Field Off Scattering Studies: Current Status

John Nugent

University of Glasgow

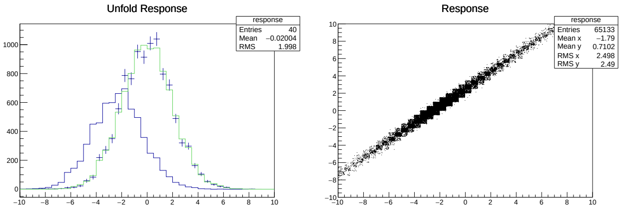
john.nugent@glasgow.ac.uk

2/8/2018

- All referee's comment from 25/6/18 completed
- Deconvolution
  - worked through Bayesian optimisation
  - added Gold Deconvolution algorithm

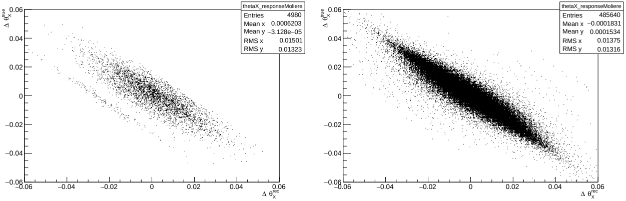
Bayesian approach first principles

- Start from simplest example provided by RooUnfold package



- Blue line measured
- Green line Truth
- Blue crosses deconvolved data

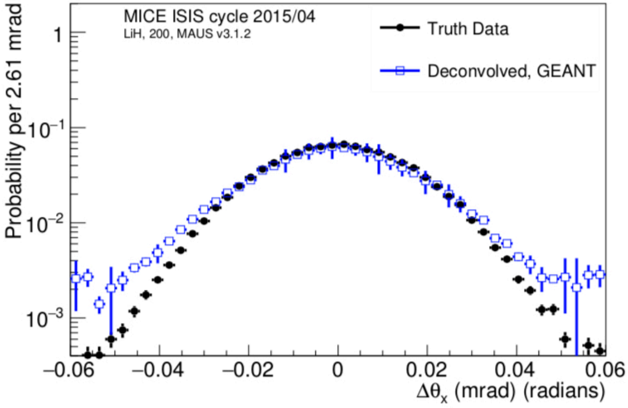
Bayesian response data vs MC



left is data and right is MC

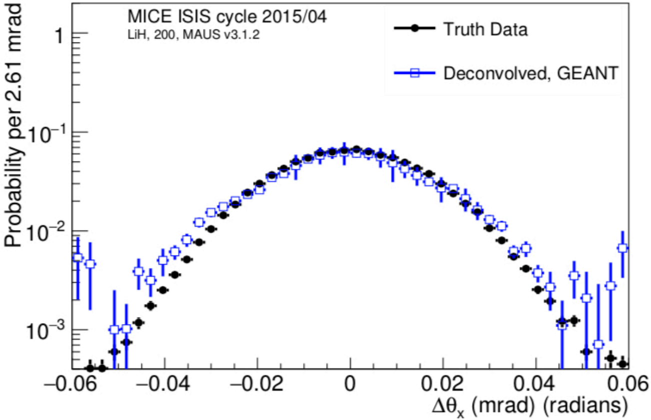
- "Measured" θ is US empty track vs DS empty convolved with model
- "Truth" θ is DS empty convolved with model vs DS empty data

MC Data comparison



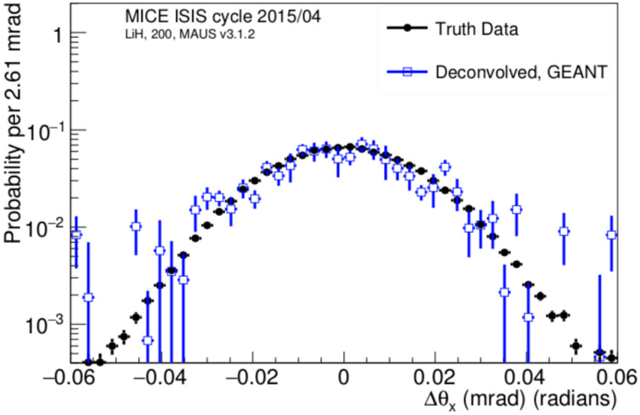
- 200 MeV/c case
- resampling 20, Baysian iter 10

