Bayesian Data Analysis in Baryon Spectroscopy

Stefanie Lewis
University of Glasgow
CHEP 13
Amsterdam, The Netherlands

University of Glasgow

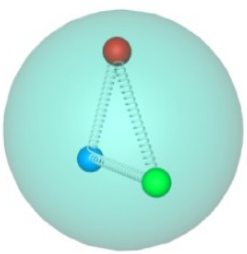
Overview

- Physics background
- Data analysis
- Data parallelism
- Results
- Summary and Future Work

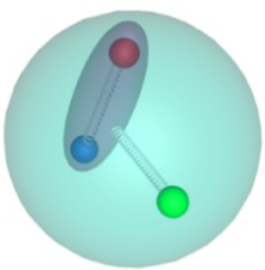
University of Glasgow

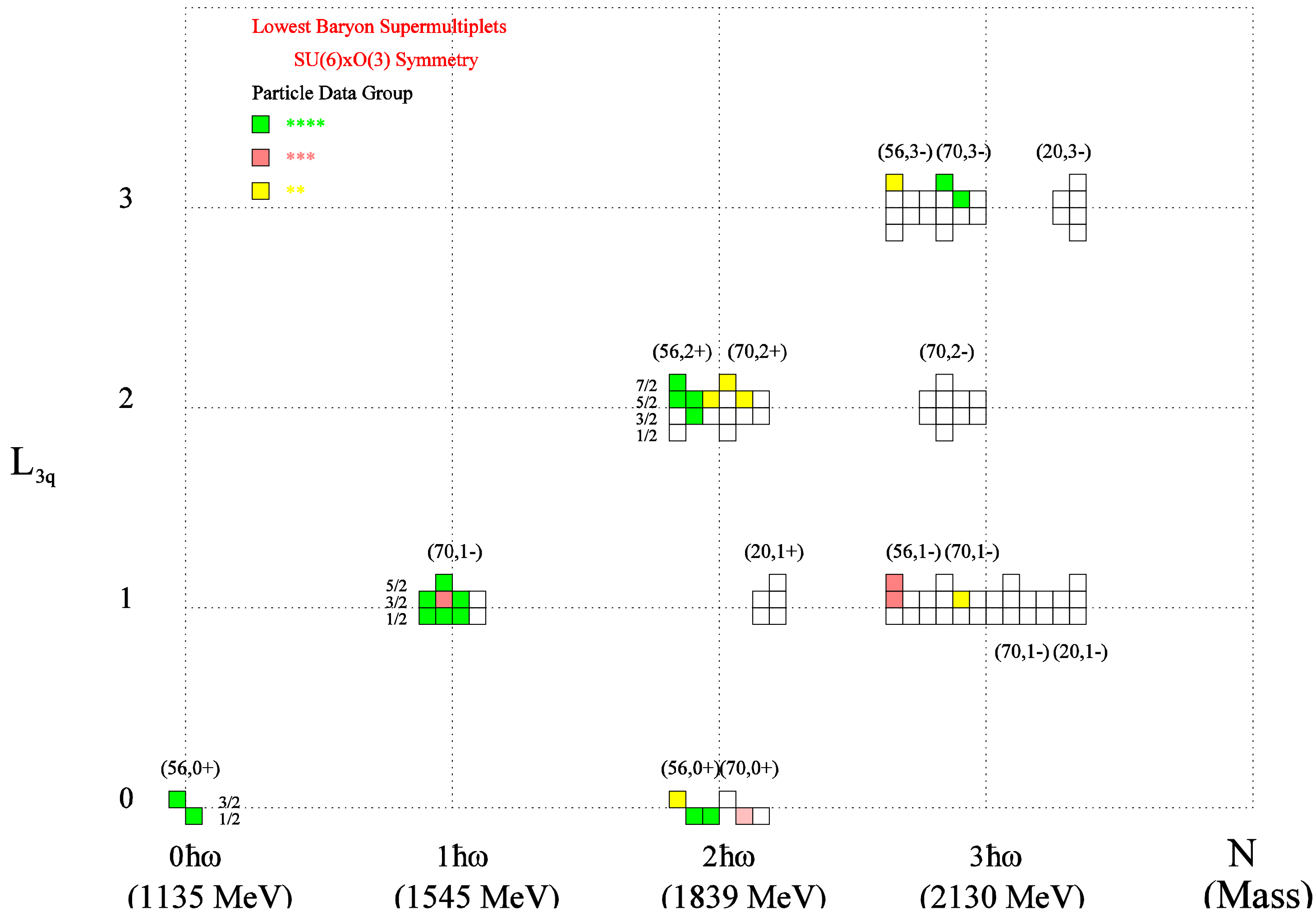
Background

Symmetric quark model

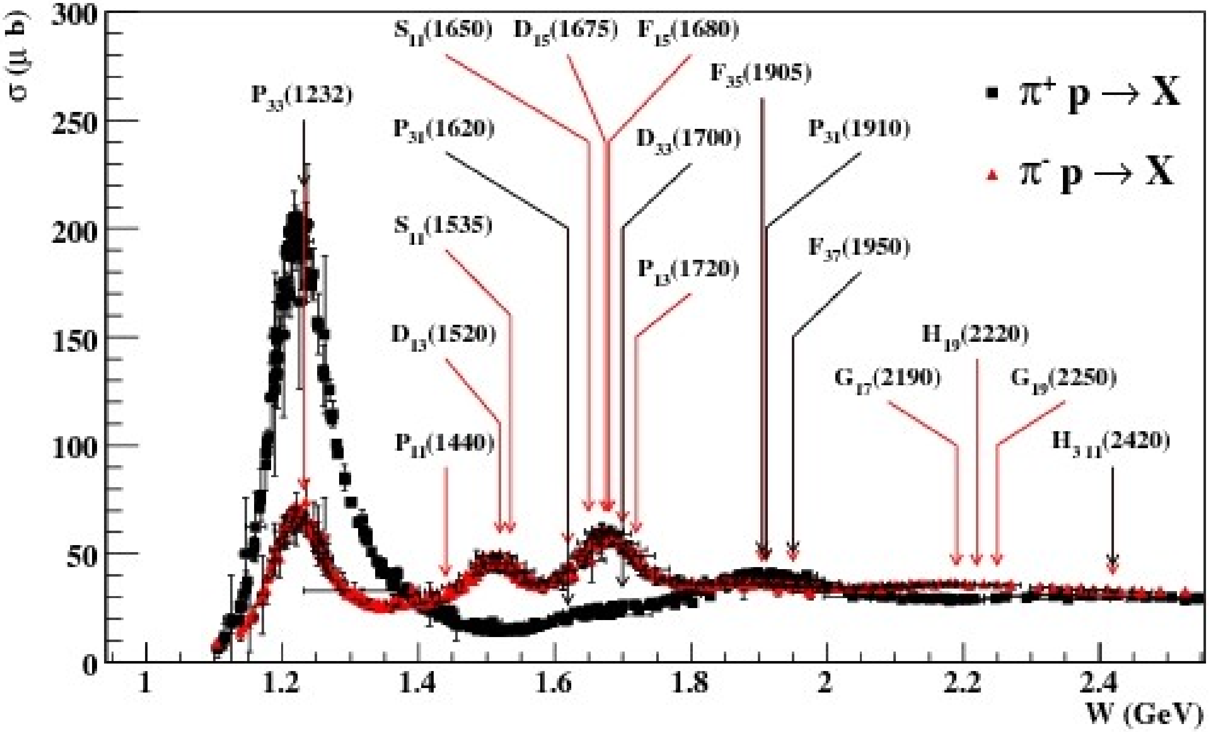


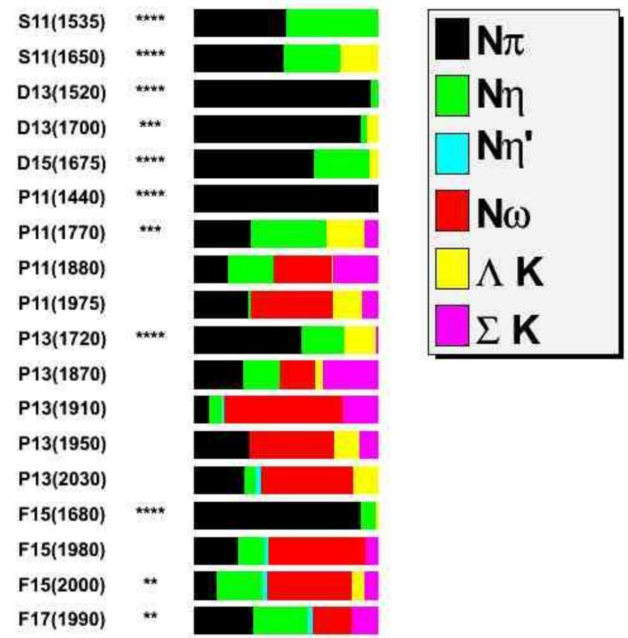
Diquark model





Background

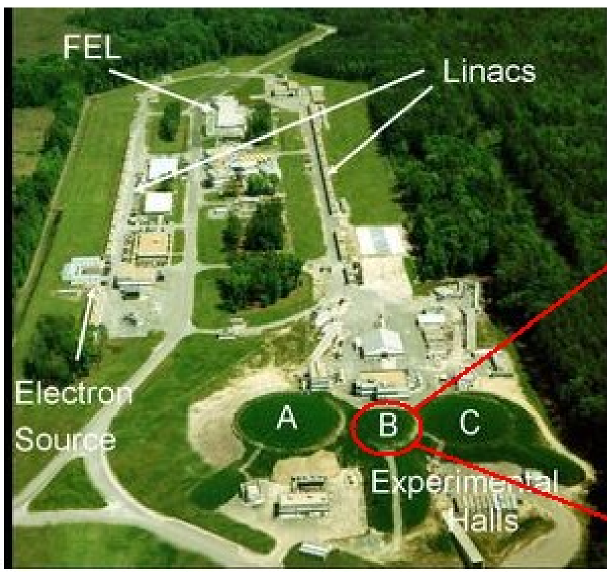


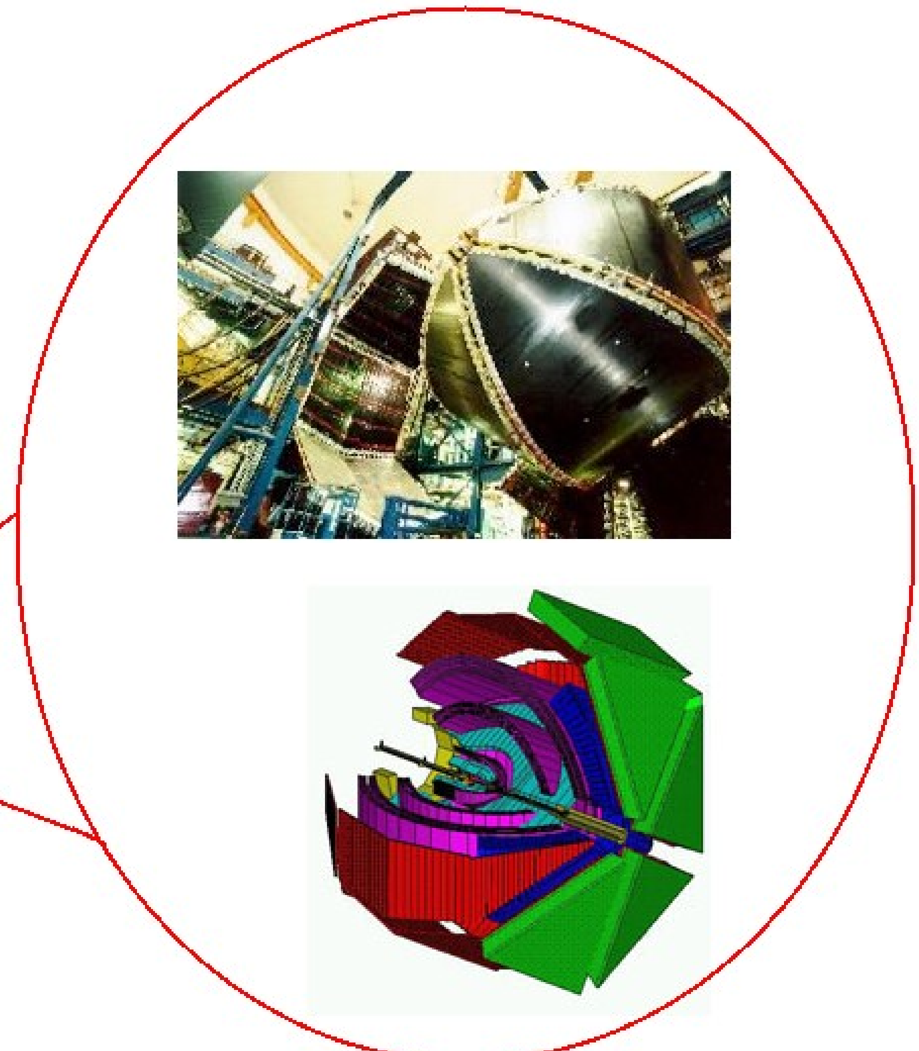


Peaks are wide and overlapping

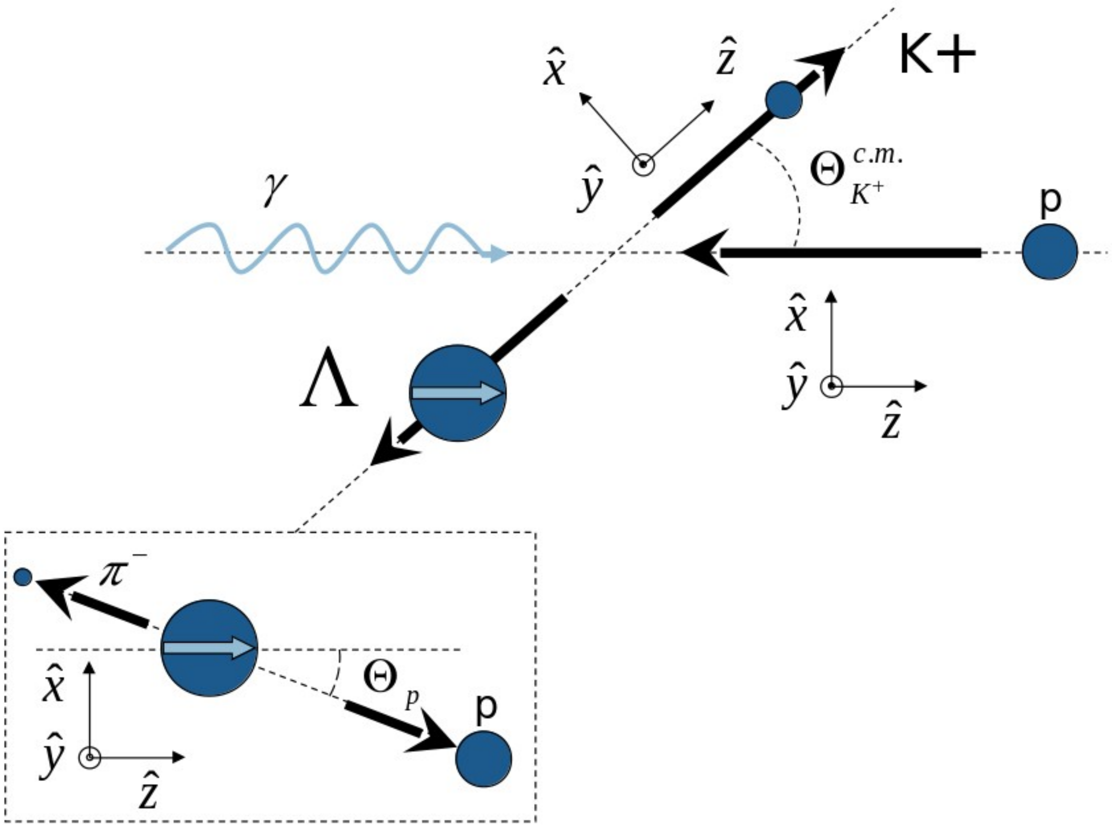
CLAS

Jefferson Lab



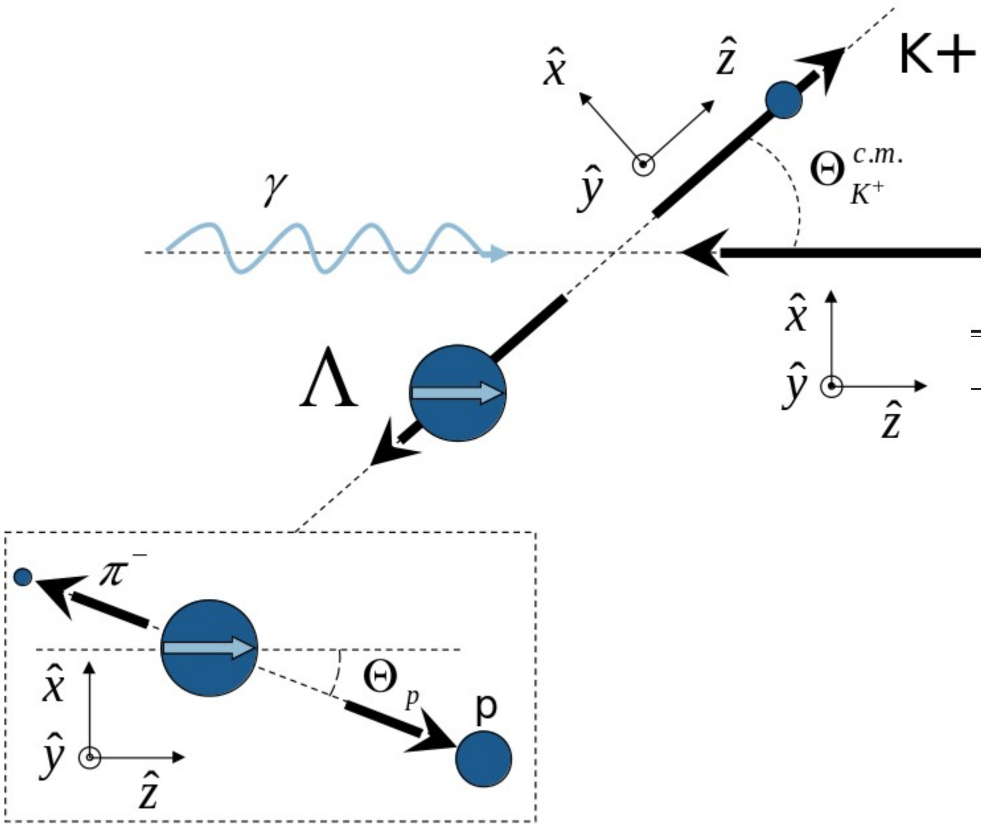


Background



$$\gamma p \longrightarrow K^+\Lambda \longrightarrow K^+\pi^- p$$

Background

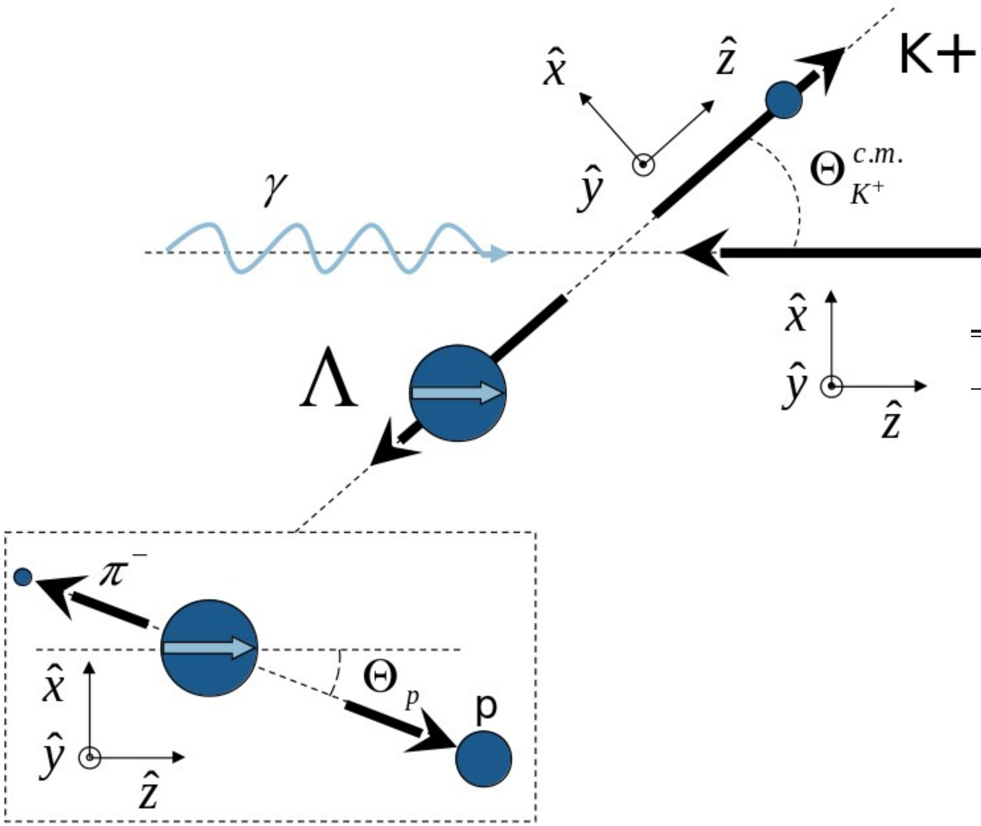


$$\gamma p \rightarrow K^+\Lambda \rightarrow K^+\pi^- p$$

Σ = beam asymmetry

