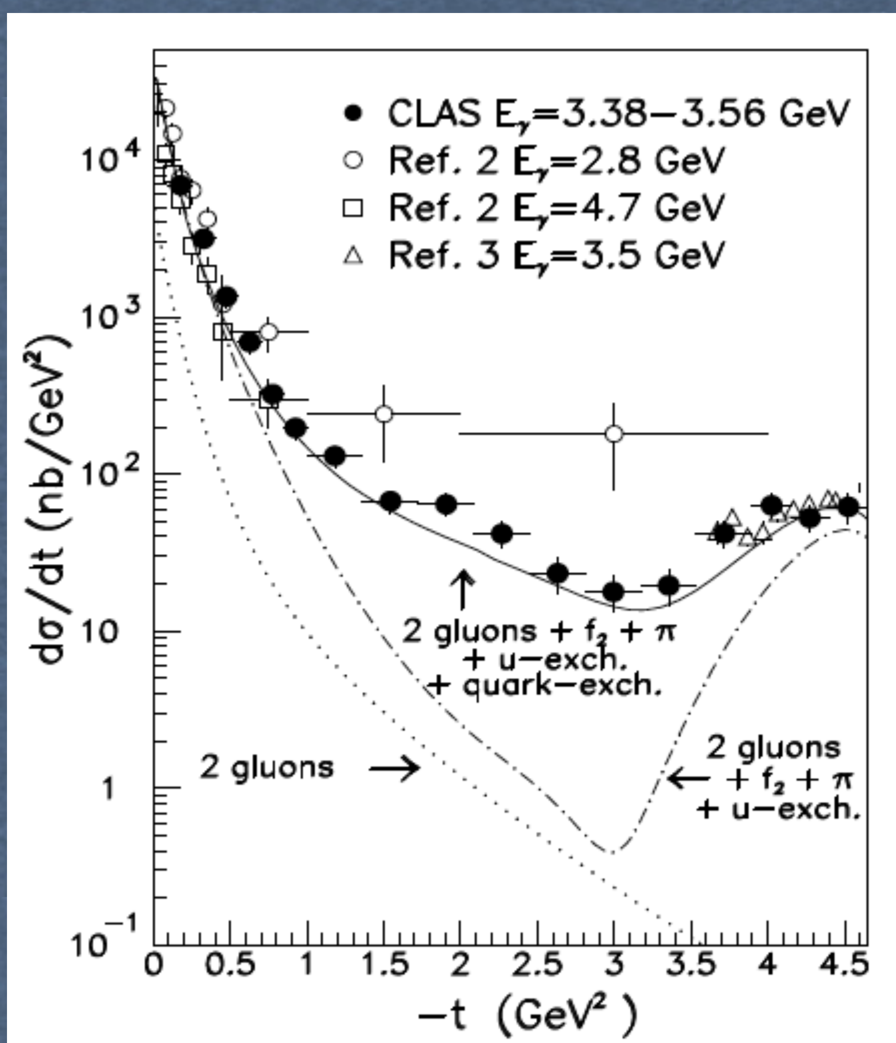
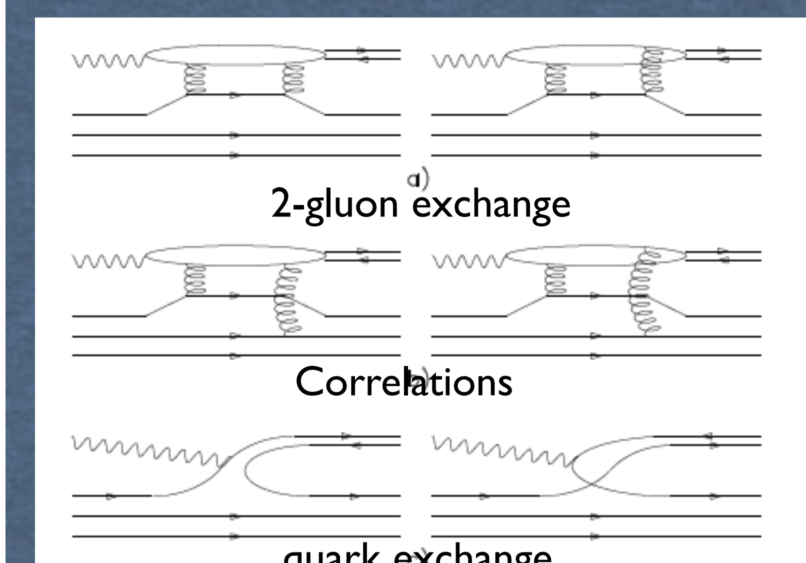
AI for data features preservation (and beyond)

Train a NN to generate events with the SAME correlations of experimental data (aka replace the xsec concept with a NN)

- Traditional approach: particles (4-momenta) measured into the detector, extract the relevant observables (e.g. cross section), extract physics mechanisms, preserve this information as replacement for the original particles



