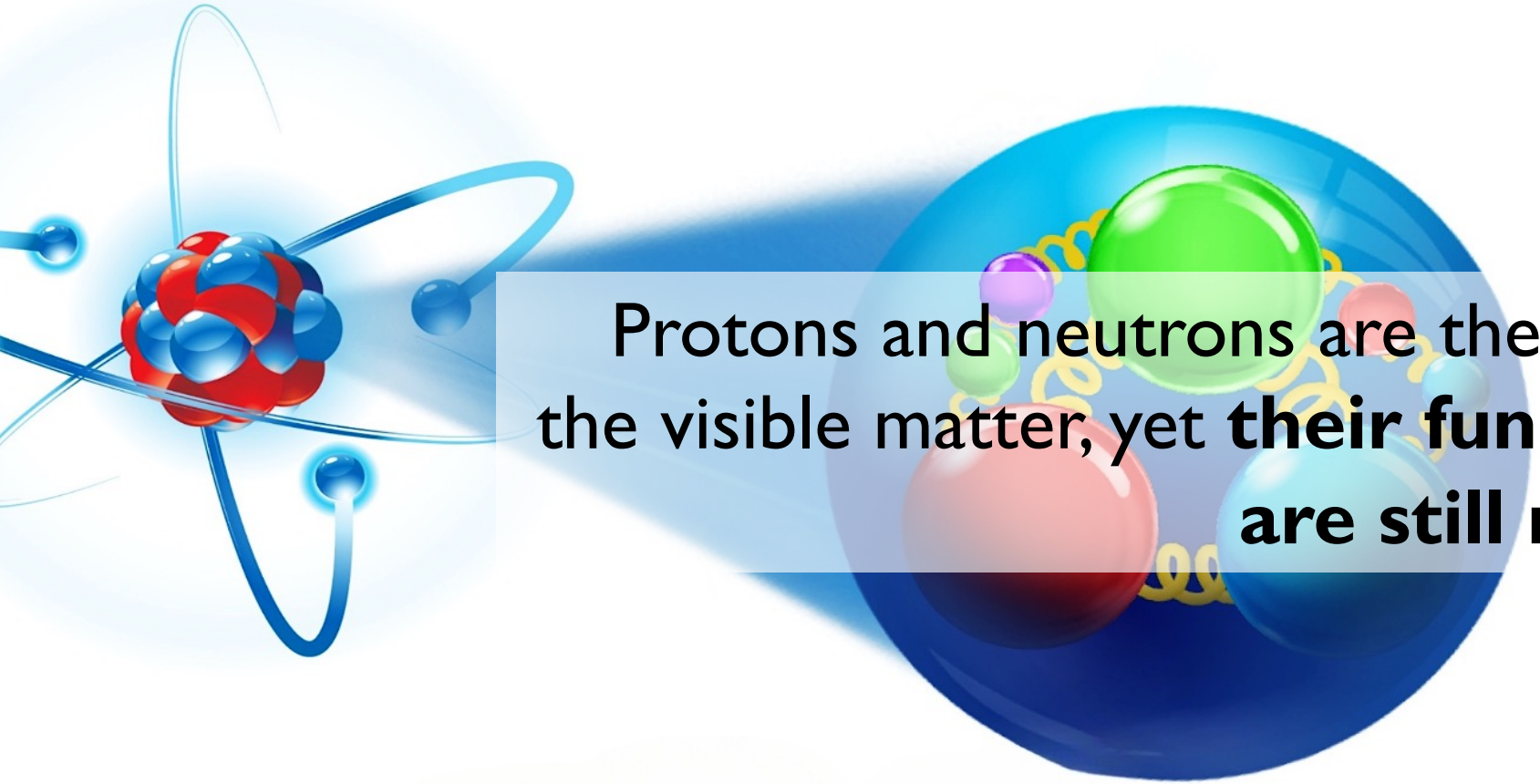
NuSKAT: A Nucleon Structure Knowledge and Analysis Toolkit

A LOI to Horizon-INFRA-2025

Coordinator: Pierre Chatagnon (DPhN, CEA Saclay)

IMT Atlantique, Nantes, France, 1st of July 2025

Fundamental properties of the nucleon



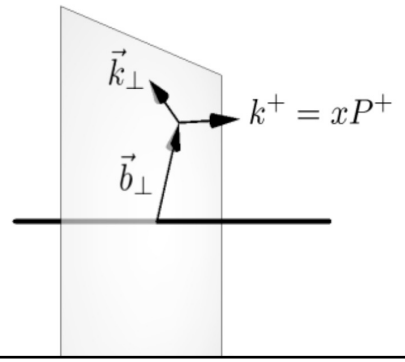
Protons and neutrons are the main building blocks of the visible matter, yet **their fundamental properties are still not fully understood.**

Generalized Parton Distributions

Generalized Parton Distributions

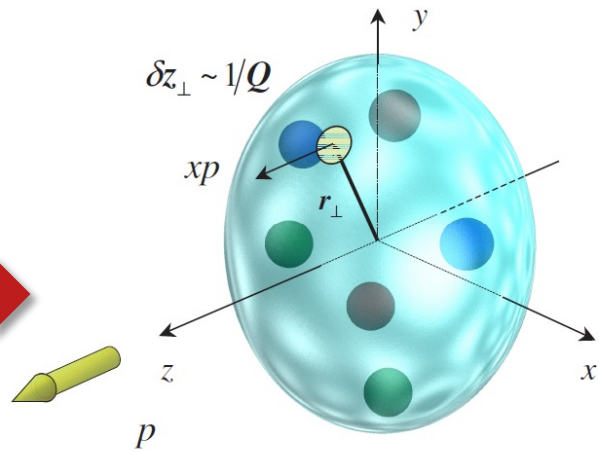
$$H, \tilde{H}, E, \tilde{E}(x, \xi, t)$$