Observable	Type	Amplitude Combination
Σ	Single	$\lvert a_1\rvert^2 + \lvert a_2\rvert^2 - \lvert a_3\rvert^2 - \lvert a_4\rvert^2$
P		$\lvert a_1\rvert^2 - \lvert a_2\rvert^2 + \lvert a_3\rvert^2 - \lvert a_4\rvert^2$
T		$\lvert a_1\rvert^2 - \lvert a_2\rvert^2 - \lvert a_3\rvert^2 + \lvert a_4\rvert^2$
E	Beam-target	$2\Re(a_1a_3^* + a_2a_4^*)$
F		$2\Im(a_1a_3^* - a_2a_4^*)$
G		$2\Im(a_1a_3^* + a_2a_4^*)$
H		$-2\Re(a_1a_3^* - a_2a_4^*)$
C_x	Beam-recoil	$-2\Im(a_1a_4^* - a_2a_3^*)$
C_z		$2\Re(a_1a_4^* + a_2a_3^*)$
O_x		$2\Re(a_1a_4^* - a_2a_3^*)$
O_z		$2\Im(a_1a_4^* + a_2a_3^*)$
T_x	Target-recoil	$2\Re(a_1a_2^* - a_3a_4^*)$
T_z		$2\Im(a_1a_2^* - a_3a_4^*)$
L_x		$-2\Im(a_1a_2^* + a_3a_4^*)$
L_z		$2\Re(a_1a_2^* + a_3a_4^*)$

Background



$$\gamma p \rightarrow K^+\Lambda \rightarrow K^+\pi^- p$$

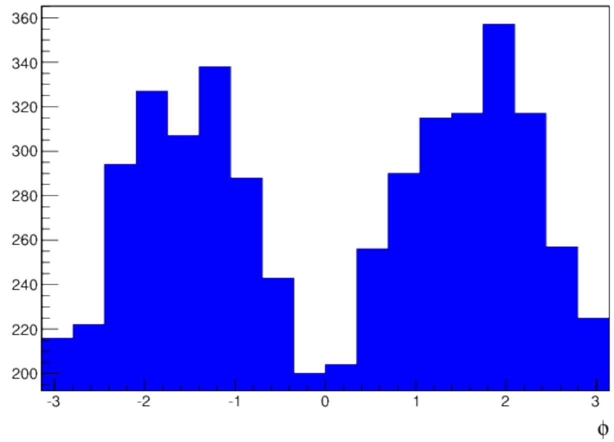
Σ = beam asymmetry

Observable	Type	Amplitude Combination
Σ	Single	$\lvert a_1\rvert^2 + \lvert a_2\rvert^2 - \lvert a_3\rvert^2 - \lvert a_4\rvert^2$