CLAS g6 ω photo production at large momentum transfer



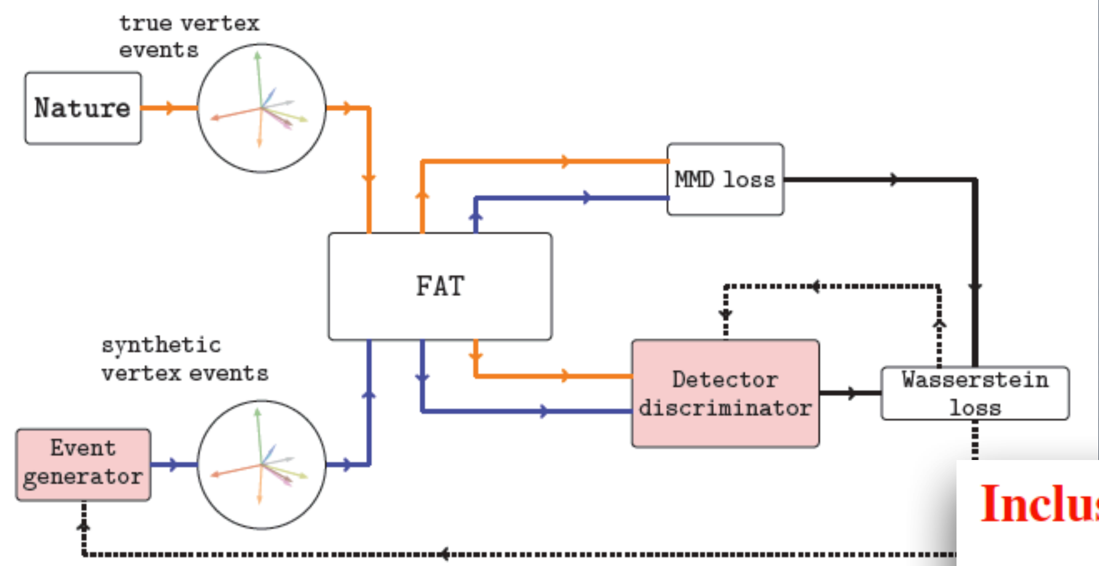
- It worked (and still works!) well if limited to channels with a single variable
- Xsec, Polarization observables, angular distribution, decay matrix, …
- It does not work (in practice) when you have several independent variables: multi-particle final states (spectroscopy) or multi-variable correlations (SIDIS)

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

AI enabled MCEG

Goals

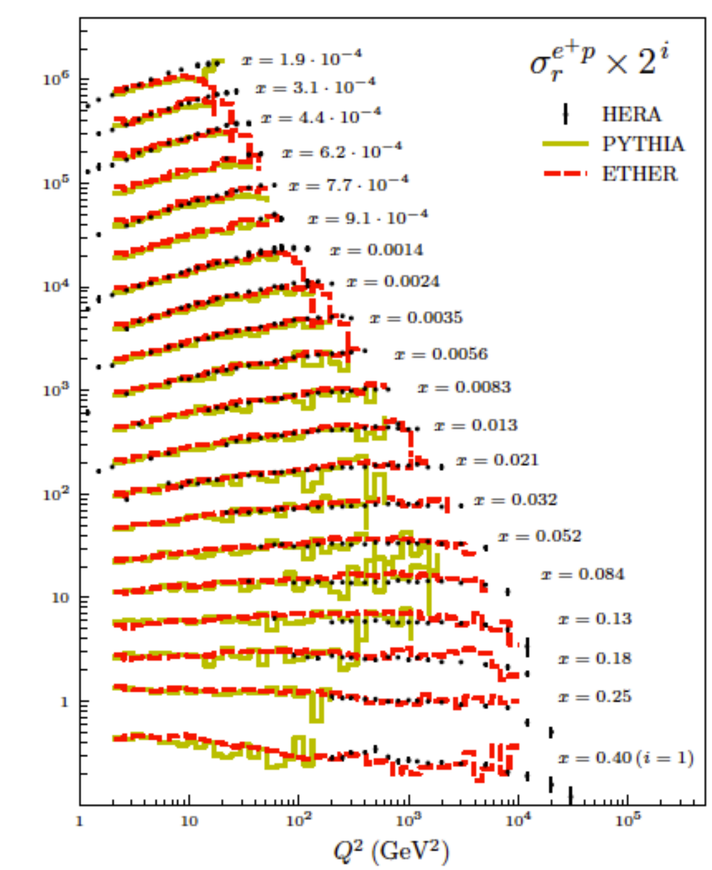
- Build a **theory-free** MCEG
- Map out particle correlations without biases from approximated theory
- MCEG as a **data storage utility**



