AI application to Hadron Spectroscopy

Marco Spreafico

on behalf of A(i)DAPT Working Group

A(i)DAPT

AI for Data Analysis and PreservaTion

A(i)DAPT motivation

NP/HEP experiments are affected by the following problems:
- Data are affected by detector's effects
- Multidimensional problems reduced in lower-dimensional spaces
- Large datasets are difficult to manipulate and preserve

Can AI support NP/HEP experiments and allow to extract physics form data in a more efficient way?

A(i)DAPT
AI for Data Analysis and PreservaTion

Develop AI-supported procedures to:
- Unfold detector effects
- Accurately fit data in multiD space
- Generate synthetic data with same properties as real data

Collaborative effort:
- ML experts (ODU, JLab)
- Experimentalists (JLab Hall-B)
- Theorists (JPAC, JAM)

Detector unfolding

Detector effects make measured observables (detector-level) different from “true” observables (vertex-level)

ACCEPTANCE

Measurement only access a limited portion of phase space.

- Interpolation: holes in phase space
- Extrapolation: border of phase space

RESOLUTION

Experimental resolution can hide or wash out effect searched for.

- Unresolved spikes
- Measurement extend to unphysical region

Mitigation strategy:

- **Acceptance**: train over *”fiducial volumes”* to exclude poorly / un-measured regions to verify training convergence
- **Resolution**: closure test with a realistic detector model using detector proxy (e.g. GEANT)

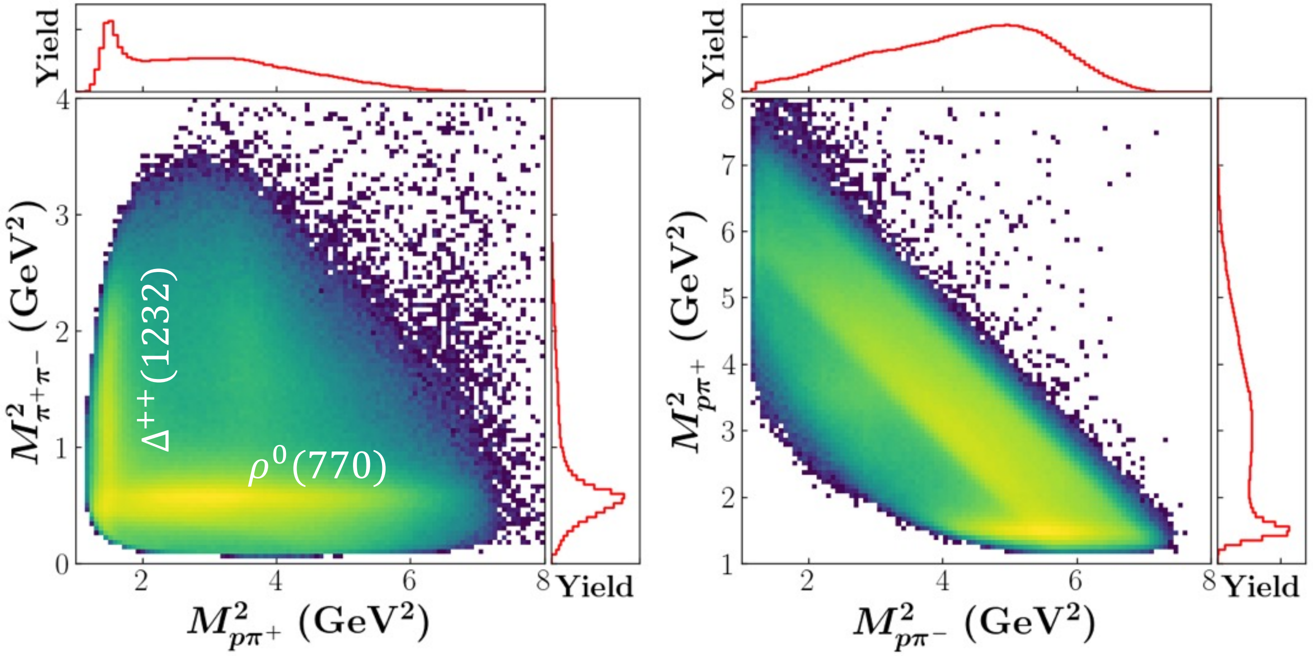
Exclusive reactions: 2 → 3

$\gamma p \rightarrow \pi^+\pi^- p$ **(unpolarized)**

- Fully known initial state
- (3x3) – 4 = 5 Independent variables
- Possible choice: $M^2_{\pi\pi}, M^2_{p\pi}, \theta_\pi, \alpha, \phi$

CLAS g11 2π photoproduction:

- $E_\gamma = (3 - 3.8)\ GeV$
- Main contribution to dynamics:
  - ρ^0photoproduction
  - Δ^{++} resonance excitation



Resonances can be hidden if the wrong set of variables is chosen

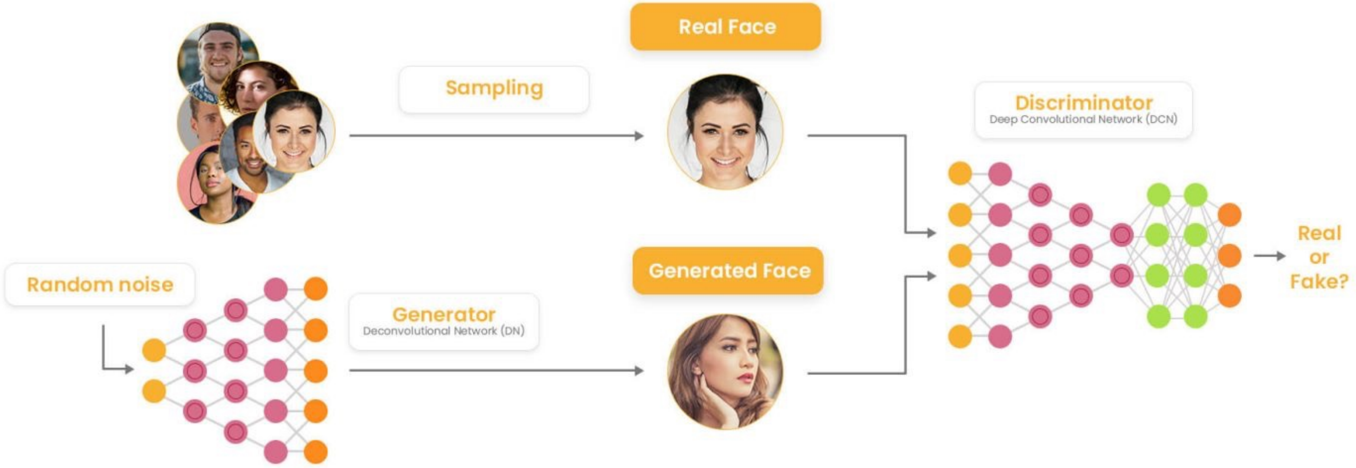
$$\frac{d^5\sigma(\gamma p \rightarrow p\pi^+\pi^-)}{dM_{\pi\pi}\, dM_{p\pi}\, d\cos(\theta)\, d\alpha\, d\phi}$$

