Dealing with high dimensional efficiency corrections in T2K cross section measurements

NuXTract Workshop 2023 @ CERN

5/10/23

Sam Jenkins
on behalf of the T2K collaboration

How do we calculate the cross section?

$$\frac{d\sigma}{dx_i} = \frac{N_i^{signal}}{\epsilon_i \Phi N_{targets}} \times \frac{1}{\Delta x_i}$$

How do we calculate the cross section?

$$\frac{d\sigma}{dx_i} = \frac{N_i^{signal}}{\epsilon_i \Phi N_{targets}} \times \frac{1}{\Delta x_i}$$

- Variable we measure in - can be done in as many dimensions as we want (single differential, double differential, etc…)

How do we calculate the cross section?

$$\frac{d\sigma}{dx_i} = \frac{N_i^{signal}}{\epsilon_i \Phi N_{targets}} \times \frac{1}{\Delta x_i}$$

- Variable we measure in - can be done in as many dimensions as we want (single differential, double differential, etc…)
- Signal we wish to measure after background subtraction, obtained via your favourite fitting method!

How do we calculate the cross section?

$$\frac{d\sigma}{dx_i} = \frac{N_i^{signal}}{\epsilon_i \Phi N_{targets}} \times \frac{1}{\Delta x_i}$$

Refer to Margherita's and Nick's talks earlier this week!

- Variable we measure in - can be done in as many dimensions as we want (single differential, double differential, etc…)
- Signal we wish to measure after background subtraction, obtained via your favourite fitting method!

How do we calculate the cross section?

$$\frac{d\sigma}{dx_i} = \frac{N_i^{signal}}{\epsilon_i \Phi N_{targets}} \times \frac{1}{\Delta x_i}$$

Refer to Margherita's and Nick's talks earlier this week!

- Variable we measure in - can be done in as many dimensions as we want (single differential, double differential, etc…)
- Signal we wish to measure after background subtraction, obtained via your favourite fitting method!
- N_i^{signal} is only the signal events we measure, inherently limited by imperfect detectors - normalise by efficiency to correct for the events we miss

How do we calculate the cross section?

$$\frac{d\sigma}{dx_i} = \frac{N_i^{signal}}{\epsilon_i \Phi N_{targets}} \times \frac{1}{\Delta x_i}$$

Refer to Margherita's and Nick's talks earlier this week!

- Variable we measure in - can be done in as many dimensions as we want (single differential, double differential, etc…)
- Signal we wish to measure after background subtraction, obtained via your favourite fitting method!
- N_i^{signal} is only the signal events we measure, inherently limited by imperfect detectors - normalise by efficiency to correct for the events we miss
- Further normalisation for flux, number of targets and width of kinematic bin we measure in

How do we calculate the cross section?

$$\frac{d\sigma}{dx_i} = \frac{N_i^{signal}}{\epsilon_i \Phi N_{targets}} \times \frac{1}{\Delta x_i}$$

Refer to Margherita's and Nick's talks earlier this week!

- Variable we measure in - can be done in as many dimensions as we want (single differential, double differential, etc…)
- Signal we wish to measure after background subtraction, obtained via your favourite fitting method!
- N_i^{signal} is only the signal events we measure, inherently limited by imperfect detectors - normalise by efficiency to correct for the events we miss
- Further normalisation for flux, number of targets and width of kinematic bin we measure in

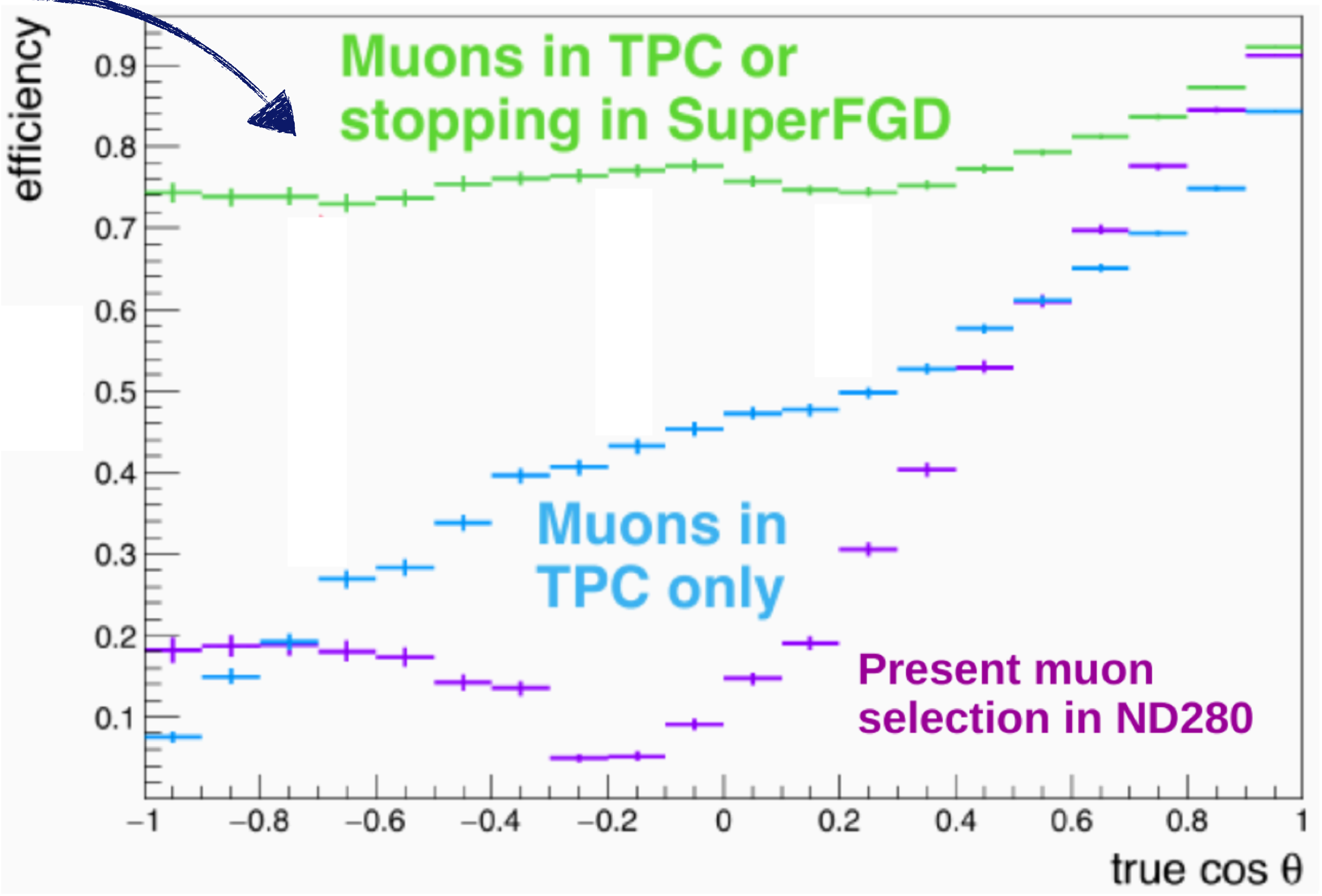
Efficiency correction

- No matter how good we make our detectors, we'll never be able to measure every event that takes place
- Define selection efficiency in the i^{th} kinematic bin as:

$$\epsilon_i = \frac{N^{signal}_{obs,\ i}}{N^{signal}_{true,\ i}}$$

- So if we normalise by ϵ_i in the cross section calculation, we recover the number of signal events we *should* have seen if our detector was perfect
- Simple, right?

Old ND280 vs ND280Upgrade detector efficiency

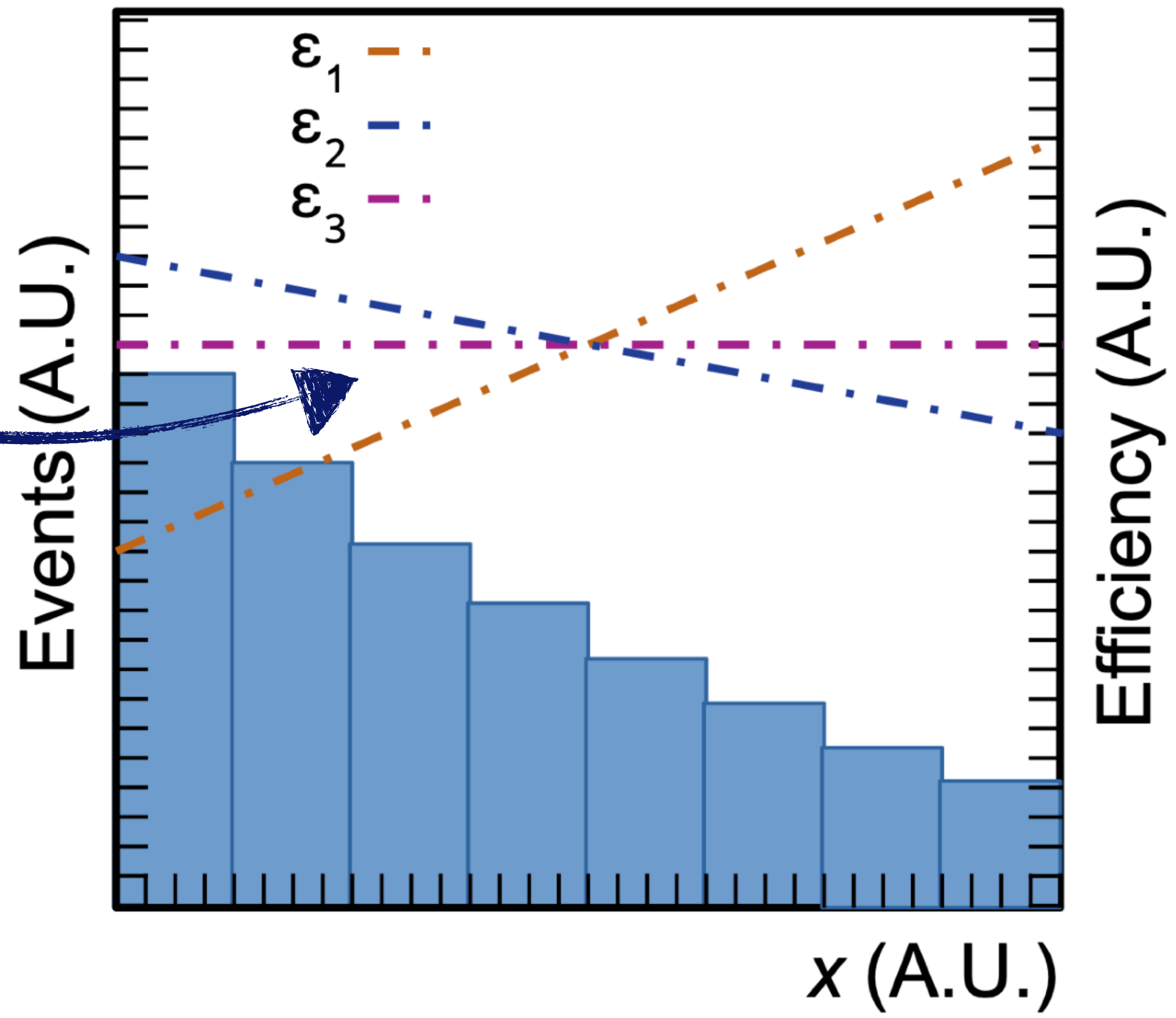


$$\frac{d\sigma}{dx_i} = \frac{N_i^{signal}}{\epsilon_i \Phi N_{targets}} \times \frac{1}{\Delta x_i}$$

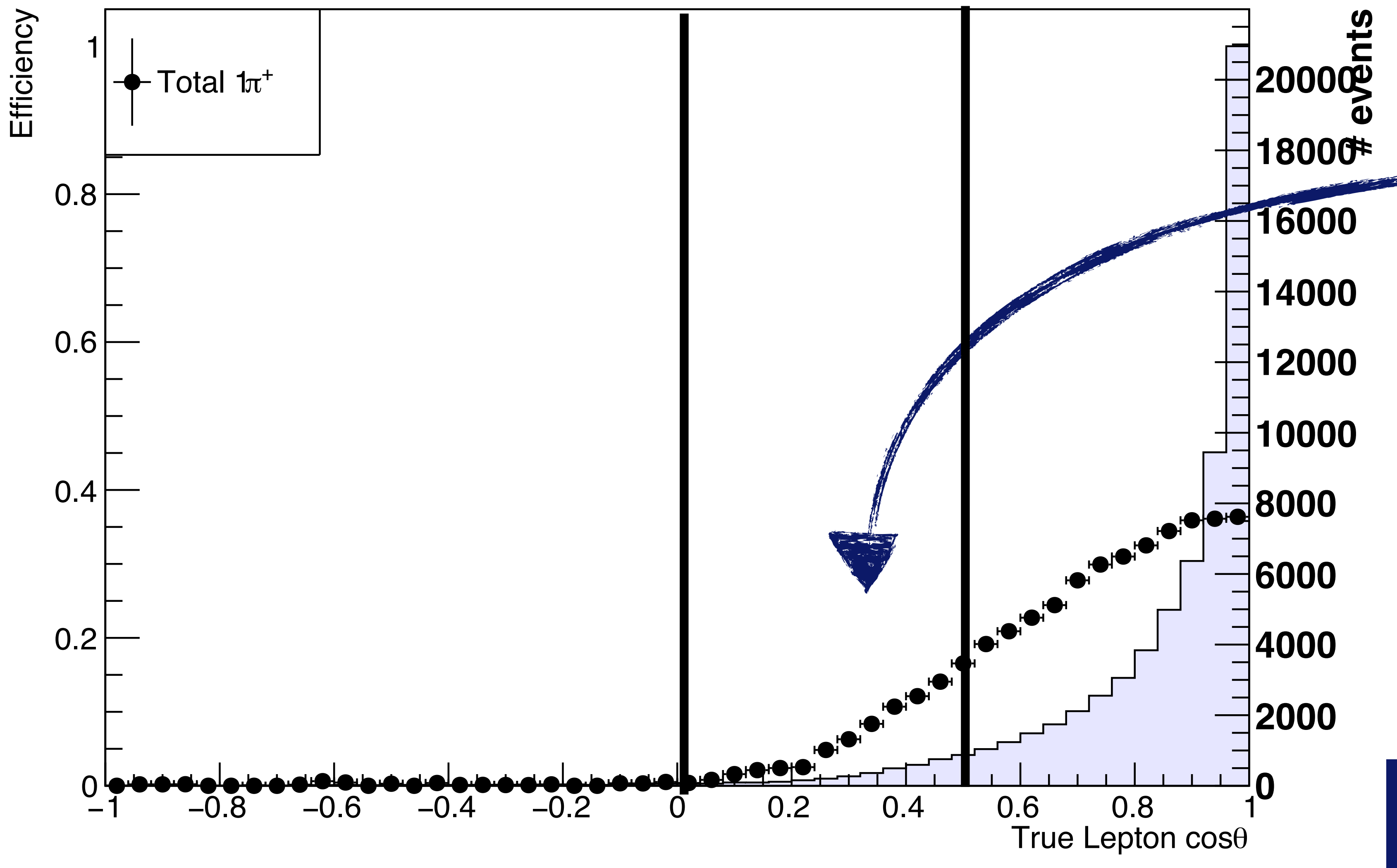
Thorsten Lux 2022 J. Phys.: Conf. Ser. 2374 012036