MC Data comparison



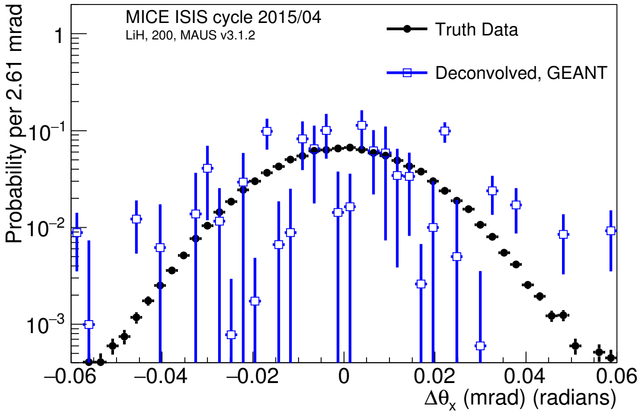
- 200 MeV/c case
- resampling 20, Baysian iter 100

MC Data comparison



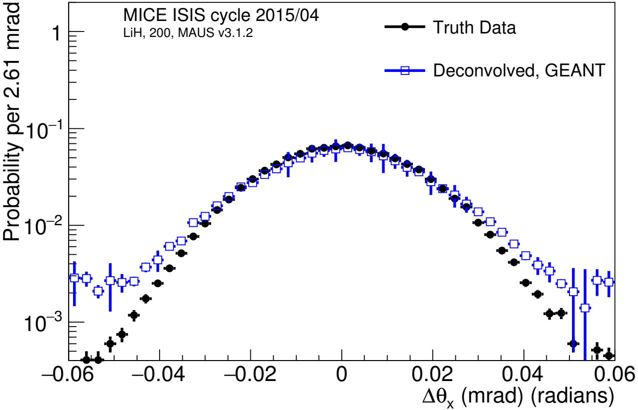
- 200 MeV/c case
- resampling 20, Baysian iter 1000

MC Data comparison



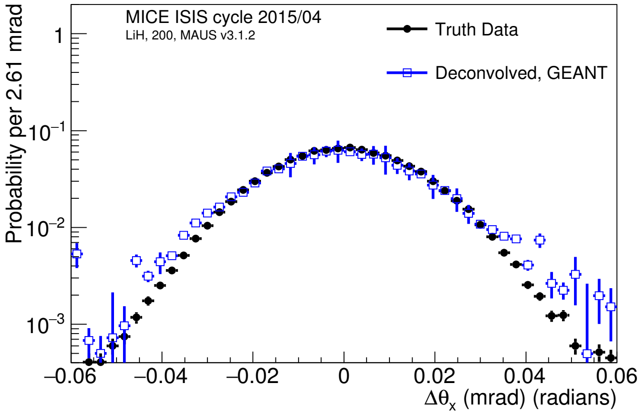
- 200 MeV/c case
- resampling 20, Baysian iter 10000

MC Data comparison



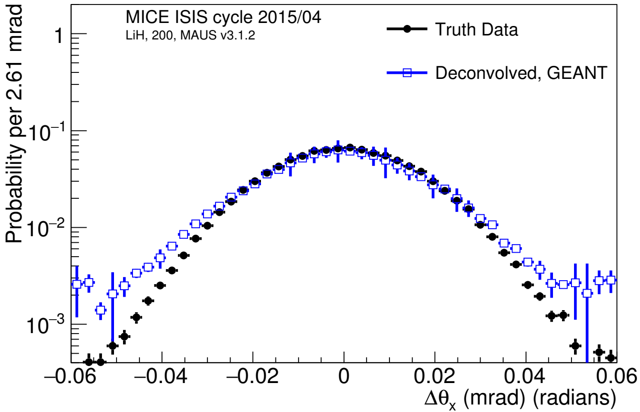
- 200 MeV/c case
- resampling 20, Baysian iter 10
- negative response matrix i.e. correlated

