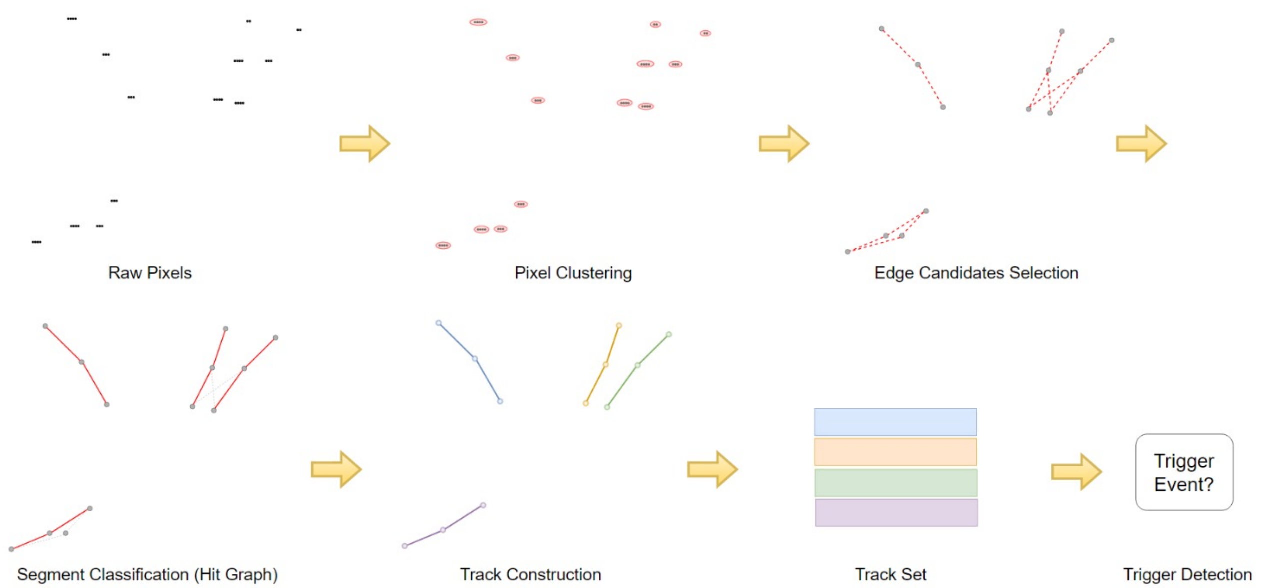


Fig. 3: Trigger Detection Pipeline with All the Steps.

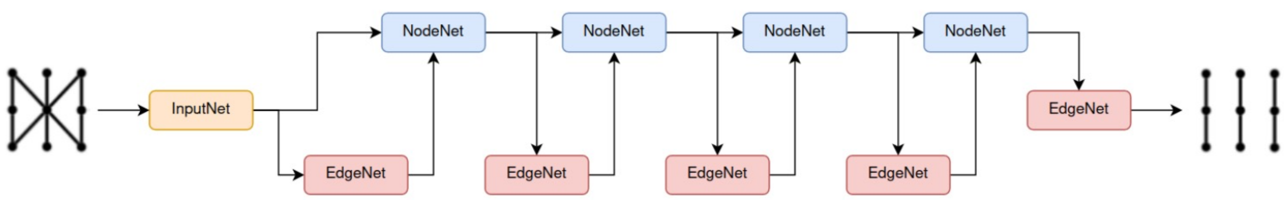


Fig. 4: Graph Neural Network for Tracking

- Graph neural network: strong capability identifying the pattern for tracking
- High accuracy with excellent reliability for signal detection

Machine Learning Model Performance

- No real experimental data from sPHENIX yet
- PYTHIA simulation to model $pp \rightarrow c\bar{c} \rightarrow D^0 \rightarrow K^-\pi^+$ collision events
- INTT and MVTX hit information saved machine learning training and testing
- Working point: signal probability > 50%

Tracking	Efficiency	Fake Rate
Performance	85%	17%

Trigger	Signal Efficiency	Background Rejection	Data Volume Reduction
Performance	90%	77%	24%

