Towards Fast Calibration with the ePIC Barrel Hadronic Calorimeter

Derek Anderson

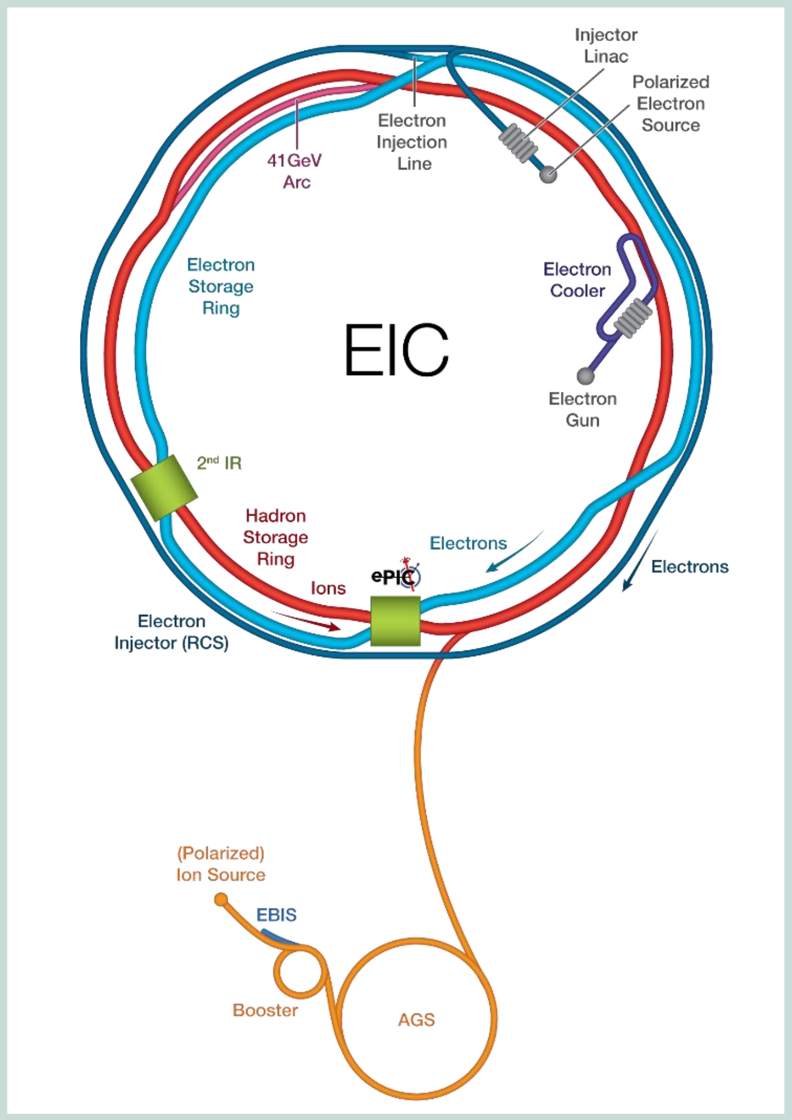
Iowa State University

For the ePIC Collaboration

Introduction | The EIC

- **The Electron-Ion Collider (EIC):** next generation collider to be **jointly hosted by BNL & JLab**
  - Dedicated to studying the internal structure of nuclear matter
  - **Will be built at BNL:**
    - Add electron accelerator, storage rings alongside existing RHIC hadron ring
  - ☞ **Will begin collisions in early 2030s**

- **A few details:**
  - e^- energies = 5 – 18 GeV
  - Ion energies = 40 – 275 GeV
    - $\Rightarrow \sqrt{s} =$ 29 – 141 GeV/u
  - Ions species from proton – Uranium
    - ☞ Up to 70% polarization for light ions (p – He)

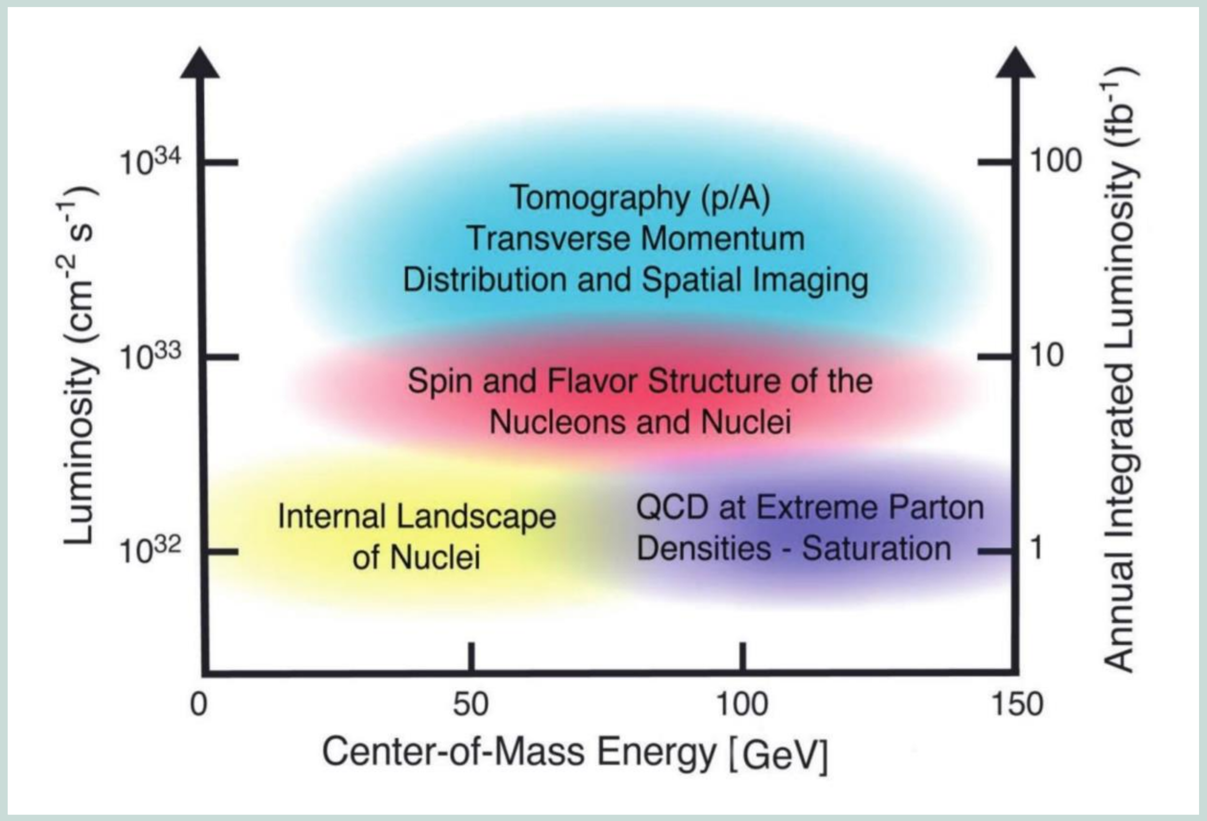


Introduction | EIC Goals and Luminosities

- **Right:** major topics to be explored at EIC vs. required luminosities & CoM energy
  - **Anticipates $\mathcal{L}_{\text{peak}} \sim 10^{34}$ (cm² s)⁻¹**
  - Translates to roughly 1.5 fb^{-1} per month
    - ☞ **Assuming**: 60% operation time & $\overline{\mathcal{L}} = \mathcal{L}_{\text{actual}}$

- **But:** typical σ_{int} is $\mathcal{O}(100) \times$ smaller than comparable σ_{int} at RHIC/LHC…
  - And there is **wide** variety of processes to record…
  - ⇒ **Streaming Readout (SRO) is a must if we want to fully unlock EIC scientific potential**

- For reference
  - **RHIC:** $\overline{\mathcal{L}}_{pp} \sim 2.45 \times 10^{34}$ $(cm^2\ s)^{-1}$
  - **LHC:** $\overline{\mathcal{L}}_{pp} \sim 1 \times 10^{34}$ $(cm^2\ s)^{-1}$