Generalized Transverse Momentum Distributions



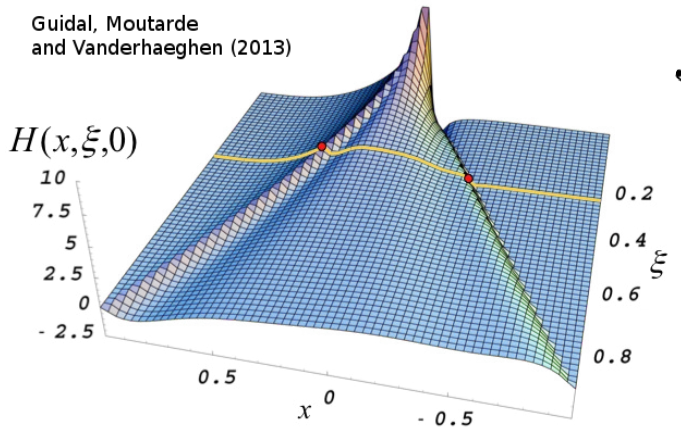
$$W(x, \xi, \vec{k}_\perp, \vec{k}_\perp \cdot \Delta_\perp, t)$$

$\int d^2\vec{k}_\perp$

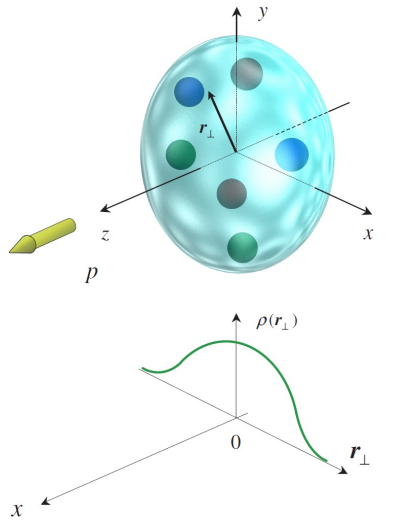


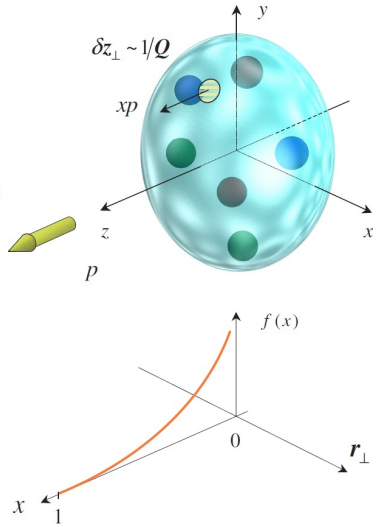
$\int dx$

$\int d^2\vec{b}_\perp$



Electro-magnetic Form Factors





Parton Distribution Functions

A.V. Belitsky, A.V. Radyushkin, "Unraveling hadron structure with generalized parton distributions", Physics Reports, 2005

Probing the fundamental properties of the nucleon…

Spin, Mass and Forces in the nucleon

→ Moments of GPDs and Gravitational Form Factors

$$\int_{-1}^{1} dx\ xH^a(x,\xi,t) = A^a(t) + \xi^2 D^a(t)$$

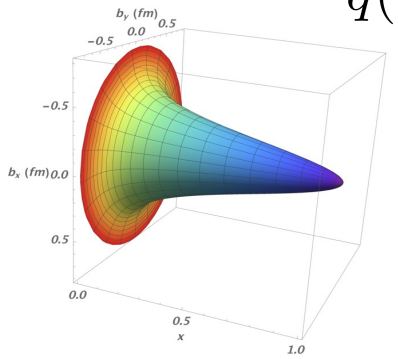
Mass — Spin — Forces

$$\int_{-1}^{1} dx\ xE^a(x,\xi,t) = B^a(t) - \xi^2 D^a(t)$$

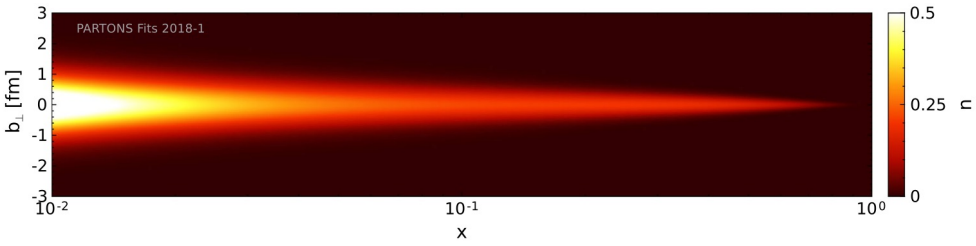
$$\frac{1}{2} = J(0) = \frac{1}{2}(A(0) + B(0)) = \frac{1}{2}\Delta\Sigma + \Delta L$$

Nucleon tomography

$$q(b_\perp, x) = \int_0^\infty \frac{d^2\Delta_\perp}{(2\pi)^2} e^{\Delta_\perp b_\perp} H(x, 0, -\Delta_\perp^2)$$



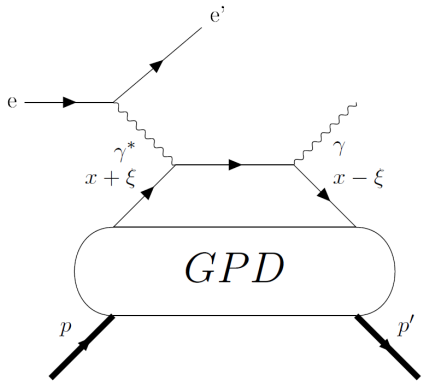
R. Dupré, M. Guidal, M.Vanderhaeghen, PRD95, 011501 (2017)



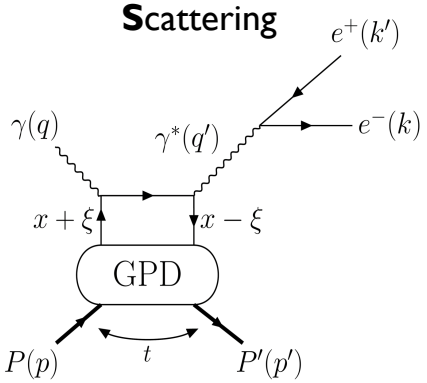
Moutarde, H., Sznajder, P. & Wagner, J. Border and skewness functions from a leading order fit to DVCS data. *Eur. Phys. J. C* **78**, 890 (2018)