Efficiency correction

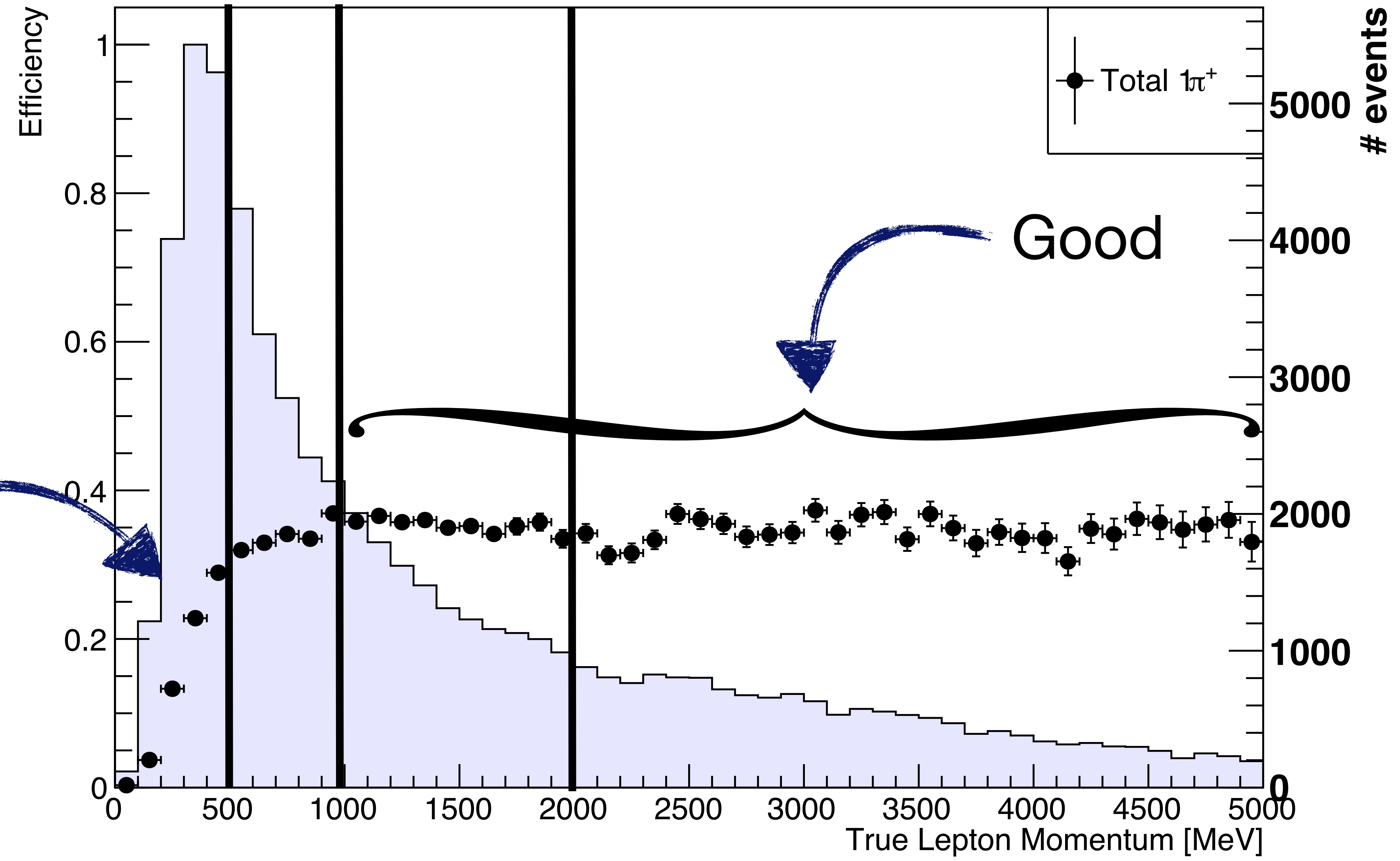
Image from
D. Cherdack



- …not quite - depending on where we place our bin edges for the variable of interest, we start to get different results
- Each ϵ_i has the same average, but bin-by-bin will yield very different corrections
- Want efficiency to be as flat as possible within a bin - if it's changing, we rely on the underlying MC → model bias
- Also have to worry about other variables we might not bin in
- So how can we deal with this?



Questionable

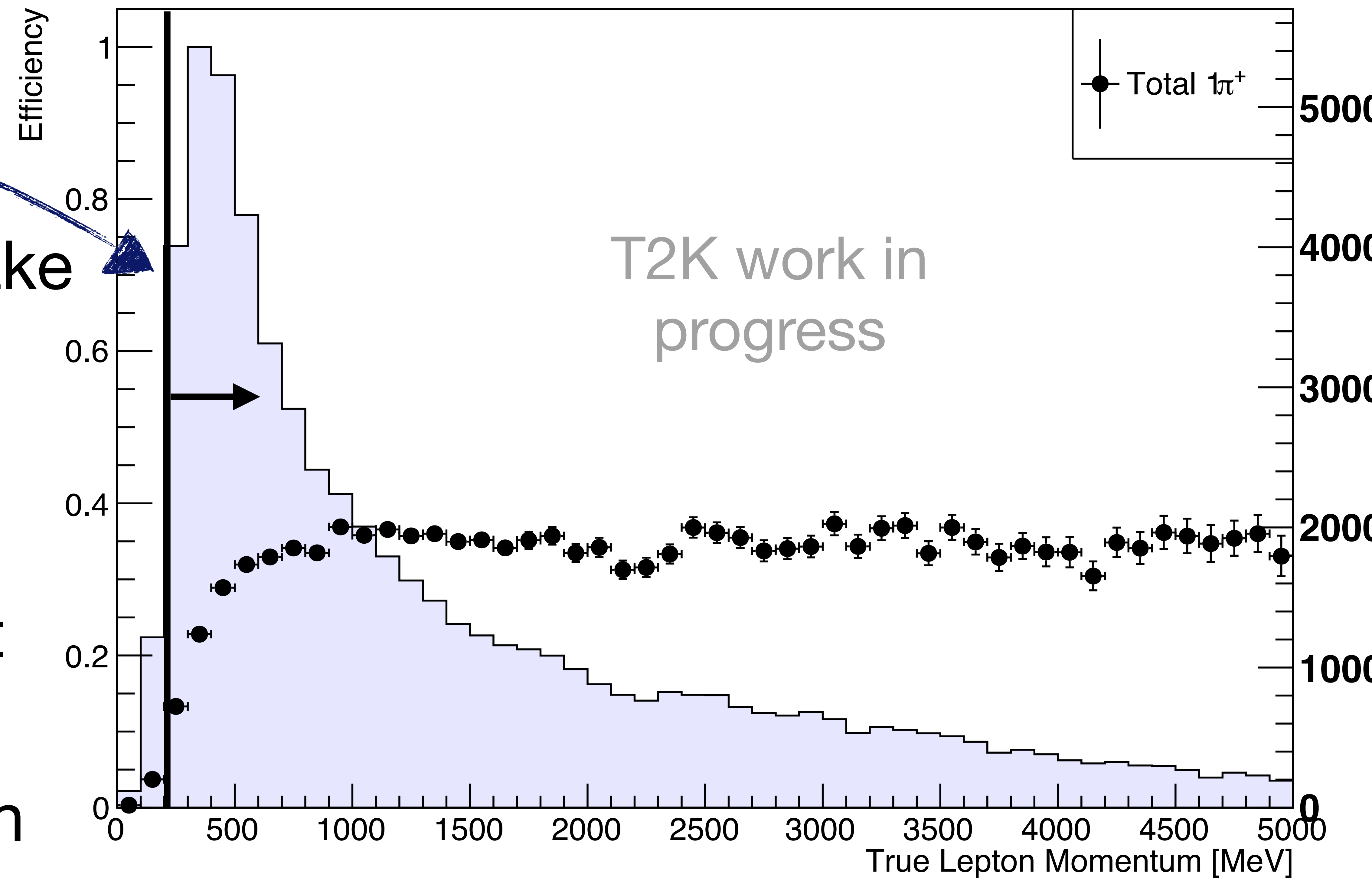


Good

Bad

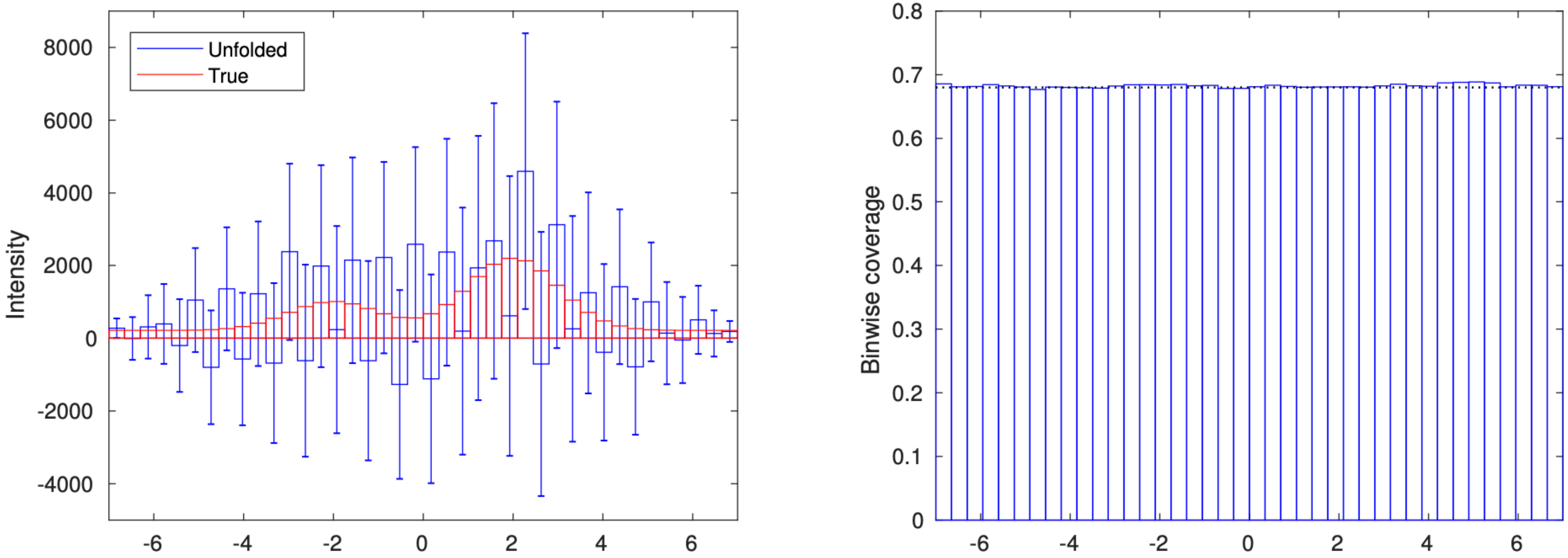
Dealing with efficiencies

- First option - accept that we can't measure what we can't measure 🤷
- Place phase space constraints on our signal, redefining the definition to regions that we know we *can* measure
- eg. muon reconstruction is less reliable at low momentum → restrict to $p_\mu > 200$ MeV
- In regions where this isn't possible (ie we want to make measurements), we simply have to place bin edges carefully
- Ensure efficiency within bins is as flat as possible
  - In theory we do this by making bins narrower, but this leads to high statistical error
  - Particular issue in high dimensional analyses, with binning in multiple variables