MC Data comparison



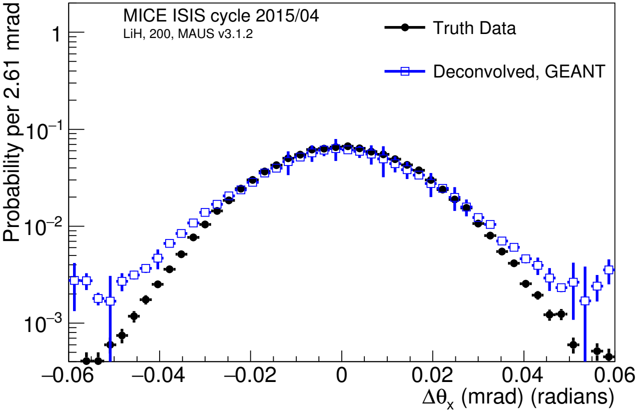
- 200 MeV/c case
- resampling 1, Baysian iter 10

MC Data comparison



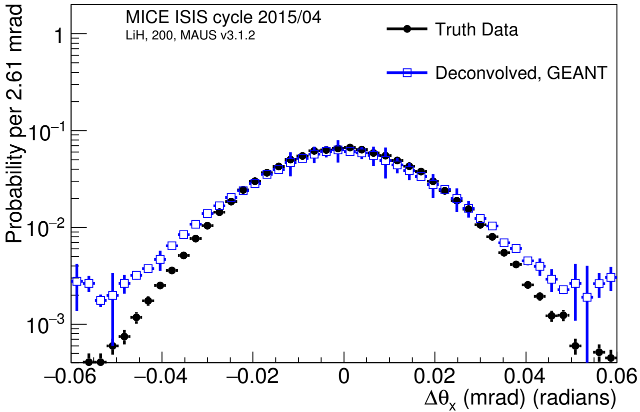
- 200 MeV/c case
- resampling 20, Baysian iter 10

MC Data comparison



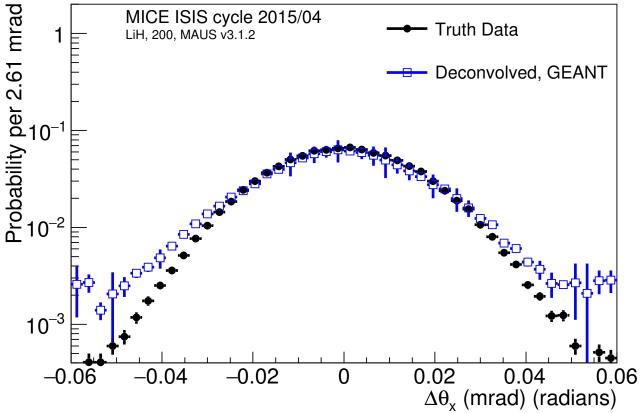
- 200 MeV/c case
- resampling 200, Baysian iter 10