AI could provide a new way to look at data and to extract observables and physics interpretation

Credit: Y.Alanazi Awadh, , P..Ambrozewicz, G. Costantini A.Hiller Blin, E. Isupov, T. Jeske, Y.Li, L.Marsicano W. Menlnitchouk, V.Mokeev, N.Sato, A.Szczepaniak, T.Viducic

Generative Adversarial Networks (GANs)

Generative model based on the competition between two Neural Network: **generator** vs **discriminator**



Generator

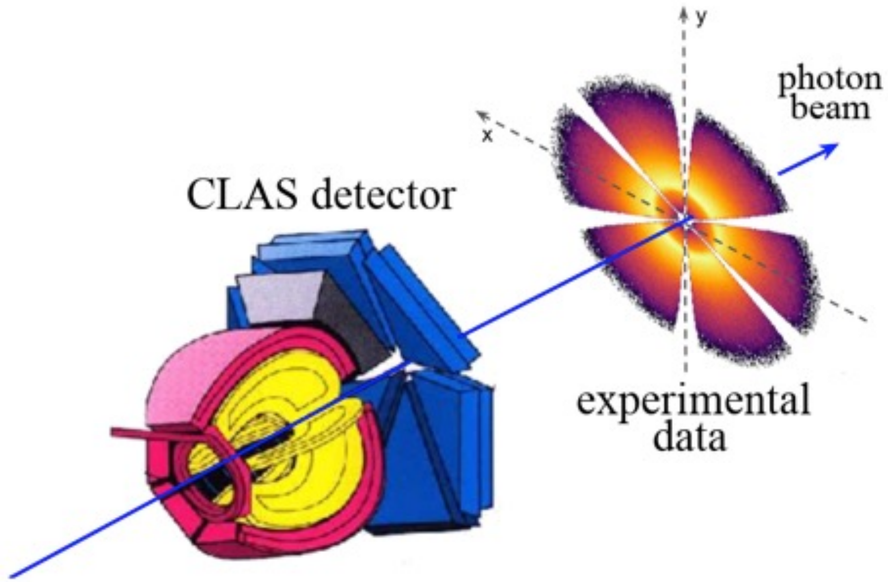
Produce synthetic data that mimic real data
- Can retain high dimensional correlations
- Can produce realistic pseudo-data quickly

Discriminator

Distinguish between synthetic and real data

MultiD case: exclusive 2π photoproduction

M. Battaglieri *et al.* (CLAS Collaboration)
Phys. Rev. Lett. 102, 102001
M. Battaglieri *et al.* (CLAS Collaboration)
Phys. Rev. D 80, 072005

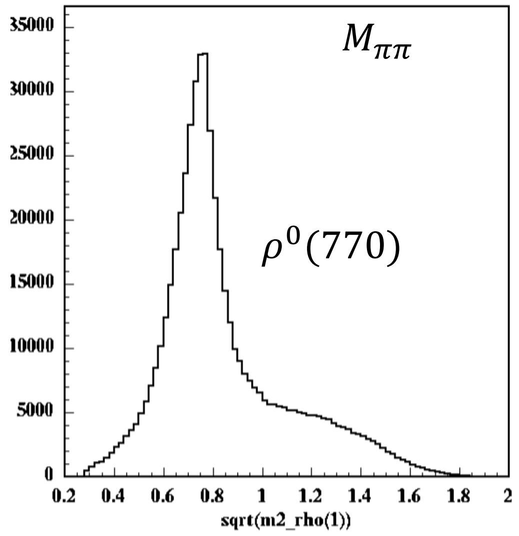


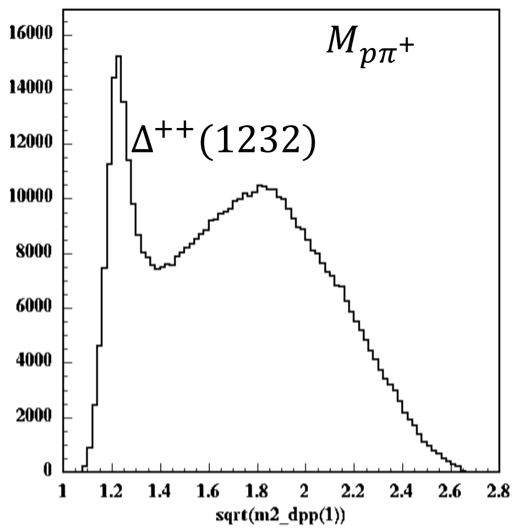
CLAS g11 kinematics

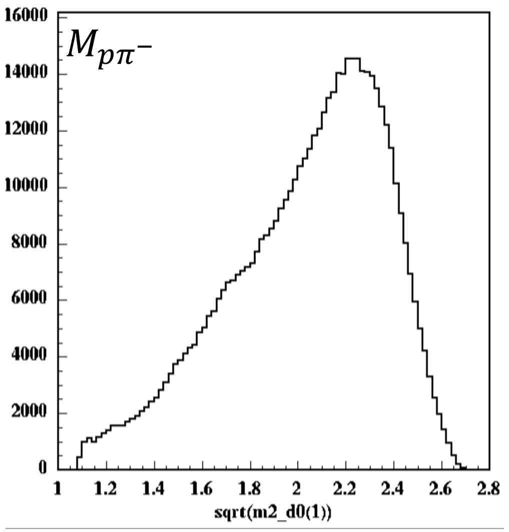
Used same fiducial cuts on (p, θ, ϕ) as in published analysis