Fine-binned efficiency correction

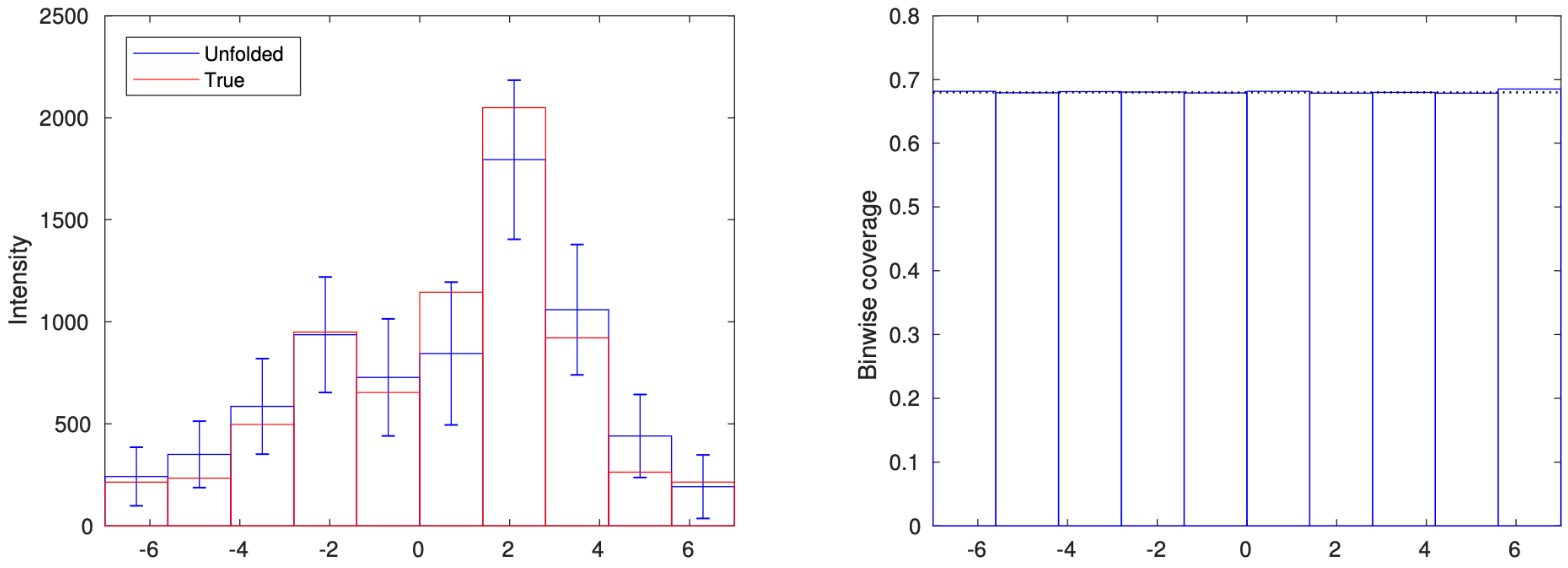
- Second option - unfold using fine binning scheme for efficiency corrections
- Ensures that ϵ_i is as flat as possible within bins, reducing model dependence and yielding correct error coverage
- But statistical errors are very high as a result…



Mikael Kuusela, Introduction to Unfolding: A Statistician's Perspective, PHYSTAT-nu 2019, CERN

Fine-binned efficiency correction

- …so then we integrate into wider bins, keeping track of correlations
- Maintains correct error coverage, but reduces statistical error to a reasonable level
- Let's see how we can apply this to a high-dimensional analysis

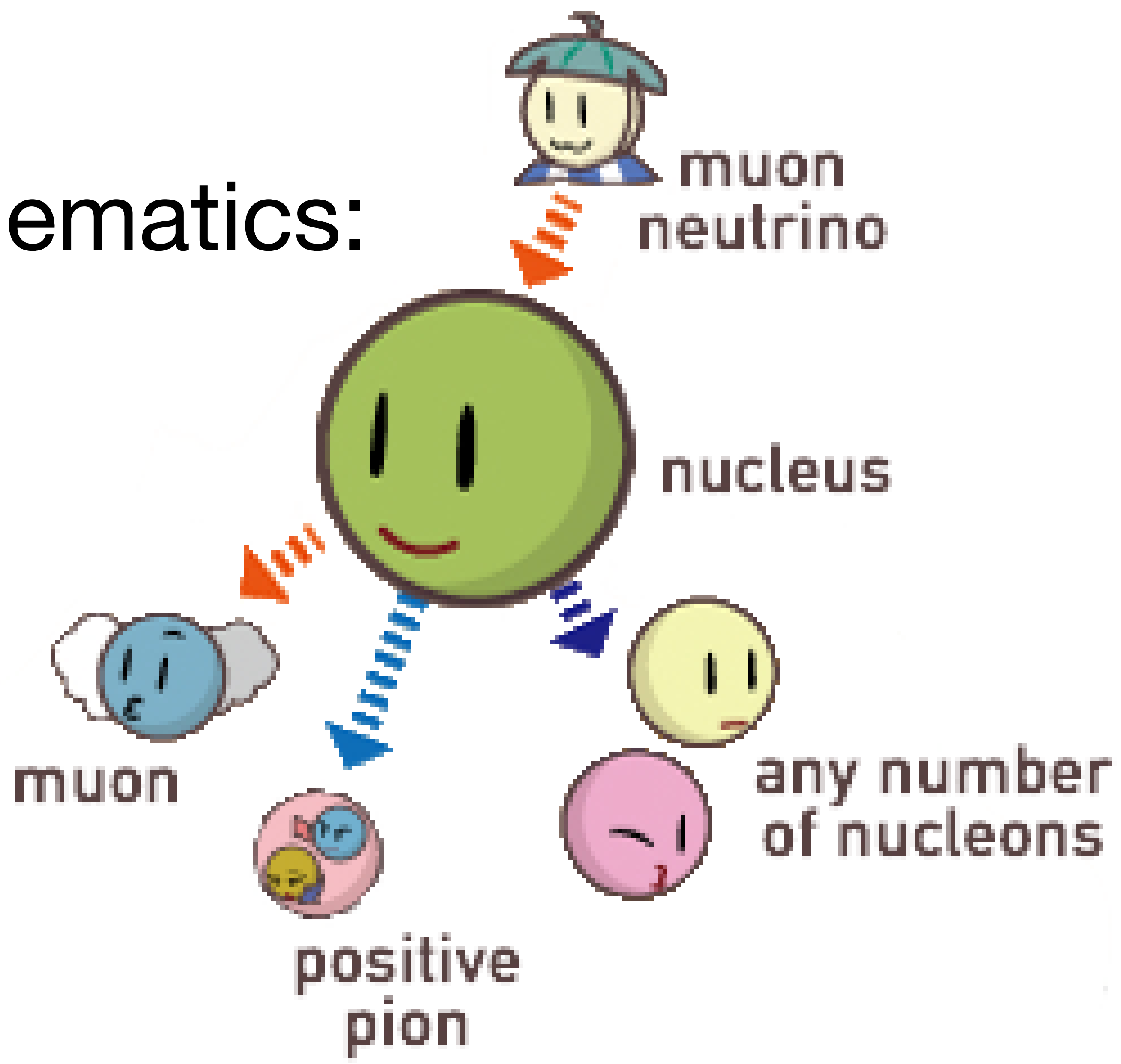


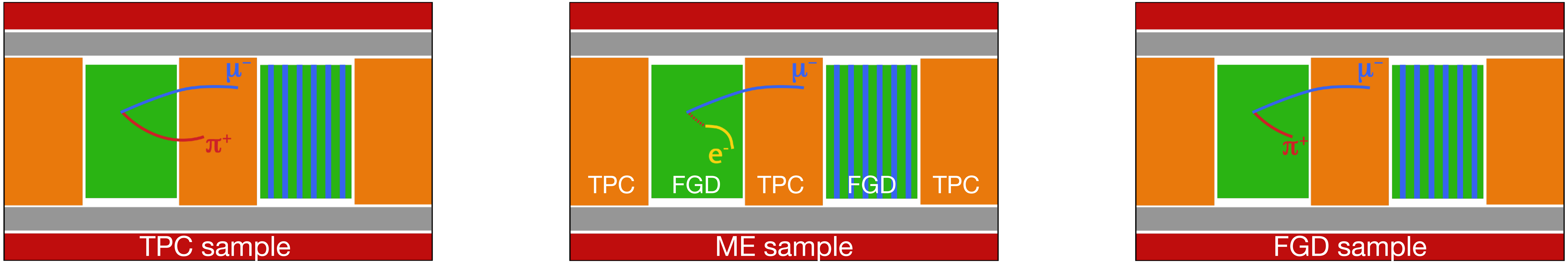
Mikael Kuusela, Introduction to Unfolding: A Statistician's Perspective, PHYSTAT-nu 2019, CERN

$\nu_\mu CC1\pi^+$ cross section

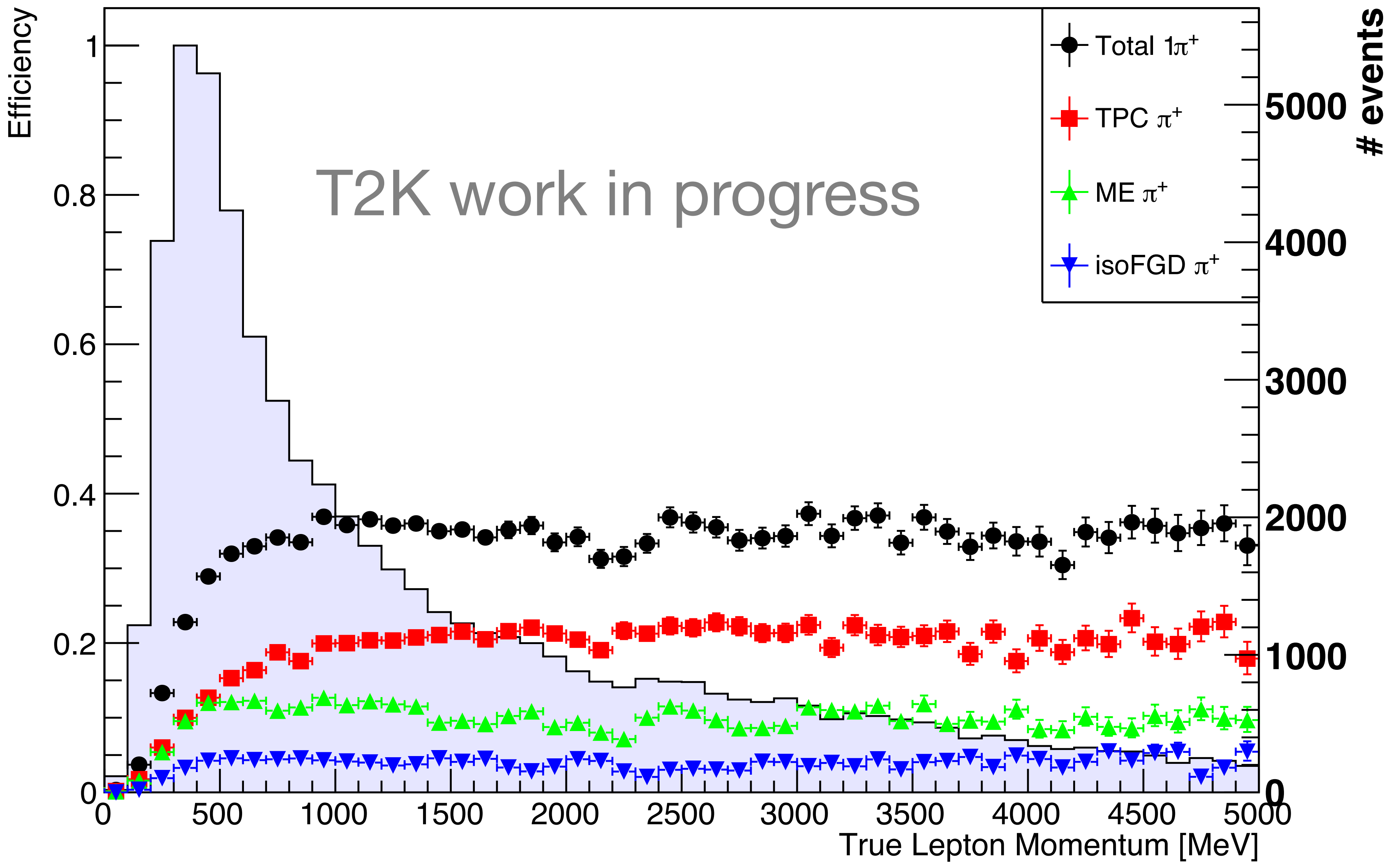
Topology diagram made using illustrations from HiggsTan (Yuki Akimoto)