- Accepted to ECML PKDD 2022 for presentation and publication

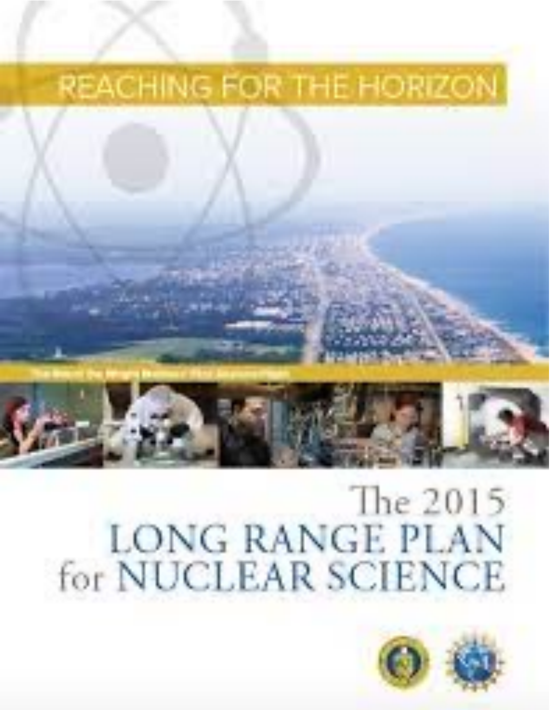
Summary

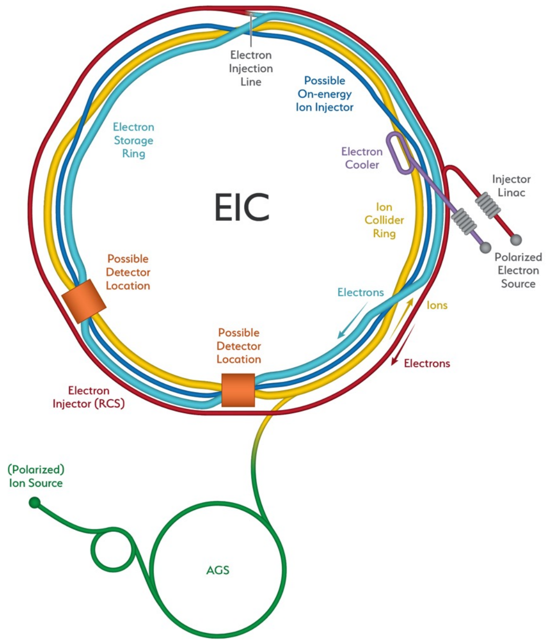
- Heavy flavor measurements crucial to achieve EIC physics goals via the studies initial state parton distribution function and final state fragmentation function
- Many challenges and opportunities for exotic hadron program at the EIC
- Success of machine learning for heavy flavor and exotic hadron studies
- Ongoing development of online fast and autonomous heavy flavor trigger with the application of novel FPGA computing and advanced machine learning techniques
- First deploy to sPHENIX and then to future EIC tracking systems in continuous streaming readout mode to trigger low p_T heavy flavor hadron events
- Application of heavy flavor trigger to potentially benefit the exotic hadron physics program at the EIC in Phase II

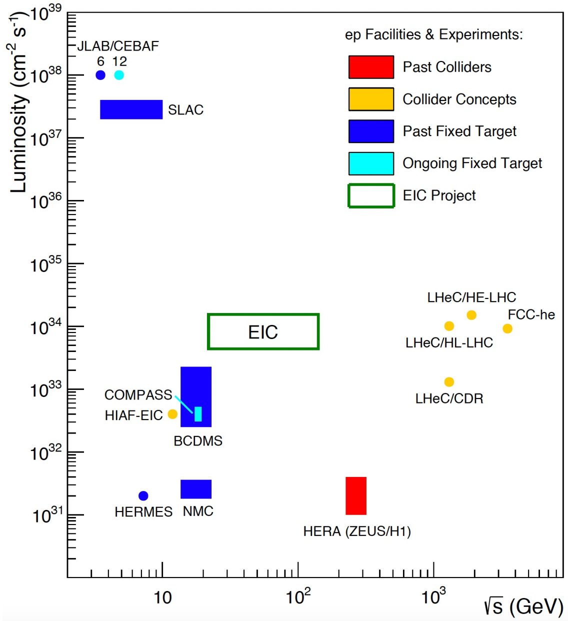
Back Up

The Electron-Ion Collider

NSAC Long Range Plan for Nuclear Science – The Electron Ion Collider (EIC)