P		$\lvert a_1\rvert^2 - \lvert a_2\rvert^2 + \lvert a_3\rvert^2 - \lvert a_4\rvert^2$
T		$\lvert a_1\rvert^2 - \lvert a_2\rvert^2 - \lvert a_3\rvert^2 + \lvert a_4\rvert^2$
E	Beam-target	$2\Re(a_1 a_3^* + a_2 a_4^*)$
F		$2\Im(a_1 a_3^* - a_2 a_4^*)$
G		$2\Im(a_1 a_3^* + a_2 a_4^*)$
H		$-2\Re(a_1 a_3^* - a_2 a_4^*)$
C_x	Beam-recoil	$-2\Im(a_1 a_4^* - a_2 a_3^*)$
C_z		$2\Re(a_1 a_4^* + a_2 a_3^*)$
O_x		$2\Re(a_1 a_4^* - a_2 a_3^*)$
O_z		$2\Im(a_1 a_4^* + a_2 a_3^*)$
T_x	Target-recoil	$2\Re(a_1 a_2^* - a_3 a_4^*)$
T_z		$2\Im(a_1 a_2^* - a_3 a_4^*)$
L_x		$-2\Im(a_1 a_2^* + a_3 a_4^*)$
L_z		$2\Re(a_1 a_2^* + a_3 a_4^*)$

Current Data Analysis

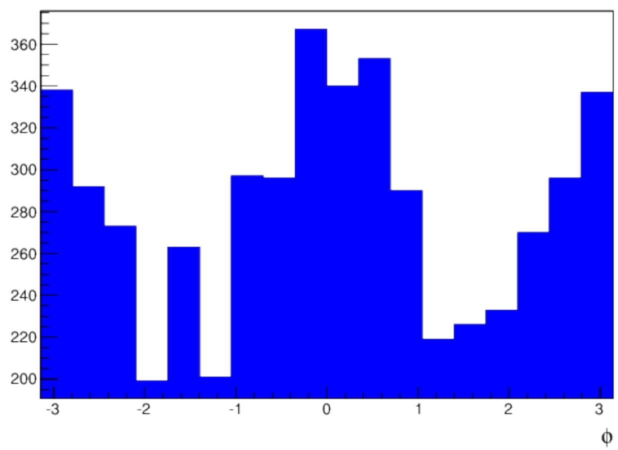
Beam

Para



Beam

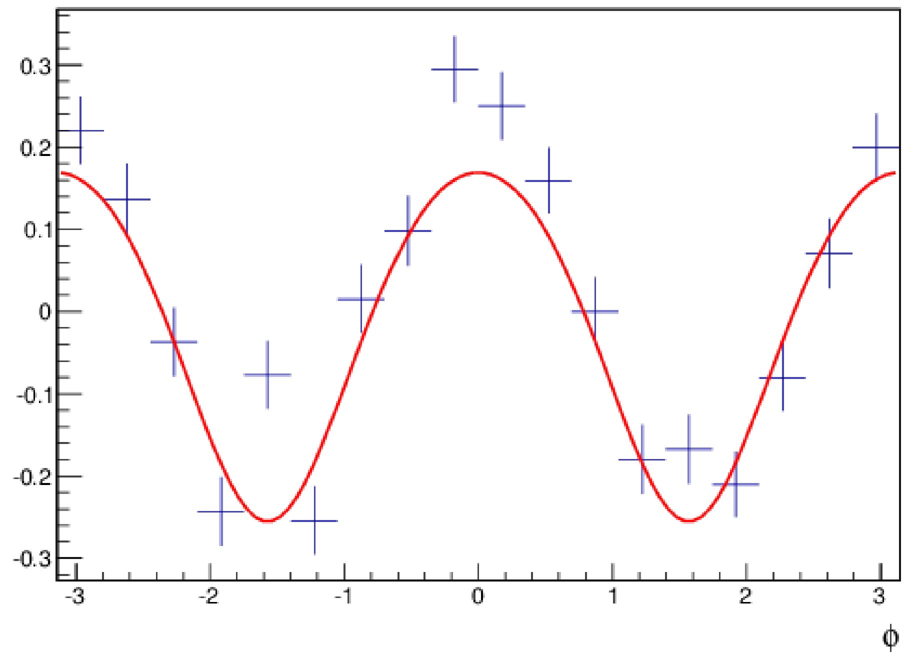
Perp



Chi-squared fitting to an asymmetry:

$$A = \frac{\sigma^{\perp} - \sigma^{\parallel}}{\sigma^{\perp} + \sigma^{\parallel}}$$

Asymmetry with sinusoidal fit function



- Bayesian Analysis!