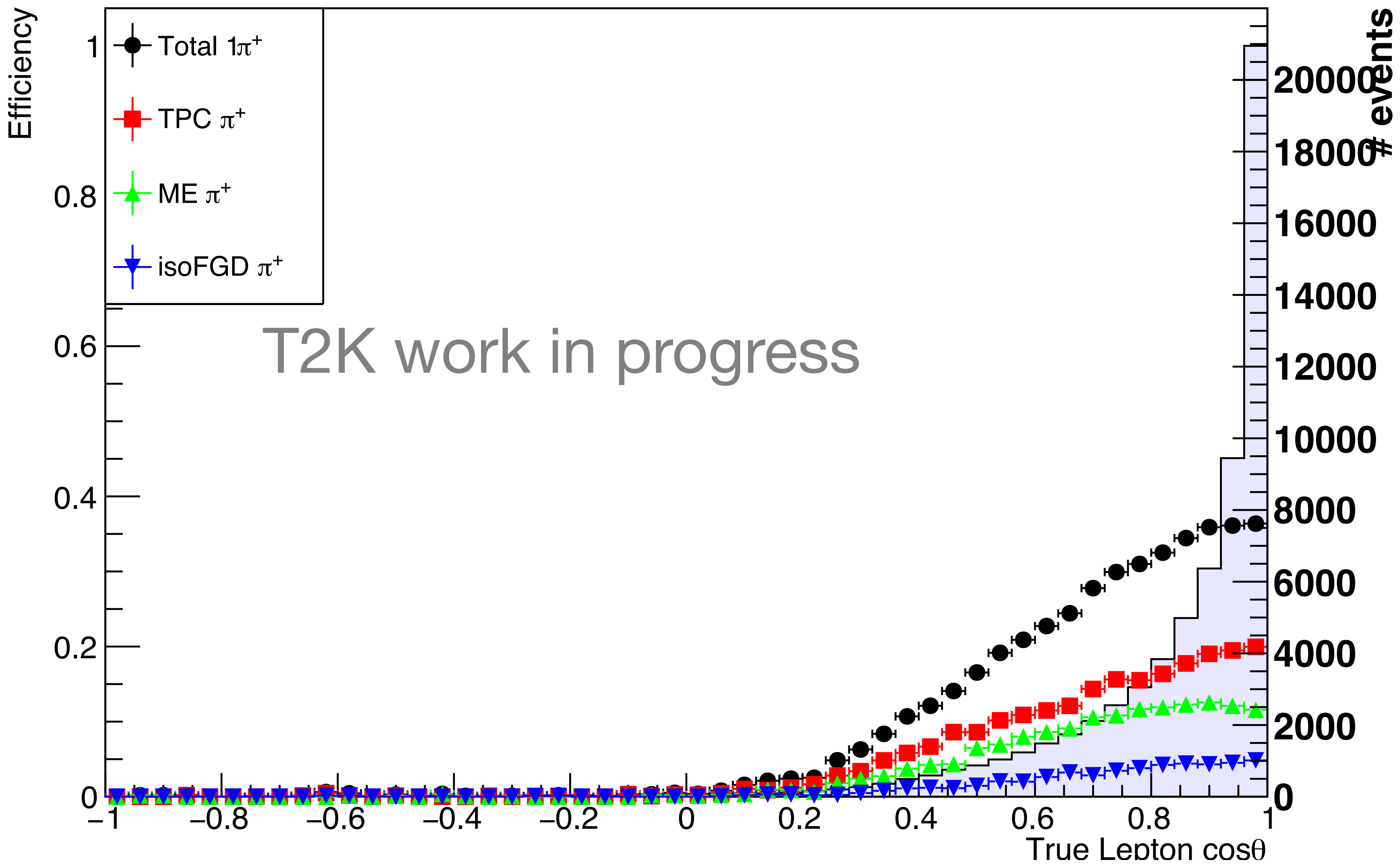
- Measuring $\nu_\mu CC1\pi^+$ cross section on both hydrocarbon and water (ratio 3:1 in ND280)
- Using three different sources of pions, covering different regions of kinematics:
  - TPC pions (higher energy, forward going)
  - Michel electron tagged pions (low energy, all angles)
  - FGD pions (mid-energy, high-angle)
- Aiming to make (statistically limited) measurement in 4D $(p_\mu, \cos\theta_\mu, p_\pi, \cos\theta_\pi)$
- Then project result down to pion kinematics