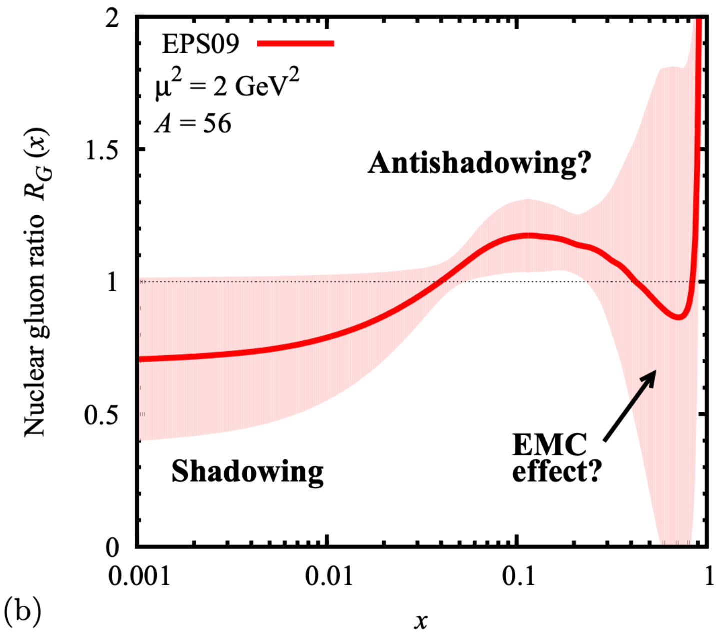


- To be built at BNL and currently moving toward CD-2 stage this year
- High luminosity: $10^{33-34} cm^{-2} s^{-1} \sim 100 - 1000 fb^{-1}$/year
- Highly polarized beam: nuclei and electron beams polarization up to 80%
- Variety of nuclei species: from p to $^{238}_{92}U$
- Broad kinematic range coverage: $\sqrt{s} = 20 - 140$ GeV
- 3D tomography of nucleon and nuclei to study non-perturbative QCD

Heavy Flavor Physics at the EIC

Quark-Gluon Structure of Nuclei

- Powerful observable complement to inclusive DIS measurements
- Clean probe: sensitive to gluon dynamics
- Probe gluon saturation regime
- Understand nuclear gluon shadowing effect
- Extract the Sivers Function with polarized DIS

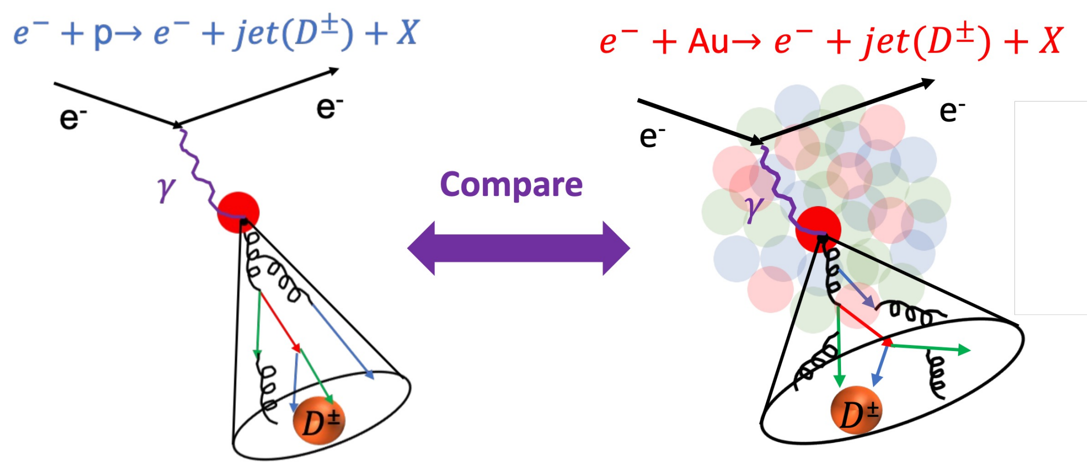


(b)