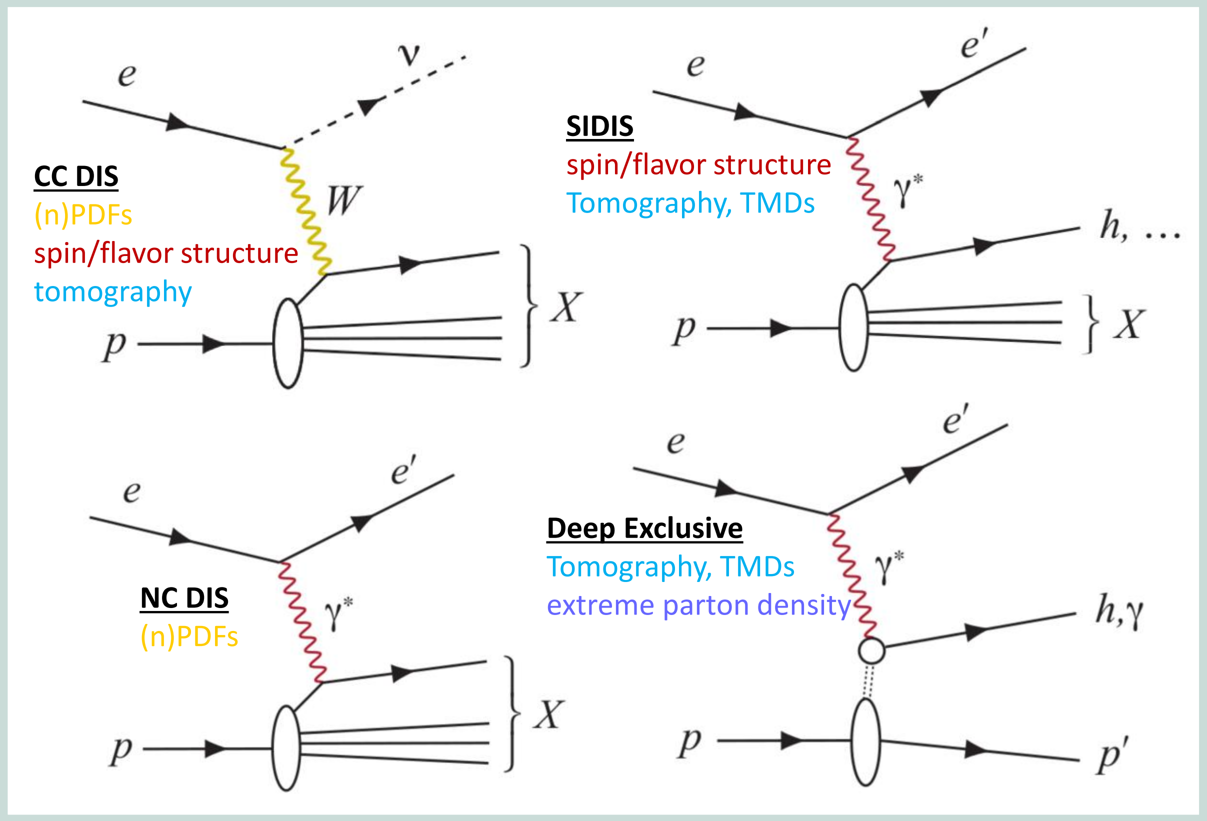


Introduction | EIC Goals and Luminosities

- **Right:** major topics to be explored at EIC vs. required luminosities & CoM energy
  - **Anticipates $\mathcal{L}_{\mathbf{peak}} \sim 10^{34}$ (cm² s)⁻¹**
  - Translates to roughly 1.5 fb⁻¹ per month
    - ☞ **Assuming**: 60% operation time & $\overline{\mathcal{L}} = \mathcal{L}_{\text{actual}}$

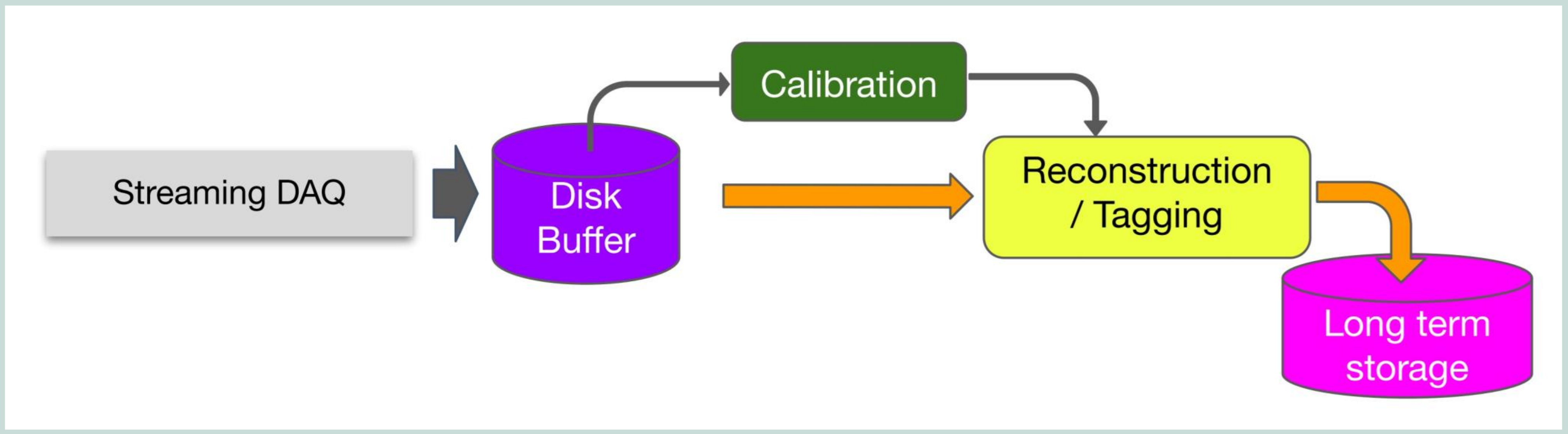
- **But:** typical σ_{int} is $\mathcal{O}(100) \times$ smaller than comparable σ_{int} at RHIC/LHC…
  - And there is **wide** variety of processes to record…
  - ⇒ **Streaming Readout (SRO) is a must if we want to fully unlock EIC scientific potential**

- For reference
  - › **RHIC:** $\overline{\mathcal{L}}_{pp} \sim 2.45 \times 10^{34}$ (cm² s)⁻¹
  - › **LHC:** $\overline{\mathcal{L}}_{pp} \sim 1 \times 10^{34}$ (cm² s)⁻¹



Introduction | Streaming Readout

From arXiv:2202.03085



- **Streaming readout (SRO):** data read out in continuous parallel streams
  - Each stream encodes when/where data was recorded
  - Data digitized at fixed rate & thresholds/zero-suppression applied locally
  - Event building, filtering, monitoring, etc. deferred until data in tiered storage
  - ☞ **See:** report on ePIC Streaming Computing Model

Current & Future Examples:

- LHCb
- Recent test implementation at JLab
- CLAS12
- **sPHENIX:** *FastML Triggering in sPHENIX*, Cameron Dean [Tuesday, 11:30 AM]
- **And ePIC!**

Introduction | SRO Advantages & Fast Calibration

- **Several advantages of SRO over traditional readout (RO)!**
  a) Enables simplified & more flexible RO hardware
     - No custom trigger hardware/firmware!
  b) Provides access to detailed knowledge of background
  c) Allows workflows to be streamlined & utilize new technologies
     - ☞ **e.g. AI/ML!**

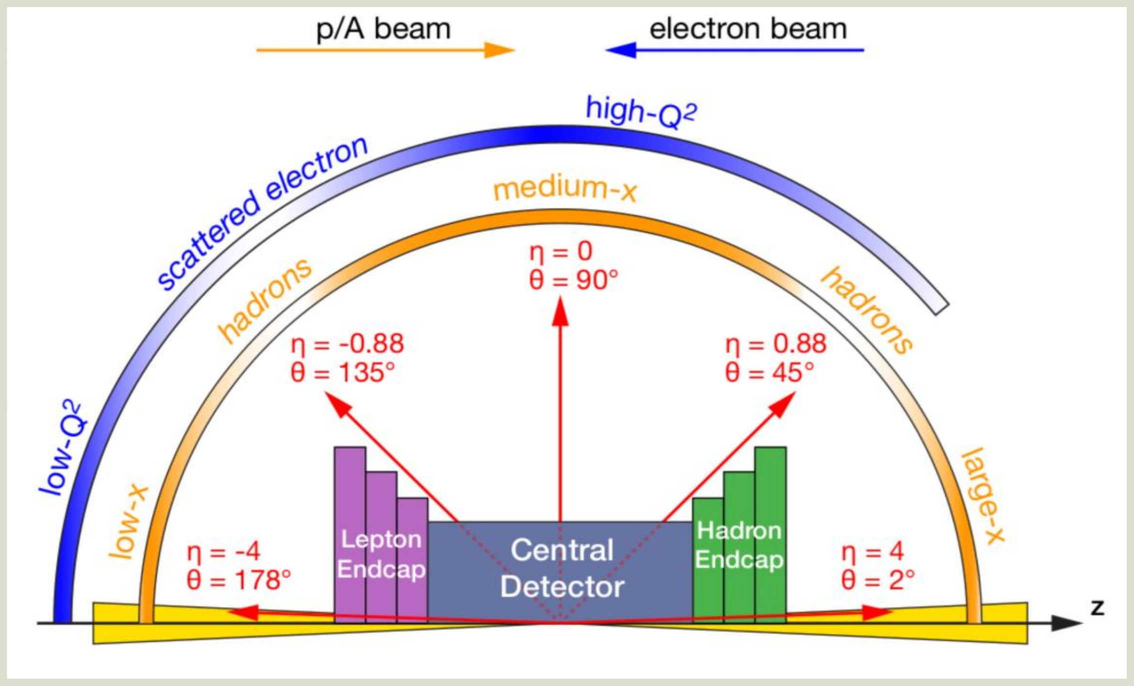
☞ **See:** arXiv:2202.03085 & ePIC Streaming Computing Model report