… via the experimental measurement of exclusive reactions

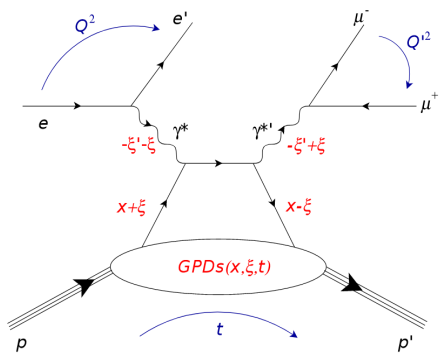
Deeply Virtual Compton Scattering



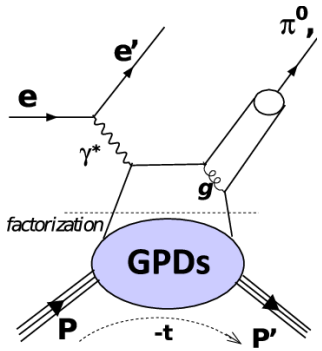
Timelike Compton Scattering



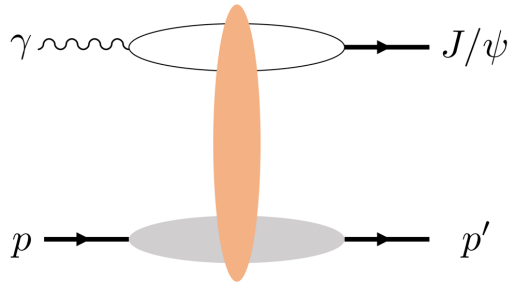
Double DVCS



Deeply Virtual Meson Production



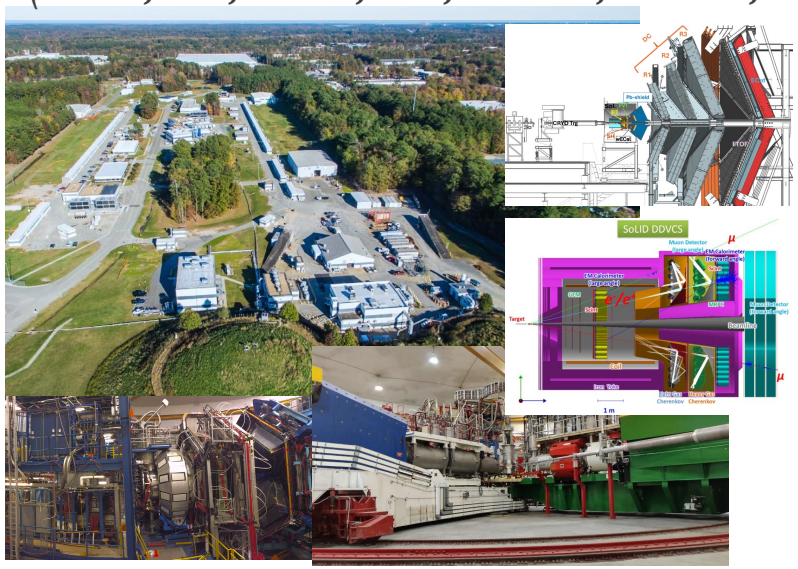
Vector Meson Production



A diverse community…

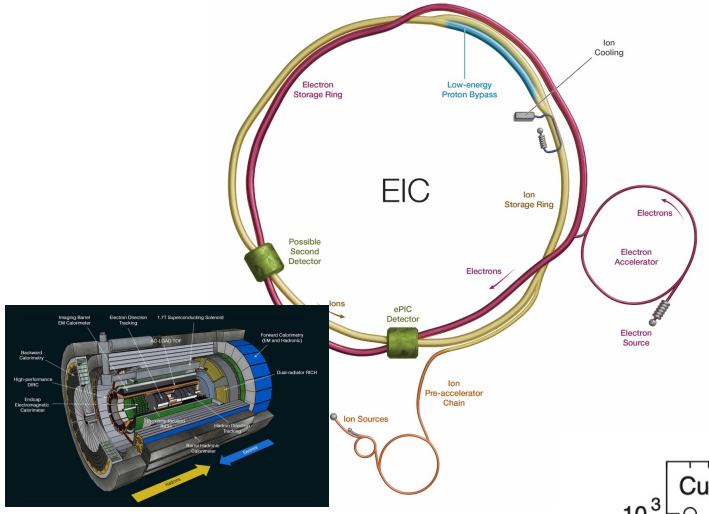
Jefferson Lab (USA)

(DVCS, TCS, DVMP, VM, DDVCS, Sullivan,…)



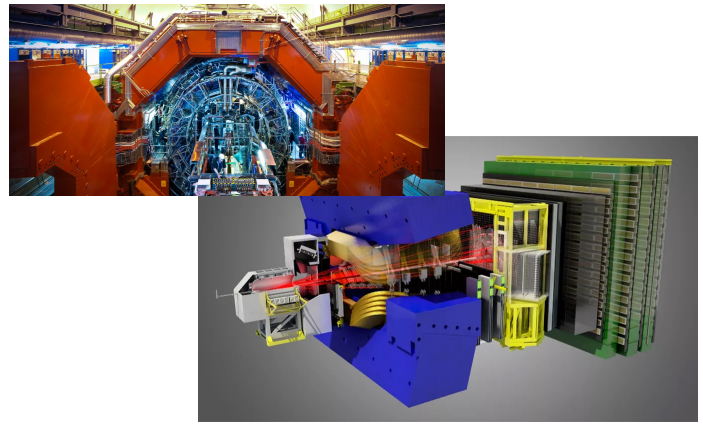
ePIC @ EIC (BNL, USA)

(DVCS, TCS, DVMP, VM production,…)