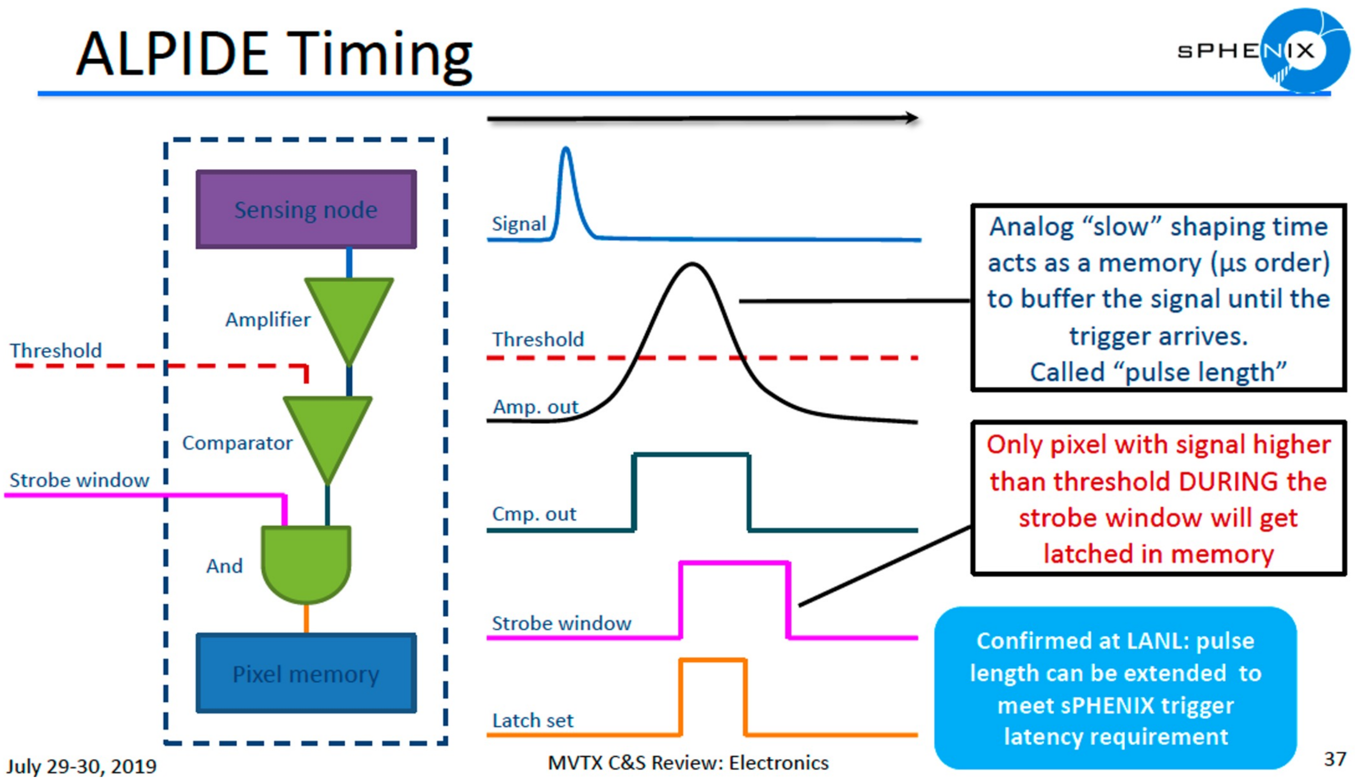
$$R_G(x) = G_A(x)/G_N(x)$$

Parton Propagation and Hadronization

- Heavy quark interaction in cold nuclear medium
- Hadronization modification from npdf measurements
- Test QCD factorization breaking in ep and eA