Progress

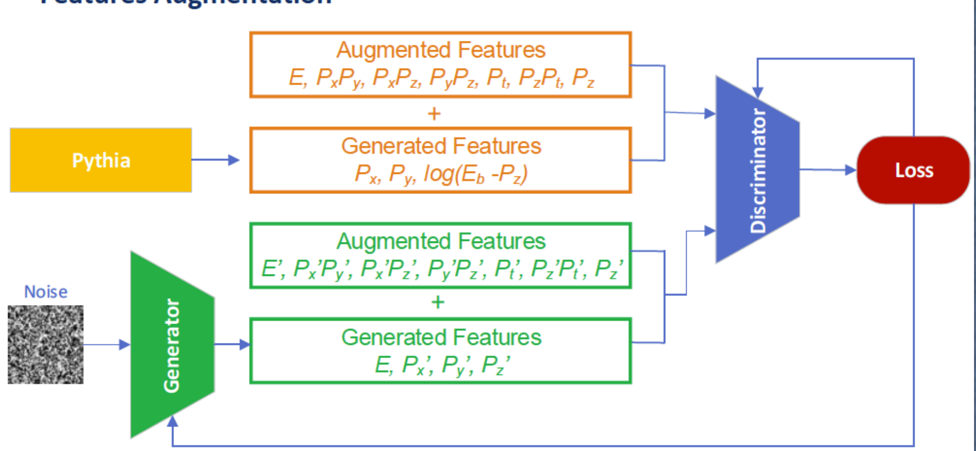
- Prototypes based on GANs
- Case study on inclusive DIS
- Case study on exclusive $\gamma + P \rightarrow P + \pi^{+} + \pi^{-}$

Inclusive DIS



Features Augmentation and Transformation GAN (FAT-GAN)