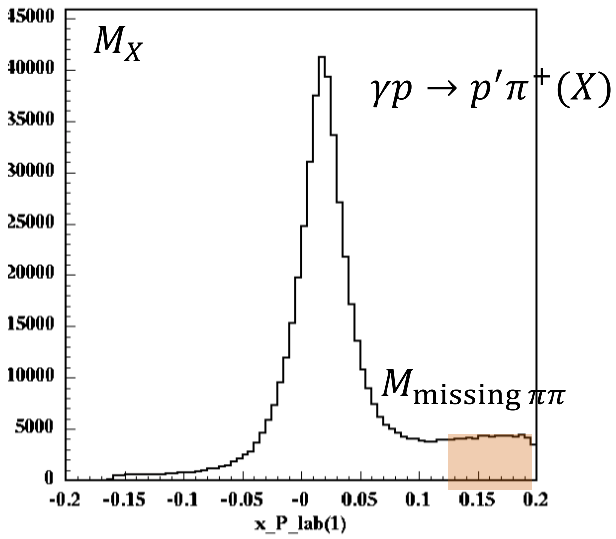
Focus on $\gamma p \to p\pi^+(\pi^-)$

- π^- momentum evaluated as missing momentum
- Multi π background from $\gamma p \to p\omega \to p\pi^+\pi^-\pi^0$
- At $E_\gamma = (3-4)$ GeV dynamics dominated by:
    - ρ^0 electroproduction: $\gamma p \to p\rho^0$
    - Δ^{++} resonance excitation: $\gamma p \to \Delta^{++}\pi^-$