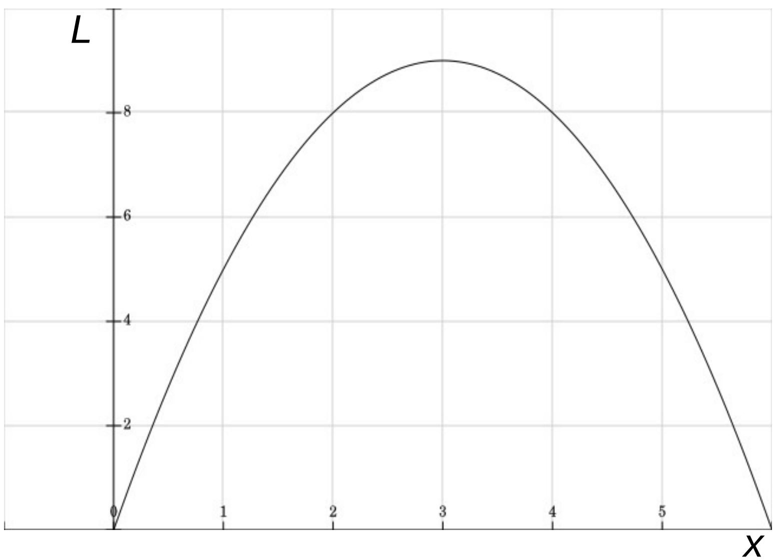
$$Posterior \propto Prior \times Likelihood$$

  - *Prior*: Initial probability distribution, contains all constraints
  - *Likelihood*: Event-by-event function
  - *Posterior*: resulting probability distribution

A New Method: Nested Sampling

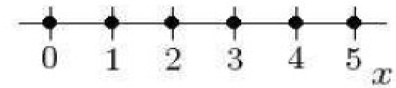
Likelihood function:

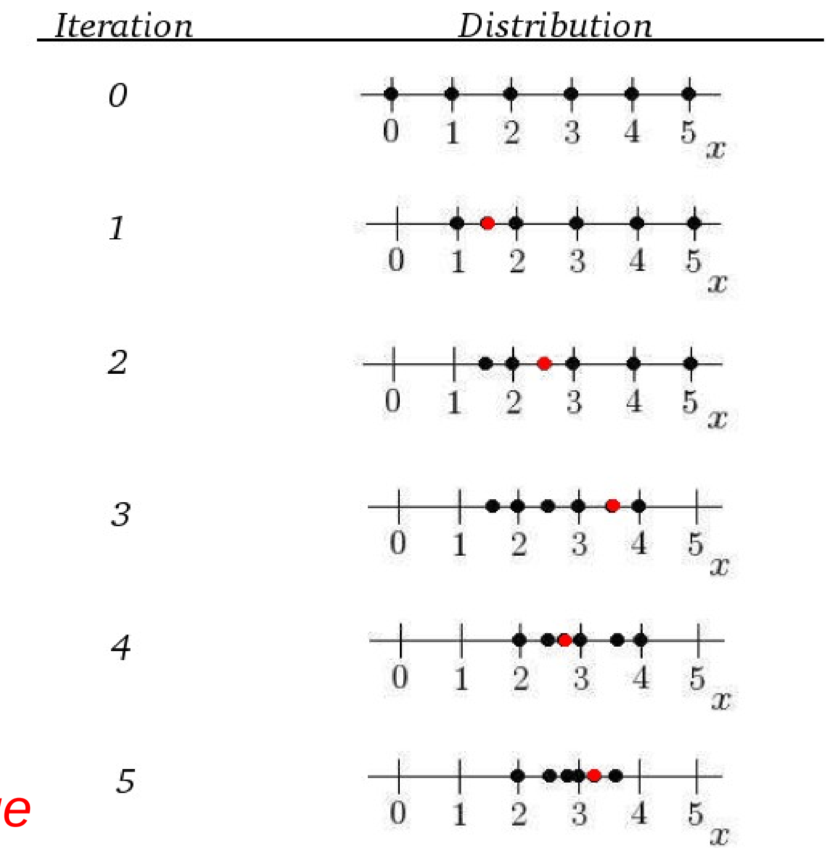
$$L = 6x - x^2$$



We want to find the maximum of this function.

Prior: uniform on (0,5):

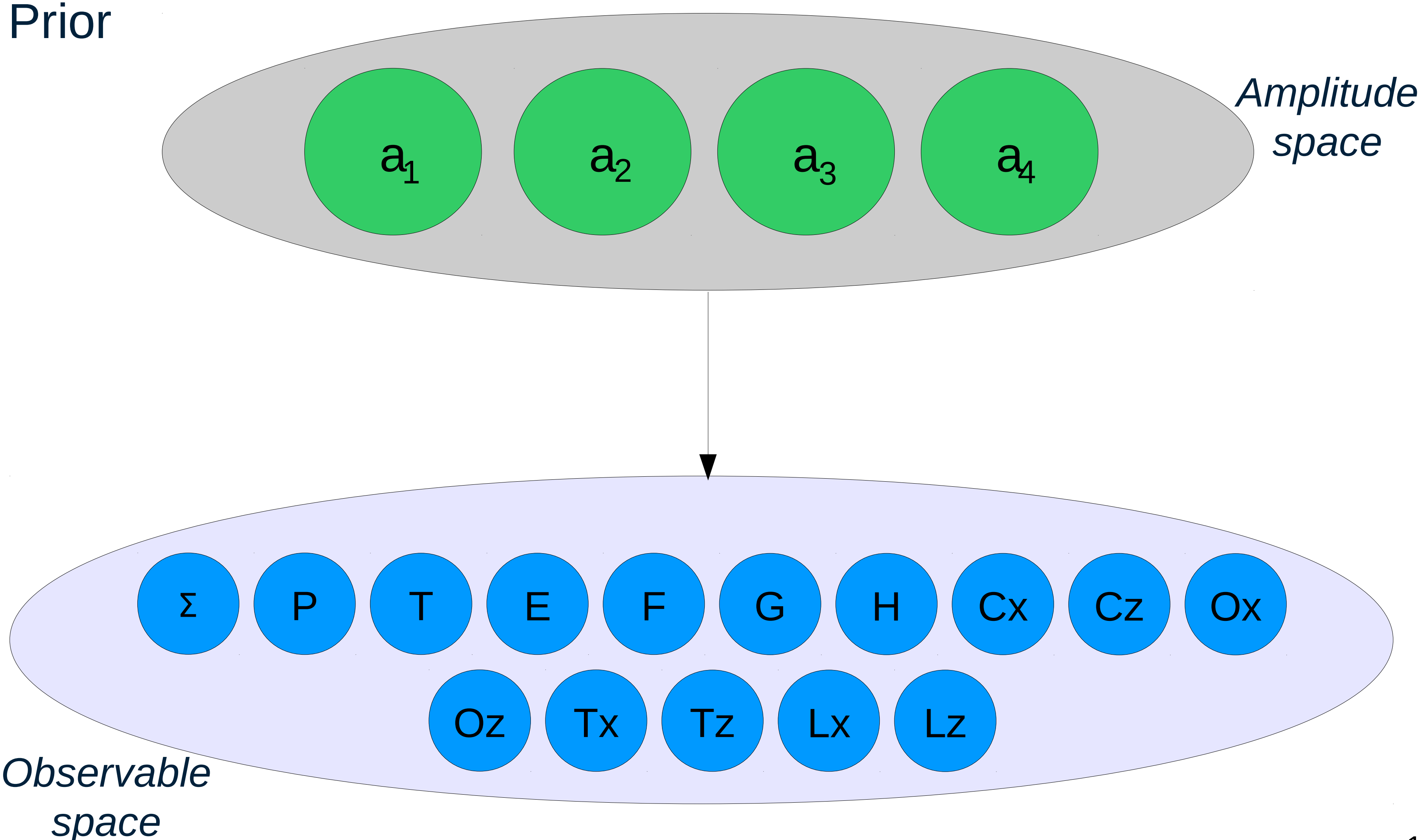




The points all converge around one value: x = 3.