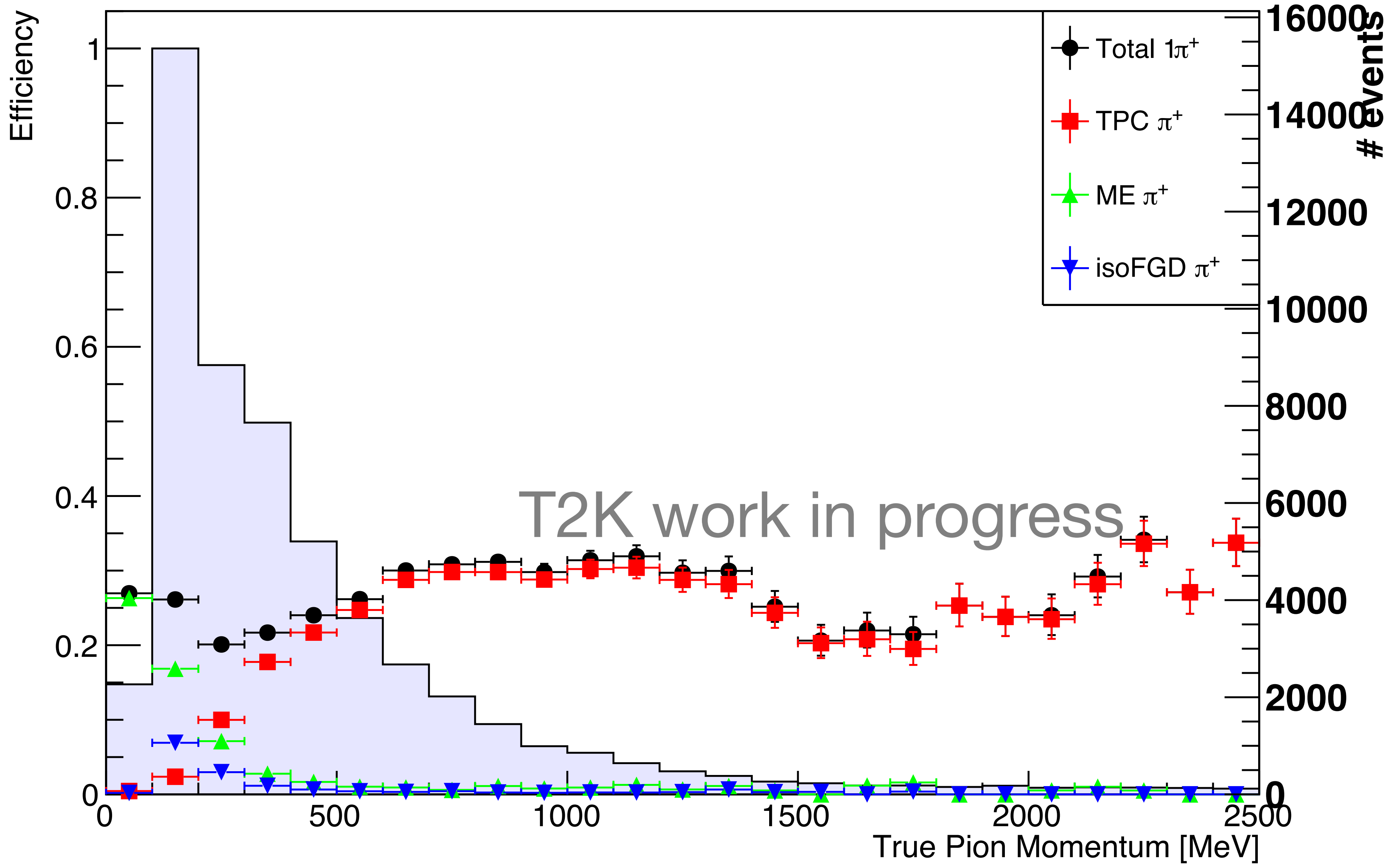


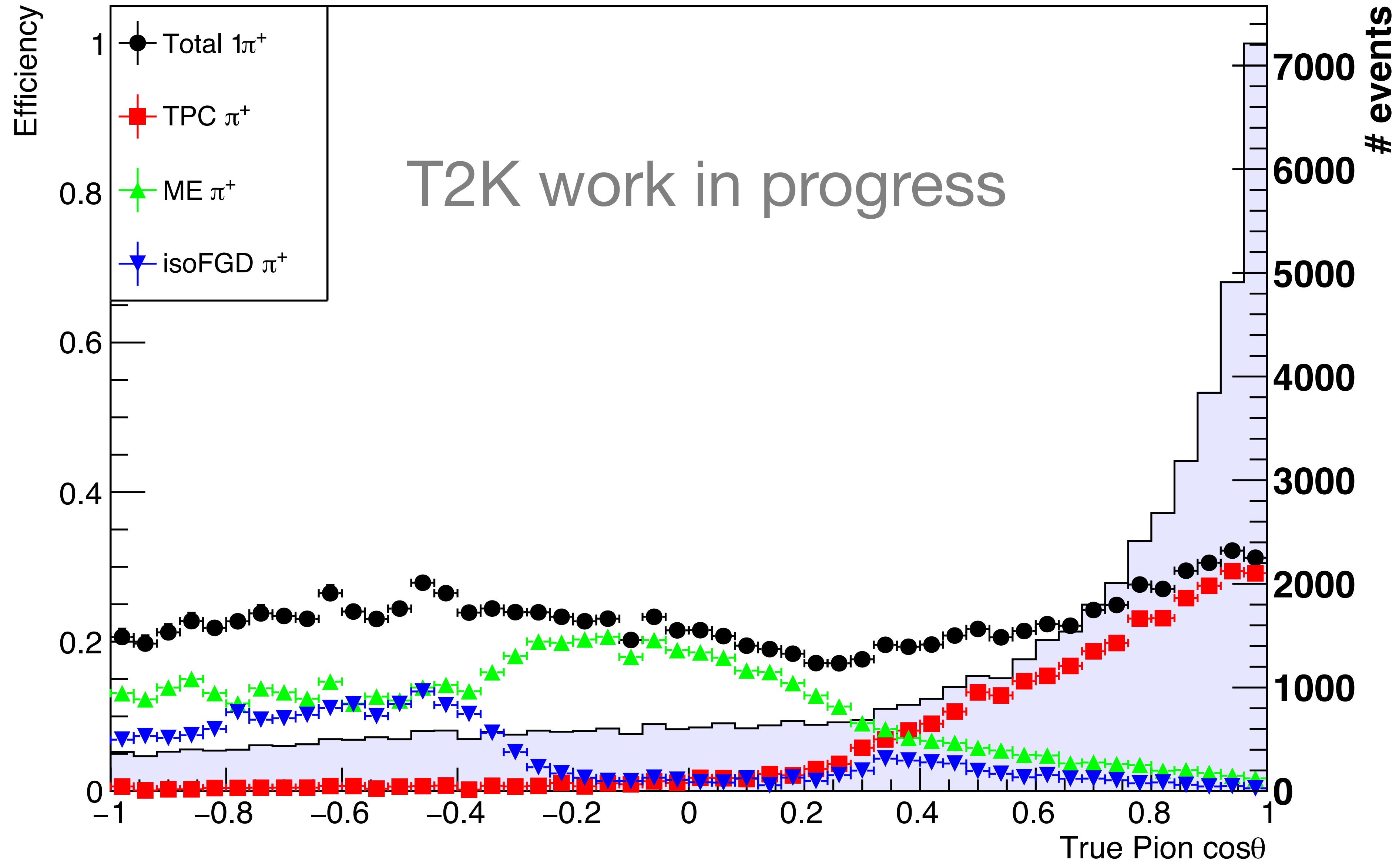
Phase space constraints



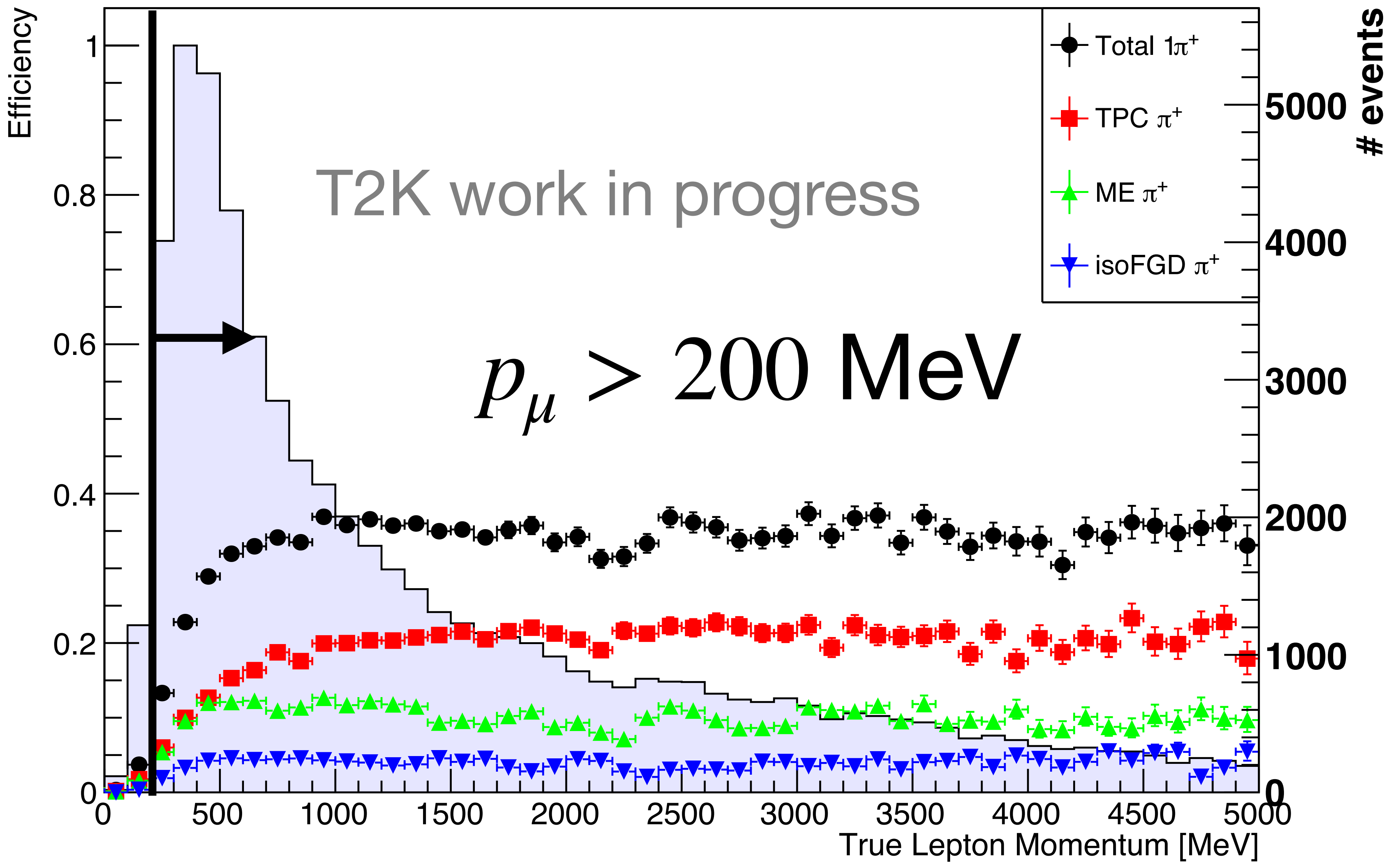


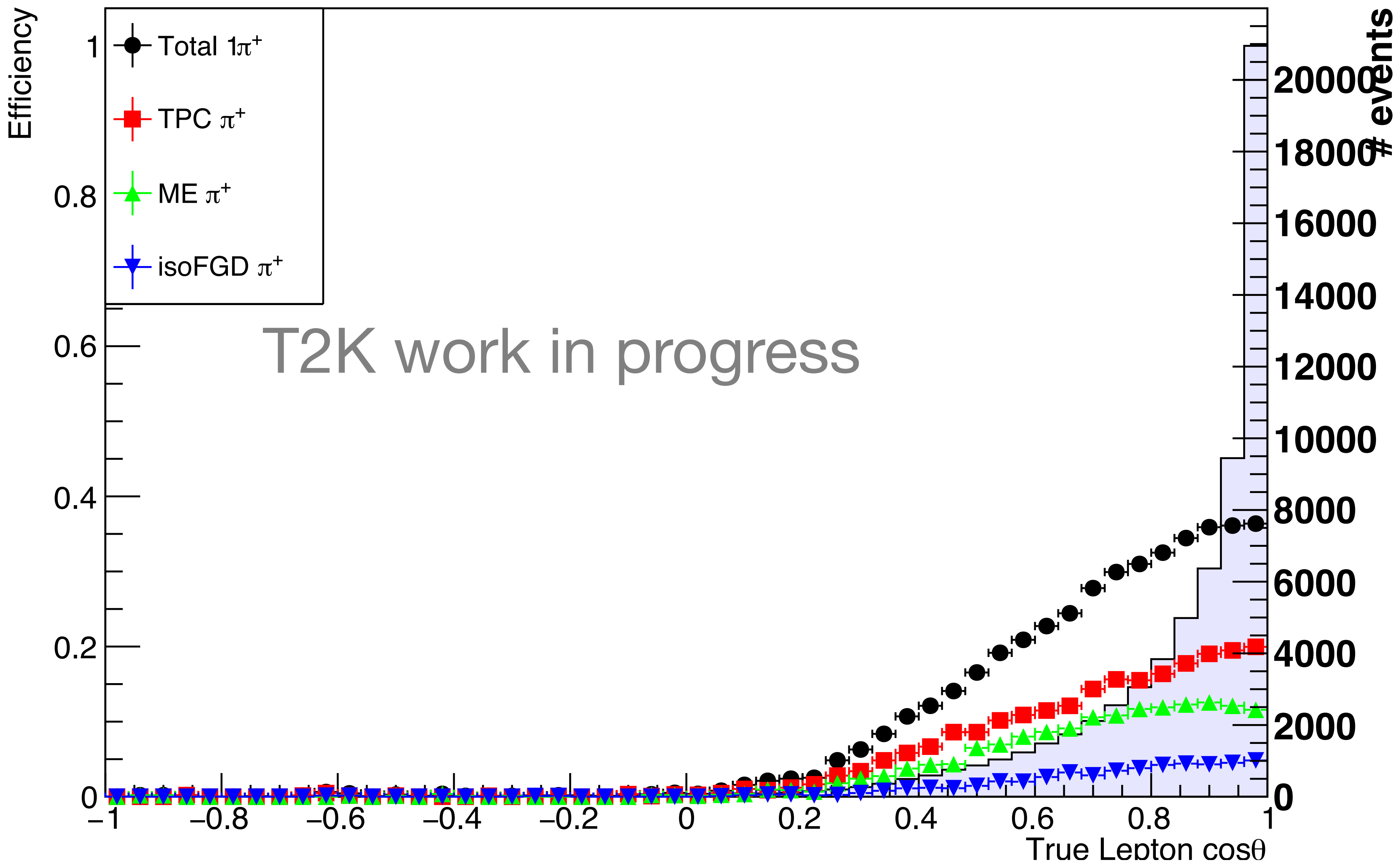
Event rates are for MC POT