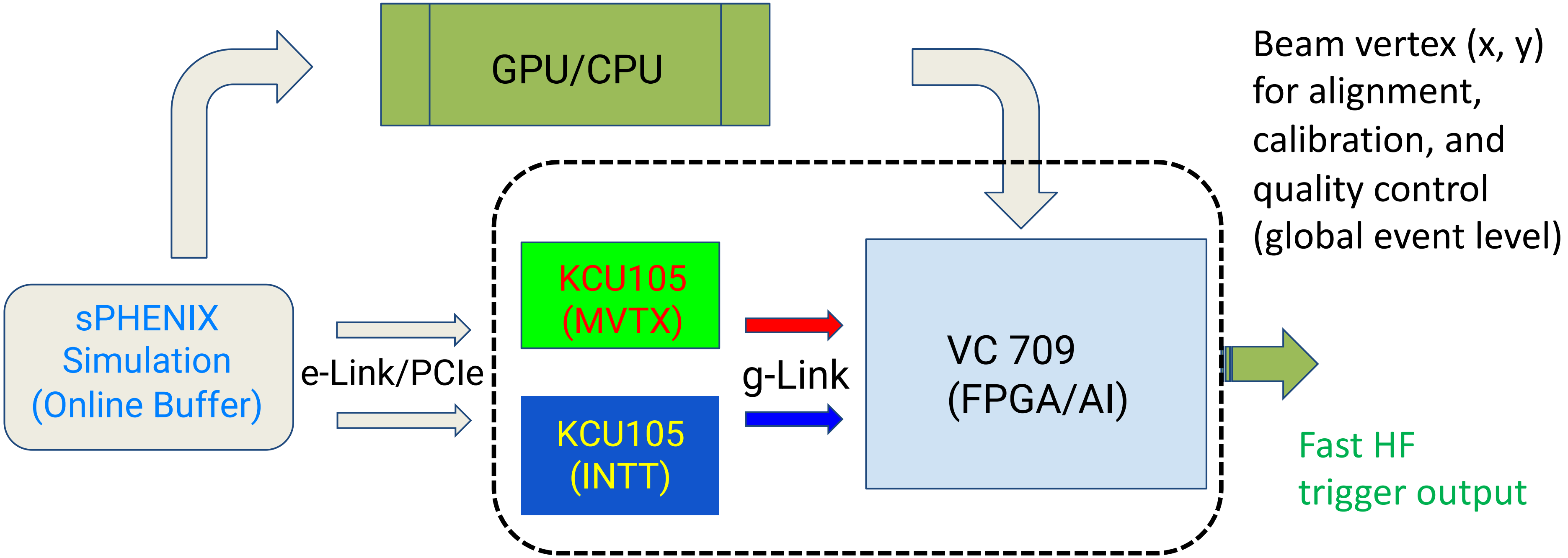


MVTX Alpide Sensor Readout



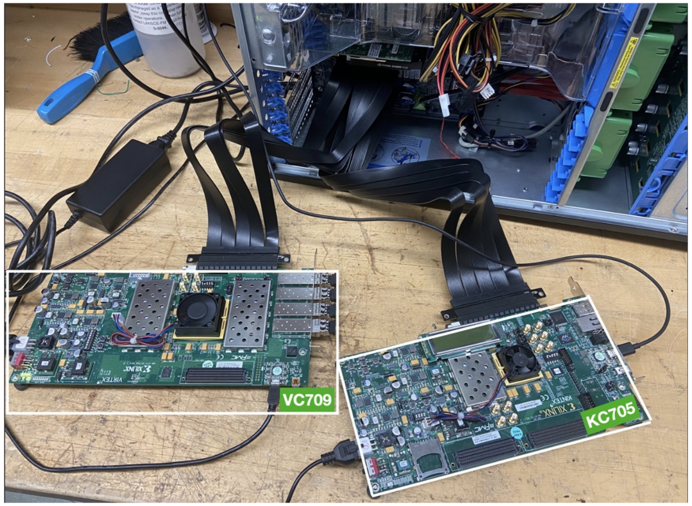
- Continuous streaming readout of strobes containing all pixel information
- Zero suppression mode

Toy Model: FPGA-Based Hardware Trigger



Toy model with sPHENIX inner tracking system as a demonstration

- Streaming readout simulation data
- 8b/10b MVTX/INTT data (KC705) to FPGA/AI Engine (VC709)
- Will test in the second year of pp data taking and third year $AuAu$ data taking, particularly for $b\bar{b}$ events