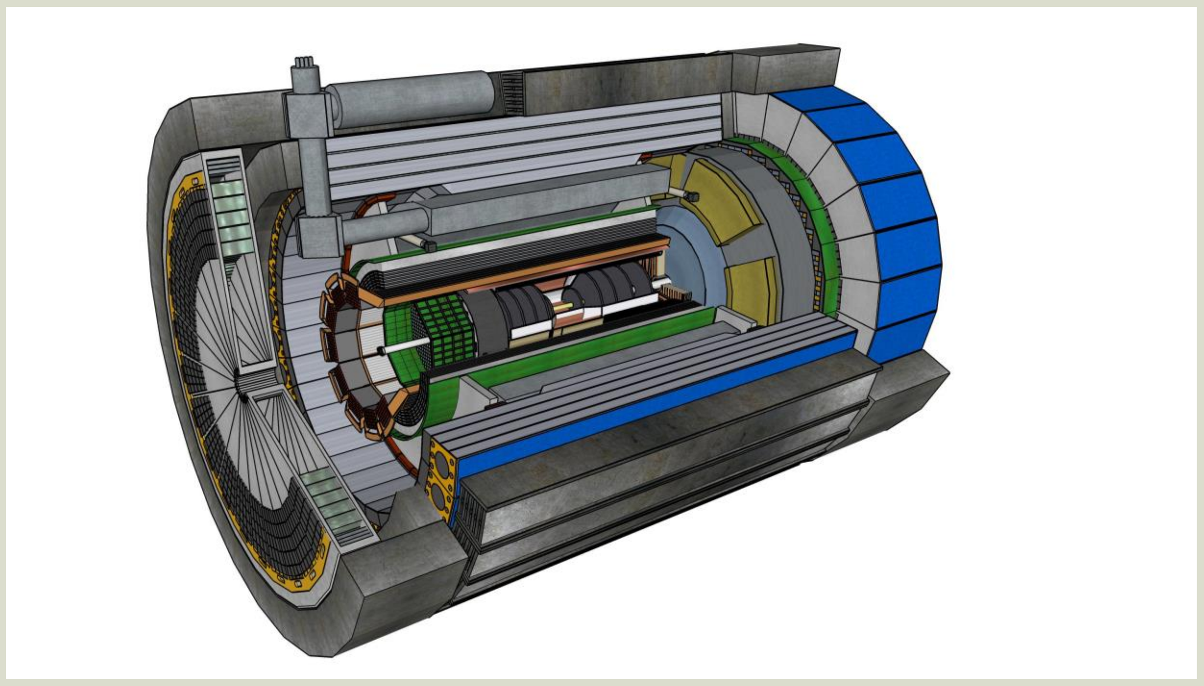
In this talk: we'll discuss initial, ongoing studies using ML to calibrate the ePIC BHCal

Building on (c):

- Typically $\mathcal{O}(1)$ year between recording data & analyzing data
  - Due to complexity of HEP/NP experiments
  - Alignment, calibration, reconstruction, & validation are costly!
- **Our goal for ePIC:** 2 – 3 weeks between recording & analyzing data!
  - Timeline driven by calibration
- Can be accomplished by integrating computing & detector, **esp. using AI/ML for:**
  - ***Autonomous alignment/calibration***
  - And rapid reconstruction + validation

ePIC | The ePIC Detector





- **ePIC:** EIC project detector to be built at IP6 (6 o' clock)
  - Fulfills EIC science mission & detector requirements
    - c.f. Yellow & NAS Reports
  - Collaboration formed in late 2022 – early 2023

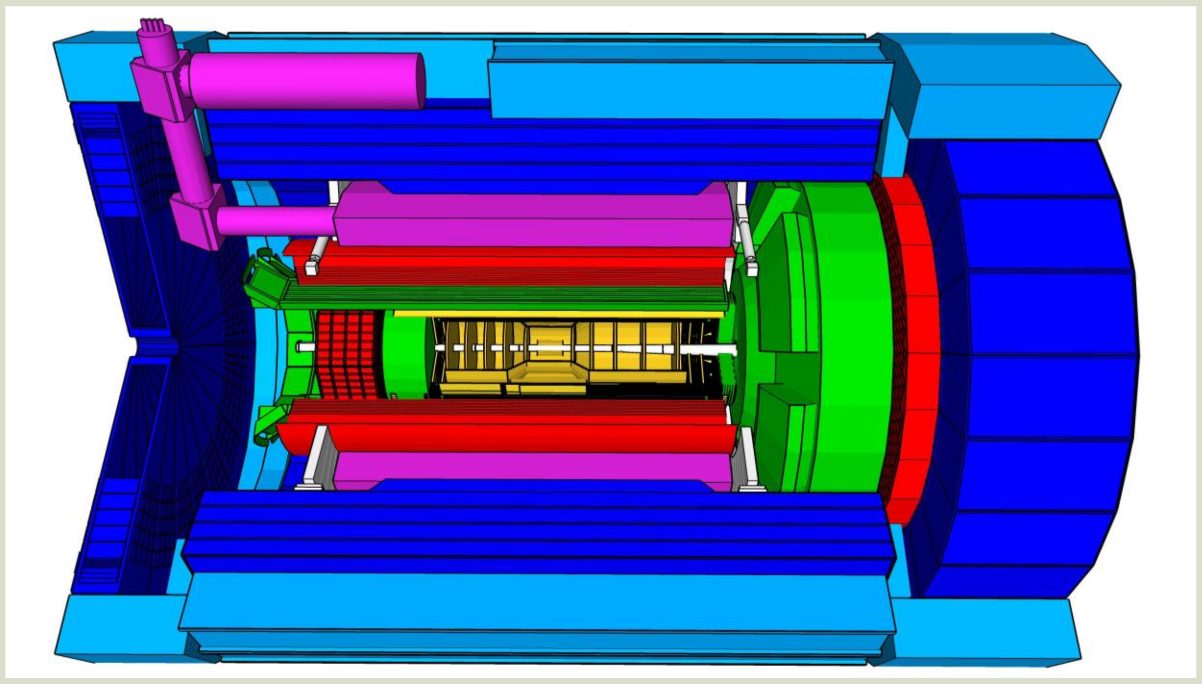
- (Almost) fully hermetic central detector
  - + Extensive coverage in forward, backward

☞ **See:** *ePIC: Status & Plans*, Zhoudunming Tu [Friday, 11 AM]

ePIC | Detector Technologies

Subsystems:

- **Tracking**
  - › **Inner layers:** MAPS detectors
  - › **Outer layers:** MPGDs (μRWELL, MMS)
- **Particle ID**
  - › **Barrel:** high-performance DIRC
  - › **Forward:** dual-radiator RICH
  - › **Backward:** proximity-focusing RICH
  - › TOF (using AC-LGADs)
- **EM Calorimetry**
  - › **Barrel:** Imaging (Si + Pb/SciFi* matrix)
  - › **Forward:** W-powder + SciFi*
  - › **Backward:** PbWO4 crystals
- **Hadronic calorimetry**
  - **Barrel:** Fe + scintillating tiles
  - **Endcaps:** Fe/W + scintillating tiles



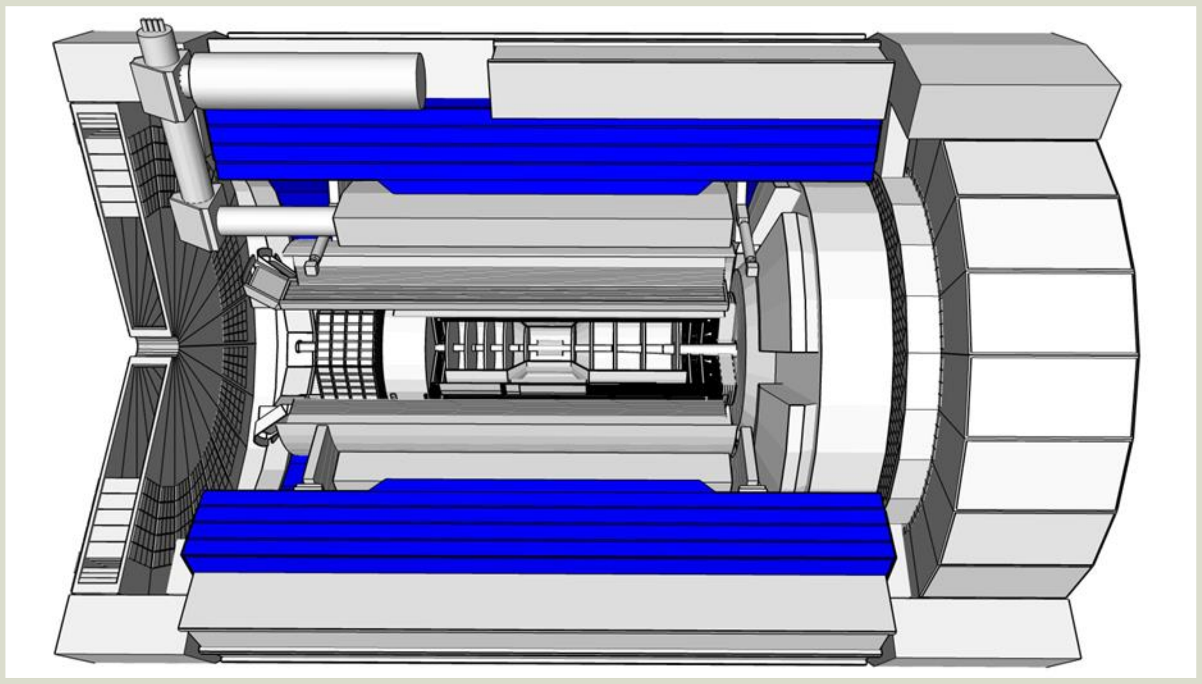
- **Note:** far-forward & backward detectors not shown
  - ☞ Full detector extends 90 m!