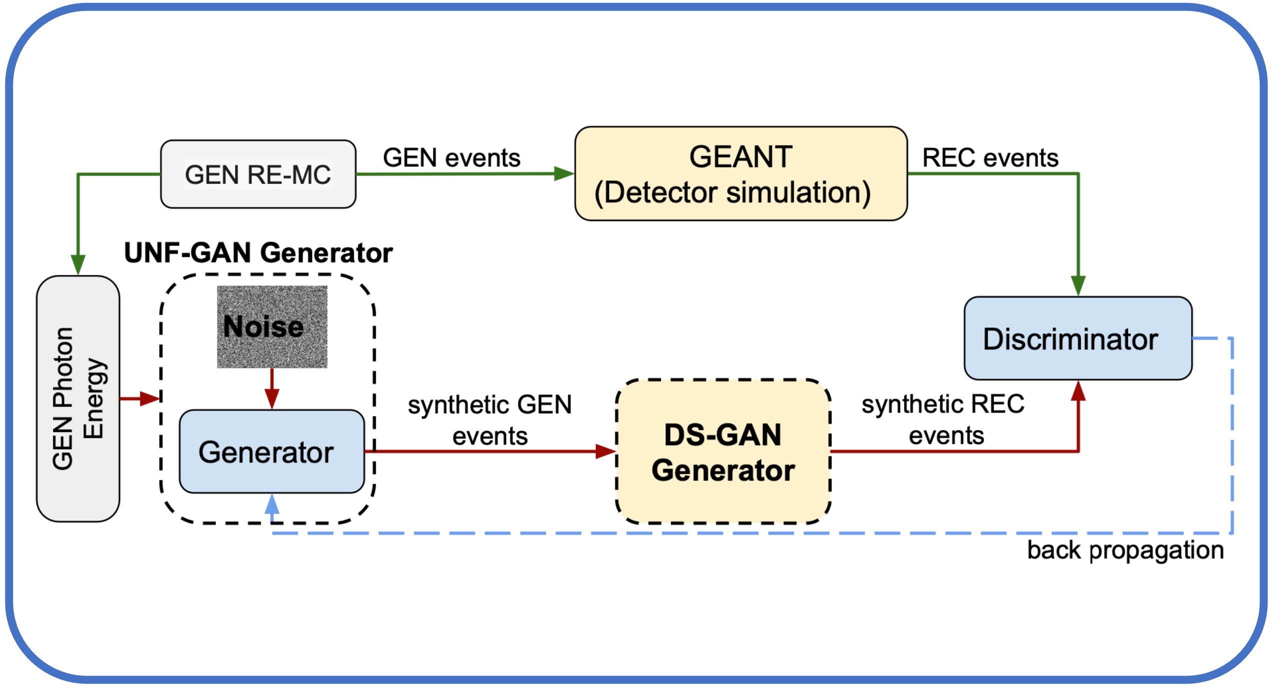


2π photoproduction closure test

CLOSURE TEST

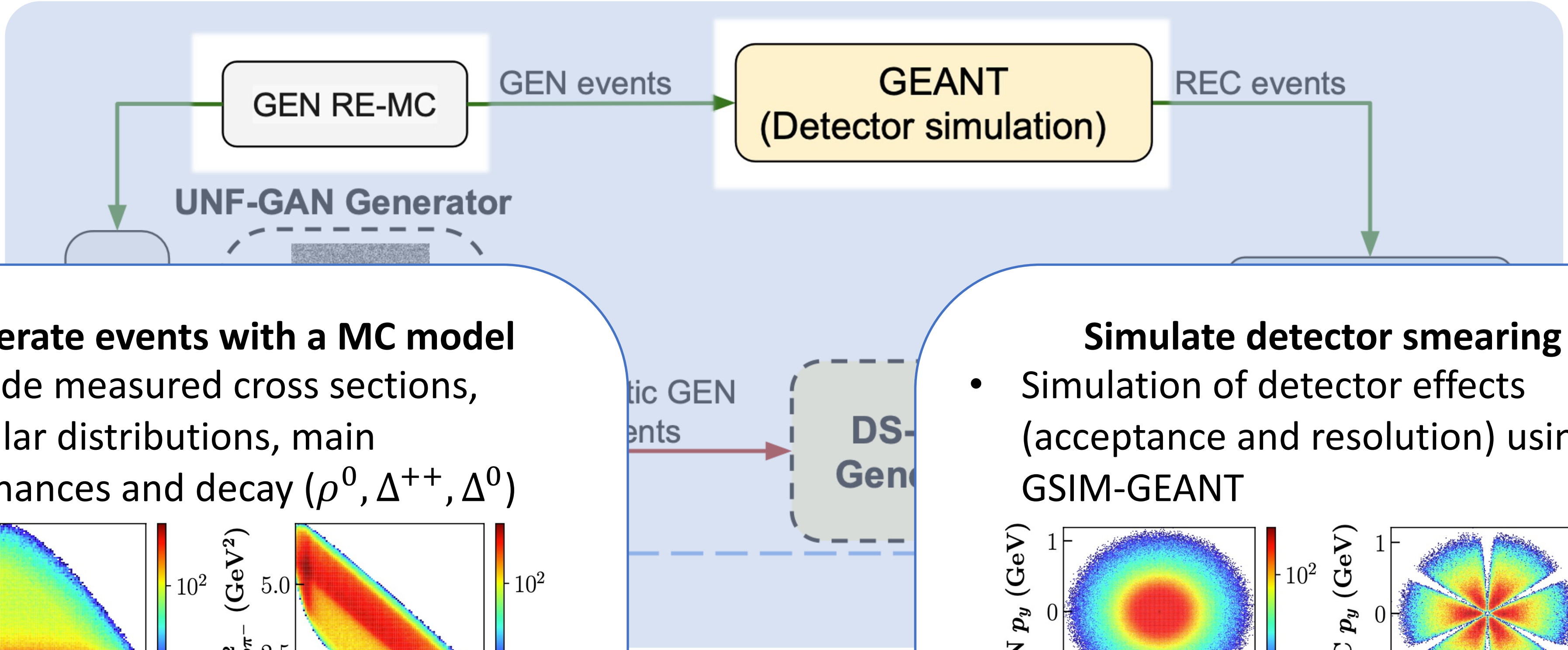
Can GANs reproduce multi dimensional correlations unfolding detector effects? How data generated by a GAN and unfolded with a GAN-based detector proxy compare to vertex-level events?

1. Generate events with a MC model
2. Simulate detector smearing using MC (GSIM-GEANT)
3. DS-GAN to simulate detector effects
   - Training on phase-space-only pseudo-data
4. UNF-GAN to generate synthetic events
   - training over MC pseudo-data
5. Compare synthetic GAN data to MC pseudo-data
6. Replace pseudo-data with CLAS data in training to unfold the vertex-level experimental distributions



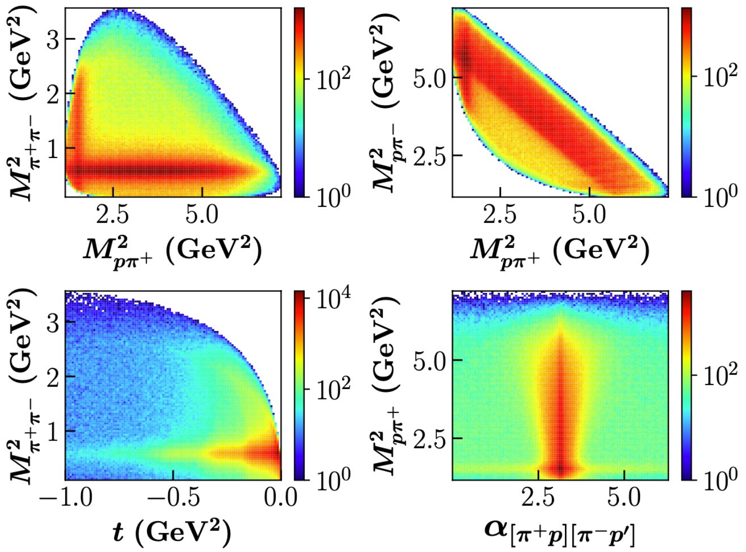
Credit: T.Alghamdi, M.Battaglieri, A.Golda, A. Hiller Blin, L.Marsicano, W.Melnitchouk, G.Montaña, E.Isupov, Y.Li, V.Mokeev, A.Pilloni, N.Sato, A.Szczepaniak, T.Vittorini, Y.Alanazi
arXiv:2307.04450

2π photoproduction closure test



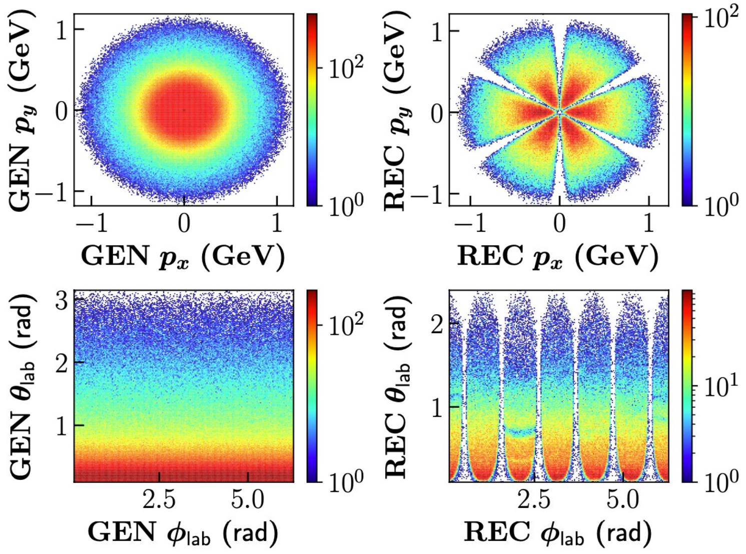
Generate events with a MC model

- Include measured cross sections, angular distributions, main resonances and decay ($\rho^0, \Delta^{++}, \Delta^0$)