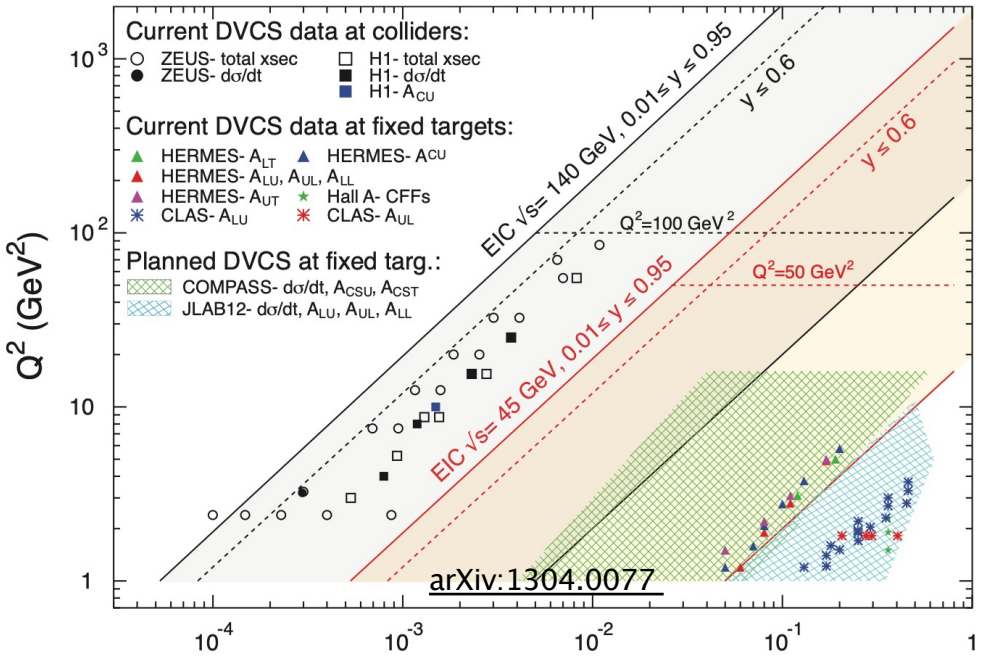
UPC programs @ LHC

(VM photoproduction, TCS)



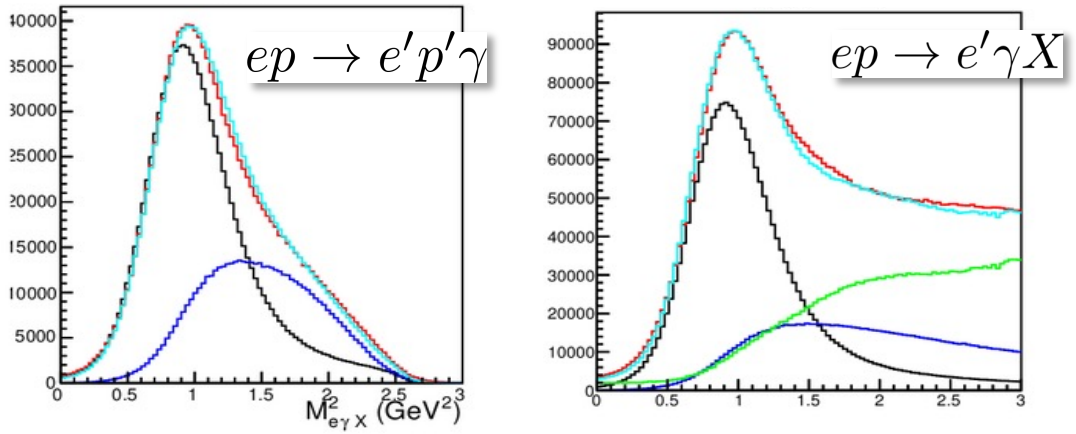
… sharing common challenges:

- Complex event topologies:
  - Particles in multiple detectors with **different efficiency, resolution, and physics background**
  - Often lead to limitation of the available phase space
- Small cross-section processes:
  - Particle Identification is critical (both **efficiency** and **purity**)
- **Large scale** simulations of complex detectors needed.



Examples of experimental challenges of exclusive measurements

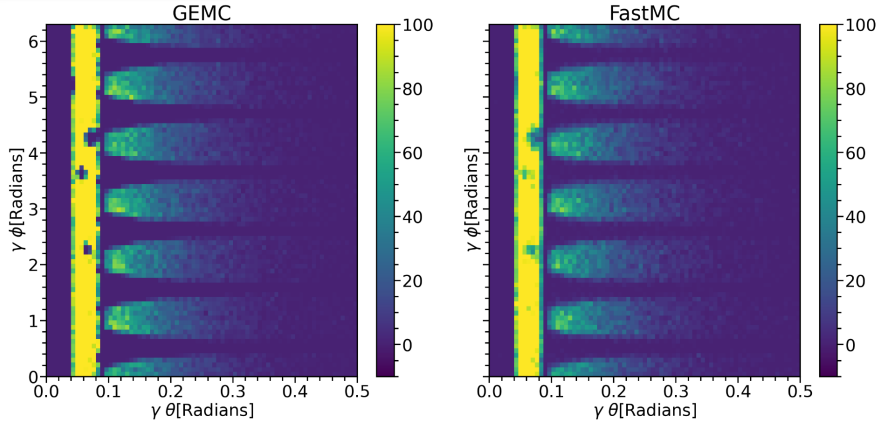
Background in proton DVCS @ CLAS12



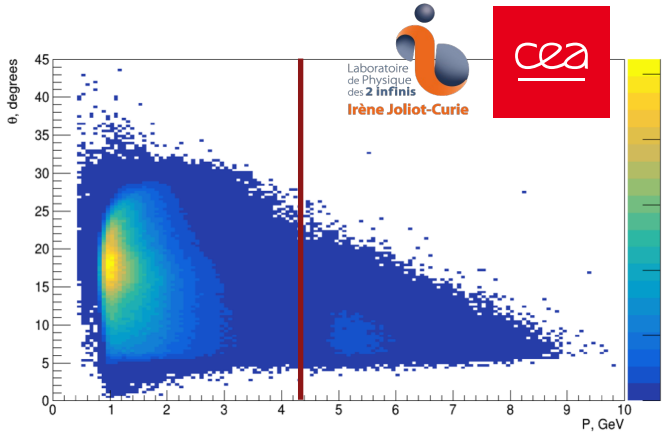
Particularly relevant for future transversely polarized experiment with CLAS12 @ Jlab