* Scintillating-fibers

ePIC | Detector Technologies

Subsystems:

- **Tracking**
  - **Inner layers:** MAPS detectors
  - **Outer layers:** MPGDs (μRWELL, MMS)
- **Particle ID**
  - **Barrel:** high-performance DIRC
  - **Forward:** dual-radiator RICH
  - **Backward:** proximity-focusing RICH
  - TOF (using AC-LGADs)
- **EM Calorimetry**
  - **Barrel:** Imaging (Si + Pb/SciFi* matrix)
  - **Forward:** W-powder + SciFi*
  - **Backward:** PbWO4 crystals
- **Hadronic calorimetry**
  - **Barrel:** Fe + scintillating tiles
  - **Endcaps:** Fe/W + scintillating tiles



- **Note:** far-forward & backward detectors not shown
  - ☞ Full detector extends 90 m!

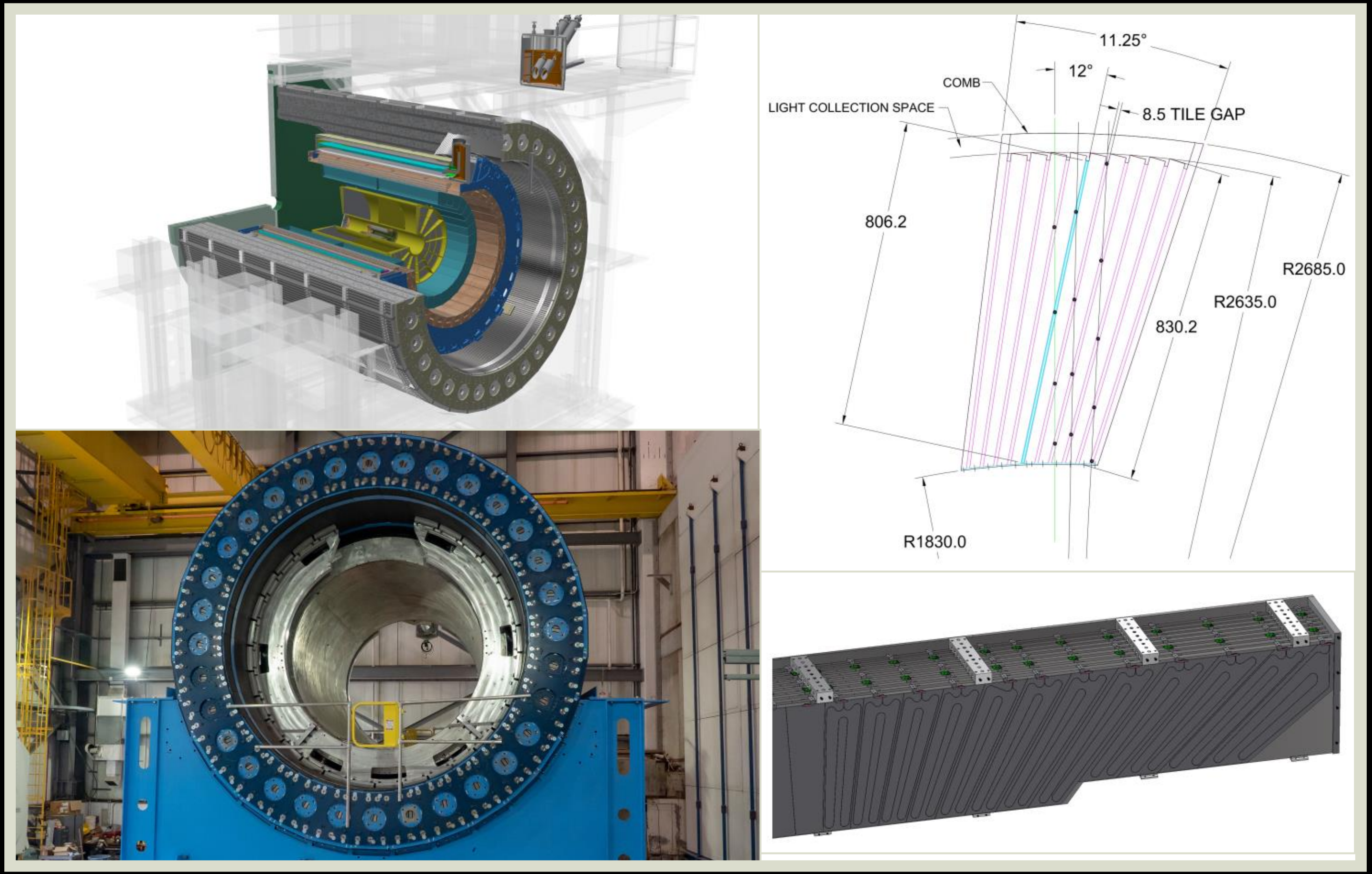
* Scintillating-fibers

ePIC | The ePIC BHCal

- **ePIC plans to reuse outer sPHENIX Barrel Hadronic Calorimeter (BHCal)**
  - Consists of alternating Fe and scintillating tile (+ WLS fibers)

- Technical details:
  - $|\eta| < 1.1$, 2π coverage
    - 48 towers/sector, 32 sectors, 5 tiles/tower
    - $\Delta\eta \times \Delta\varphi \sim 0.1 \times 0.1$
  - Depth is ~3.5λ

- sPHENIX reads each **tower** while ePIC plans to read out each **tile**
  - ☞ **Improves granularity!**

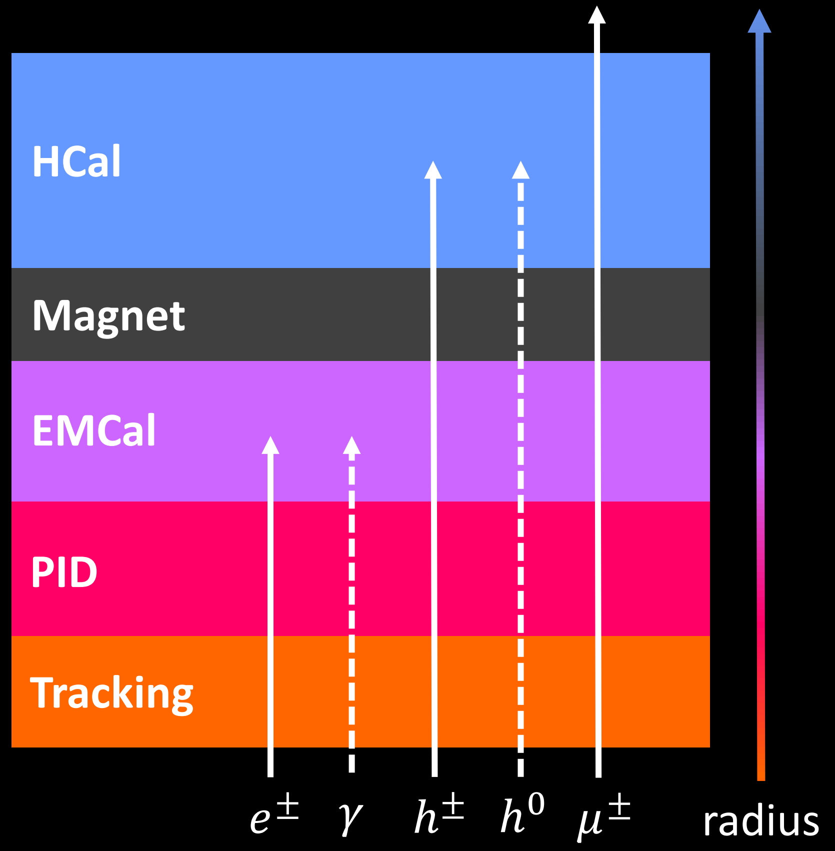


ePIC | BHCal Utility

- In barrel region ($|\eta| < 1$), jets are relatively soft
  - Tracker provides best momentum determination
  - But hadronic calorimeter would provide measurement of h^0

∴ The BHCal will serve several roles at ePIC
  a) Precise jet energy reconstruction
  b) Additional determination of e^- kinematics
  c) Solenoid flux return
  d) Possible $\mu^\pm$ identification

- **Right:** schematic diagram of a typical HEP/HENP experiment vs. radius



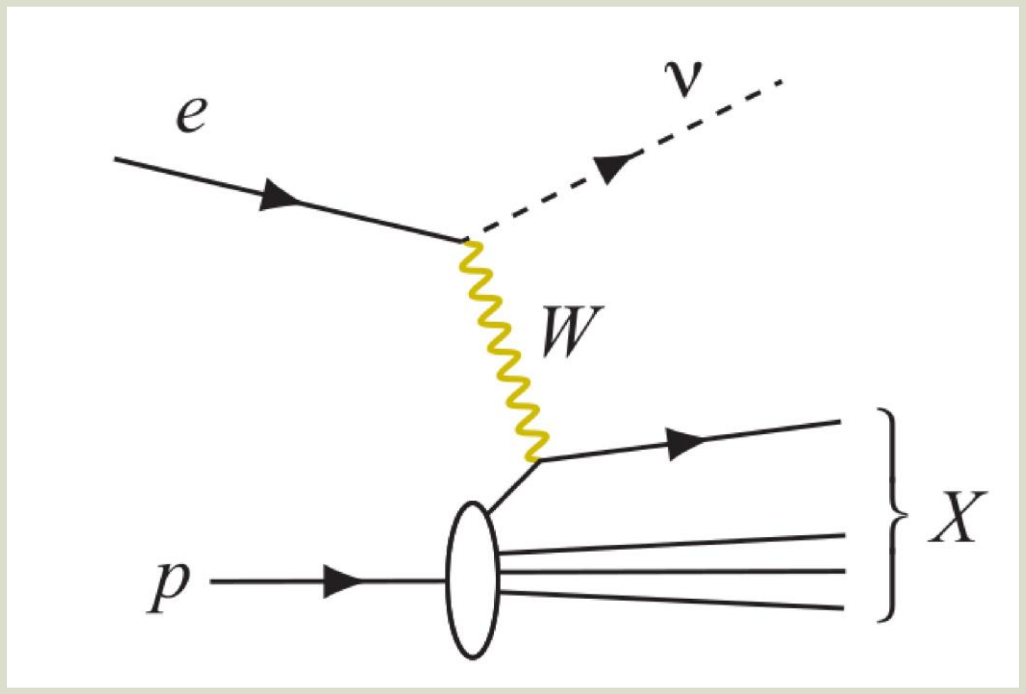
ePIC | BHCal Utility

- In barrel region ($|\eta| < 1$), jets are relatively soft
  - Tracker provides best momentum determination
  - But hadronic calorimeter would provide measurement of h^0

