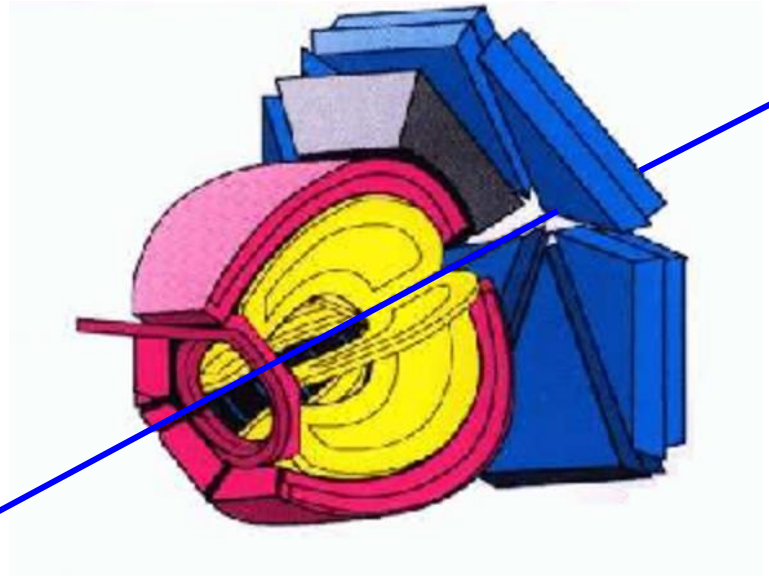
GAN uncertainty quantification in CLAS 2π channel data analysis

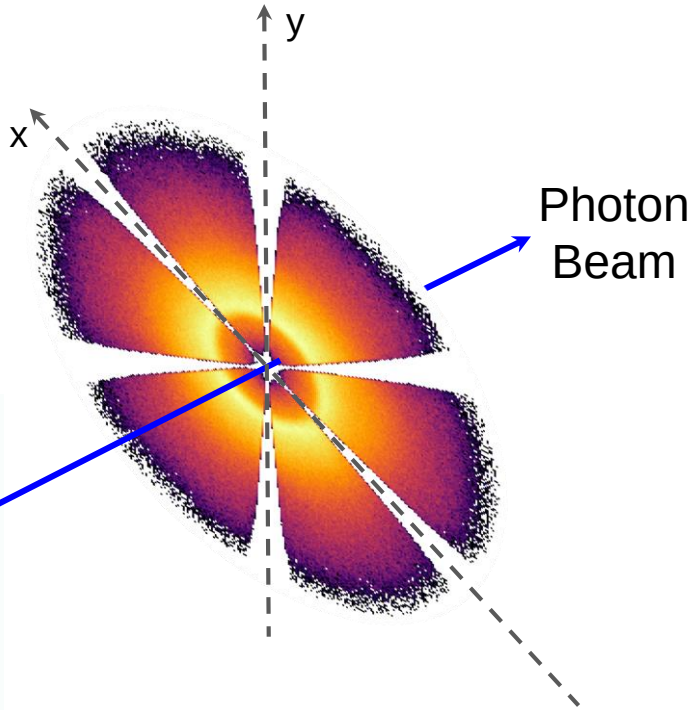
Pawel Ambrozewicz (JLab)

GAN uncertainty quantification in CLAS 2π channel data analysis

Reaction channel:

- Initial state - photon + proton
- Final state: p' + π^+ + X
- X - the undetected, "missing", particle(s)





Measured quantities: momenta of detected particles, proton and π^+

Uncertainty quantification:

Use bootstrap approach as a single sample inference method

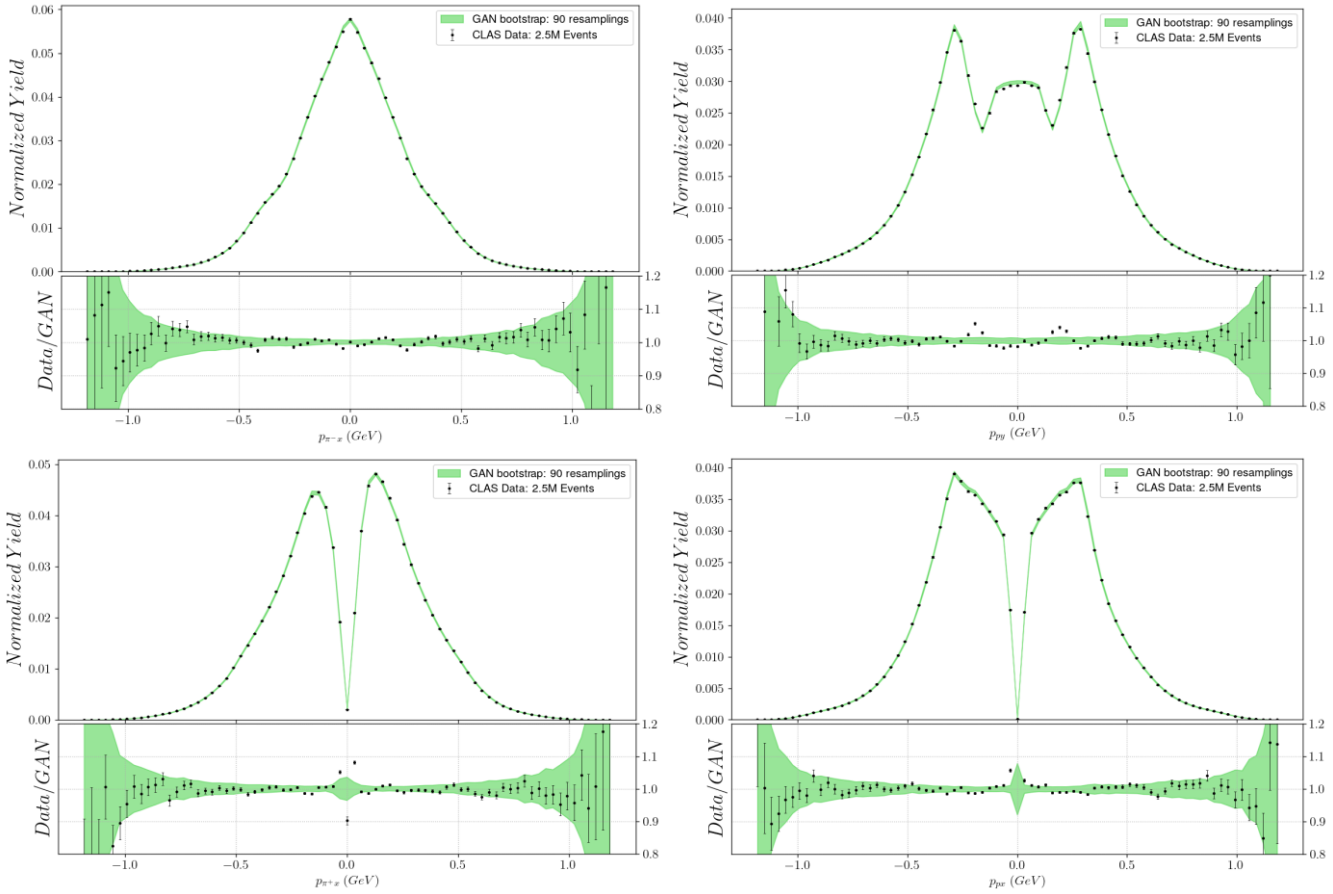
Data sample (2.5M of CLAS events) is resampled with replacements.