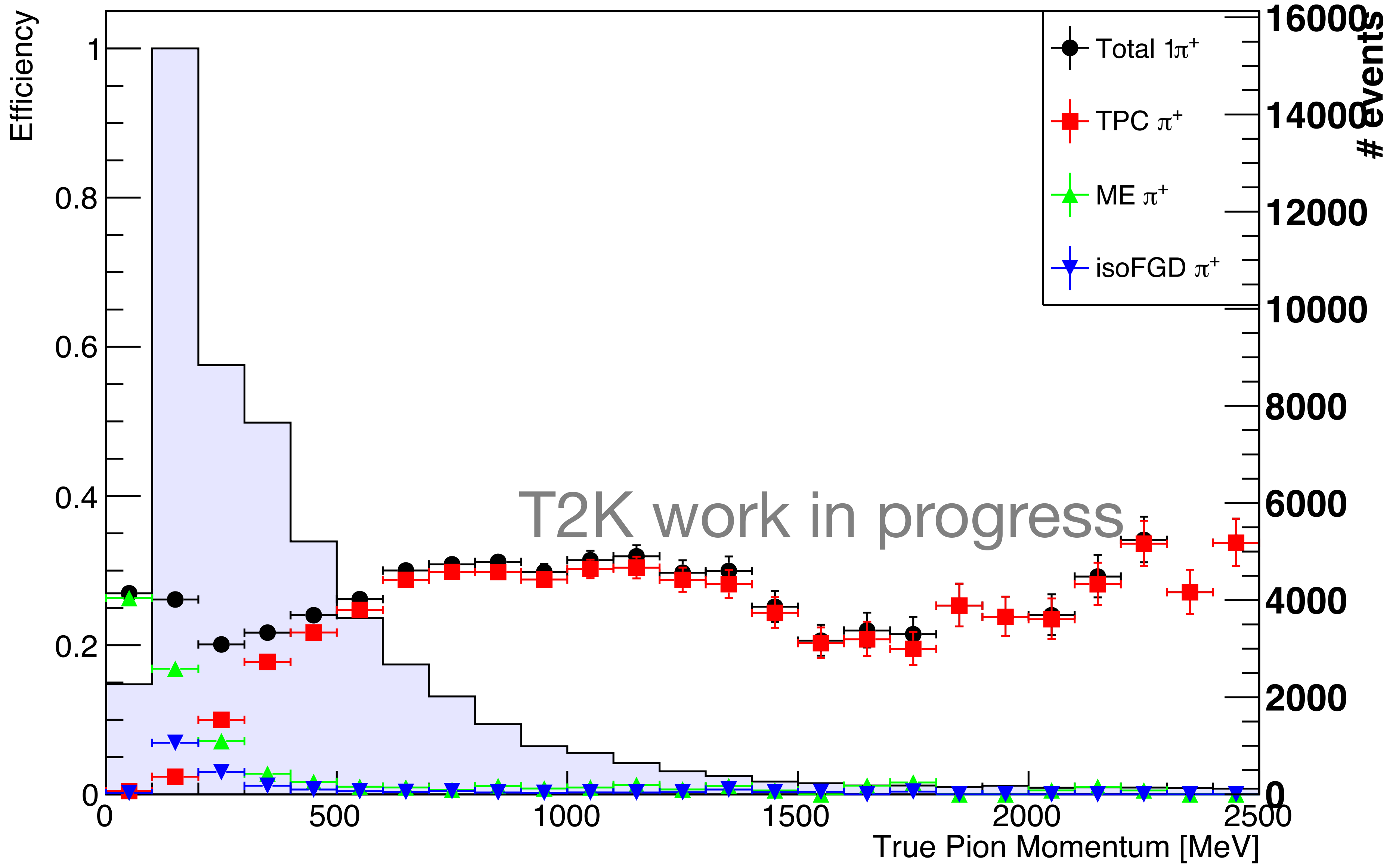


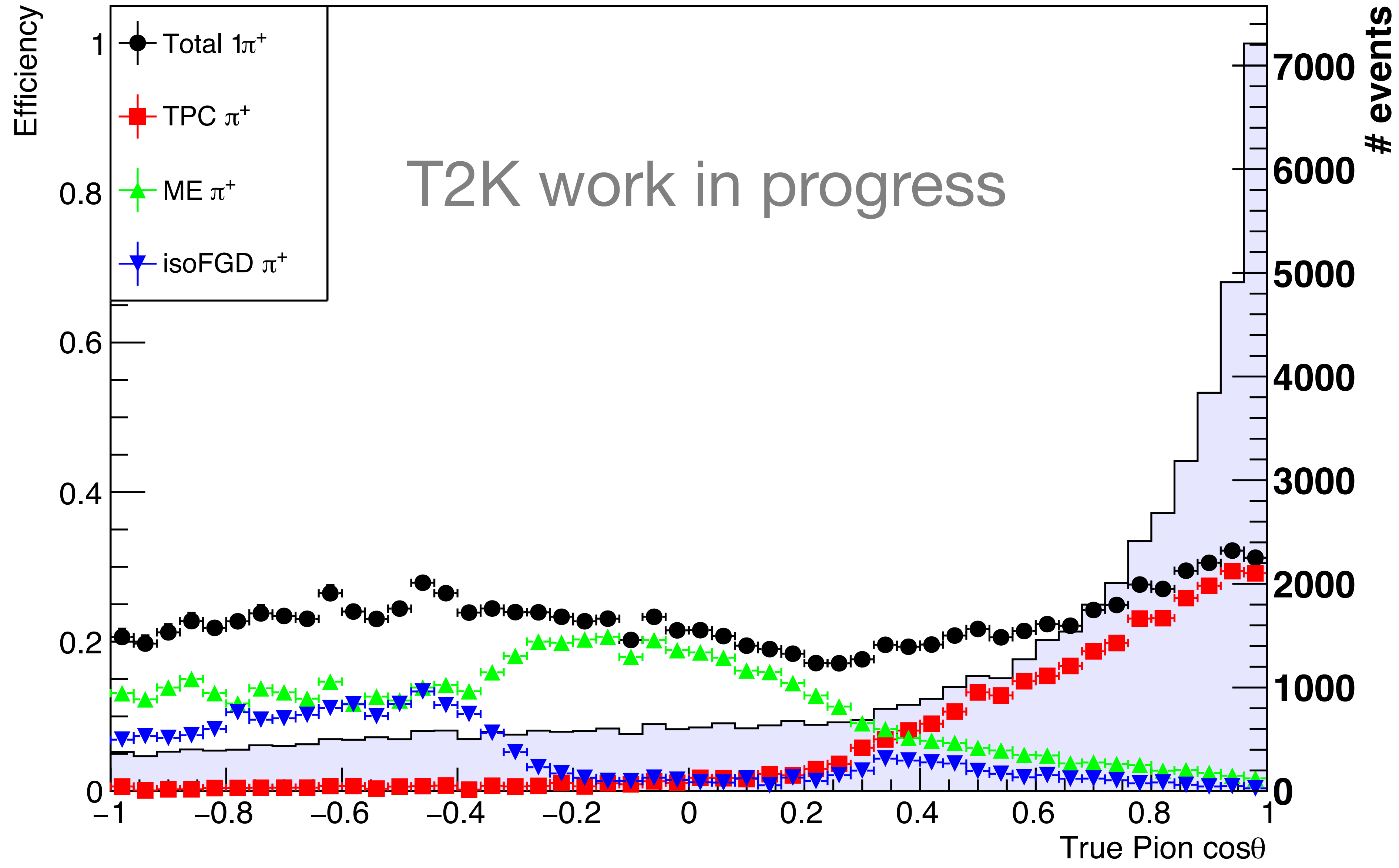
Phase space constraints



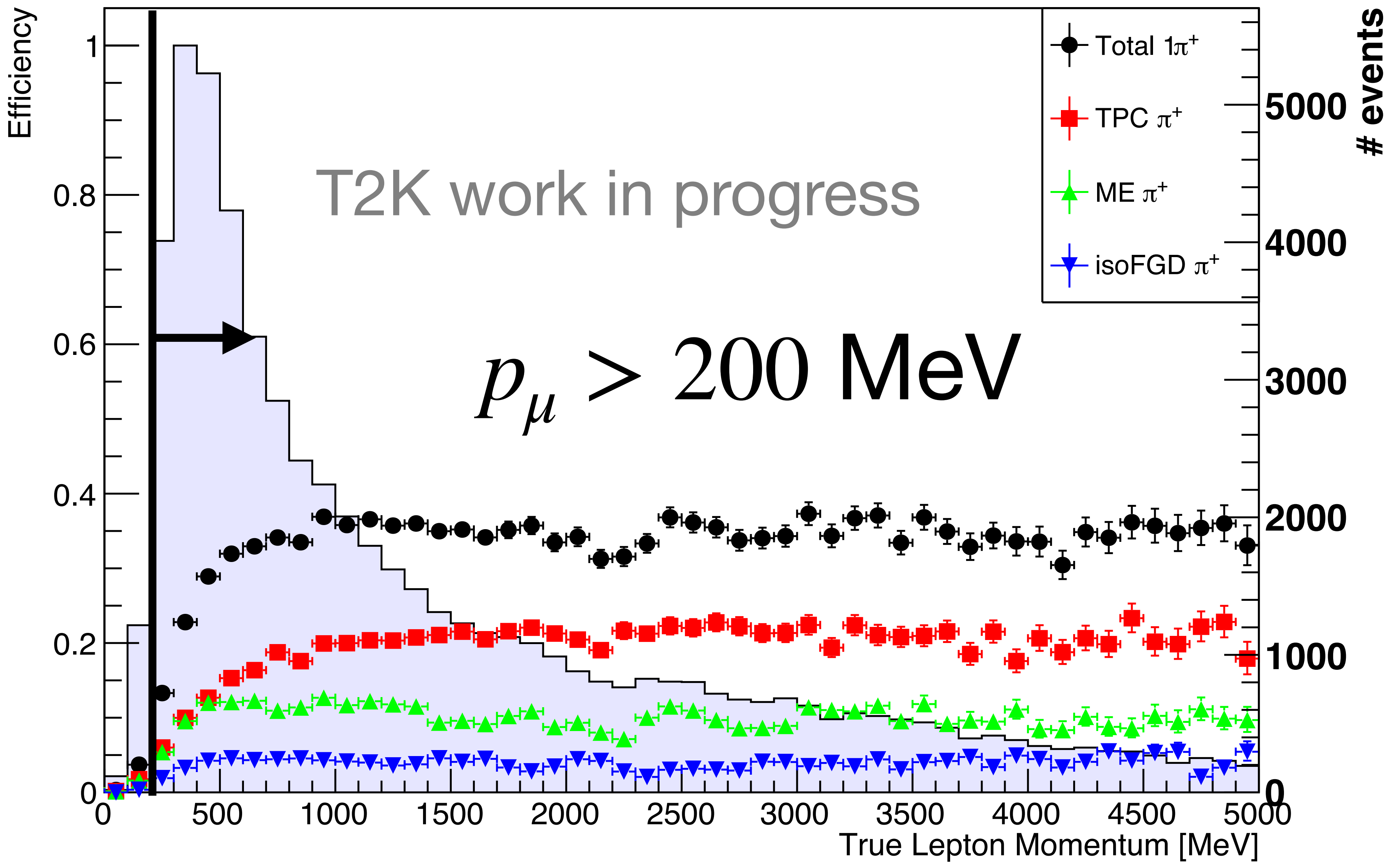


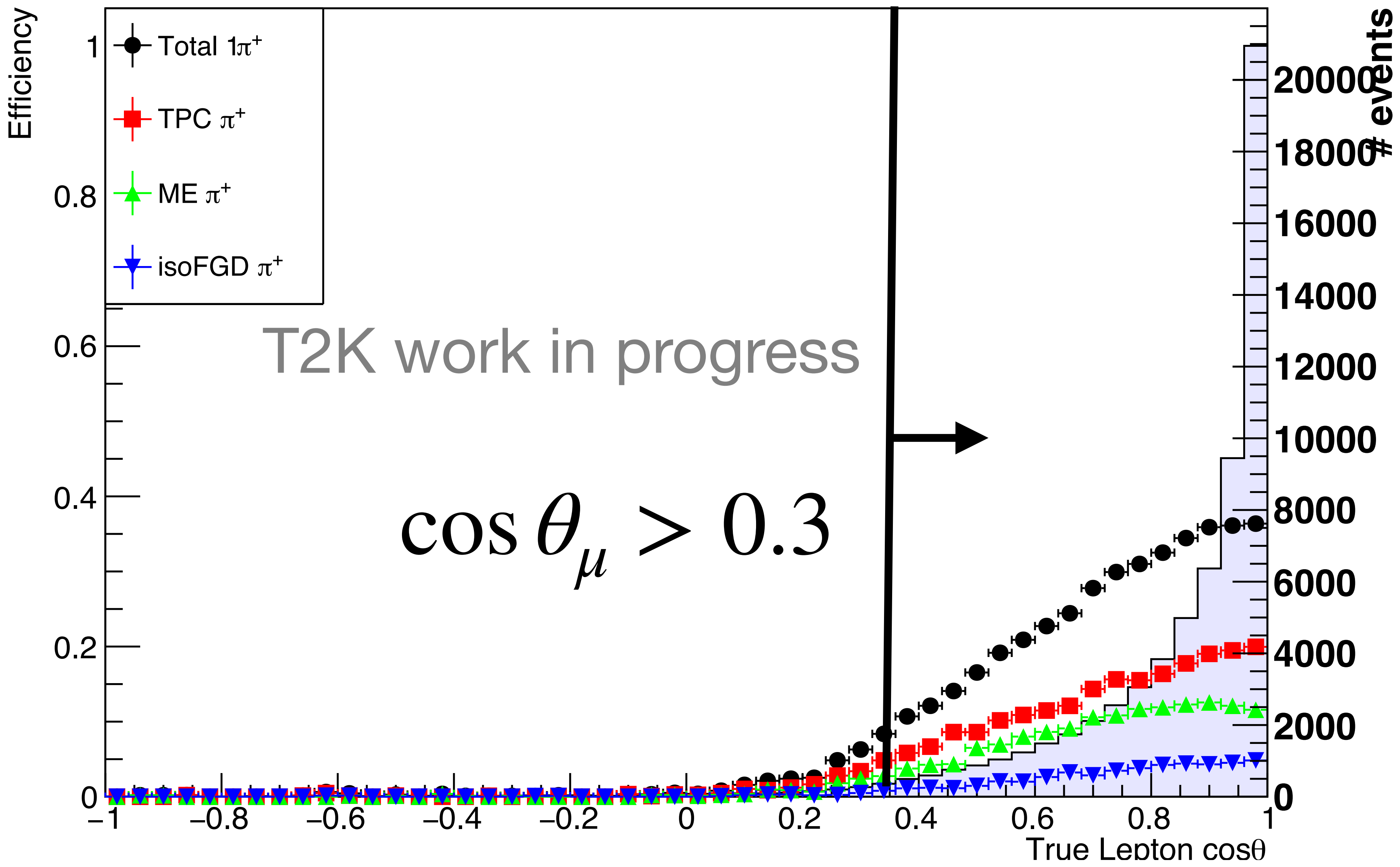
Event rates are for MC POT