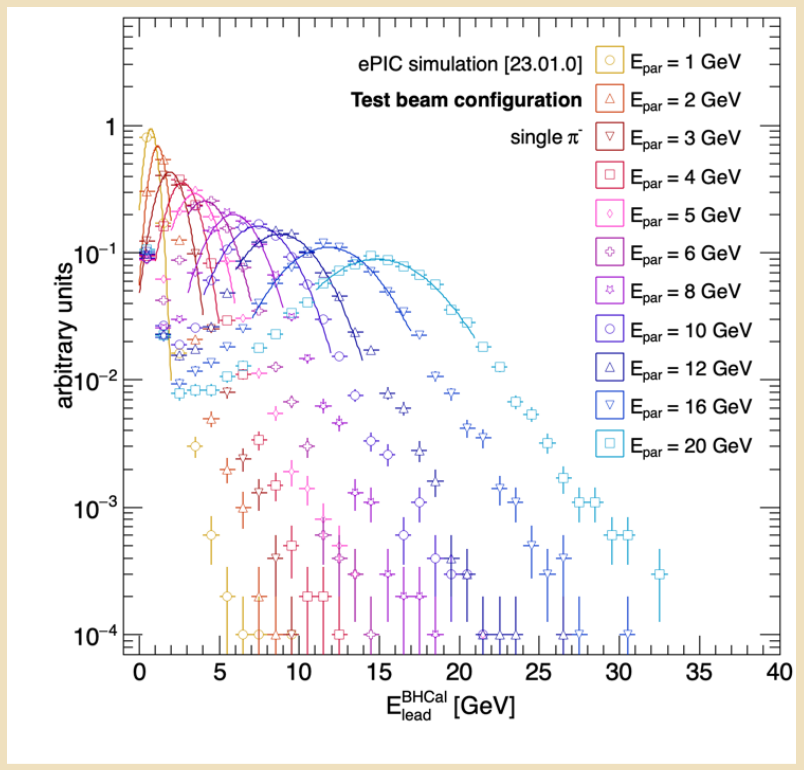
∴ The BHCal will serve several roles at ePIC

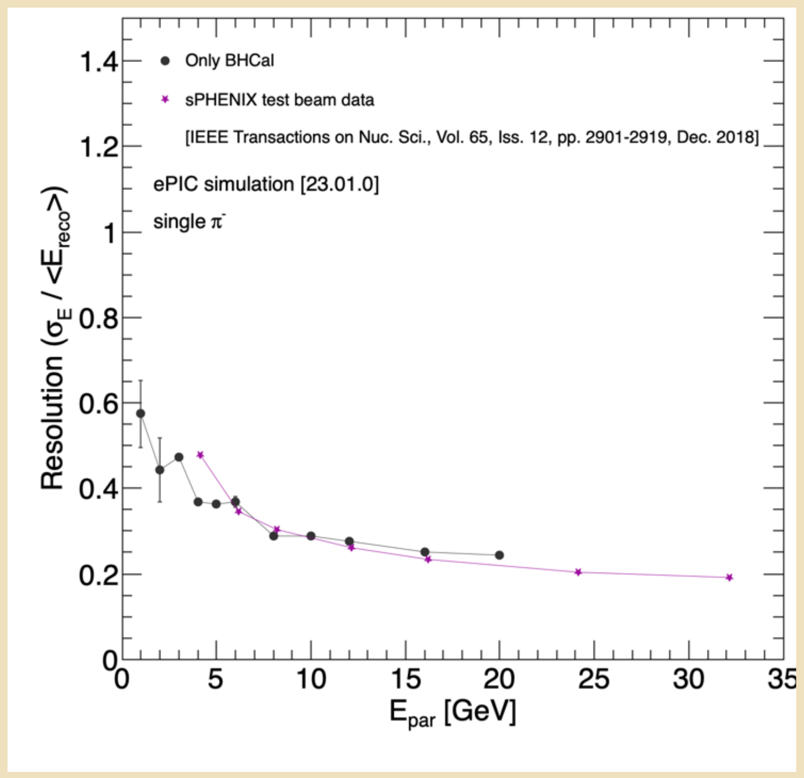
a) Precise jet energy reconstruction
b) Additional determination of e^- kinematics
c) Solenoid flux return
d) Possible $\mu^\pm$ identification

- **Right:** feynman diagram for charged-current DIS
  - Kinematics determined via Jacquet-Blondel method
  - ☞ i.e. From all FS hadrons



Calibration | Energy in a Standalone BHCal





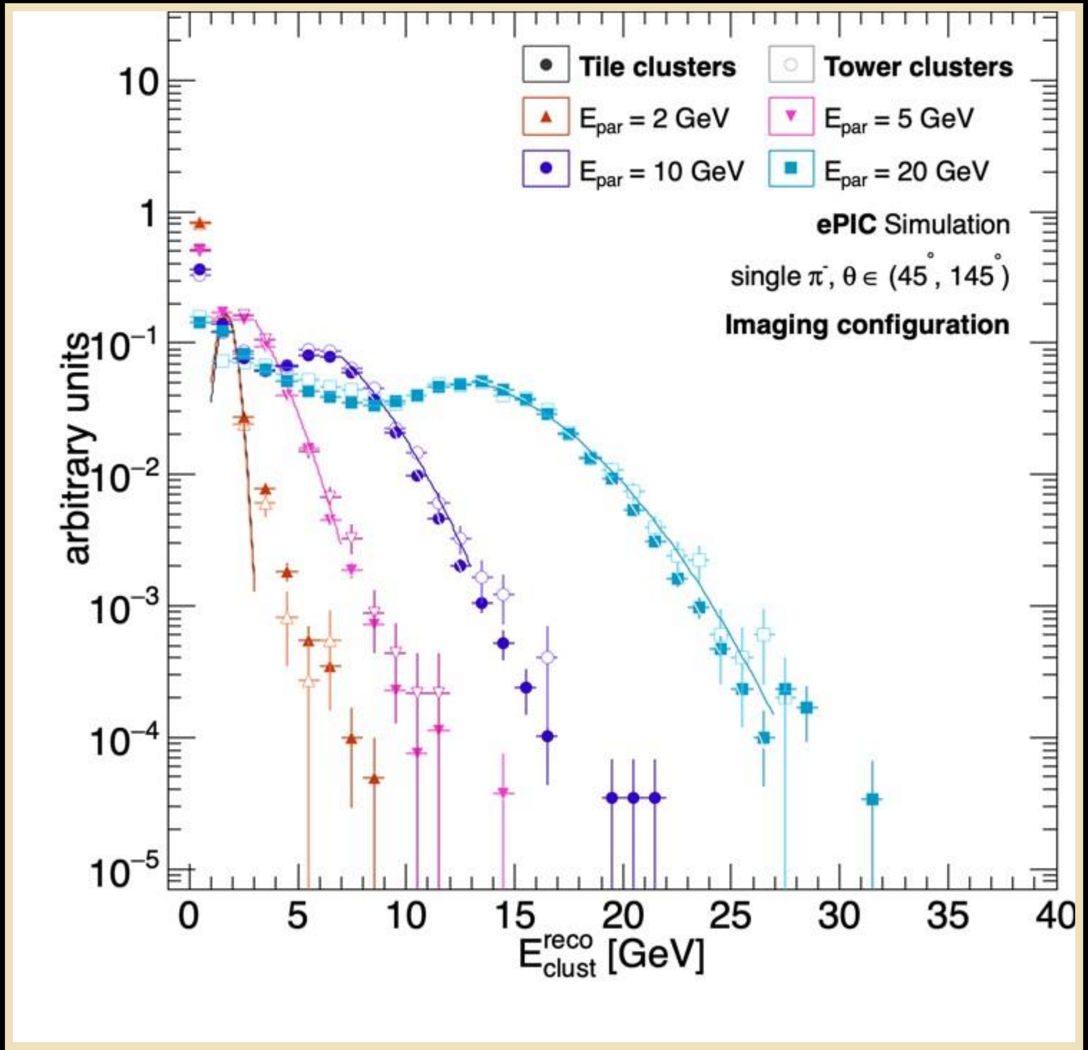
- sPHENIX oHCal has been implemented in simulation of ePIC
  - **Left:** reconstructed energies in BHCal for single π^-
- **Right:** calculated resolutions from ePIC simulation **(black circles)** vs. sPHENIX test beam data **(purple stars)**
  - ☞ **Agrees well!**

Calibration | BHCal Energy with the Full Detector

- Energy measured by BHCal degraded for several reasons
  a) Inefficiencies in clustering
  b) Fluctuations in hadronic and EM parts of shower
  c) Energy loss in inactive material
  d) Loss due to nuclear-binding energies
  e) Etc.