LSGAN is then trained on such a sample and also produces 2.5M events

The results are then combined and the mean and standard deviation computed.

Results, for some of the training features, *transverse momentum components of p' and* π^+, are shown on the right.

Training features: detected particle momenta components



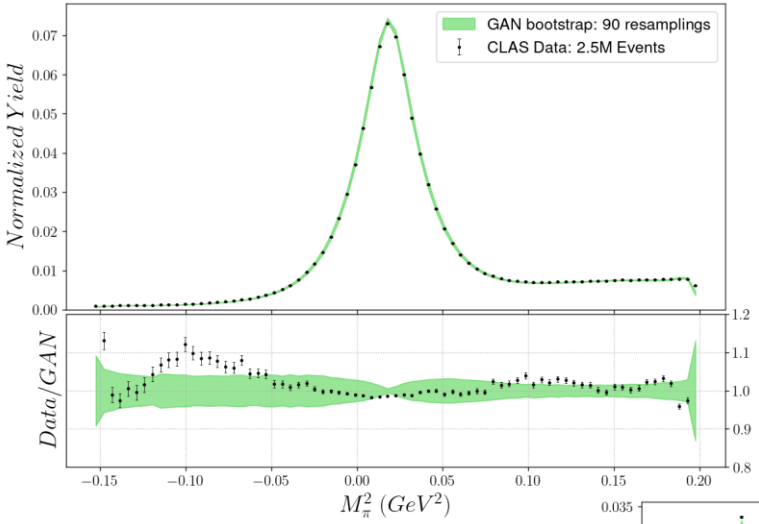
Uncertainty quantification:

Results, for some of the augmented (calculated) features, is shown on the right.

The GAN band is calculated relative to the mean, and is spanned by minimal and maximal replica values.

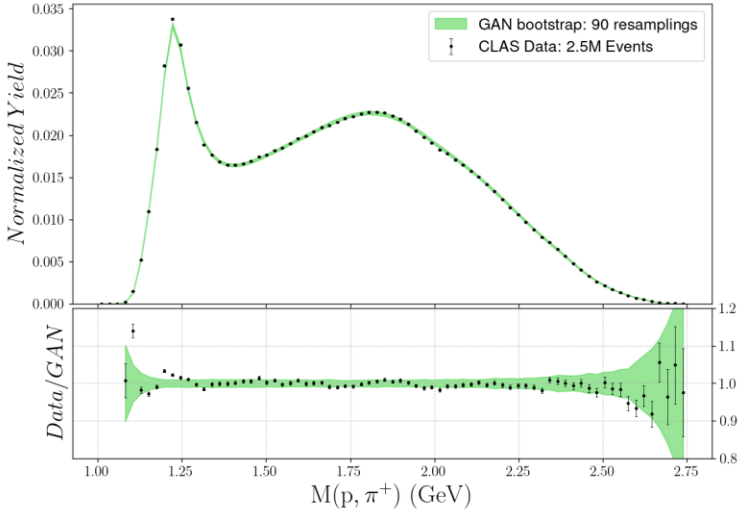
The width of the GAN band is related to the number of generated events, which is the same as in the original data sample.

The larger the experimental uncertainty the wider the GAN band.



Augmented feature:
Missing Mass Squared, M_X^2

Reconstructed variable:
Mass of p'π^+system, M(p'π^+)



Bootstrap method check

Uncertainty quantification:

Usually bootstrap method requires resampling many more times than the sample size. This is not feasible here.

Plotting GAN results per bin provides a check of the method.

On the plot it can be seen that with significant number of replicas the distribution per bin approaches a normal distribution, which allows simplifications in further inference.

An effort is being made to determine at which point resamplings could be terminated.