MC Data comparison



- 200 MeV/c case
- resampling 2000, Baysian iter 10

MC Data comparison

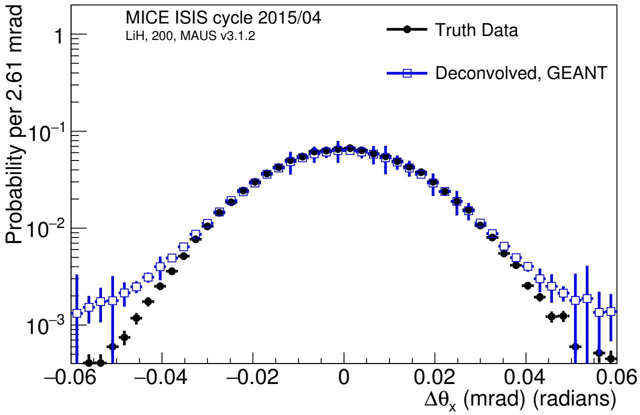


- 200 MeV/c case
- resampling 20, Baysian iter 10
- Cut on Truth P 198-202 MeV/c

Truth Response Matrix

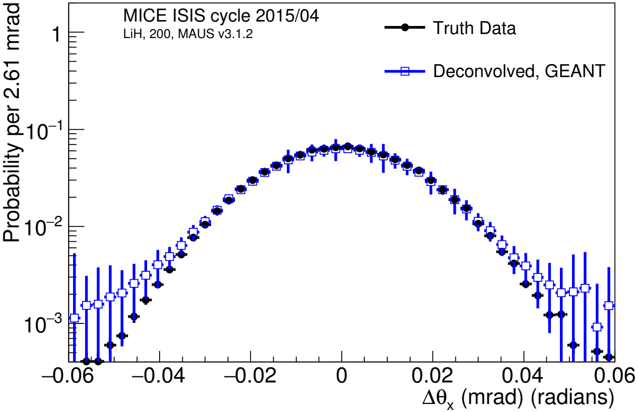
- Previous approach used the forward convolution to relate the measured track to the true scattering due to model
- Want a simplier approach similar to example shown on slide 3
- Use MC Truth and recon to build response matrix

MC Data comparison



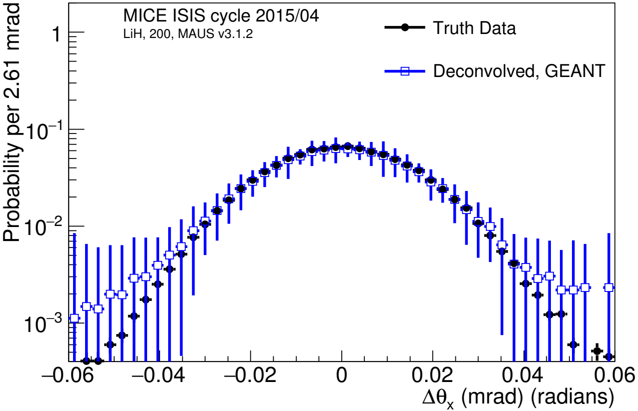
- 200 MeV/c case
- Baysian iter 10
- Use Truth response matrix AKA simplest possible RooUnfold example - no forward convolution

MC Data comparison



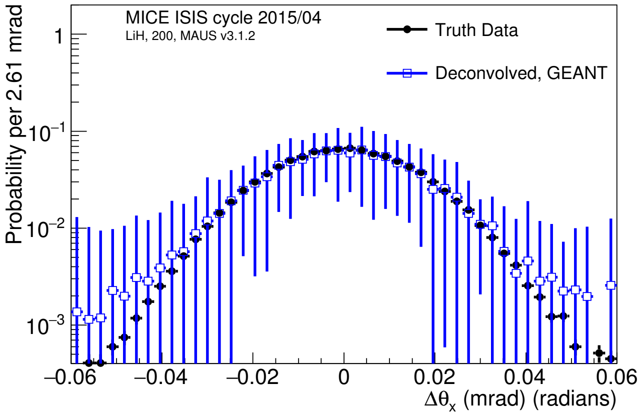
- 200 MeV/c case
- Baysian iter 100
- Use Truth response matrix AKA simplest possible RooUnfold example - no forward convolution

MC Data comparison



- 200 MeV/c case
- Baysian iter 1000
- Use Truth response matrix AKA simplest possible RooUnfold example - no forward convolution

MC Data comparison



- 200 MeV/c case
- Baysian iter 10000
- Use Truth response matrix AKA simplest possible RooUnfold example - no forward convolution

Gold Algorithm

- Looking to extract true scattering distributions from raw distributions. Remove effects of interstitial scattering, tracker resolution.
- Technique employed in nuclear γ-ray spectroscopy and image restoration.
- Does not rely on MC Truth or scattering models, purely data driven technique The scattering distribution that is measured by MICE can be stated as:

$$x'(\theta) = \int_{-\infty}^{t} x(\Theta)h(\theta - \Theta)d\Theta + n(\theta) = x(\theta) * h(\theta) + n(\theta), \quad (1)$$

where $x'(t)$ is the raw LiH scattering distribution measured. $h(t)$ is empty channel data includes the interstitial material + tracker resolution. $x(t)$ is scattering distribution due only to the absorber material. $n(t)$ is additive noise and the $*$ denotes the convolution operator.