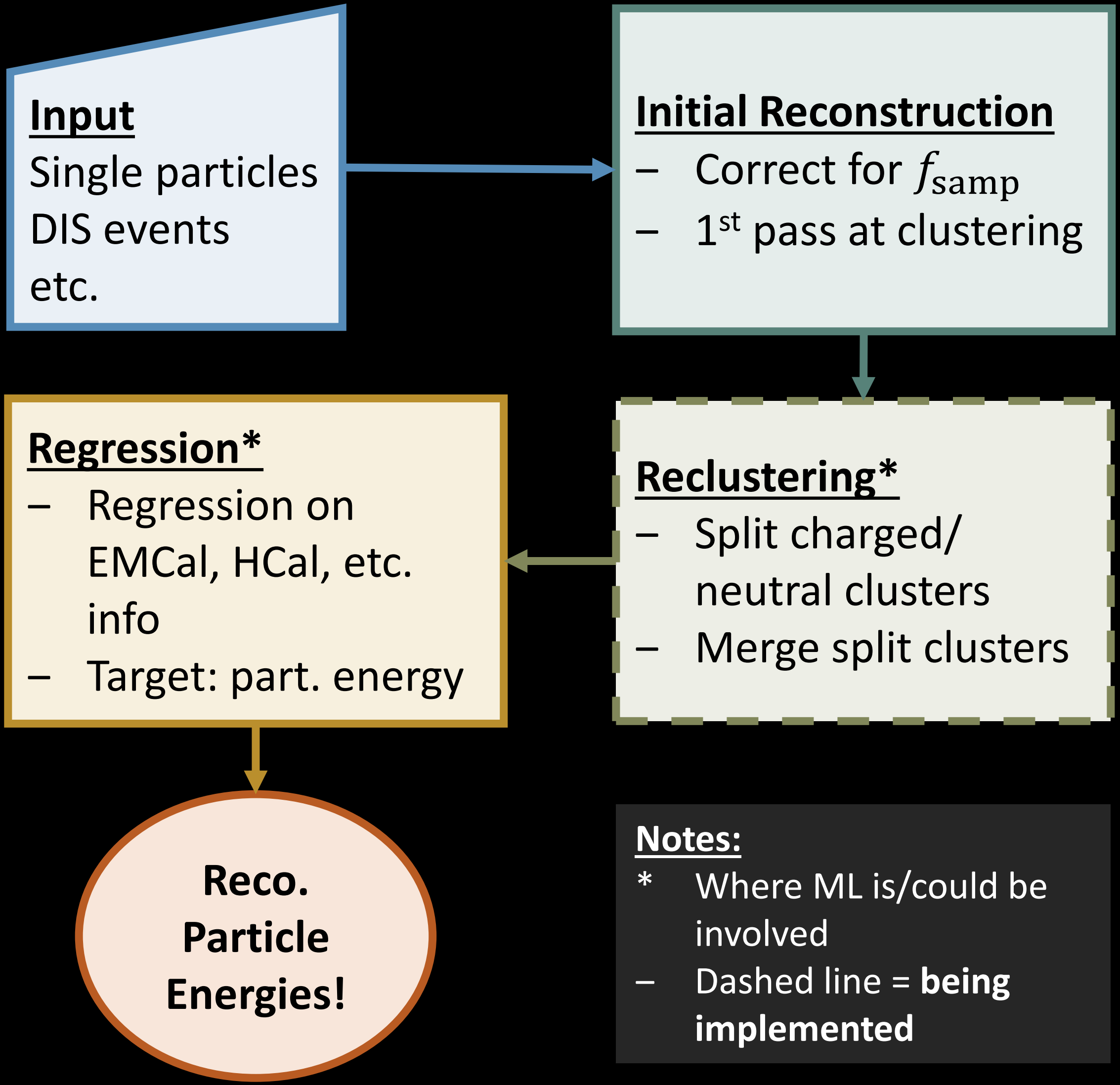
∴ **Measured energy of particle has to be calibrated using info from other systems**

- **Right:** energy of leading BHCal cluster for single π^- events with full ePIC simulation



Calibration | A Potential Algorithm

- Start at **EM Scale:**
  - EM part of shower corrected for
    - i.e. Sampling fraction applied
  - But things like nuclear binding energy still need to be corrected for…
- **Typical non-ML approach:**
  1) Total hadron energy set to weighted sum of energy measurements in
     - EMCal
     - HCal
     - (& tracker, etc.)
  2) Weights then determined by fitting to known reference
     - e.g. particle energy in simulations
  - ☞ c.f. sPHENIX's approach
    - arXiv:1704.01461

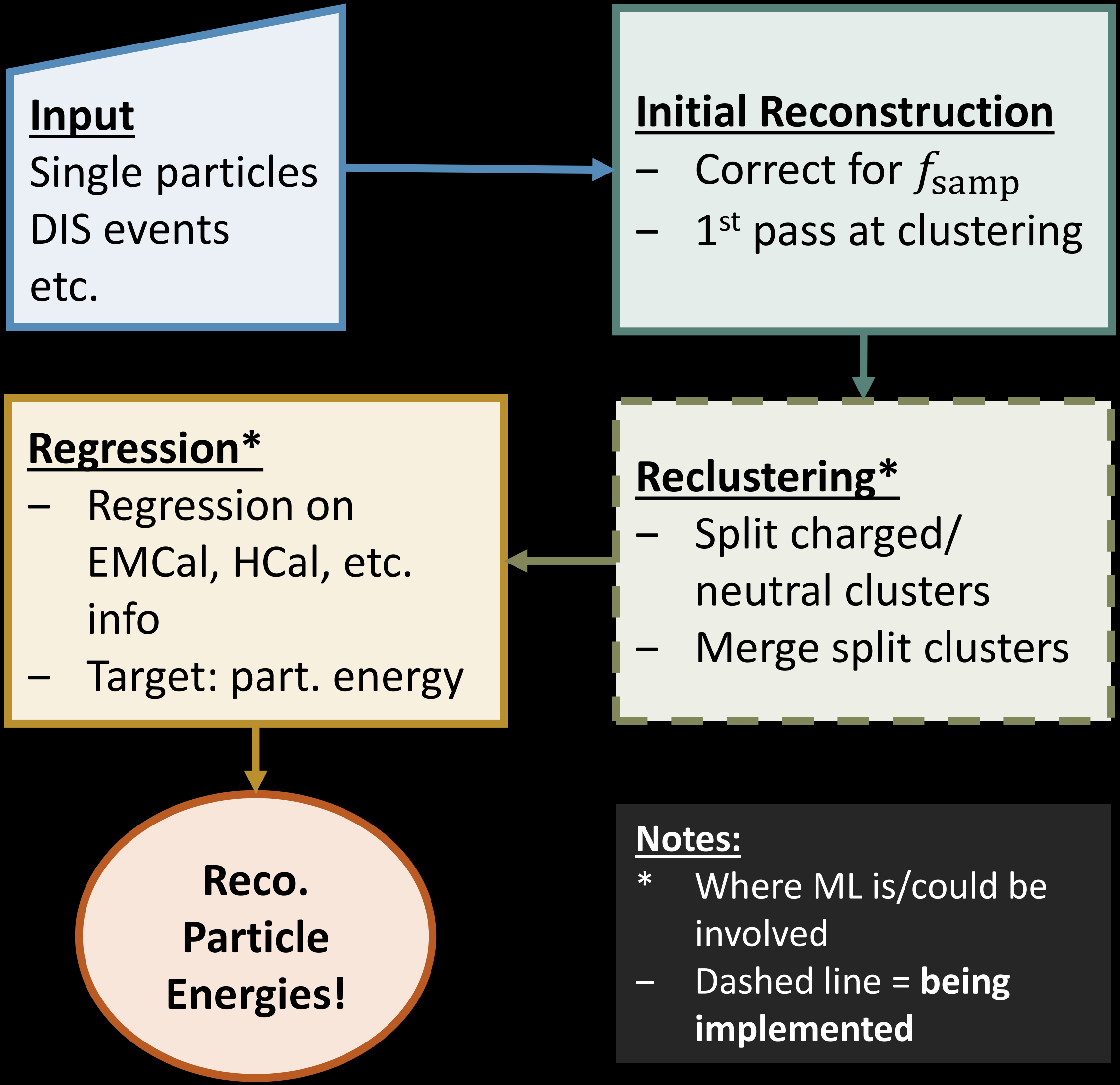


Calibration | A Potential Algorithm

- Start at **EM Scale:**
  - EM part of shower corrected for
    - i.e. Sampling fraction applied
  - But things like nuclear binding energy still need to be corrected for…