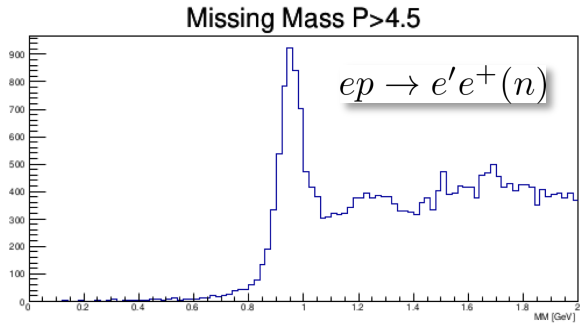
Photon simulation in DVCS @ CLAS12

$ep \to e'p'\gamma$



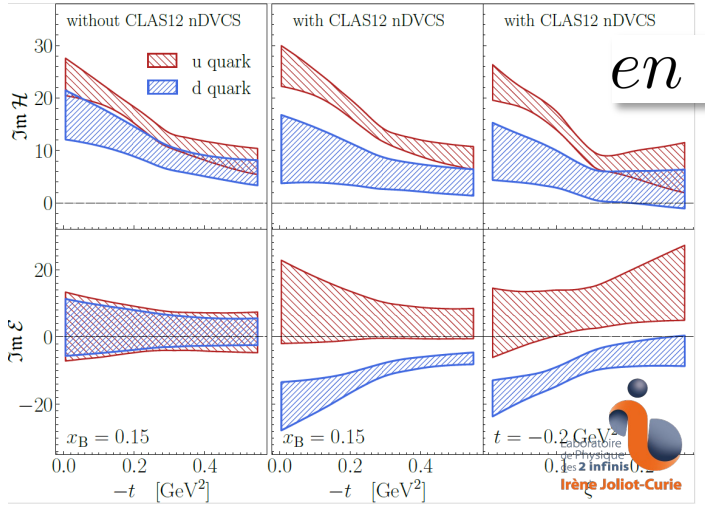
Lepton identification @ CLAS12





→ Solved using calorimeter segmentation and ML algorithms

Interpretation of the data



$en \to e'n'\gamma$

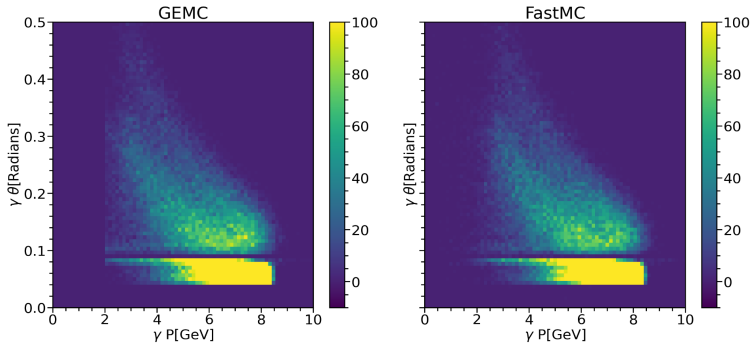
Strong need to quantify the impact of current data, new analysis techniques and future measurements.

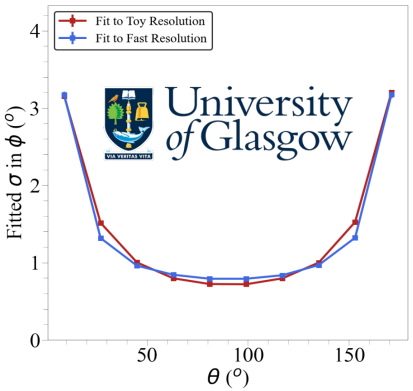
Phys.Rev.Lett. 133 (2024) 21, 211903

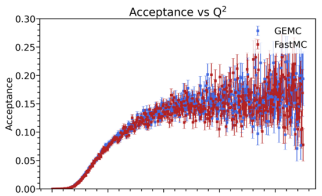
The NuSKAT packages

Efficient data processing

Fast MC & unfolding





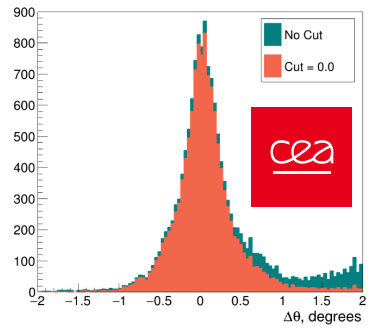


Laboratoire de Physique des 2 infinis Irène Joliot-Curie

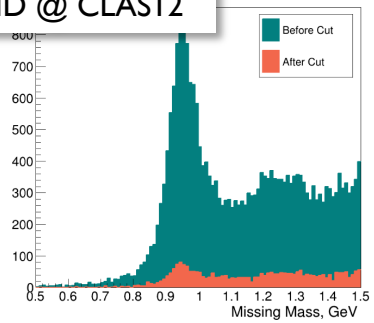
cea

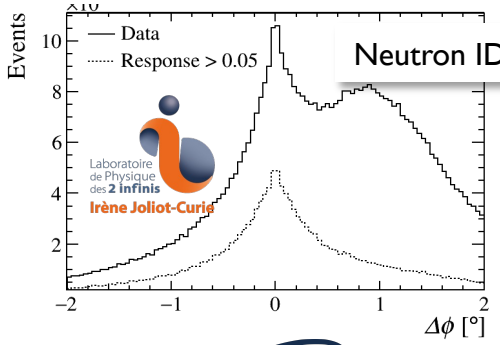
- Reliable fast MC, including realistic resolution, and fast unfolding.
- Reduced computing resources usage.