- **Features Transformation**
- **Features Augmentation**

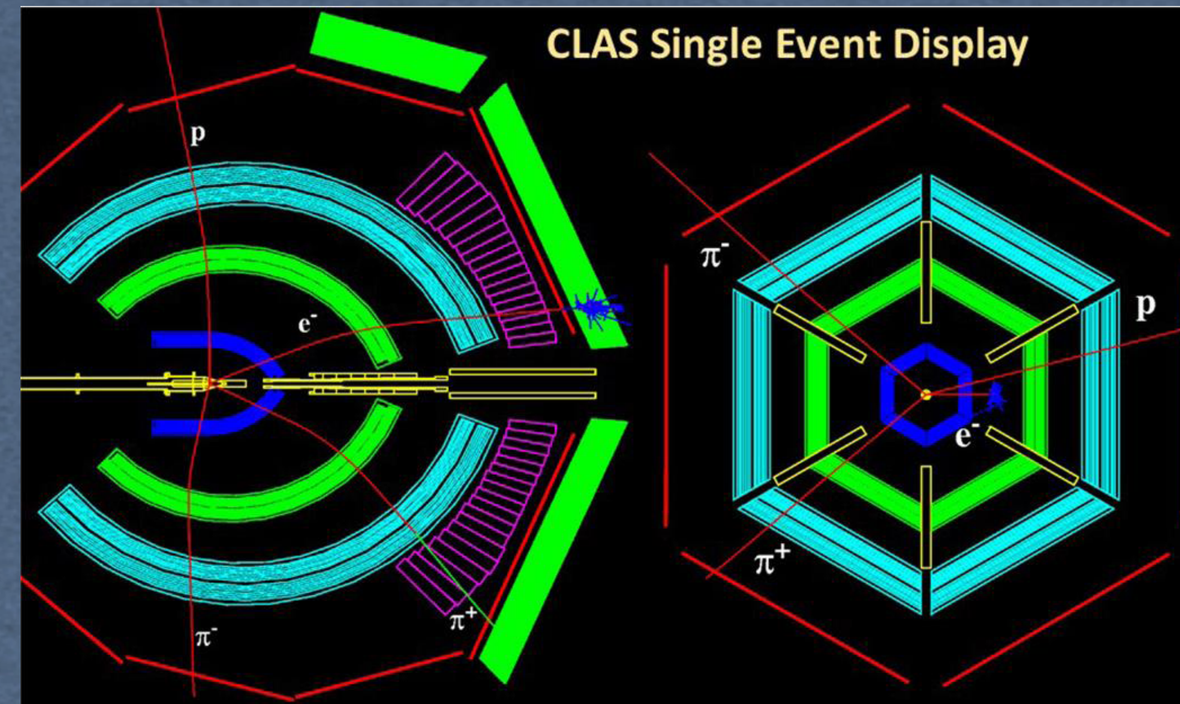


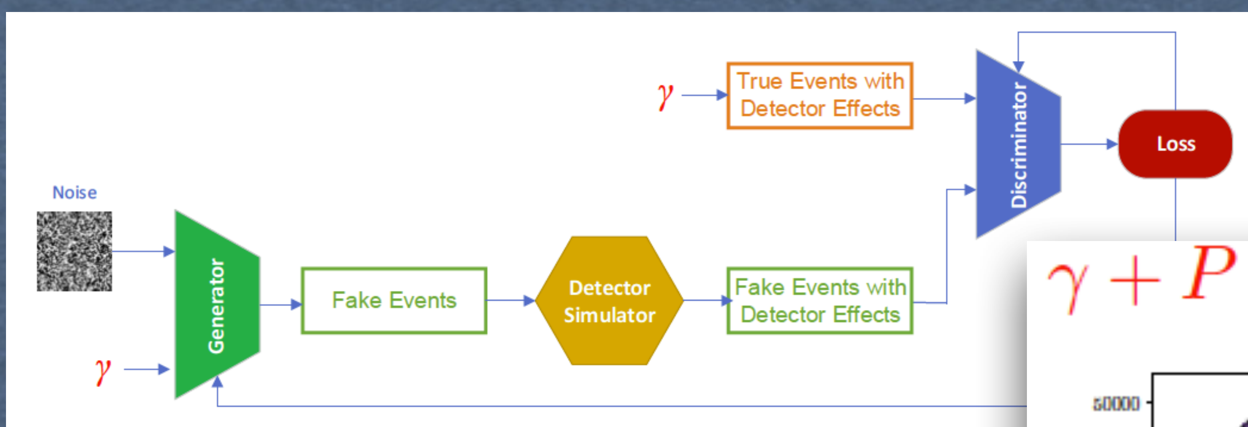
Y.Li

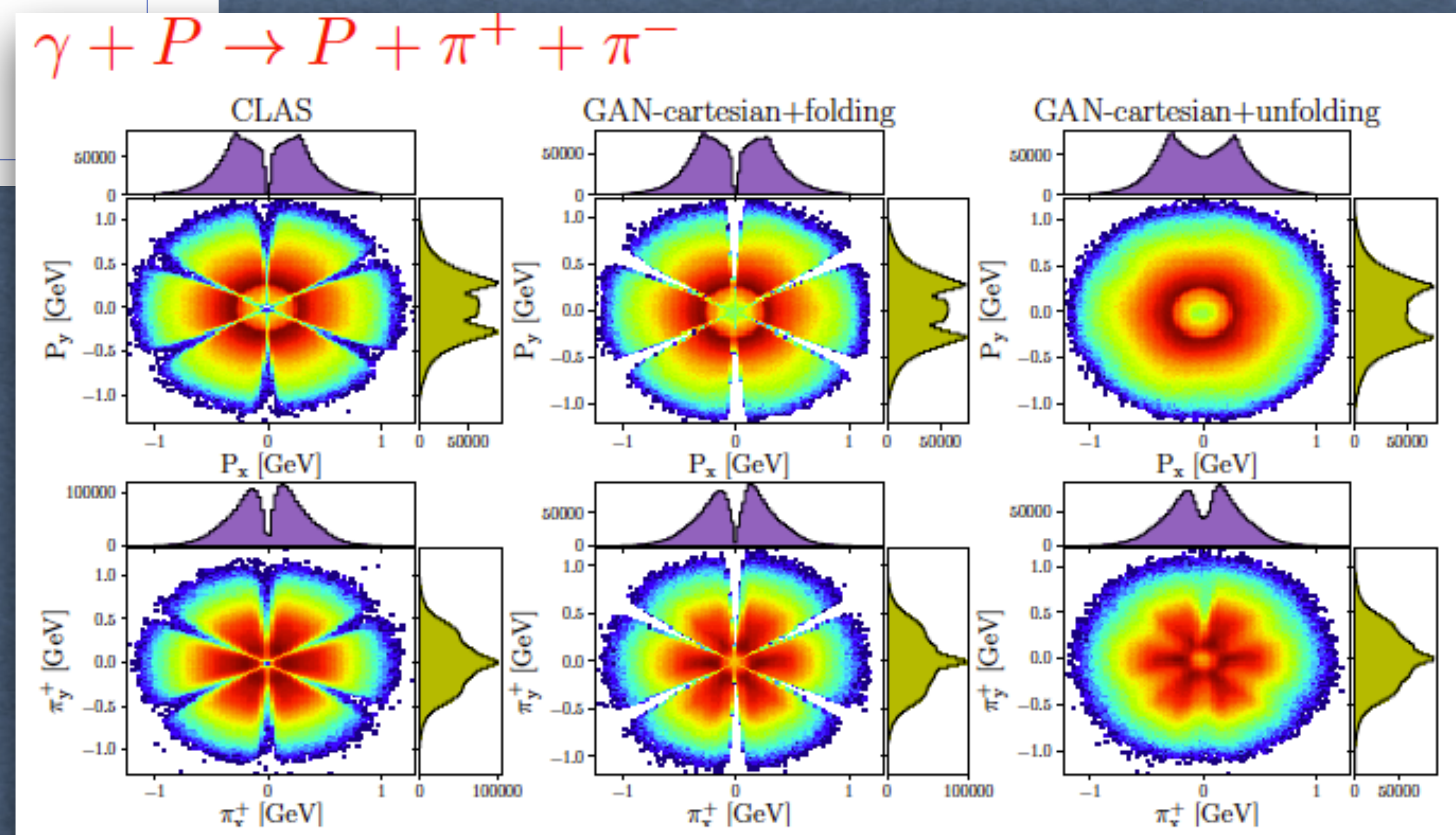
N.Sato,W.Melnichouk

Exclusive channel analysis

- CLAS g11 photo production data - 2pi channel
- Available a huge stat with different combination of missing particle in final state
- We started from $\gamma\, p \rightarrow p\ \ \pi^+ (\pi^-)$
- A small contamination from 3π channel (dominated by ω)
- Independent variable distributions