Nested Sampling Program

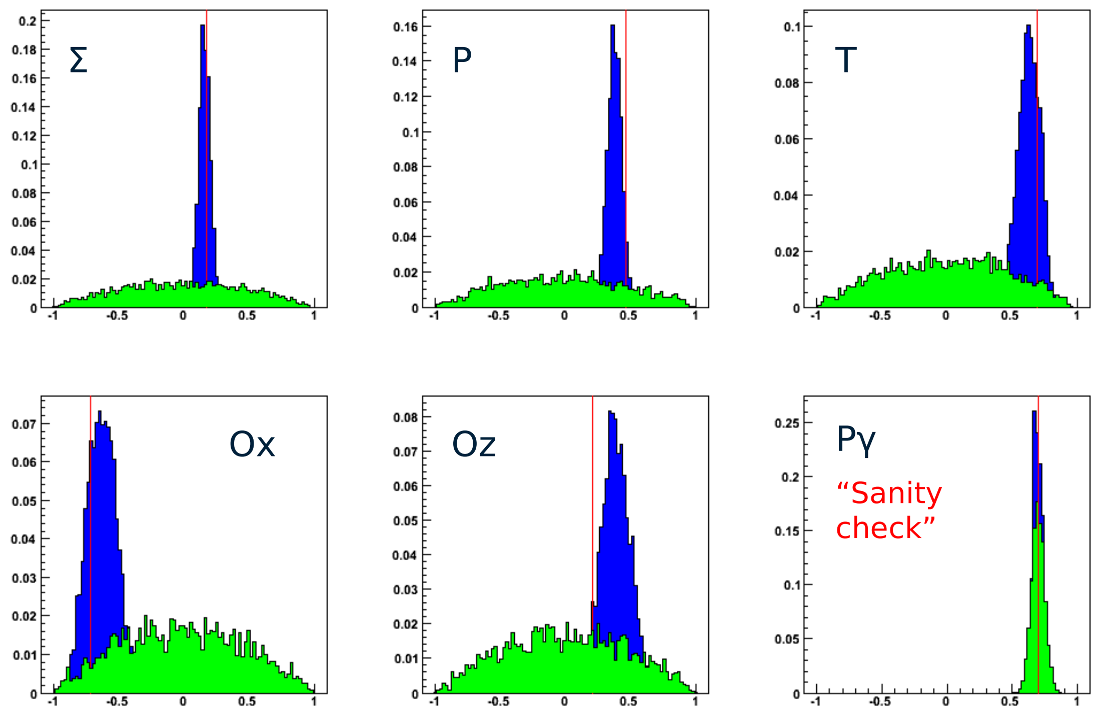
Prior



University of Glasgow

Data Analysis

- New method based on Bayesian analysis:
  - Can extract multiple observables simultaneously
  - Event-by-event, no information loss due to binning
  - Information (e.g. constraints) taken into account



1000 Events

Probability density functions

Green: prior

Blue: posterior

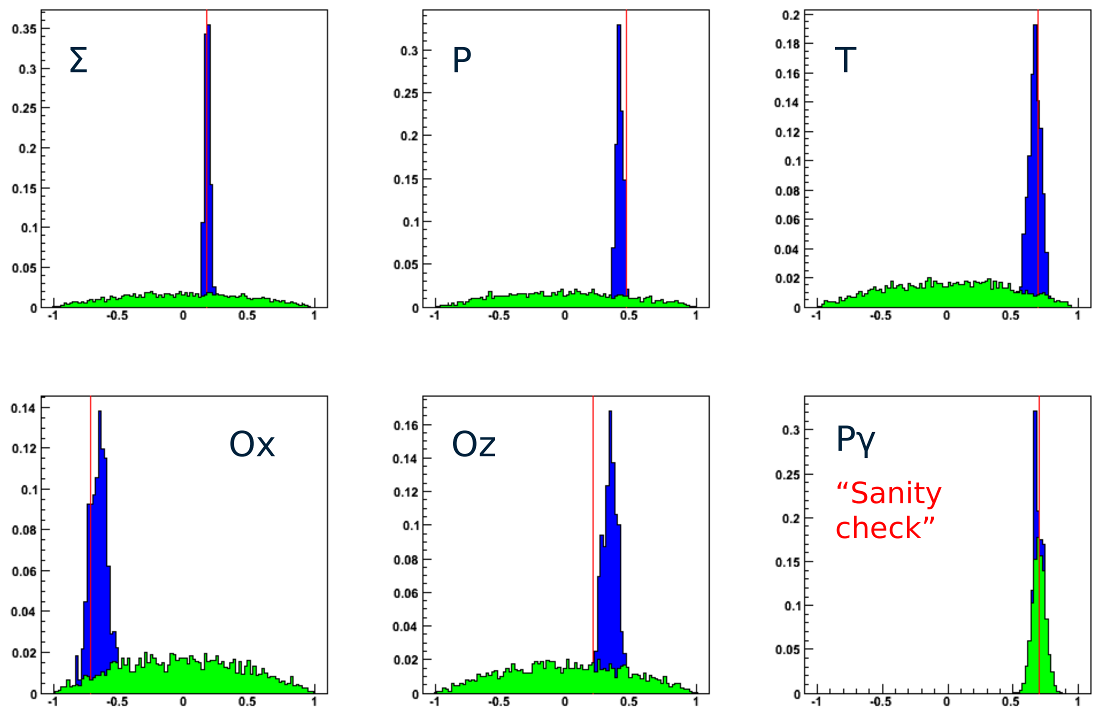
Simulated data

Data Analysis

- New method based on Bayesian analysis:
  - Can extract multiple observables simultaneously
  - Event-by-event, no information loss due to binning
  - Information (e.g. constraints) taken into account

5000 Events

Probability density functions



Green: prior

Blue: posterior

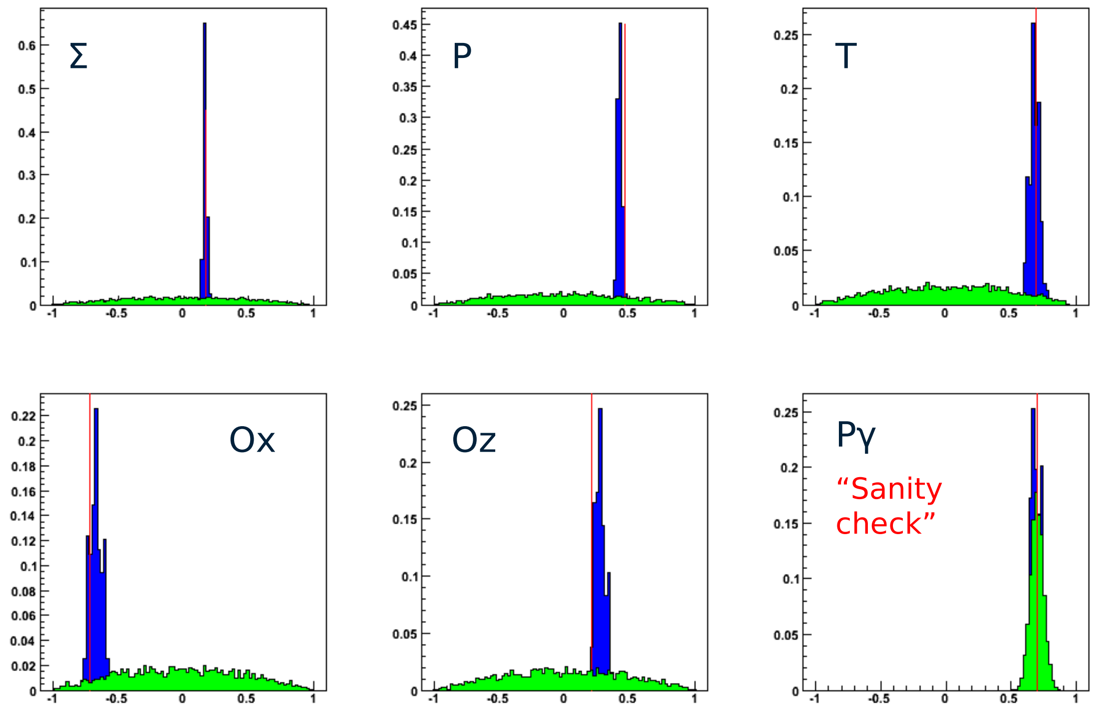
Simulated data

Data Analysis

- New method based on Bayesian analysis:
  - Can extract multiple observables simultaneously
  - Event-by-event, no information loss due to binning
  - Information (e.g. constraints) taken into account