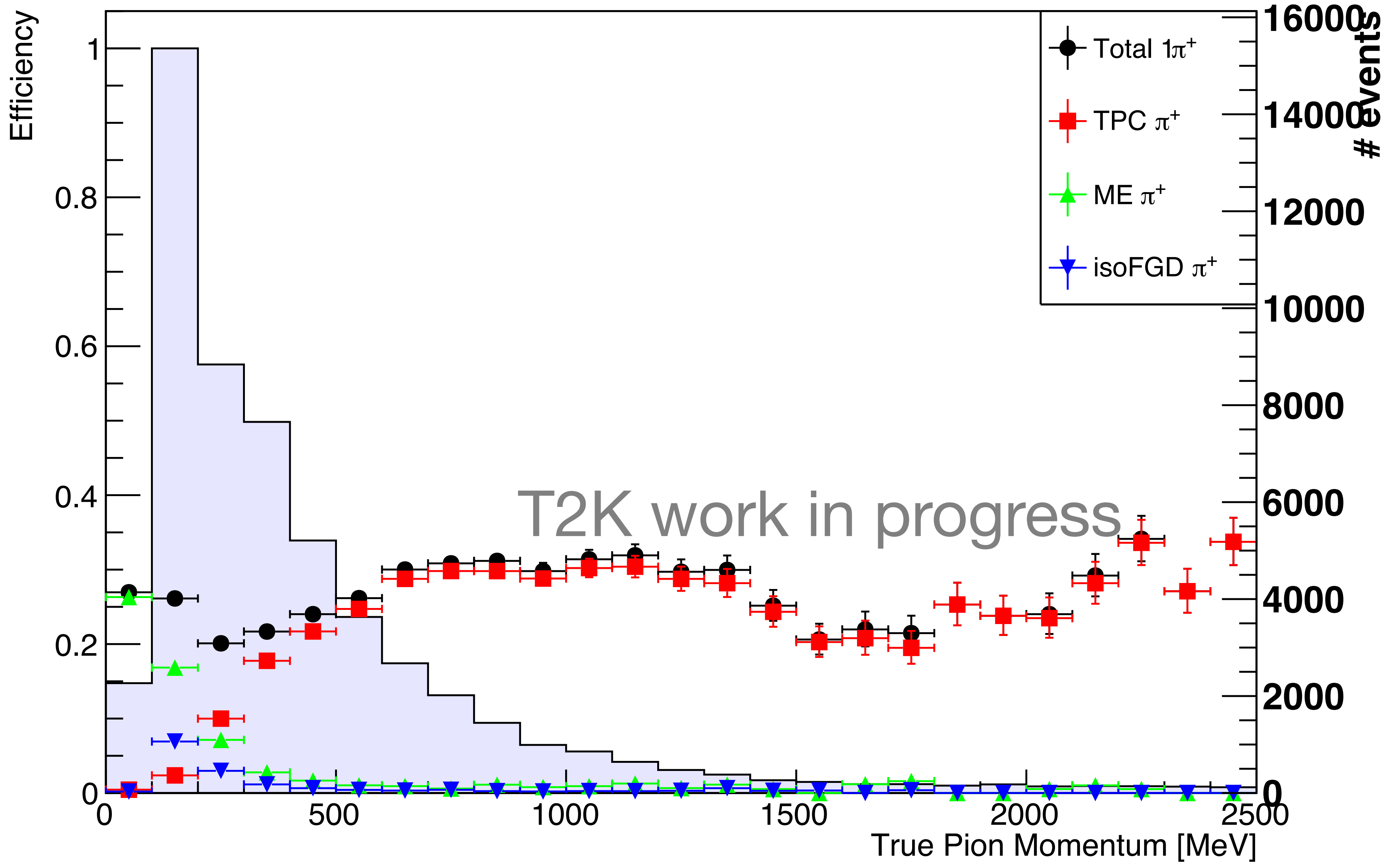


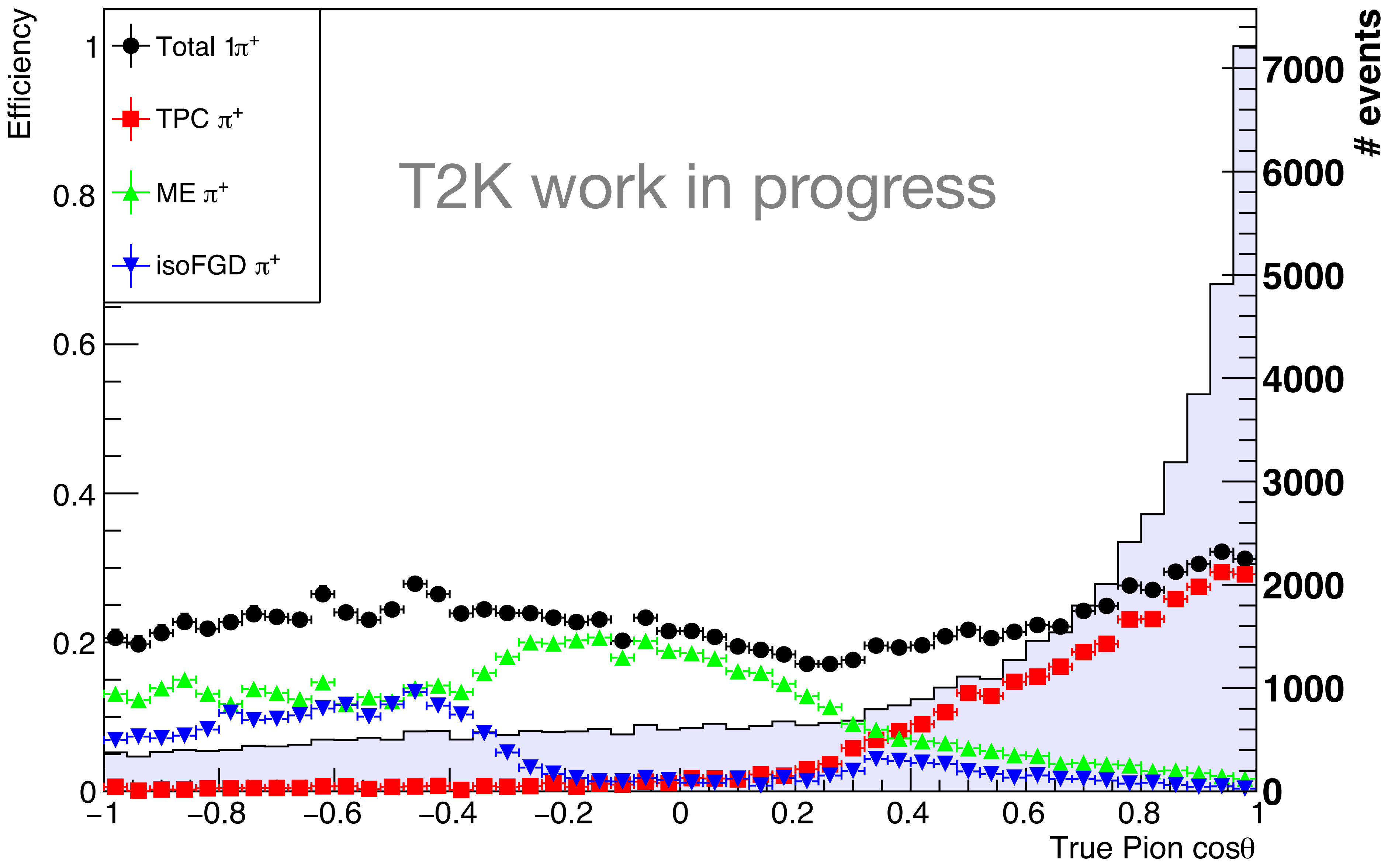
Phase space constraints



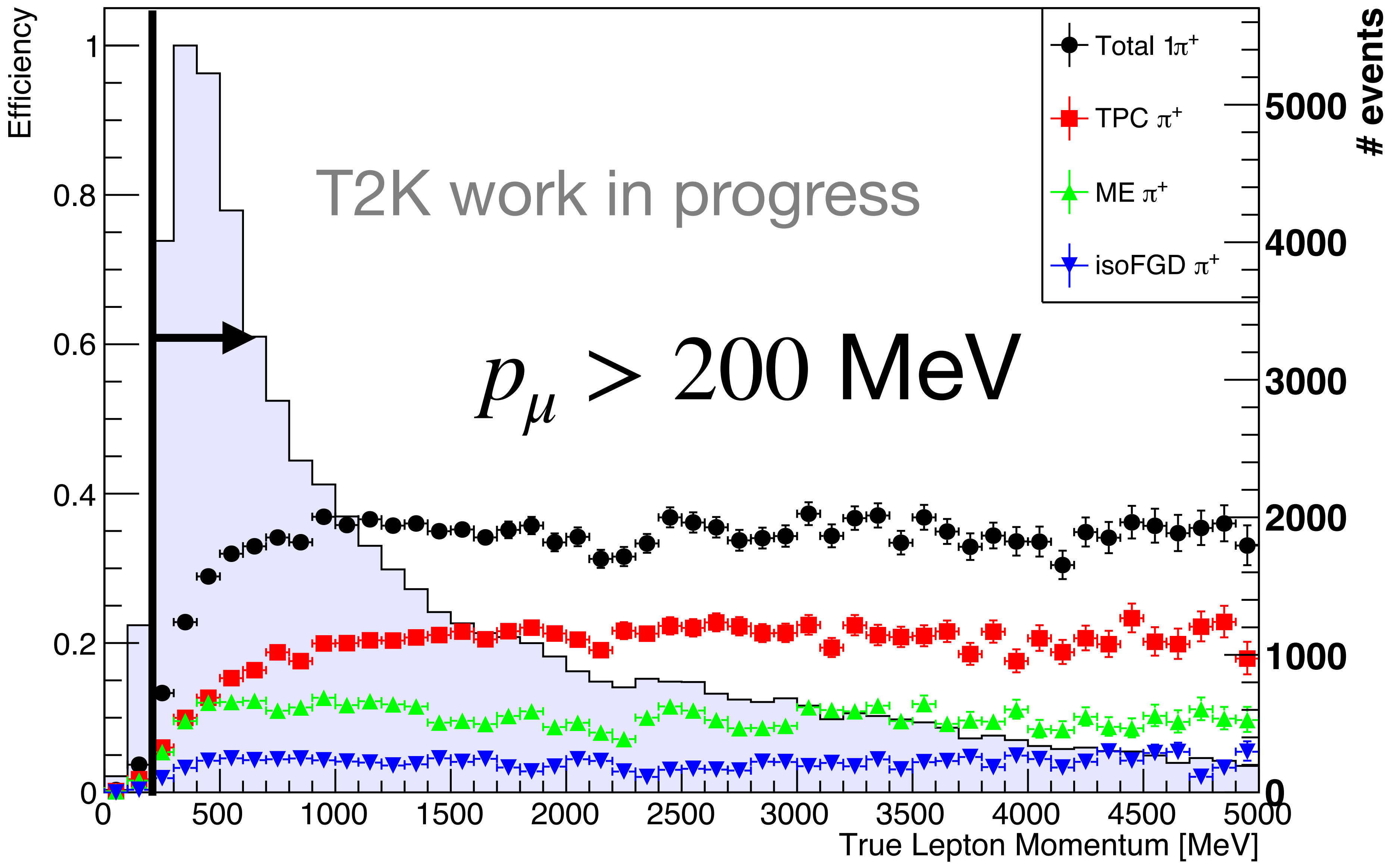


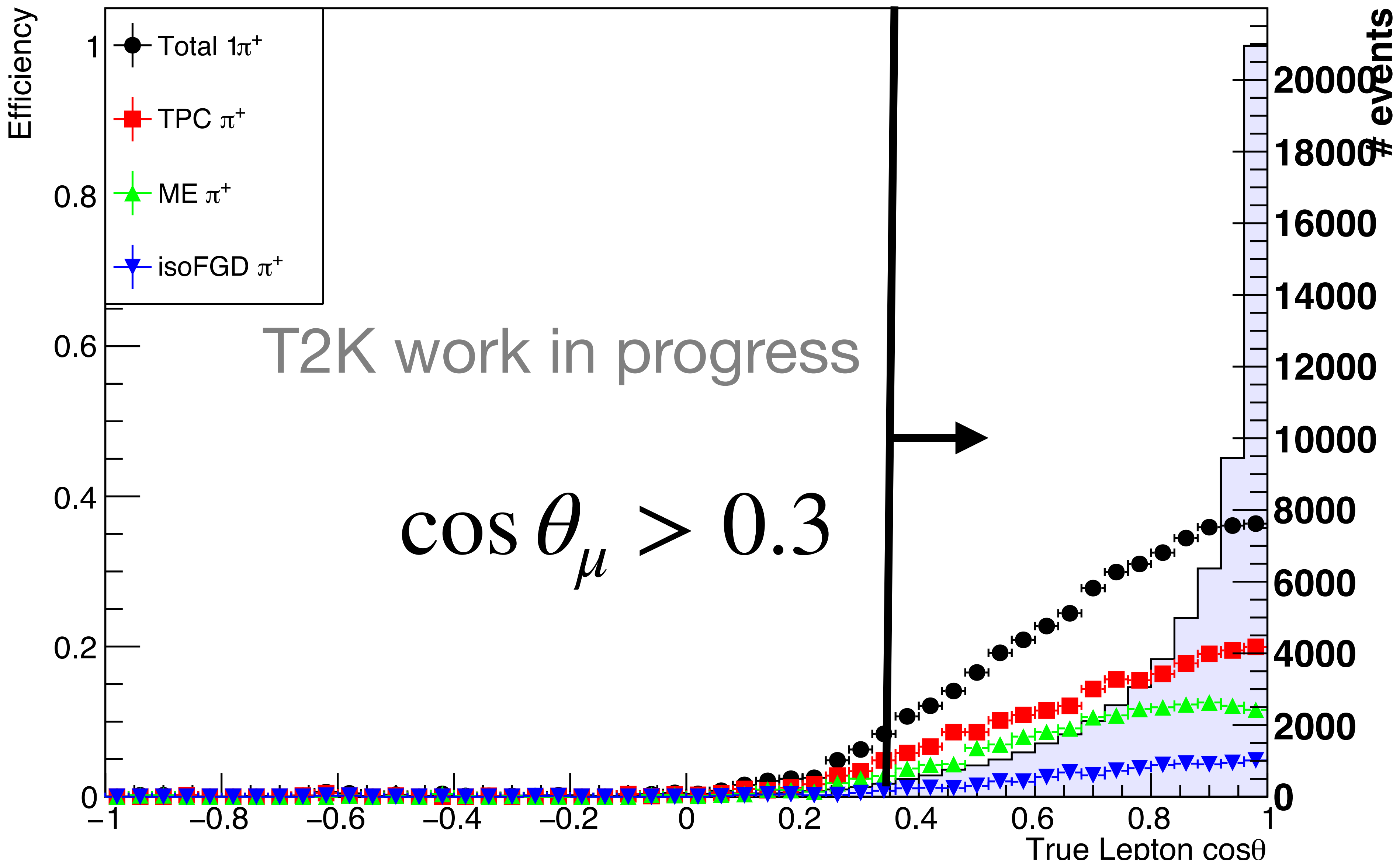
Event rates are for MC POT



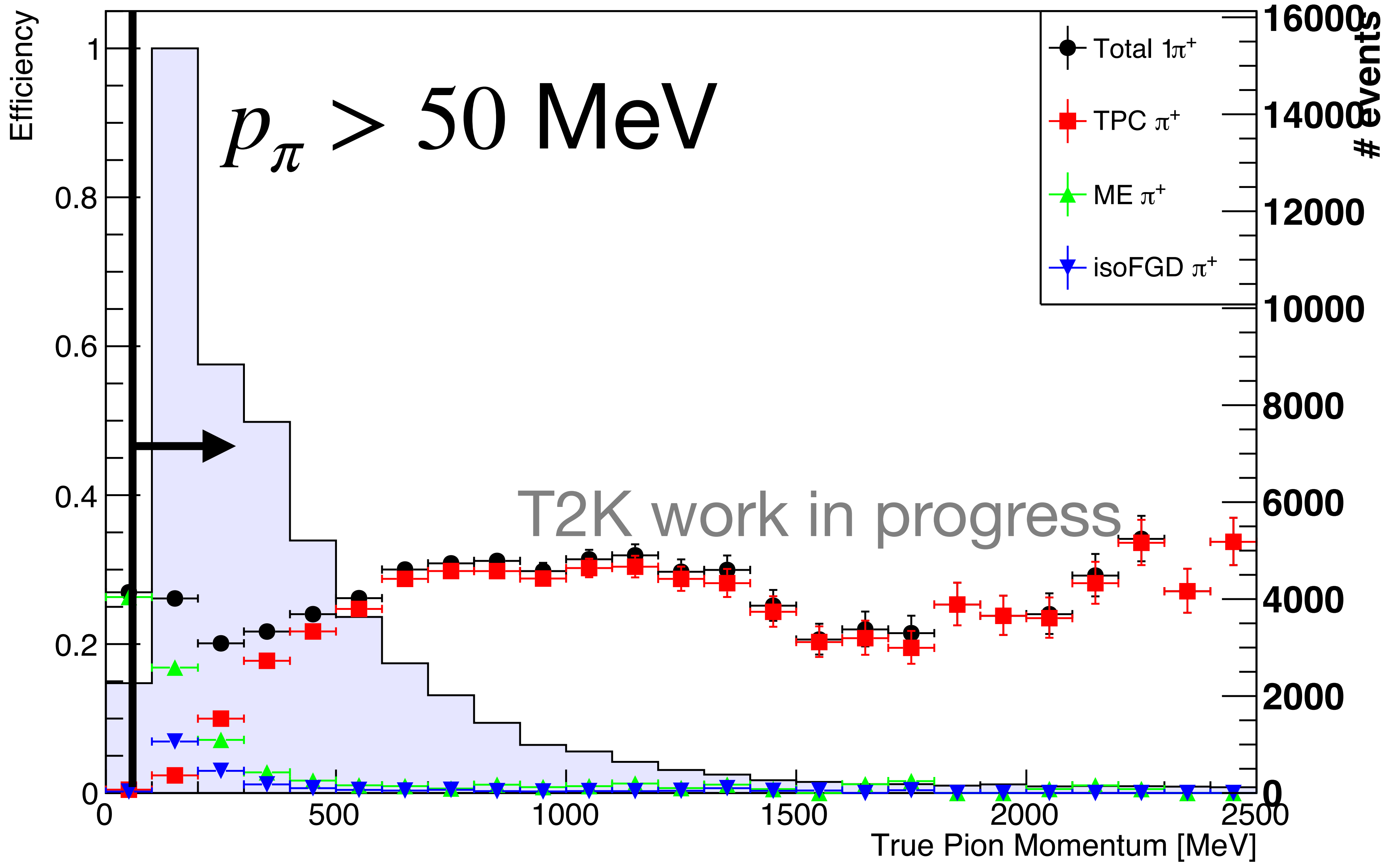