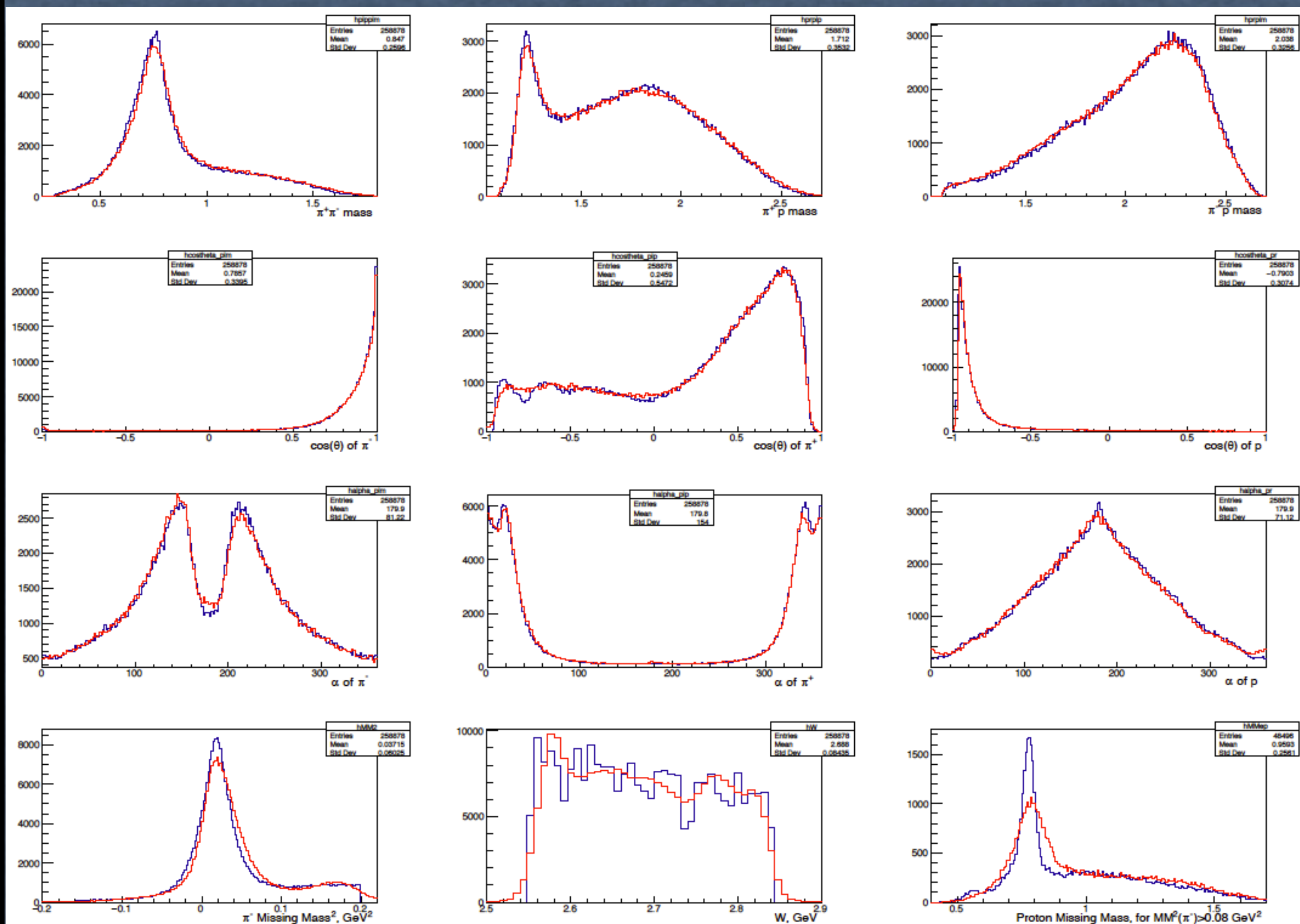






Y.Li, N.Sato

Results and Validation



- How to be sure syntetic-data represent data?
- 1d comparison is (reasonably) easy but multi-dim?
- Binned or un-binned observable comparison?
- Procedure that dos not assume the number of particle in the final state?