Heavy Flavor Trigger Workflow

1. Fetch events from event buffer (from PYTHIA Monte Carlo simulation to raw data)

2. Data Pre-processing Clustering (on FPGA low latency implementation)

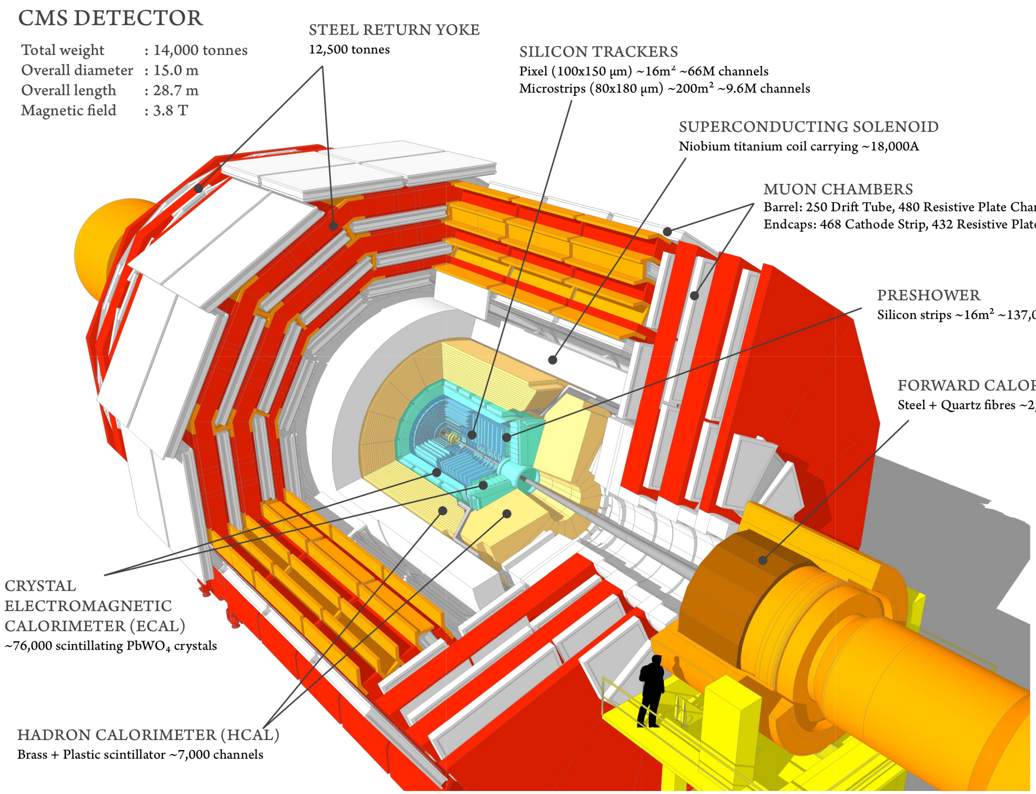
3. Clustering: outlier hits Removal (Done in FPGA with MVTX and INTT only)

5. Tracking: connecting the clusters and forming the primary vertex (Done in FPGA)

4. Triggering based on tracks features (Done in FPGA)

5. Triggers on TPC (Interface and integration with SPHENIX Detector)

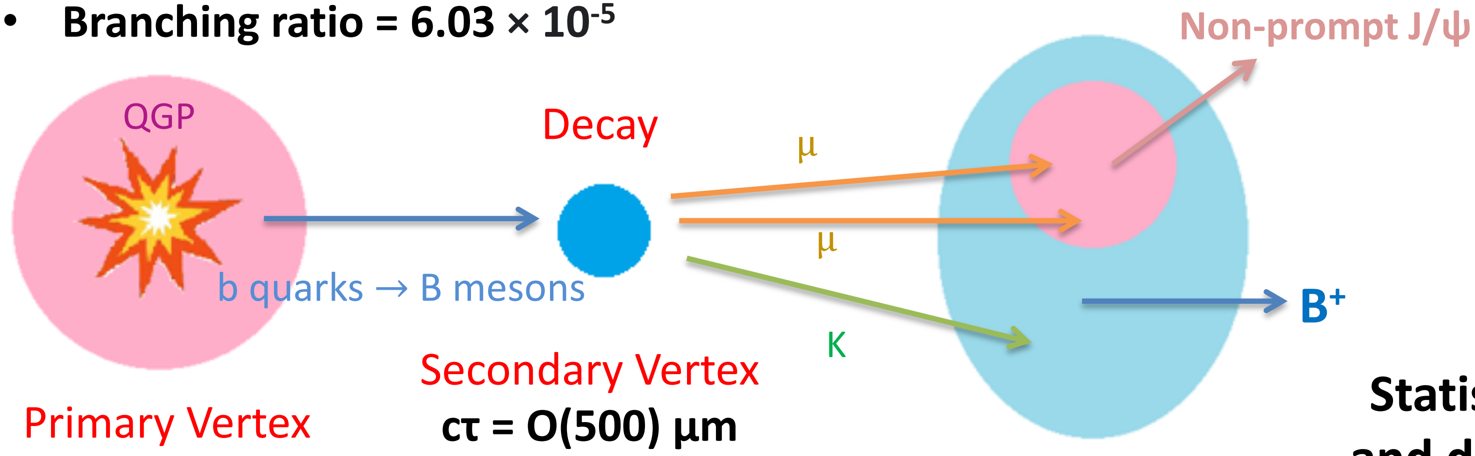
Fully Reconstructed B_s^0 with CMS at the LHC