10000 Events

Probability density functions



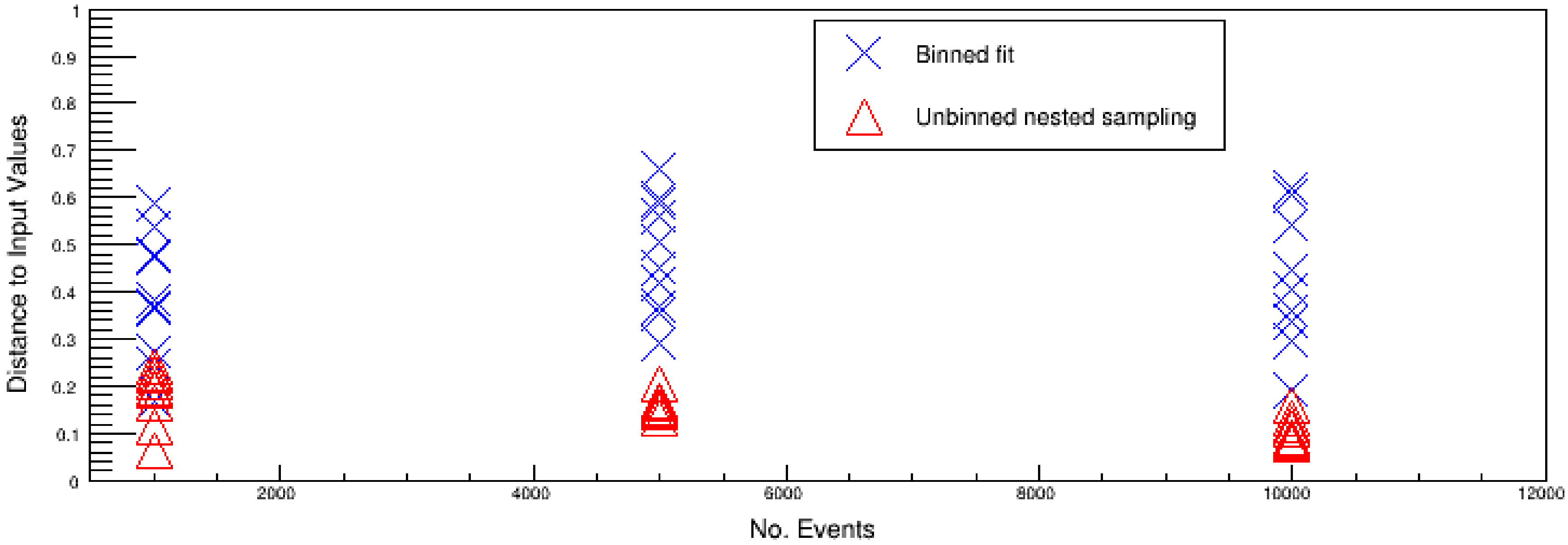
Green: prior

Blue: posterior

Simulated data

Quantitative comparison to current chi-squared method

$$D^2 = \sum_i (\mathcal{O}_i^{true} - \mathcal{O}_i^{calc})^2$$

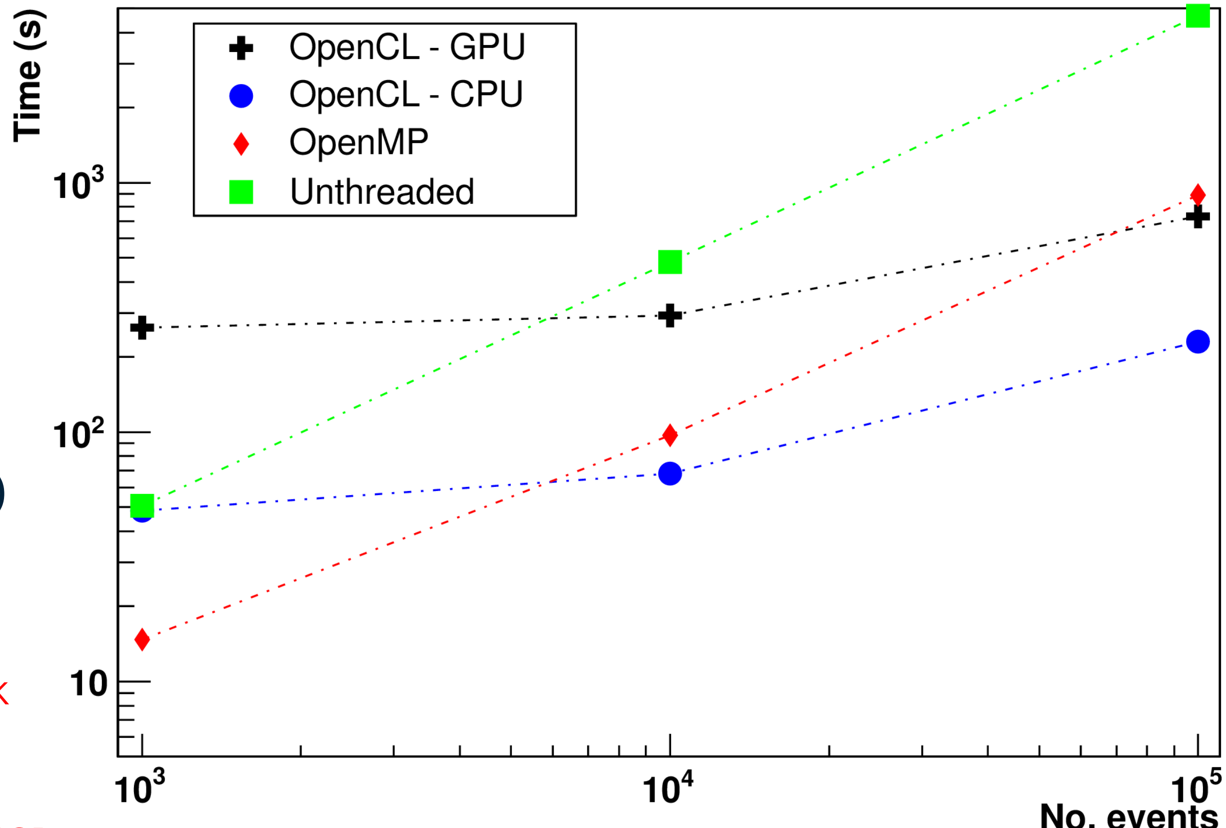


Data Parallelism

Goal: To speed up run-time of analysis program

Implementations:

- OpenCL on CPU
- OpenCL on GPU
- OpenMP
- Original, unthreaded on CPU (after optimisation)



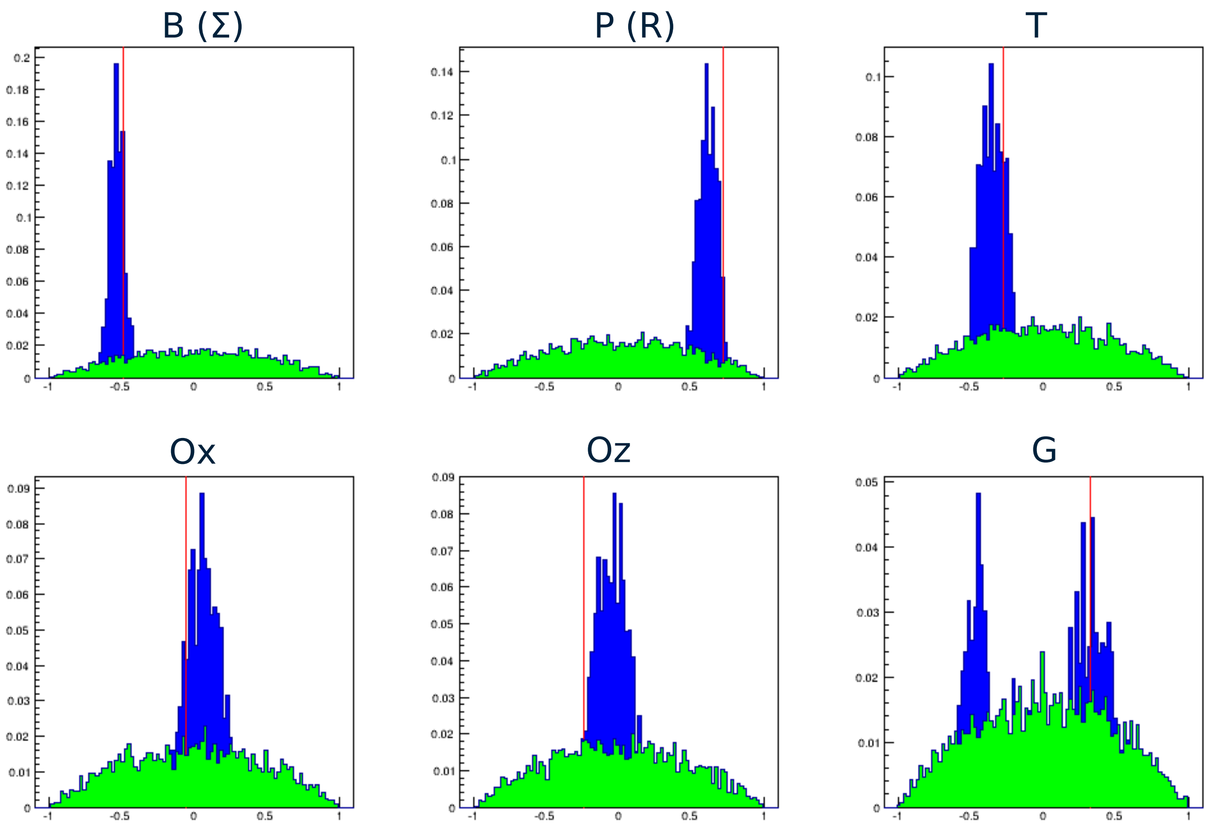
Run on an Intel® Core™ i7-2700K CPU @ 3.50GHz,