$$\frac{d\sigma}{dt dM_{\pi\pi} d\Omega_{\pi}}$$

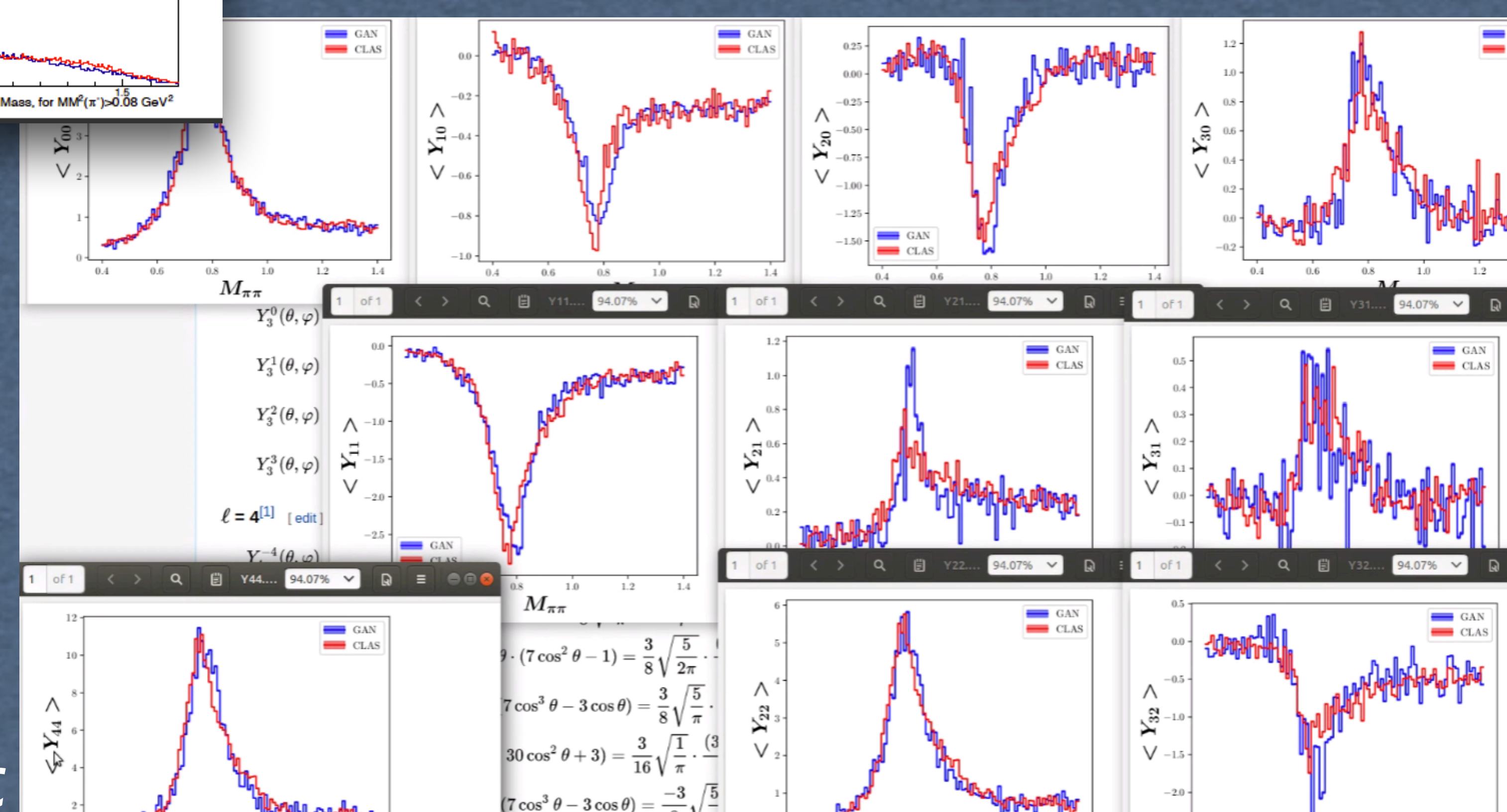
1-d comparison of projected 5-differential xsec