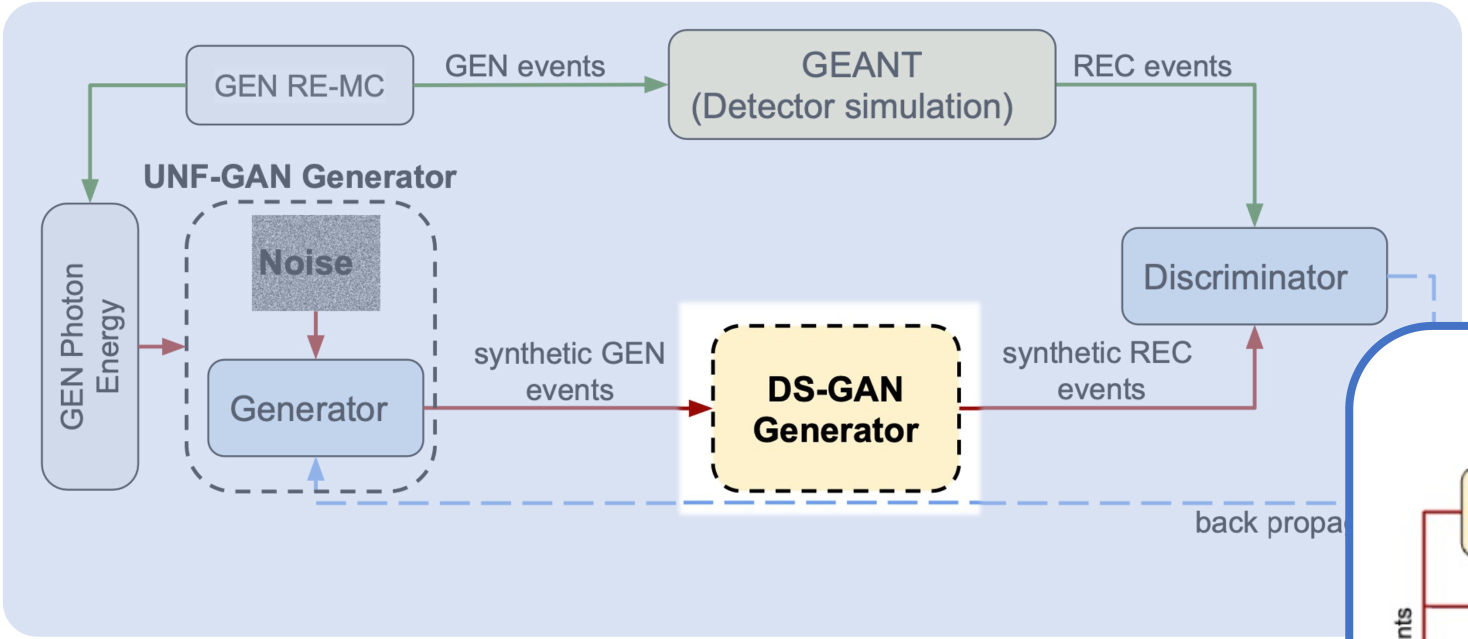


Simulate detector smearing

- Simulation of detector effects (acceptance and resolution) using GSIM-GEANT

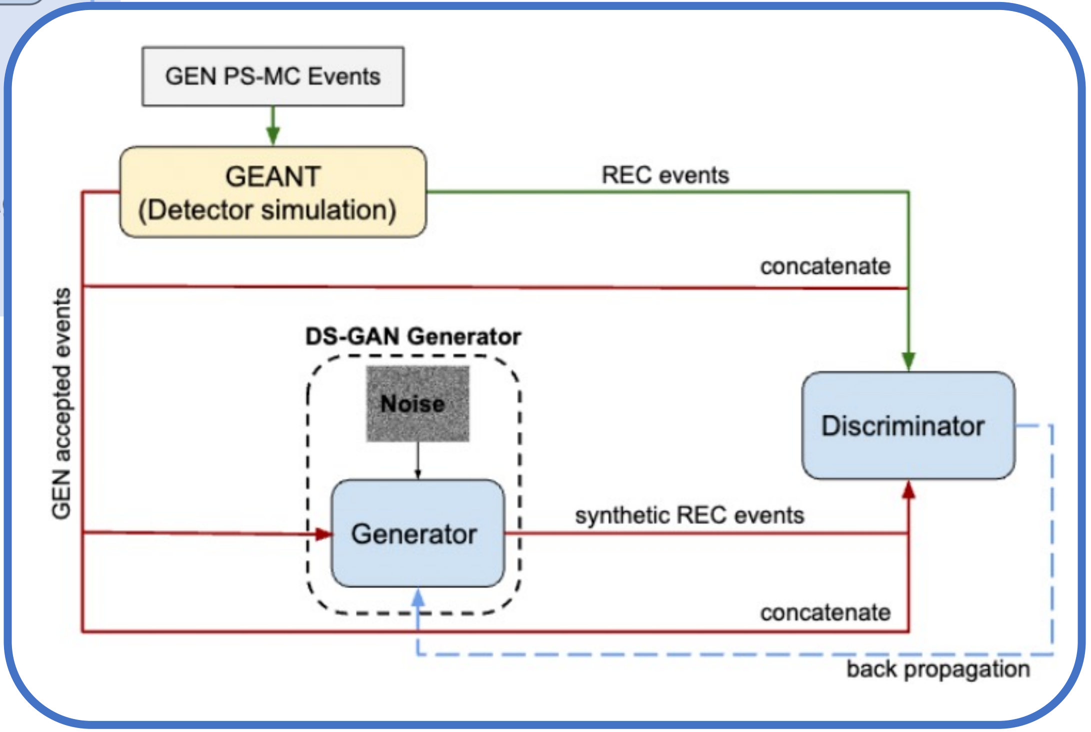


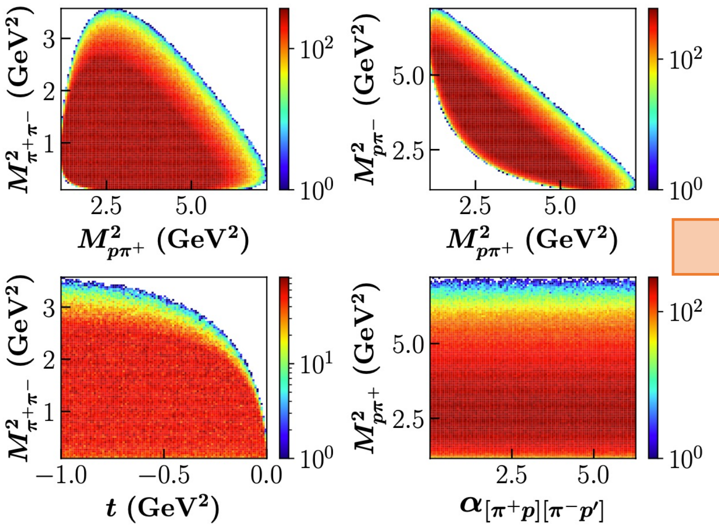
2π photoproduction closure test

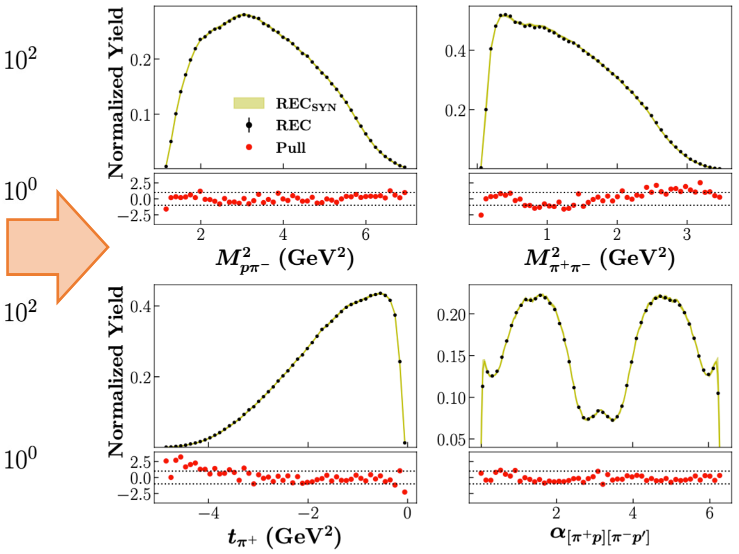