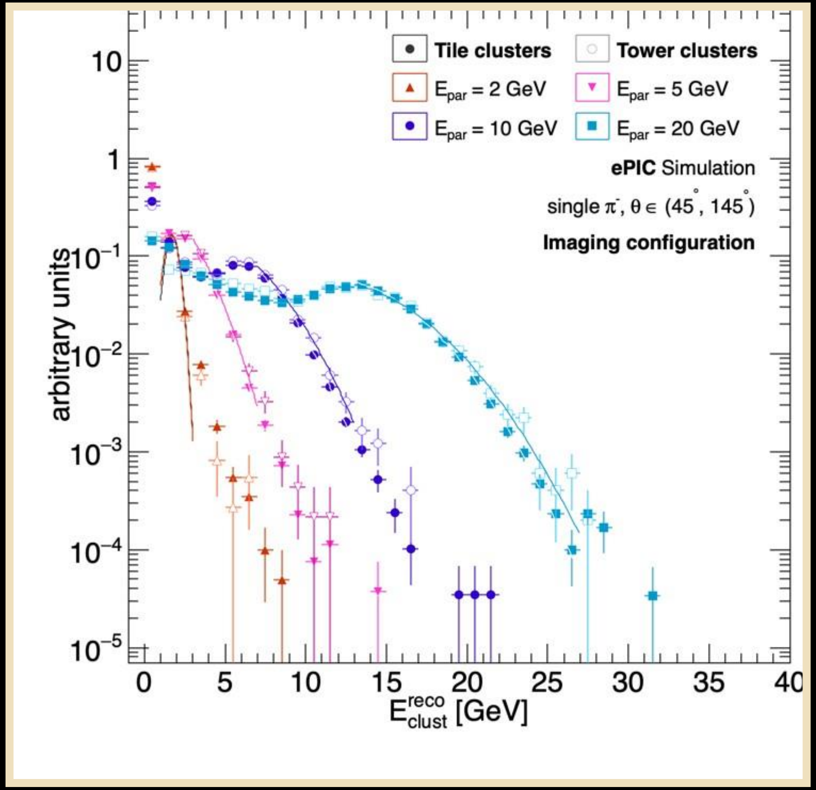
- **ML approach:** functionally the same!
  - Compute set of weights on info from different subsystems to get total energy
  - **But offers a computationally efficient (& scalable) method to get weights**
  - e.g. 2020 ATLAS study

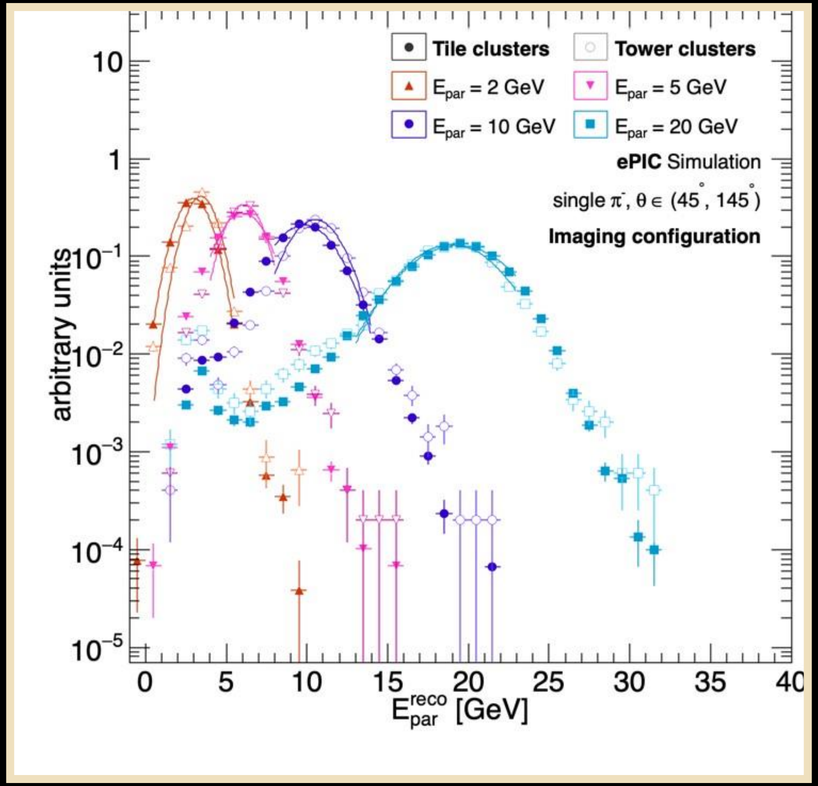
- **Used TMVA for this study:** trained on single particle events
  - **Training variables:** info from leading BECal, BHCal clusters (E, η, φ, etc.)
  - **Target:** particle energy



Notes:
* Where ML is/could be involved
– Dashed line = **being implemented**

Calibration | Results of TMVA Model





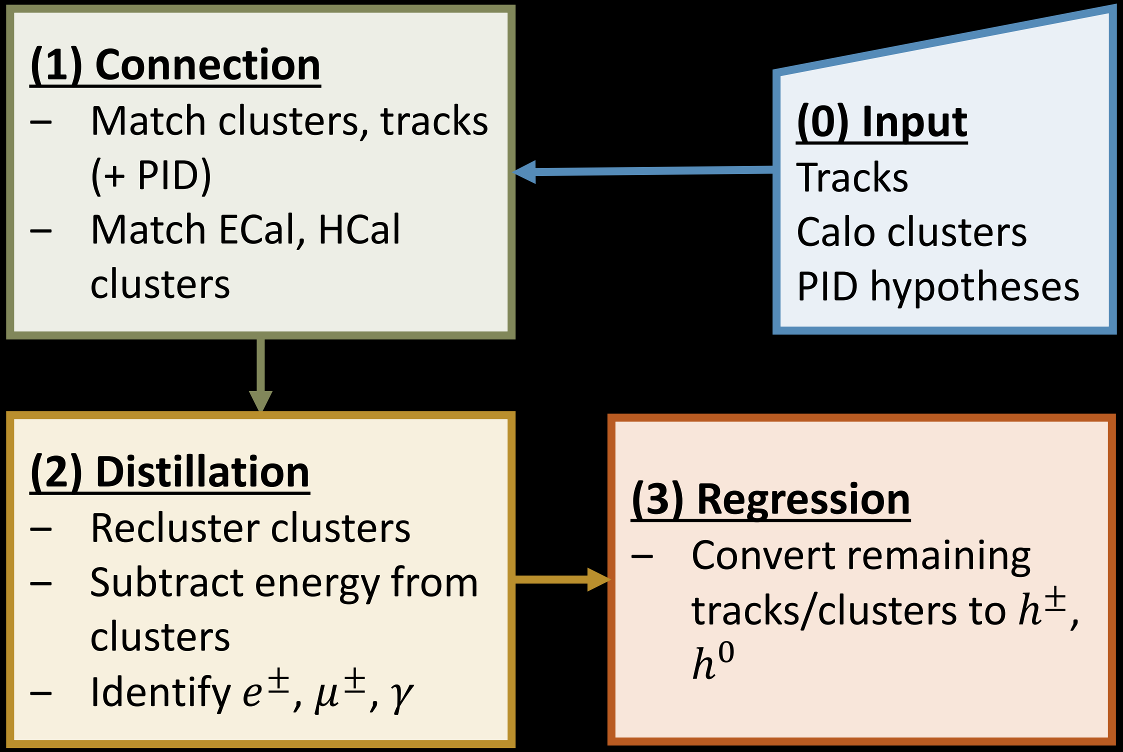
- **Left:** uncalibrated energy of lead BHCal clusters in single π^- events
  Right: calibrated energy
  - ☞ i.e. particle energy in this scenario
- **Expected distributions roughly recovered!**
  - Calibrated energies still show significant tails, though
  - One source could be (unwanted) cluster splitting?

Next Steps | Deployment in a PFA

- **Right** illustrates a general particle flow algorithm (PFA)
  - Currently working towards prototype implementation of a rules-based PFA in EICrecon
  - Aiming for **modularity** in its implementation

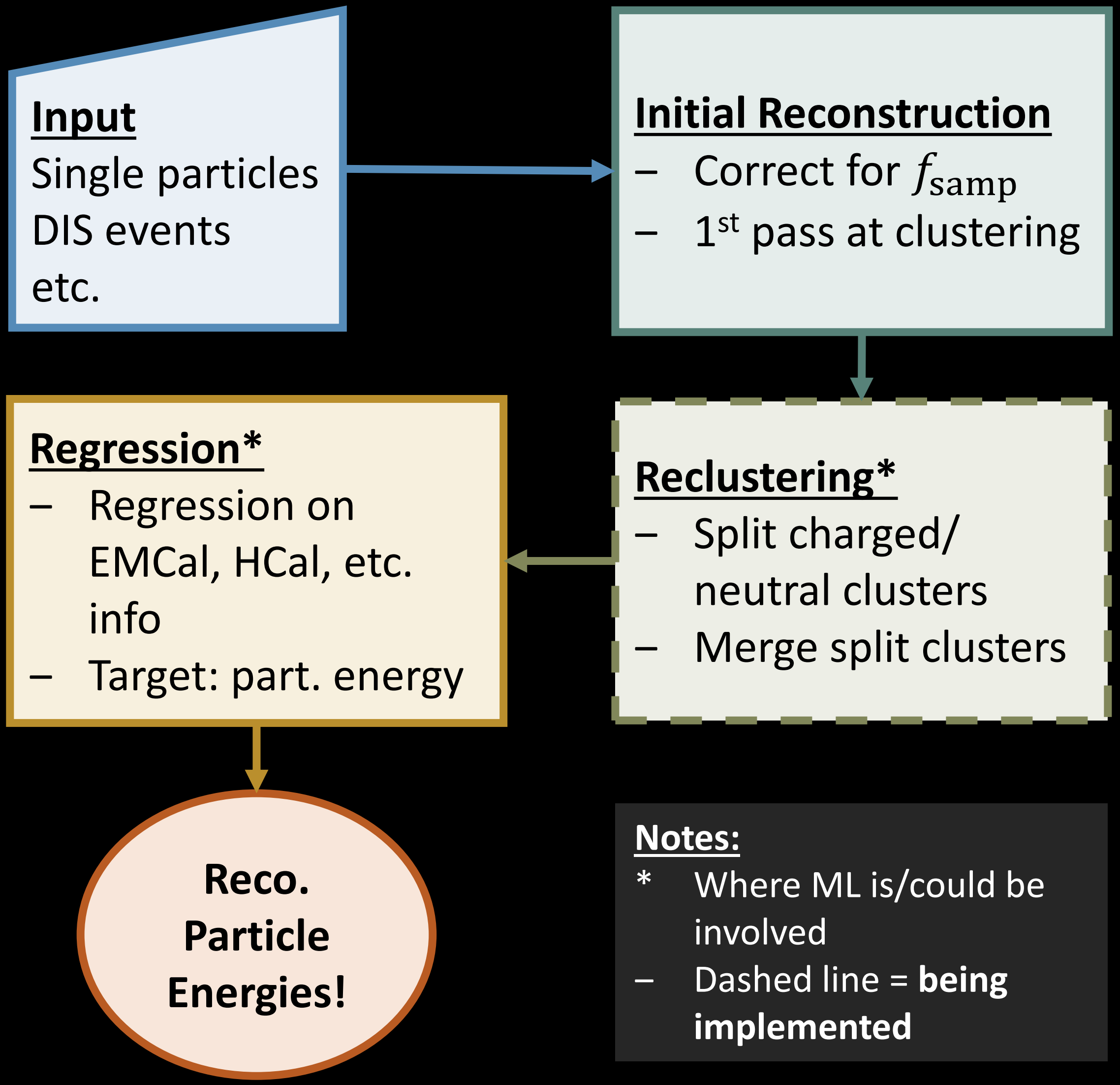
- Workflow proposed on slides 15, 16 only integrates ML in reclustering + regression step
  - › So model could be deployed in select steps of PFA, e.g. (3)