- Fully reconstructed B mesons from decay chain involve J/ψ
- Take advantage of excellent muon and tracking capabilities
- Not using hadronic particle identification or calorimeter information
- Constrain intermediate state resonances to improve the results
- Apply multivariate approach with machine learning techniques

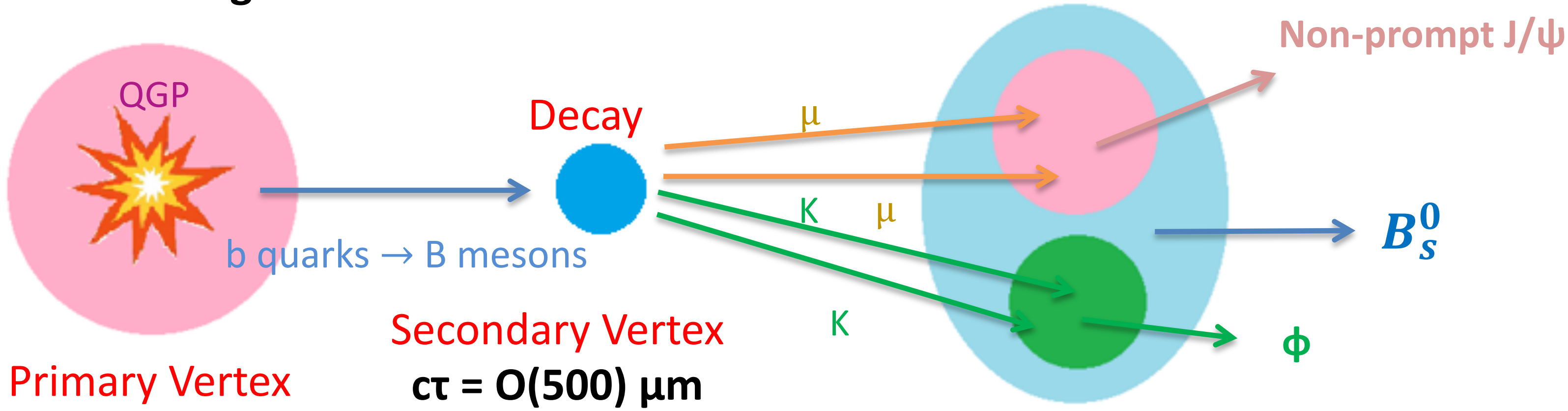
Analysis Strategies

- B^+: via the decay channel $B^+ \rightarrow J/\psi K^+ \rightarrow \mu^+\mu^- K^+$
- Branching ratio = 6.03 × 10^{-5}