Detector simulator GAN

- Secondary GAN to learn detector effects
- Trained on phase space MC events
- Uncertainty quantification via pull calculation





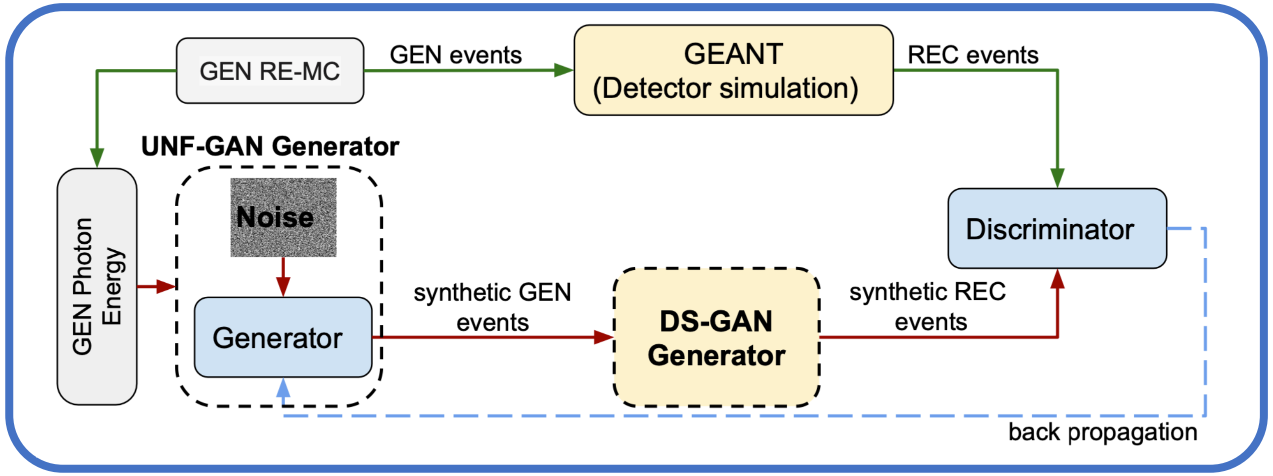


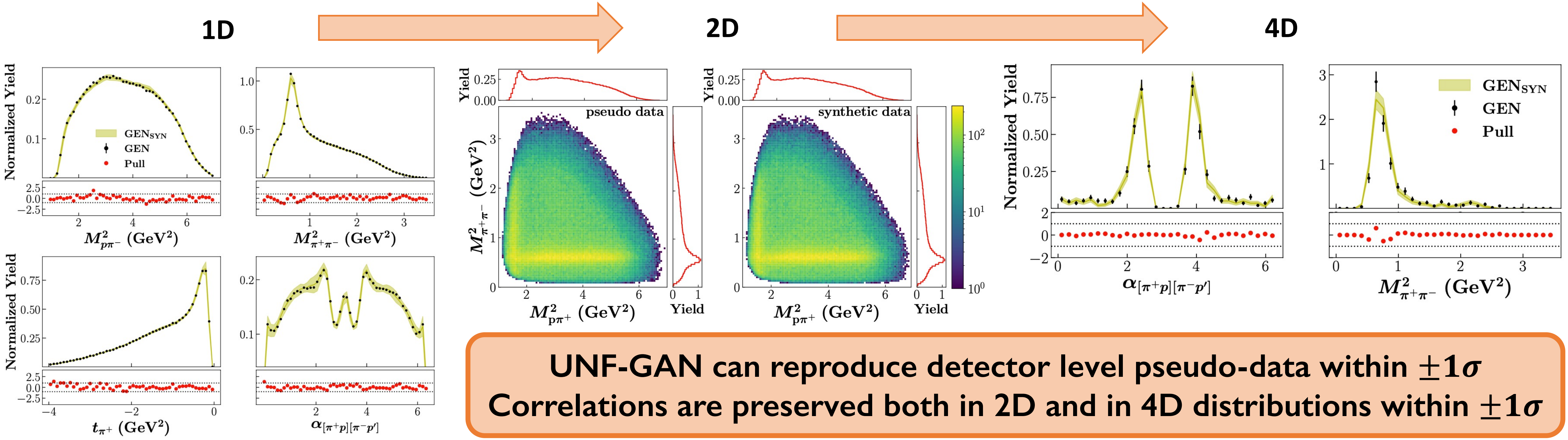
DS-GAN learned detector effects!