E.Isupov, V.Mokeev

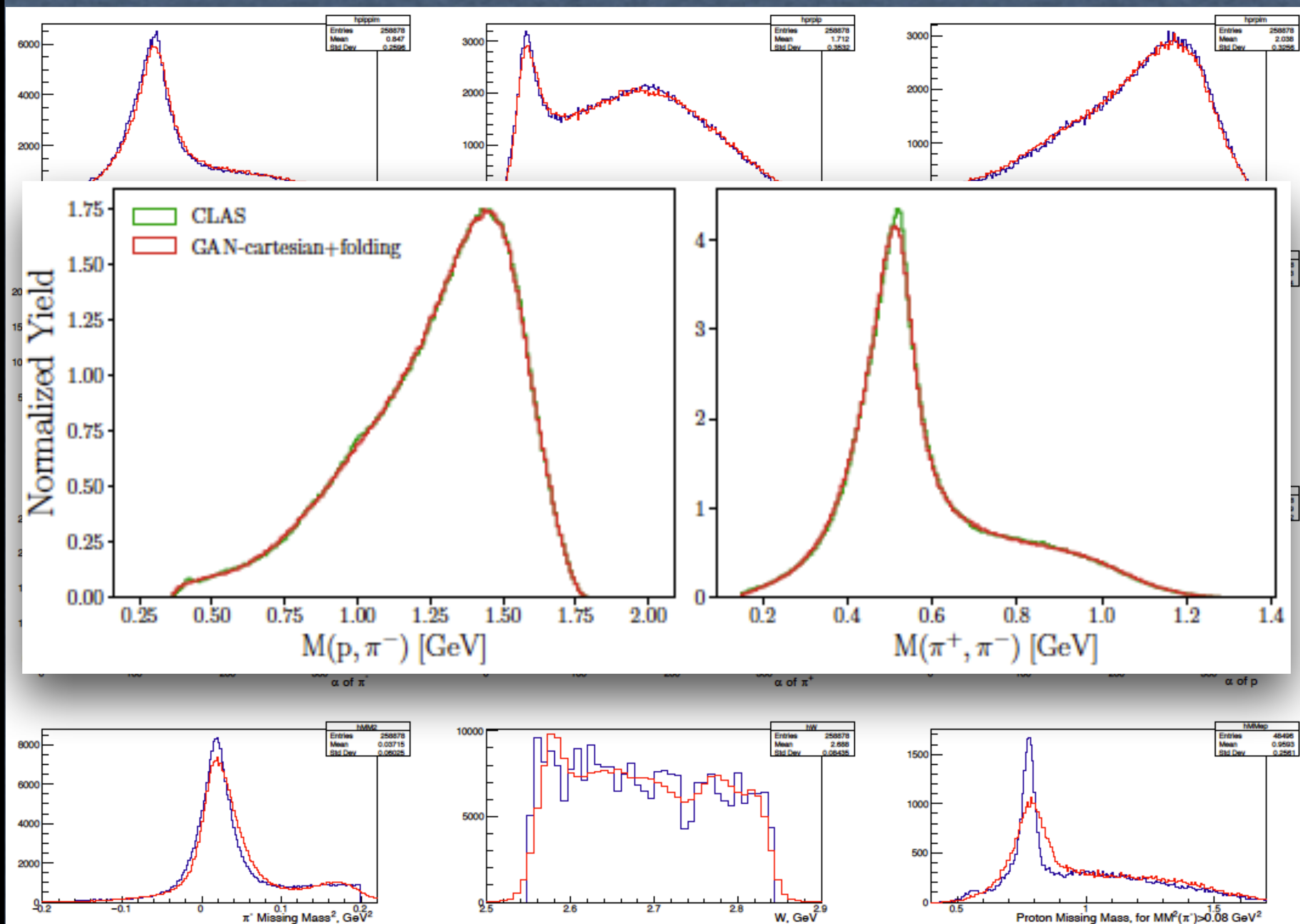
Moments of angular distributions

$$\langle Y_{LM}\rangle(E_\gamma, t, M_{\pi\pi}) = \sqrt{4\pi}\int d\Omega_\pi \frac{d\sigma}{dtdM_{\pi\pi}d\Omega_\pi} Y_{LM}(\Omega_\pi),$$