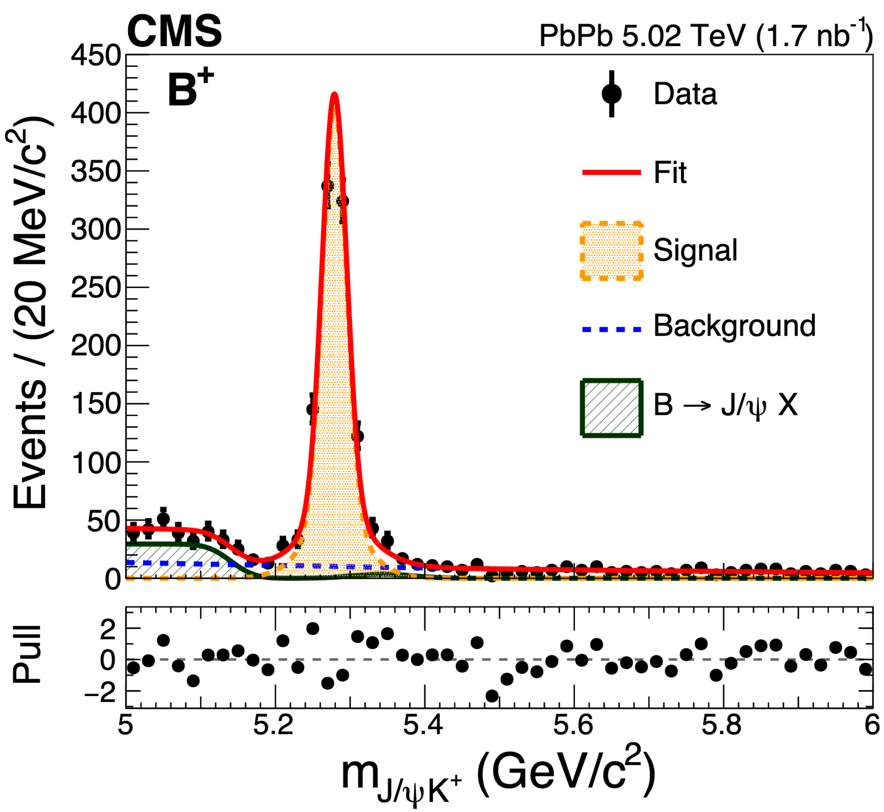


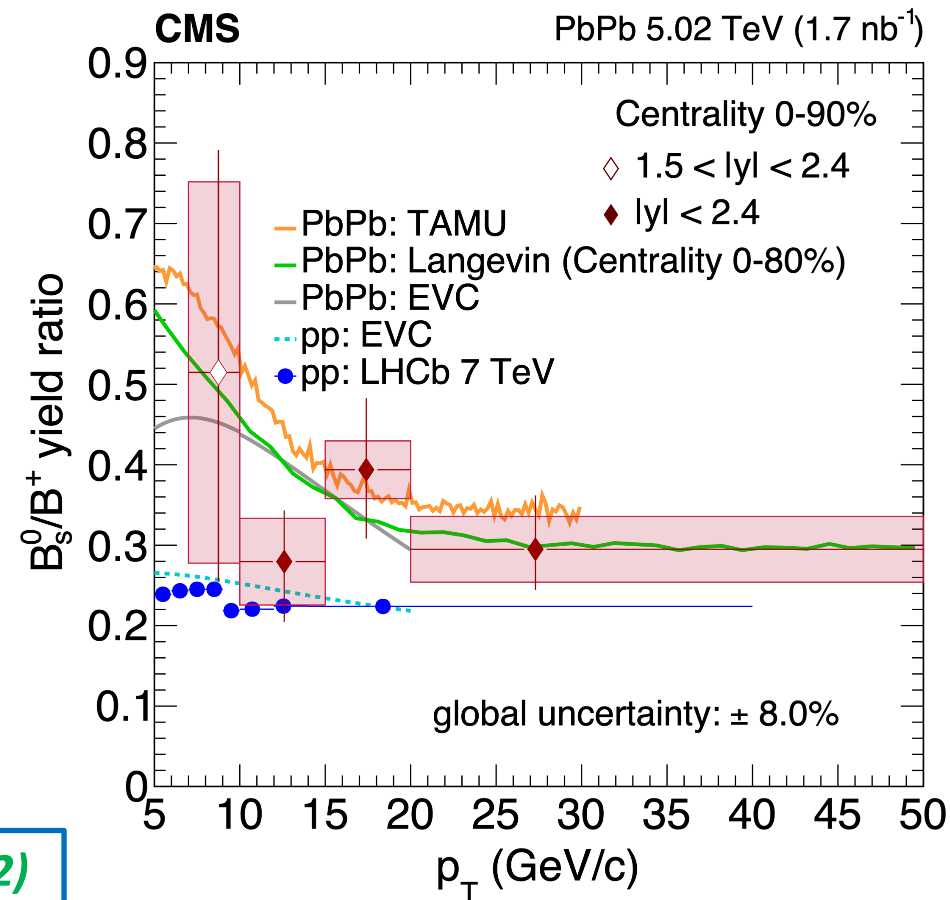
Statistically enriched and dedicated dimuon triggered datasets from 2018 LHC PbPb run

- B_s^0: using the decay channel $B_s^0 \rightarrow J/\psi \phi \rightarrow \mu^+\mu^- K^+ K^-$
- Branching ratio = 3.17 × 10^{-5}



Physics Results





Phys. Lett. B 829 (2022)

- First observation of fully reconstructed B_s^0 down to $p_T = 7$ GeV/c with greater than 5σ significance in heavy-ion collisions
- Beauty hadronization mechanism: quark coalescence required to describe the data
- Success of machine learning in heavy flavor and exotic hadrons