2π photoproduction closure test

Training of the UNF-GAN with pseudo-data

- Trained on MC pseudo-data
- Generated synthetic vertex-level data
- Detector effects applied with DS-GAN
- Uncertainty estimated with pull quantification



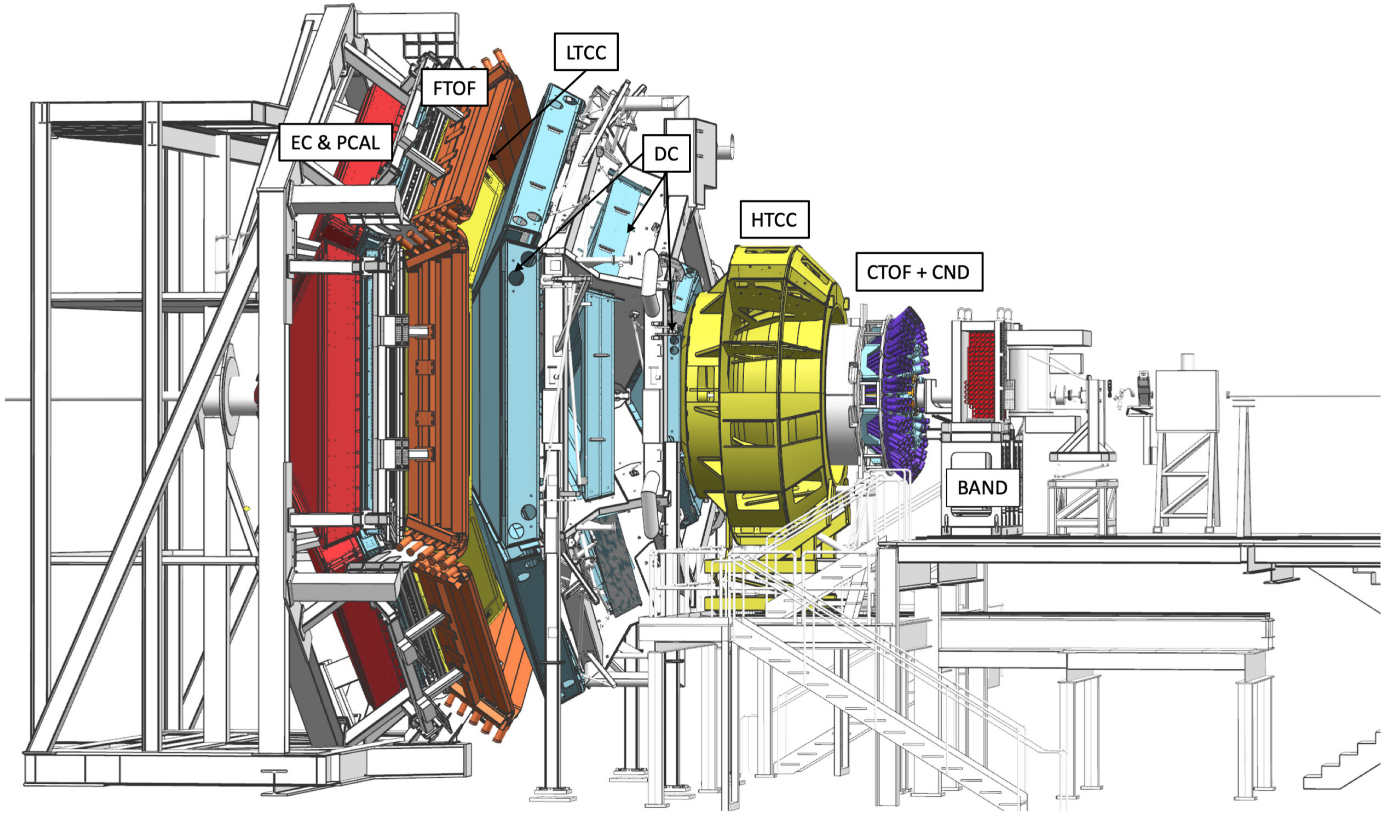


UNF-GAN can reproduce detector level pseudo-data within $\pm1\sigma$
Correlations are preserved both in 2D and in 4D distributions within $\pm1\sigma$

Next step

The next step to achieve A(i)DAPT goals:

- **Application to real data**
  - Train UNF-GAN using CLAS g11 data
  - Assess GAN capability to mimic real data
- **Application to CLAS12 detector and physics**
  - Train DS-GAN on CLAS12 pseudo-data
  - Apply UNF-GAN to electroproduction data
- **Extrapolation of scattering amplitudes**
  - Extract amplitudes from differential cross-sections exploiting theoretical constraints
  - Test on elastic scattering $\pi^+\pi^- \rightarrow \pi^+\pi^-$
  - Extend to multi-particle exclusive channels