ML-based Particle identification



Lepton ID @ CLAS12

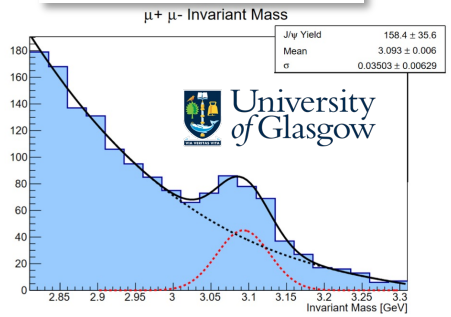




Neutron ID @ CLAS12

Laboratoire de Physique des 2 infinis Irène Joliot-Curie

Muon ID @ CLAS12

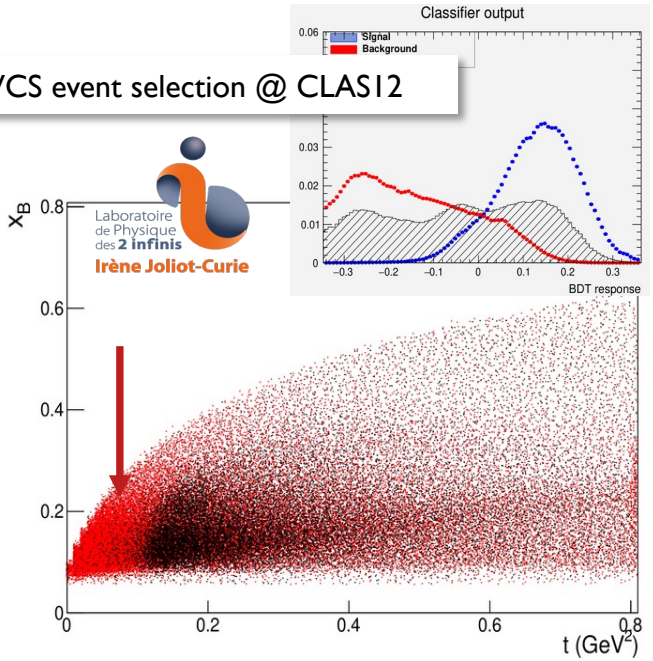


INFN

- Usage of ML have been a proven strategy to identified neutrons, and leptons including muons
- Mutualization of the ML-PID knowledge
- Common and baseline tools for future experiments

ML-based event selection

DVCS event selection @ CLAS12



INFN

University of Glasgow

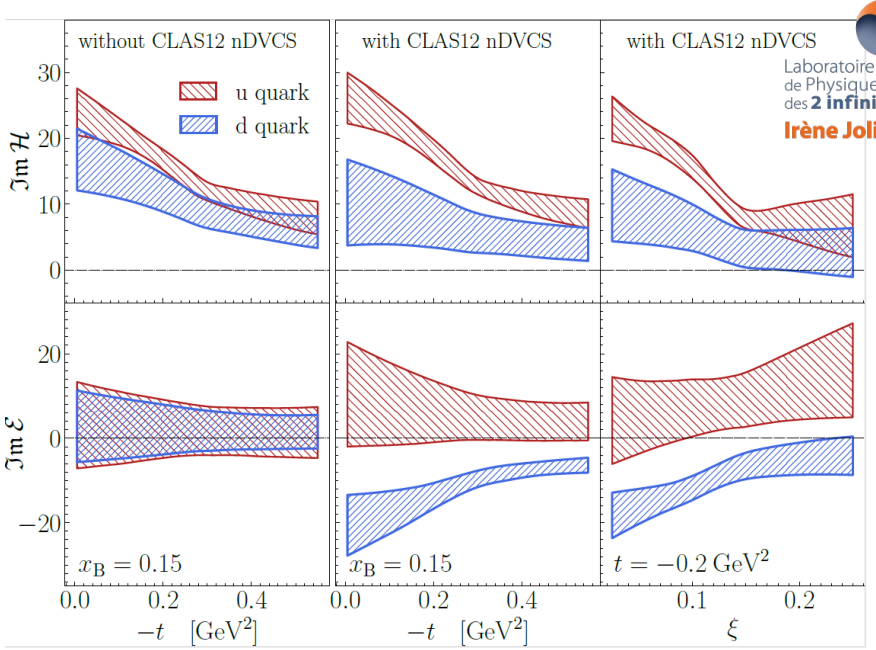
- Exclusive reactions analysis often relies on exclusivity cuts
- Use of ML leads to more efficient cut, potentially allowing new phase spaces

The NuSKAT packages

Interpretation of data

CFFs extraction

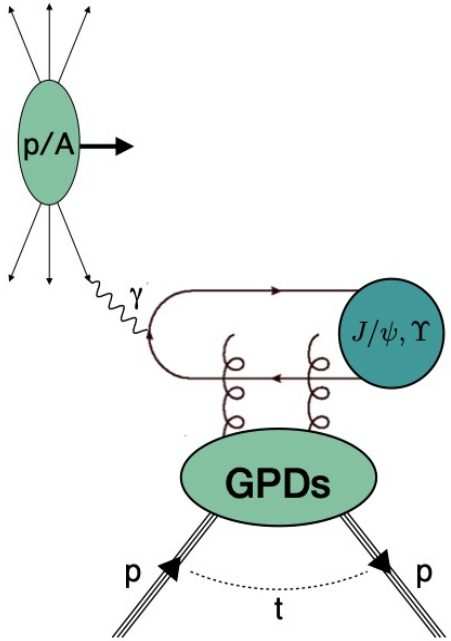
- Existing Neural Network procedure to fit DVCS data.
  - K. Kumericki et al., JHEP 07, 073531 (2011)
  - M. Cuic, K. Kumericki, et al., Phys. Rev. Lett. 533 125, 232005 (2020)
- Need to include new data and future experiments (TCS, DDVCS, UPCs, VM) → PARTONS, GEPARD