Gold Algorithm

- For discrete systems, this statement can be expressed as:

$$x'(i) = \sum_{k=0}^{i} x(k)h(i-k) + n(i) = x(i) * h(i) + n(i), \tag{2}$$

an expression which represents a general system of linear equations that can be written in matrix form as:

$$x' = Hx + n \tag{3}$$

where the matrix H has dimension $N \times M$, the vectors x' and n have N elements and the vector x has M elements, while $N \geq M$. To find a least squares solution of the system of linear equations given in 3

$$||Hx - x'||^2 \tag{4}$$

must be minimised.

Gold Algorithm

$$x' = H'x \tag{5}$$

- where $H' = H^{\boldsymbol{T}} H H^{\boldsymbol{T}}$ and $H^{\boldsymbol{T}}$ is a Toeplitz matrix[1]. x' is known from data, and the method iterates over:

$$x_i^{(k+1)} = \frac{x_i'}{\sum_{m=0}^{N-1} H'_{im} x_m^{(k)}} x_i(K) \tag{6}$$

where

$$\begin{aligned} i &= 0, 1, ..., N-1, \\ k &= 1, 2, 3, ..., L, \\ x^0 &= [1, 1..., 1]^T \end{aligned} \tag{7}$$

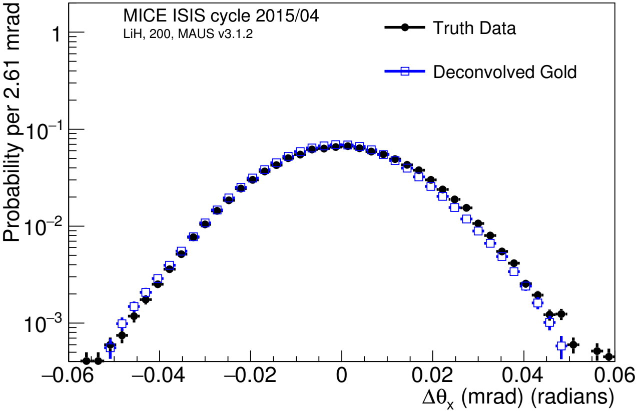
where L is the number of iterations.

[1] A Toeplitz matrix is an $n \times n$ matrix $T_n = [t_{k,j}; k, j = 0, 1, ..., n-1]$ where $t_{k,j} = t_{k-j}$

Gold Algorithm

- This method is encapsulated in a ROOT class TSpectrum
- The ROOT class accepts histograms as input and the scattering distributions for the two cases, with and without absorber, were used as input with the output being the final measured scattering distribution.
- Full details are availble: `https://root.cern.ch/doc/v608/classTSpectrum.html`

MC Data comparison

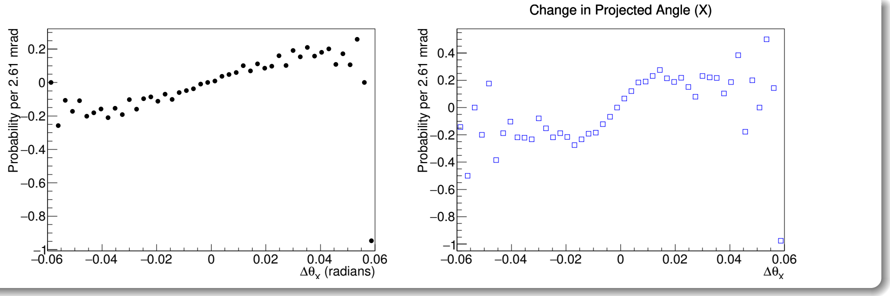


- 200 MeV/c case
- trkr acceptane + Gold iter 10
- Blue squares are symmetrically distributioned, black dots are asymmetric

Asymmetry

$$A_i = \frac{h_1 - h_2}{h_1 + h_2} \tag{8}$$

left is LiH MC recon & right is empty MC recon



- At Warwick workshop it was shown that if such an asymmetry is present in the data then there is a misalignment in the geometry
- A similar misalignment is in the MC geometry, as shown above
- When the batch system at Glasgow is working I will scan and correct

Job List

- Correct asymmetry in MC recon scattering distributions
- MC Truth selection

Bayesian Convergence

$$\chi^2 = \sum \frac{(P_1 - P_2)^2 N}{(P_1 + P_2)}$$