32GB RAM

GPU: NVIDIA® Tesla® S2075 6GB GDDR5

University of Glasgow

Combining Datasets



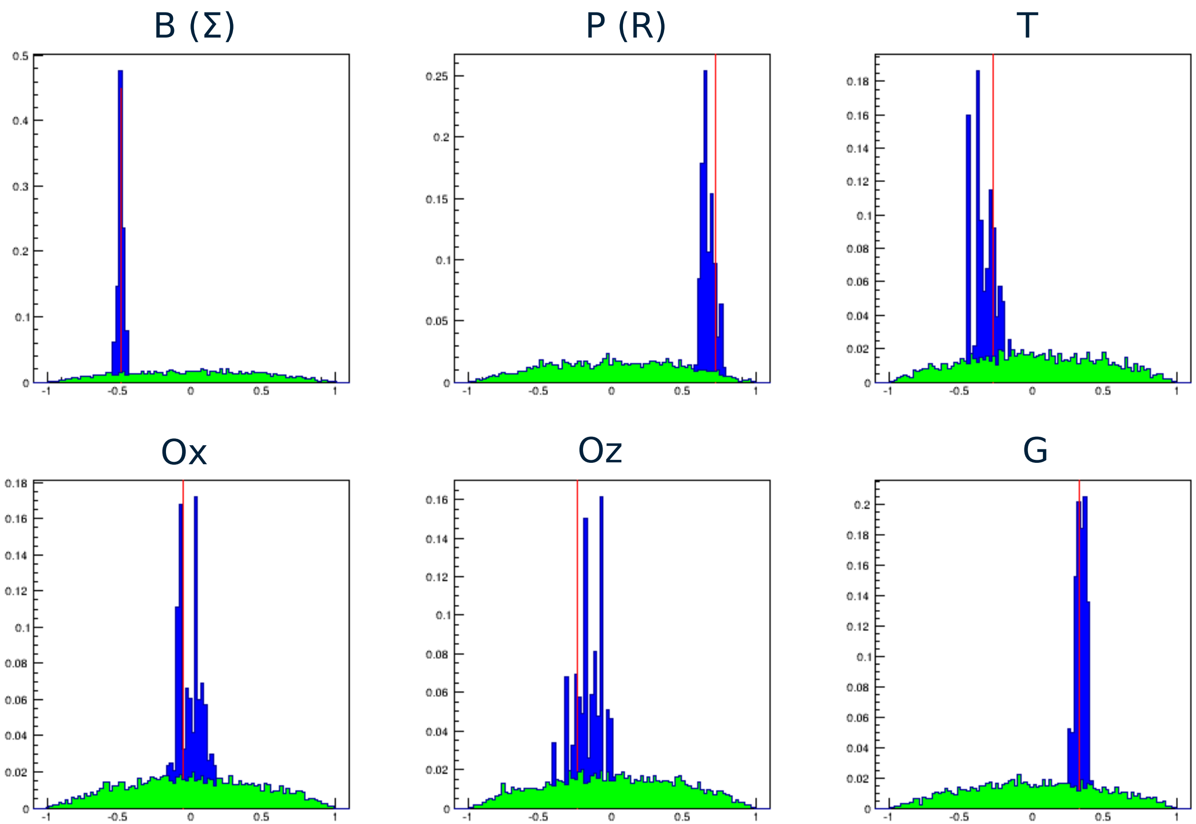
- Currently proof-of-concept using simulated data
- Can combine datasets from different experiments with consistent results!
- Exp 1 yields Σ, P, T, Ox, Oz.
- Exp 2 alone yields Σ, G.

Exp 1: *Linearly polarised beam, unpolarised target, recoil detected*

Exp 2: *Linearly polarised beam, longitudinally polarised target, no recoil information*

University of Glasgow

Combining Datasets



- Currently proof-of-concept using simulated data
- Can combine datasets from different PMP experiments with consistent results!
- Exp 1 yields Σ, P, T, Ox, Oz.
- Exp 2 alone yields Σ, G.

Exp 1: *Linearly polarised beam, unpolarised target, recoil detected*

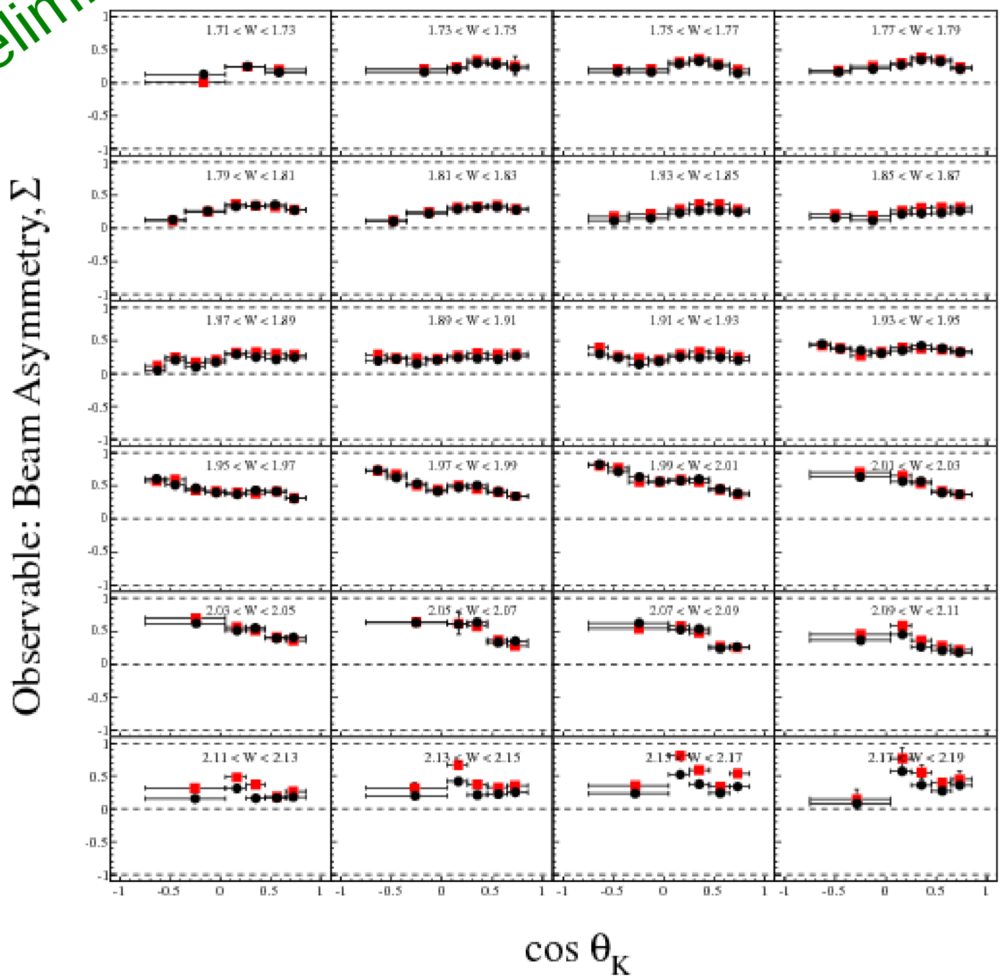
Exp 2: *Linearly polarised beam, longitudinally polarised target, no recoil information*

Results from data

Nested sampling results compared to Maximum Likelihood (ML) results

$$\gamma + p \rightarrow K^+ \Lambda$$

Preliminary



- Linear beam, unpolarised target experiment
- Liquid hydrogen target
- Photon energy 1.3 – 2.1 GeV

Black circles = nested sampling

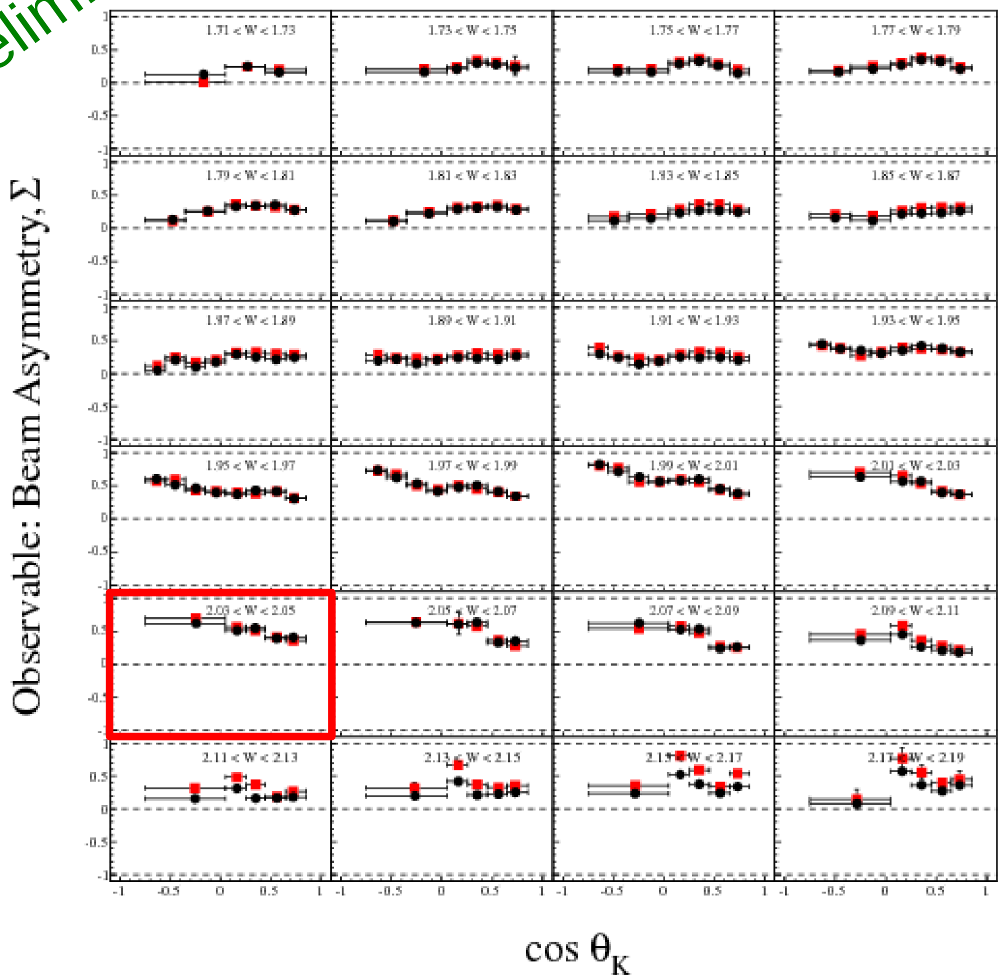
Red squares = ML

Results from data

Nested sampling results compared to Maximum Likelihood (ML) results

$$\gamma + p \rightarrow K^+ \Lambda$$

Preliminary



- Linear beam, unpolarised target experiment
- Liquid hydrogen target
- Photon energy 1.3 – 2.1 GeV