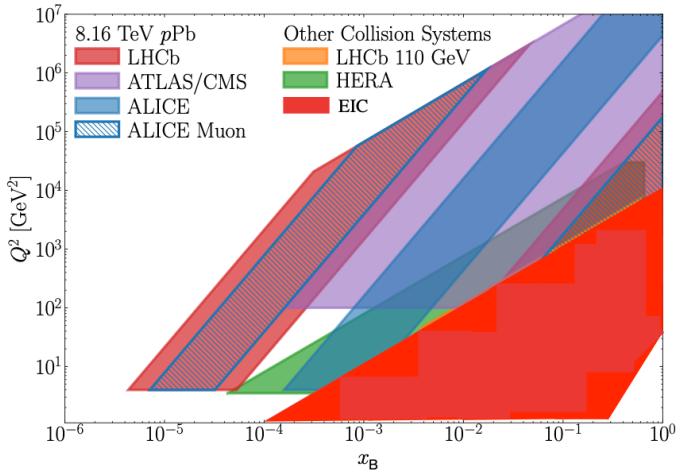


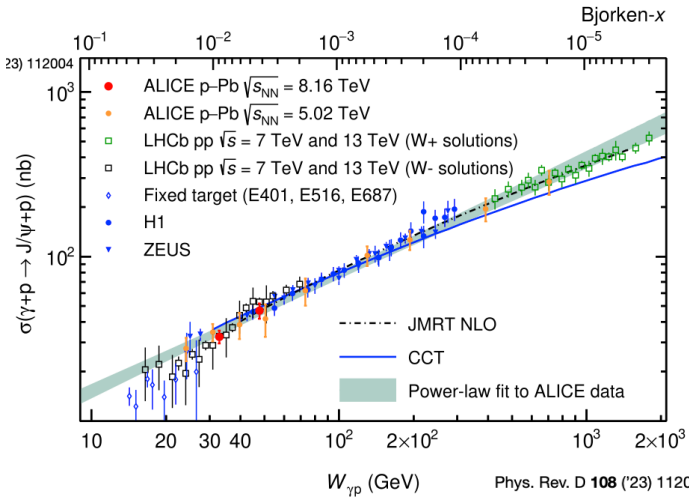
Extraction of quark-CFFs using the CLAS12 p,n DVCS data and neural networks (*Phys.Rev.Lett.* 133 (2024) 21, 211903)

Ultra-Peripheral Collision & extension of 3DPartons

- UPC measurements at the LHC can constrain the GPDs at very low x.
- Need for event generator and phenomenology tools
  → PARTONS, EPIC, GEPARD
- Complementarity with the EiC.







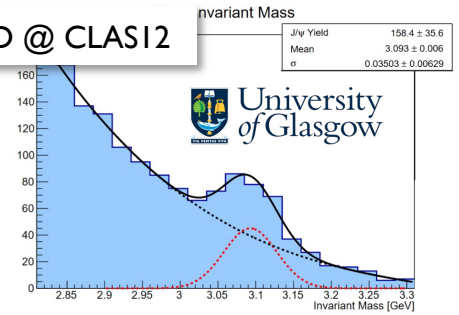
Deliverables and milestones

New tools to facilitate current and future exclusive measurements

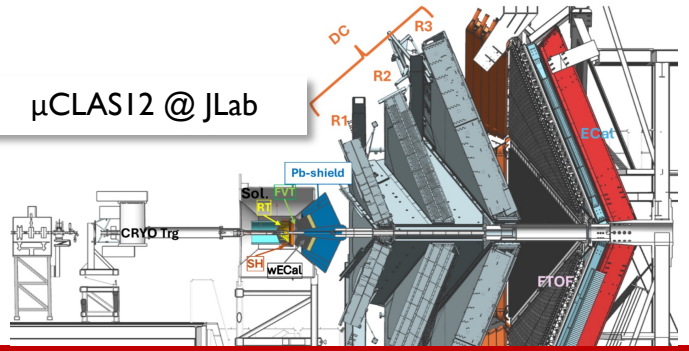
- Capitalize on existing tools and methods developed in the community to build a common framework and new tools

An example: Muon ID @CLAS12 and @ePIC

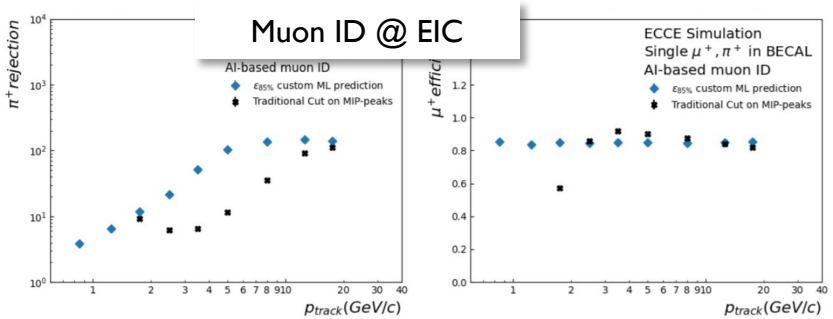
Muon ID @ CLAS12



μCLAS12 @ JLab



Muon ID @ EIC



Allaire, C., Ammendola, R., Aschenauer, EC. *et al.* Artificial Intelligence for the Electron Ion Collider (AI4EIC). *Comput Softw Big Sci* **8**, 5 (2024)