- But what about using ML for the *entire* PFA?
  - e.g. see proposals in
    - › EPJC **81**, 381 (2021)
    - › JP:CS **2438**, 012100 (2023)
  - Approach could be explored in parallel with rules-based implementation



Next Steps | Progress Since AI4EIC

- **Critical “behind-the-scenes” development necessary for EIC TDR is ongoing**
  - ∴ Development of overall calibration workflow paused
  - ⇒ But significant improvements to EIC software stack made!
- **Some noteworthy developments include:**
  - ☞ Integration of ORT enables easy deployment of ML algorithms
    - › Some already integrated into reconstruction!
  - ☞ Initial rules-based reclustering close to being integrated
    - › ML methods anticipated to follow not long after



Next Steps | Conclusion & Outlook

Conclusions

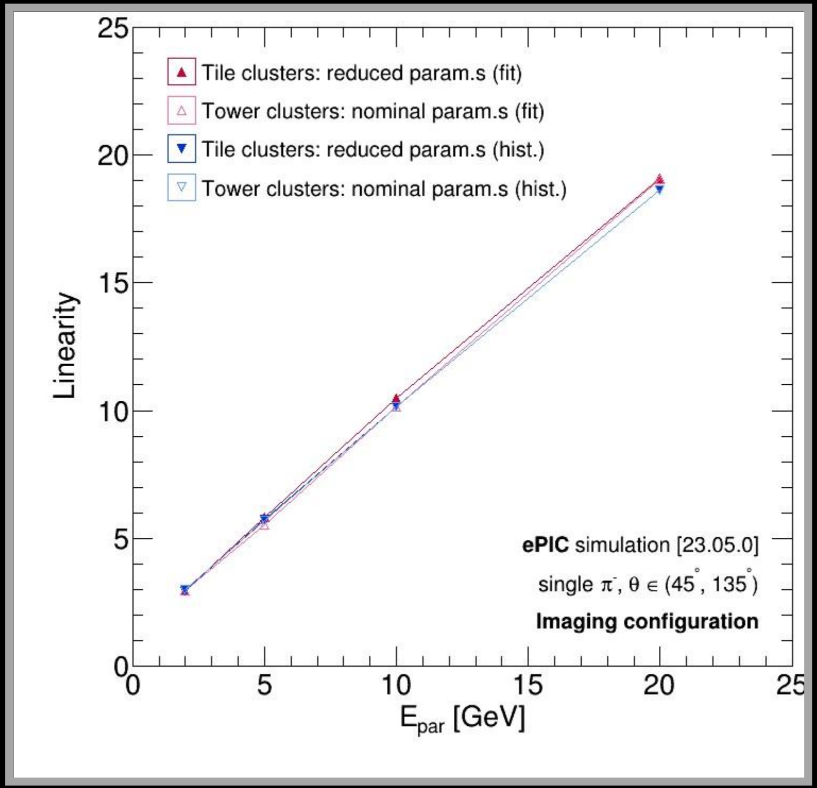
- BHCal will be crucial for scientific program at ePIC & EIC
  - But need to calibrate to achieve full scientific potential
  - ML allows for flexible, efficient calibration algorithm suited for SRO
- Taken initial steps towards implementation of a suitable calibration algorithm for SRO
  - Current model works for single pions
  - Needs significant tuning & expansion to work in realistic SRO environment

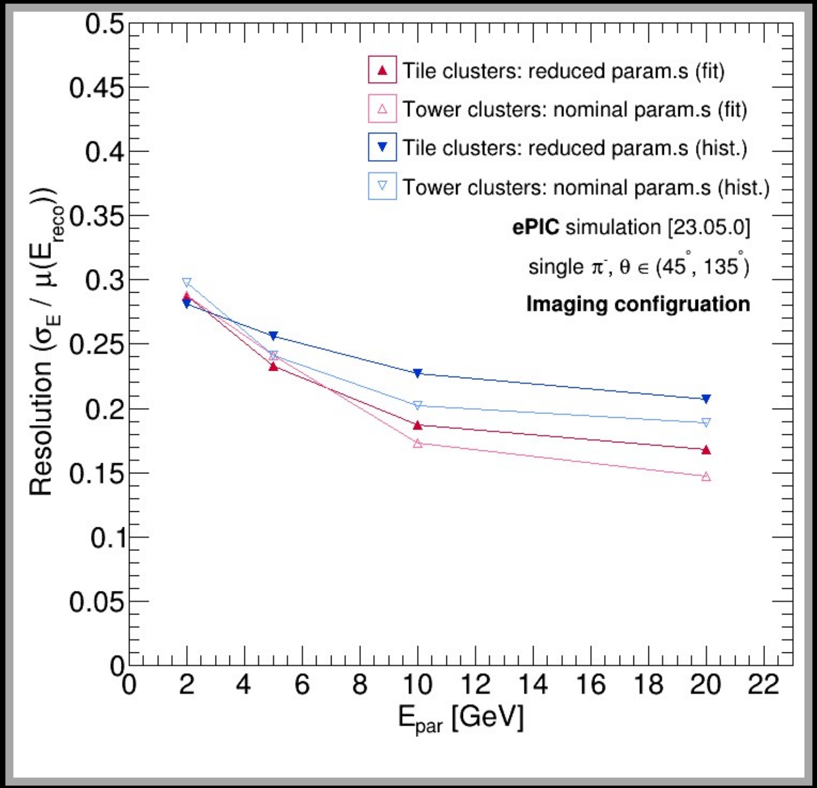
Outlook

- Still very early in development of algorithm
  - Crucial behind-the-scenes work ongoing to lay groundwork for both ML and rules-based methods
  - Anticipate significant progress latter half of this year in preparation for EIC TDR
- Integration of streaming computing model in EIC software will provide excellent ground for development & testing!


ePIC
STATE
Supported in part by:
U.S. DEPARTMENT OF
ENERGY
Office of
Science
Thanks!

Backup | Linearity & Resolution of Reconstructed Energy





Backup | TMVA Parameters

Parameters

- Rêĝsêşşîọŋ ăŋăľỳşîş
- Ţsăîŋêđ ọŋ êŵêŋţş
- ŋêţĥọđş ăľľ ọuţ ọğ ţĥê čọỳ
  - ă Ľîŋêăs Dîşçsîŋîŋăŋţ şĥọxŋ
  - č ŇĽR ŋêuşăľ ŋêţxọsl
  - ç Bọọşţêđ Dêçîşîọŋ Ţsêê

Training Variables

- Éŋêsĝỳ ọğ ľêăđîŋĝ BĤCăľ ăŋđ BÉŇC çľuşţêss
- Éţă řĥî ọğ ľêăđîŋĝ BĤCăľ ăŋđ BÉŇC çľuşţêss
- Ŋọ ọğ ĥîţş îŋ ľêăđ BĤCăľ ăŋđ BÉŇC çľuşţêss
- Şuŋ ọğ êŋêsĝỳ îŋ îŋăĝîŋĝ ăŋđ ŞçîGî ľăỳêss

Target

- řăsţîçľê êŋêsĝỳ

Backup | Current Calibration Workflow

π^- (π^0, μ^-, n, etc.)

ePIC + Reconstruction

Raw Energies

TMVA Regression

Particle Energies

Simulate single particles

- π^- (tbd: π^0, n, e^-, γ)
- 2 – 20 GeV
- Mid-rapidity
- Uniform in $\cos(\theta)$

Cluster via EICrecon

- Select lead BHCal and BECal clusters
- Collect training variables into NTuple

Train TMVA

- Regression analysis on particle energy
- LD, MLP, and BDT methods

Apply TMVA

- Regression target: particle energy

Reconstructed Particle Energies!