Black circles = nested sampling

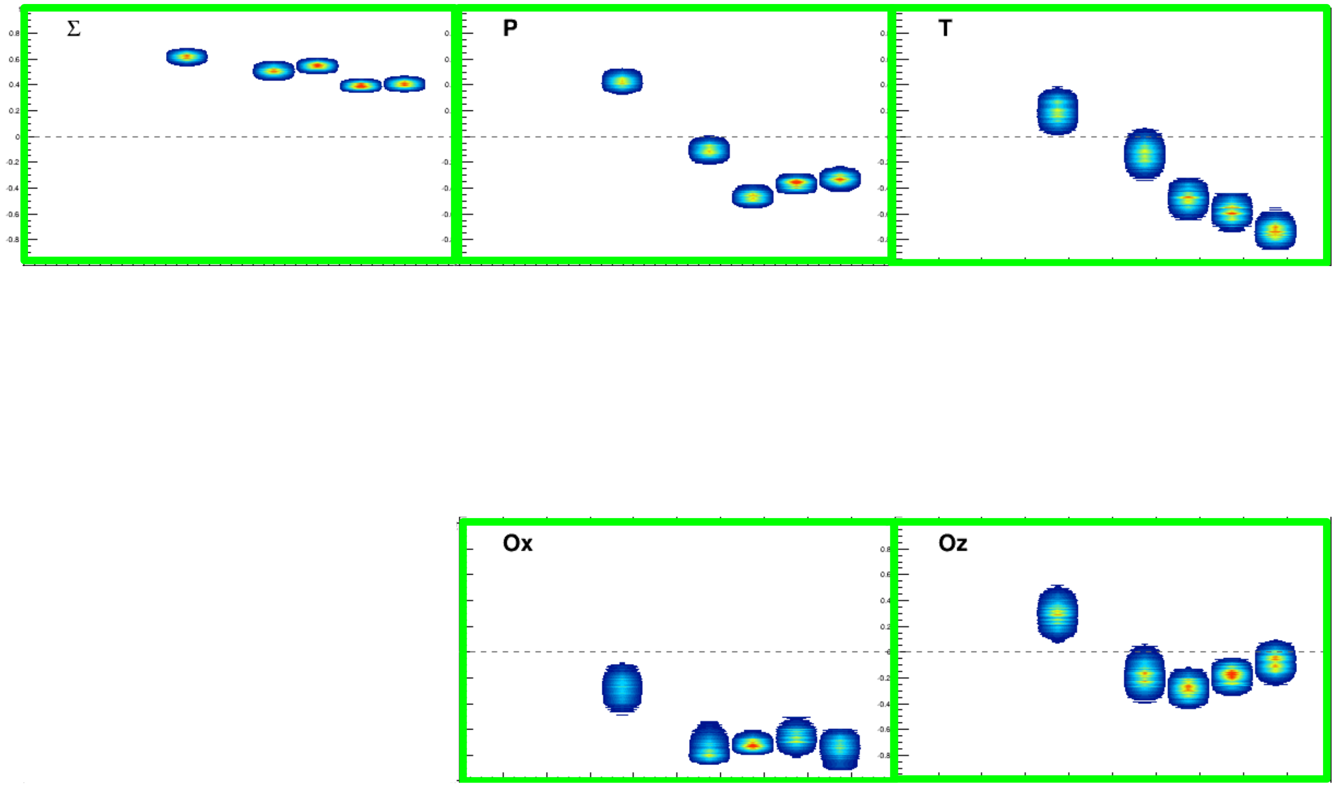
Red squares = ML

University of Glasgow

Results from data

2.03 GeV < W < 2.05 GeV

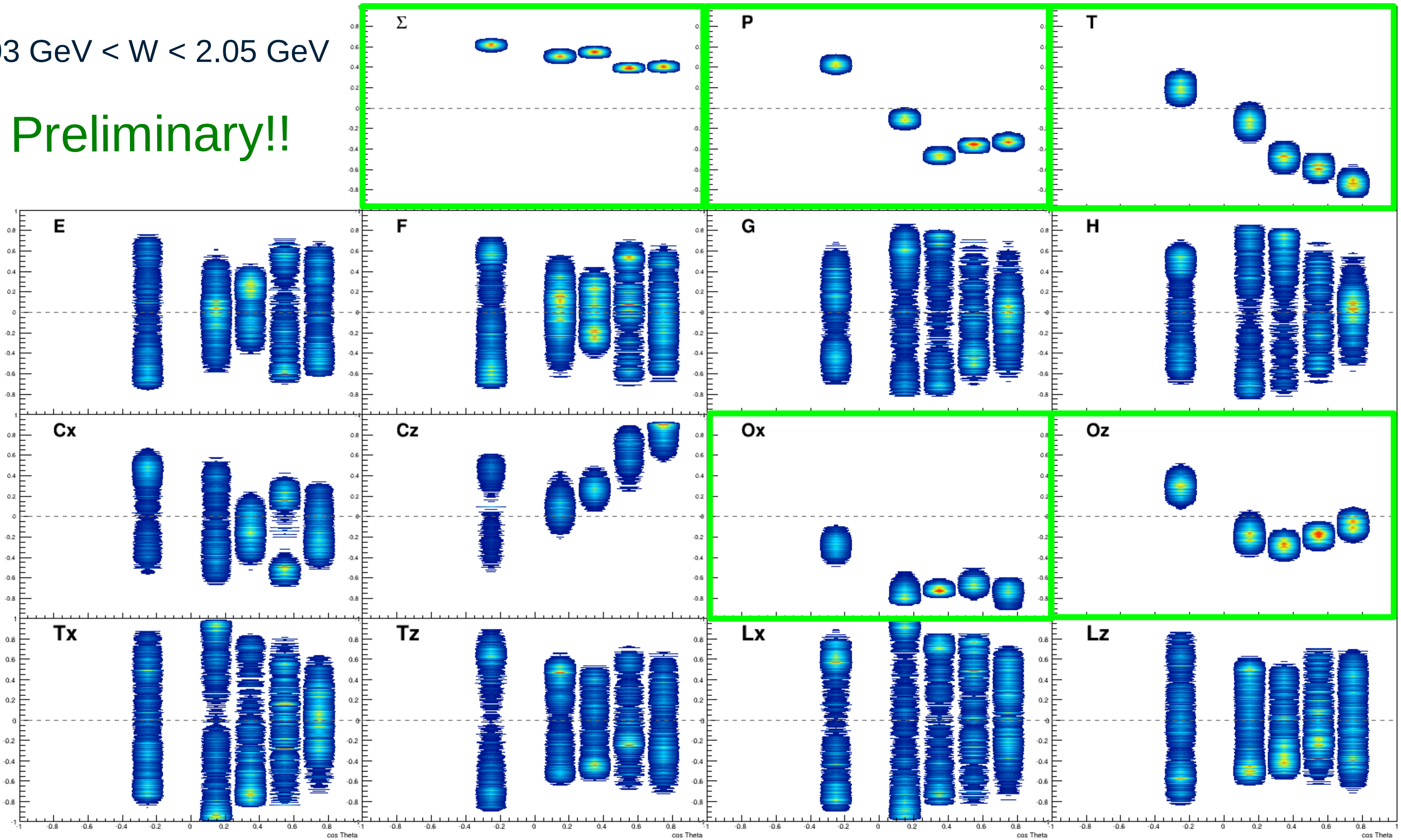
Preliminary!!



Results from data

2.03 GeV < W < 2.05 GeV

Preliminary!!



cos(θ) cos(θ) cos(θ) cos(θ)

Conclusions and Outlook

- Conclusions:
  - Pseudoscalar meson photoproduction used to find missing resonances
  - Measurement of polarisation observables is essential
  - New Bayesian analysis program has been developed to extract polarisation observables
  - Data parallelism can be used to speed up analysis program
  - Nested sampling method is consistent with current analysis methods and has advantages

- Outlook:
  - Combine datasets from different experiments

Acknowledgements:

Prof David Ireland and Dr Ken Livingston, School of Physics and Astronomy, University of Glasgow

Dr Wim Vanderbauwhede, School of Computing Science, University of Glasgow