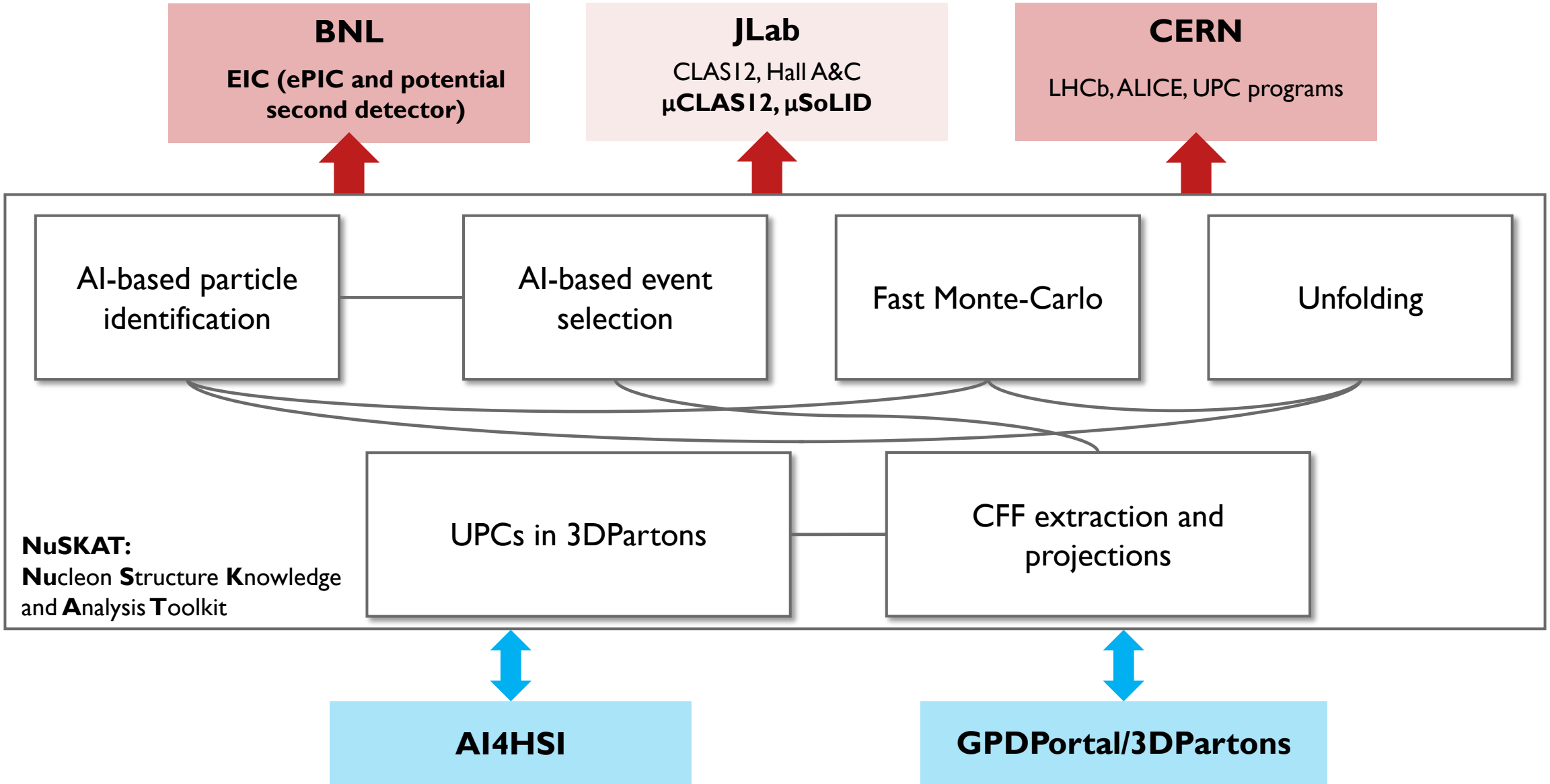
Promote the use of these tools

- Publicly available tools on a **dedicated website**.
- Enforce FAIR (Findable, Accessible, Interoperable, Reusable) principles within the NuSKAT project.

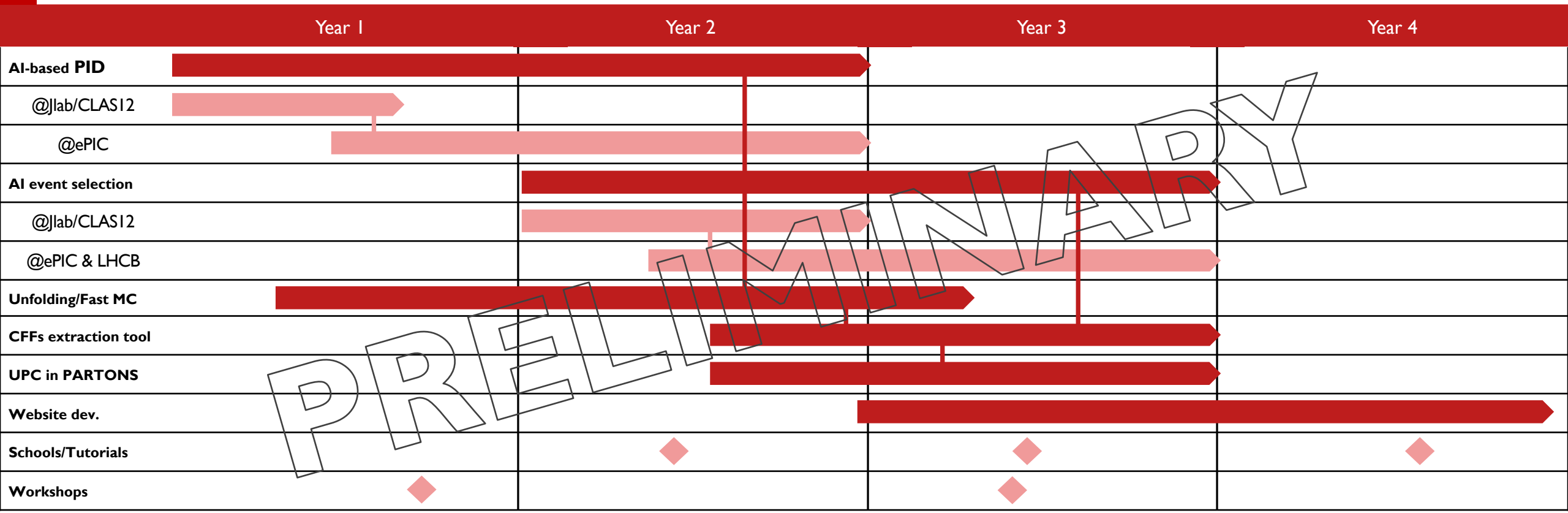
A strong community engagement, with emphasis on tutorials

- Strengthen the community by organizing multiple workshops (such as AI4EIC) and tutorials.
- Dedicated budget for collaborative work.

Synergies


BNL
EIC (ePIC and potential second detector)
JLab
CLAS12, Hall A&C
μCLAS12, μSoLID
CERN
LHCb, ALICE, UPC programs
AI-based particle identification
AI-based event selection
Fast Monte-Carlo
Unfolding
UPCs in 3DPartons
CFF extraction and projections
NuSKAT:
Nucleon Structure Knowledge and Analysis Toolkit
AI4HSI
GPDPortal/3DPartons

Timeline & budget



Cost breakdown

Total requested budget: 760k€

- 6 years of postdocs: 630k€
- Travels support (conferences and collaborative work): 80k€
- Workshop, schools and tutorial organization: 50k€
  2 schools, 2 workshops and 1 tutorials over 4 years

Key takeaways

Experimental measurements of exclusive reactions are crucial to understand the **fundamental properties of the nucleons**.

Current and future measurements need to fully **exploit the possibilities provided by ML**, aiming for **precision physics with large data sets**.

NuSKAT aims at providing a **framework and collaborative effort** to address clearly defined challenges with: impact in multiple experiments, links with other infrastructures, and a **strong community engagement**.