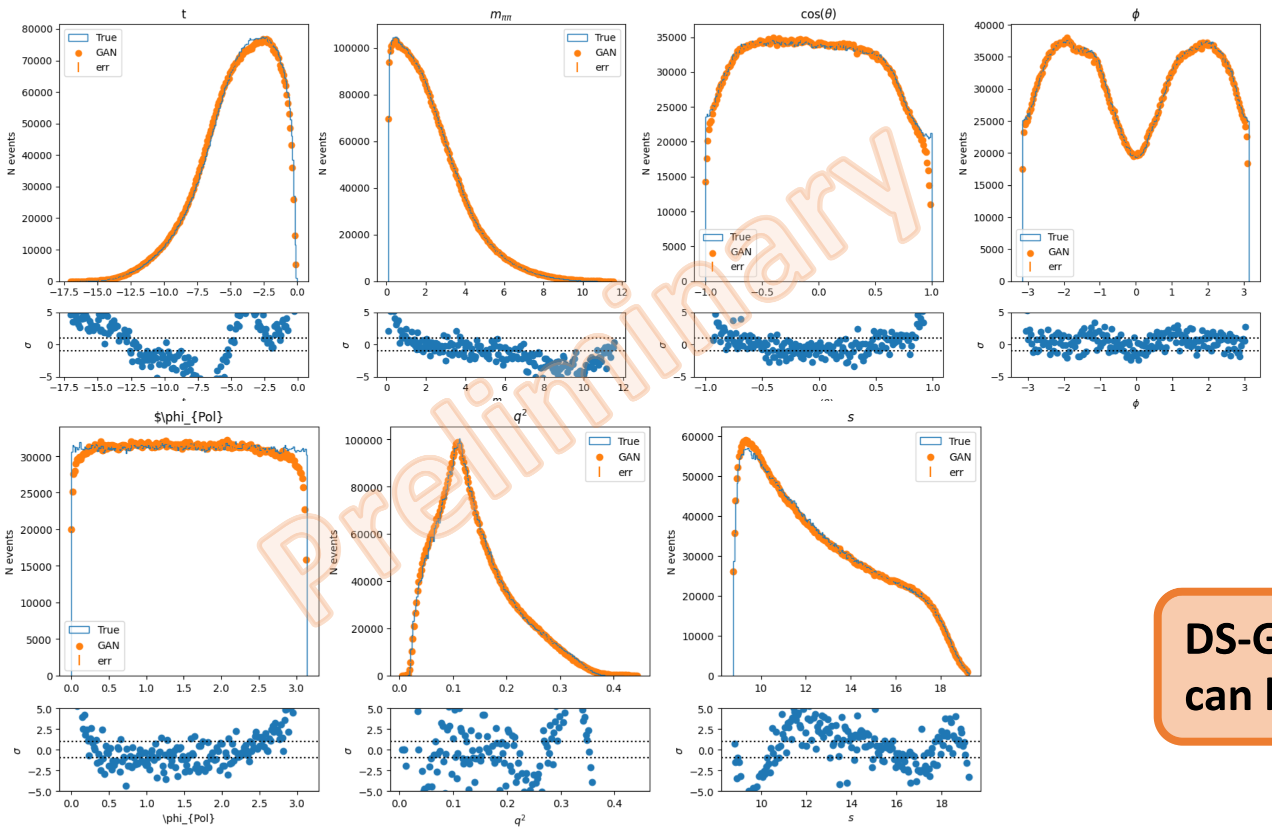
Our goal is to develop a new tool accessible to everyone and that can be used to improve any analysis

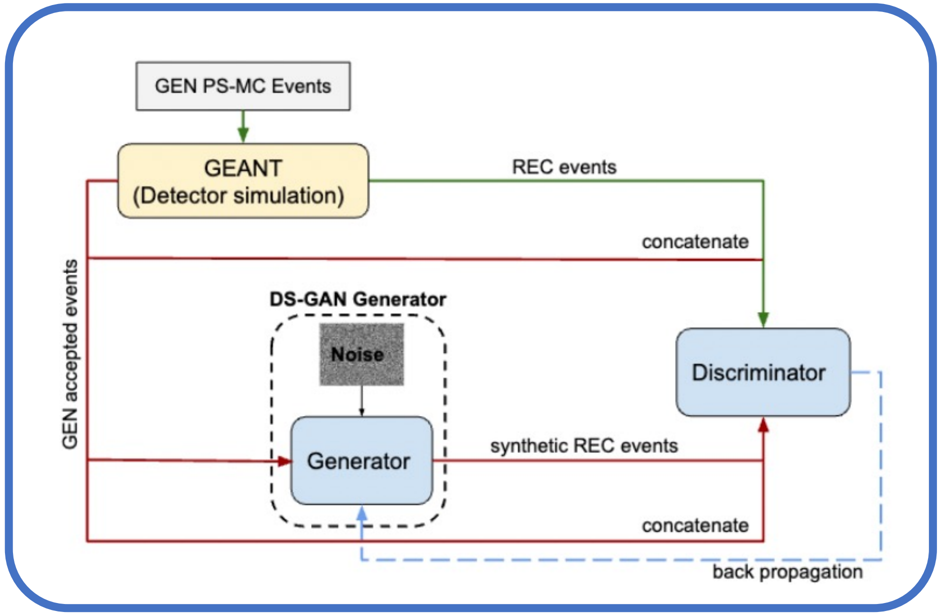
CLAS12 Electroproduction - $ep \rightarrow e\pi^+\pi^-p$

DS-GAN application to CLAS12 pseudo-data

- Different detector layout
- 7 independent variables

→ Robustness test





DS-GAN can reproduce different detector layouts and can be generalized to different sets of variables

Summary

A(I)DAPT program aims to demonstrate a novel way to extract and interpret physics observables

- We aim at creating AI-powered algorithms to address NP/HEP challenges:
  - Unfold detector effects
  - Preserve data in an alternative and efficient way

- Performed closure test on 2π photoproduction :
  - GAN can mimic realistic pseudo-data
  - GAN synthetic data retain multi-dimensional correlations

- Proven algorithm robustness:
  - It can reproduce different detector layouts (CLAS, CLAS12)
  - It can simulate different processes (photoproduction, electroproduction)

We are working on:
- Quantification of systematic error introduced by detector acceptance
- Application on real data (CLAS and CLAS12 2π data)
- Evaluation of scattering amplitude to generalize results

Thank you!