A.Blin, A.szczepaniak and JPAC



Results and Validation

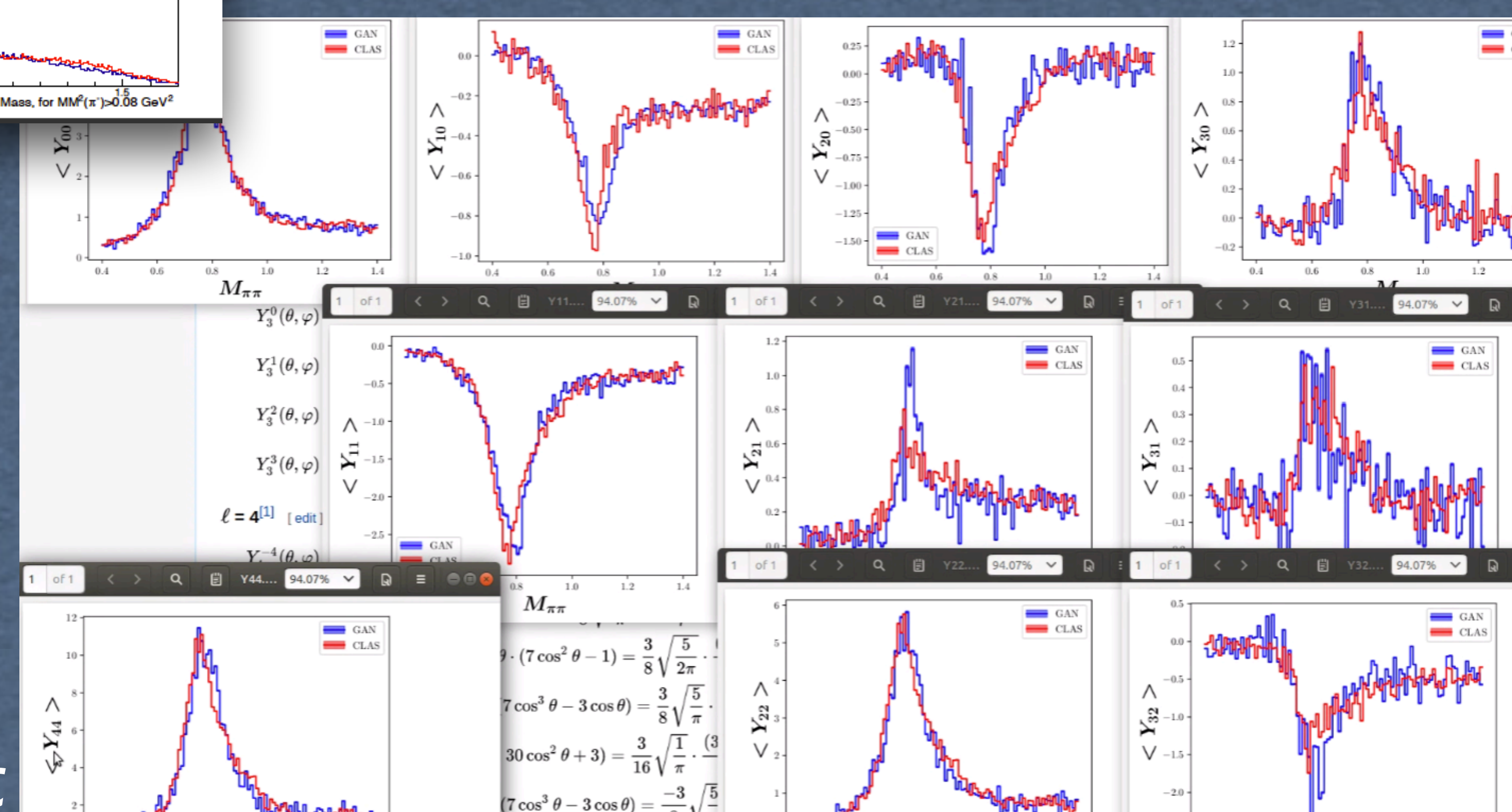


- How to be sure syntetic-data represent data?
- 1d comparison is (reasonably) easy but multi-dim?
- Binned or un-binned observable comparison?
- Procedure that dos not assume the number of particle in the final state?

$$\frac{d\sigma}{dt dM_{\pi\pi} d\Omega_\pi}$$

1-d comparison of projected 5-differential xsec

E.Isupov, V.Mokeev



Moments of angular distributions

$$\langle Y_{LM}\rangle(E_\gamma, t, M_{\pi\pi}) = \sqrt{4\pi}\int d\Omega_\pi \frac{d\sigma}{dt dM_{\pi\pi} d\Omega_\pi} Y_{LM}(\Omega_\pi),$$

A.Blin, A.szczepaniak and JPAC

AI for data features preservation (and beyond)

Train a NN to generate events with the SAME correlations of experimental data (aka replace the xsec concept with a NN)

- Preserve data in an alternative form (to be applied to current physics program)
- Extract data features from pseudo-data
- Statistics: can we use the NN to determine the necessary statistics for a given analysis?
- Statistics: if the NN does capture the essence, use it to generate a larger data sample (super resolution)
- Detector Efficiency: can we use different final states to extract a Single Particle Efficiency from data?
- Physics I: sub-leading features are reproduced?
- Physics II: define other NNs to reproduce the elementary process and use them to 'fit' the original
- Physics III: use elementary NNs to fit new/different data (initial/final state) to incorporate Universality
- Physics IV: study NNs properties (poles, cuts, dynamics, …)
- Physics V: fit a model to the NN to extract physics features

- **Collaborative effort**
- **Sub-group:**
  - **ML**
  - **Data manipulation**
  - **Validation**
  - **Unfolding**
  - **Theory**
- **Regular weekly meeting**
- **Wiki page**

LDRD program + ODU ML
Y. Li (co-PI) (ODU, CS prof.)
Y. Alanazi (ODU, CS grad.)
M. P. Kuchera (Davidson College, CS prof.)
R. Ramanujan (Davidson College, CS prof.)
L. Velasco (U. Dallas, Physics undergrad)
P. Ambrozewicz (JLab, post-doc)
T. Liu (JLab, Theory post-doc)
W. Melnitchouk (PI) (JLab, Theory)
N. Sato (co-PI) (JLab, Theory)

Hall-B Data Preservation Task Force
H. Avakian (PI)
G. Gavalian (core)
M. Ungaro (core)
V. Mokeev (core)
N. Baltzell (external)
A. Vossen (external)

Other collaborators
A. Blin (JLab)
A. Szczepaniac (IU/JLab)
A. Pilloni (ECT*)
C. Fanelli (MIT)
G. Costantini (UniBs)
L. Biondo (UniME)
L. Marsicano (INFN-GE)
T. Roark (ODU)
T. Viducic (ODU)
N. Tyler (SCU)