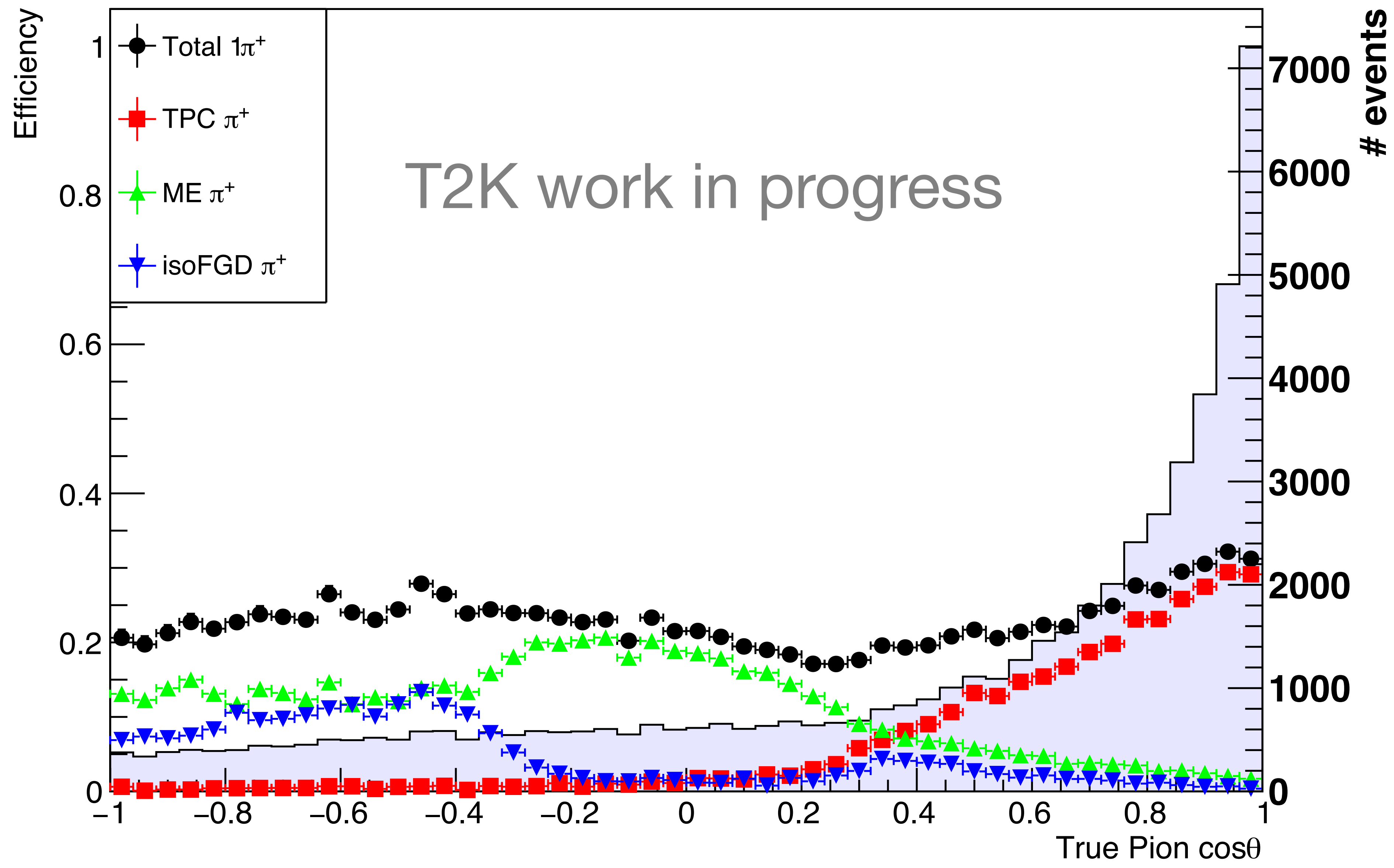
Phase space constraints



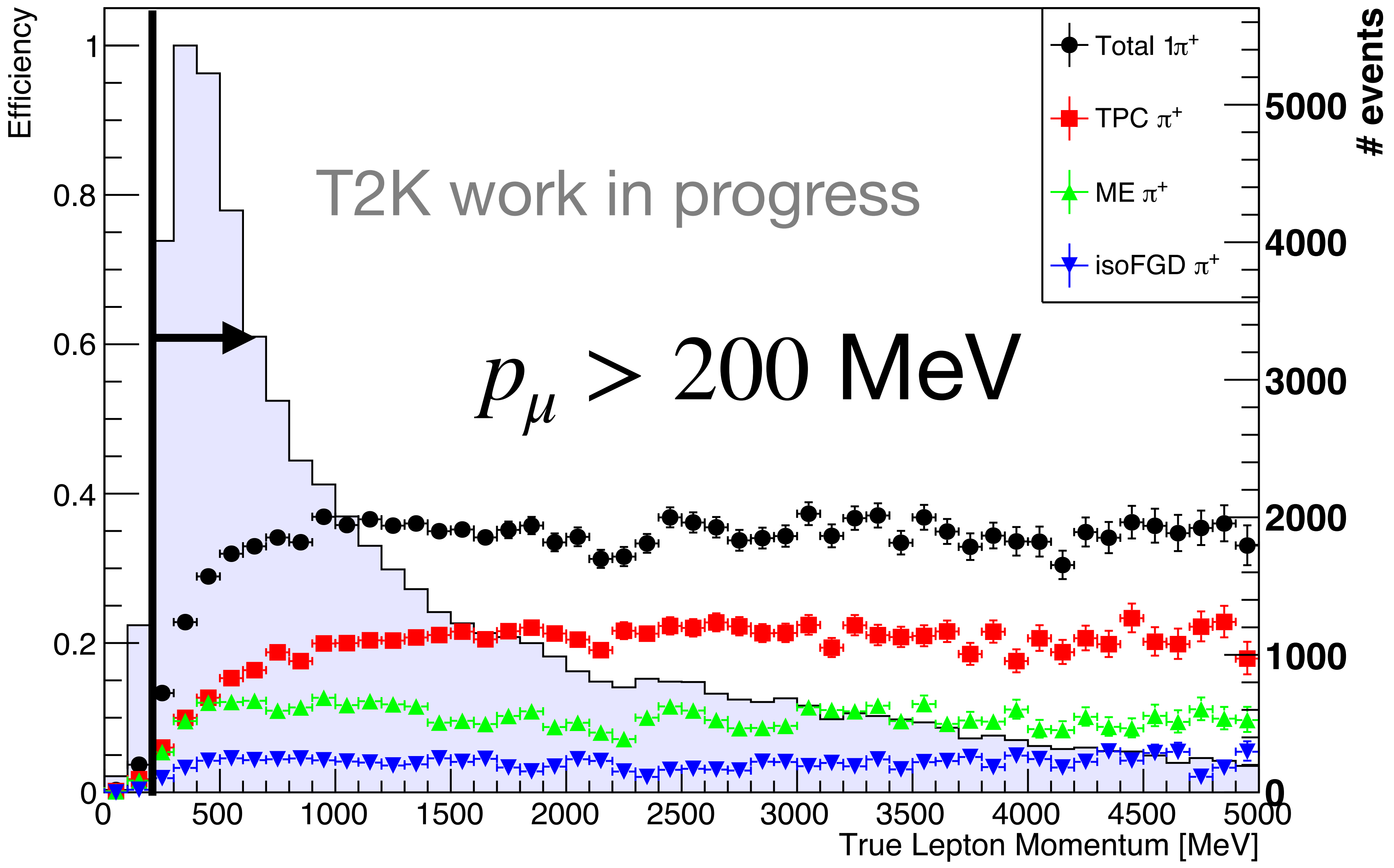


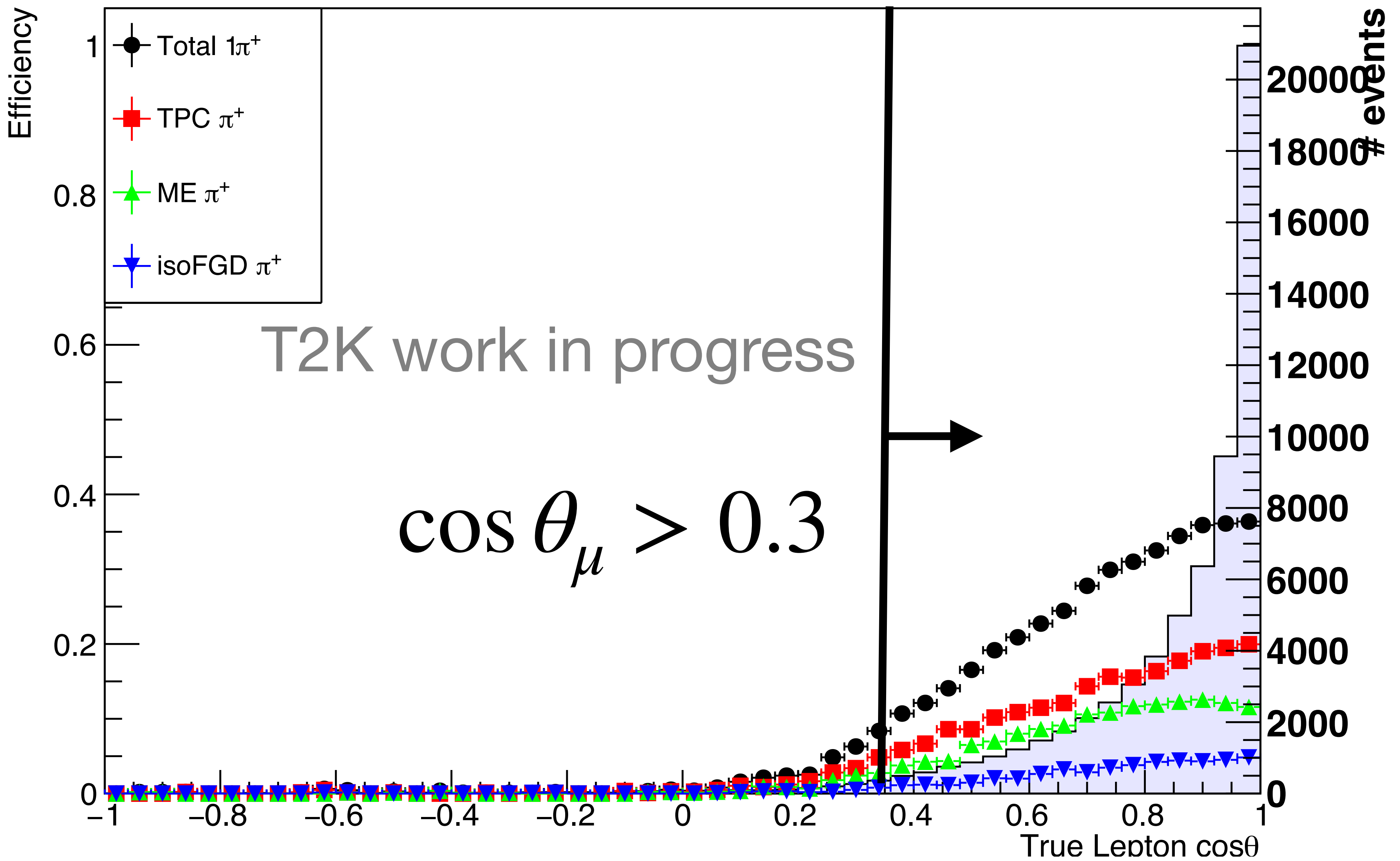
Event rates are for MC POT



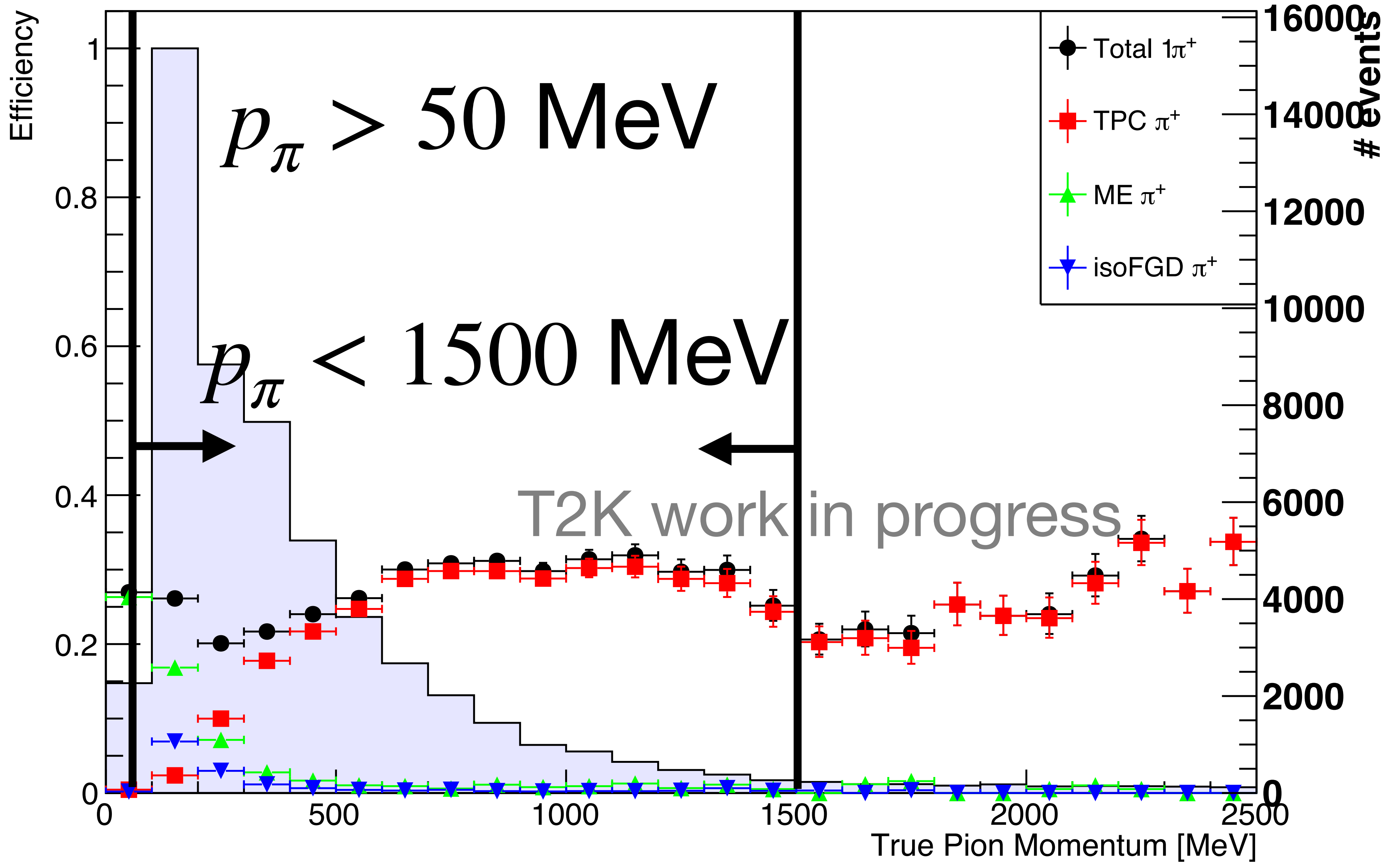


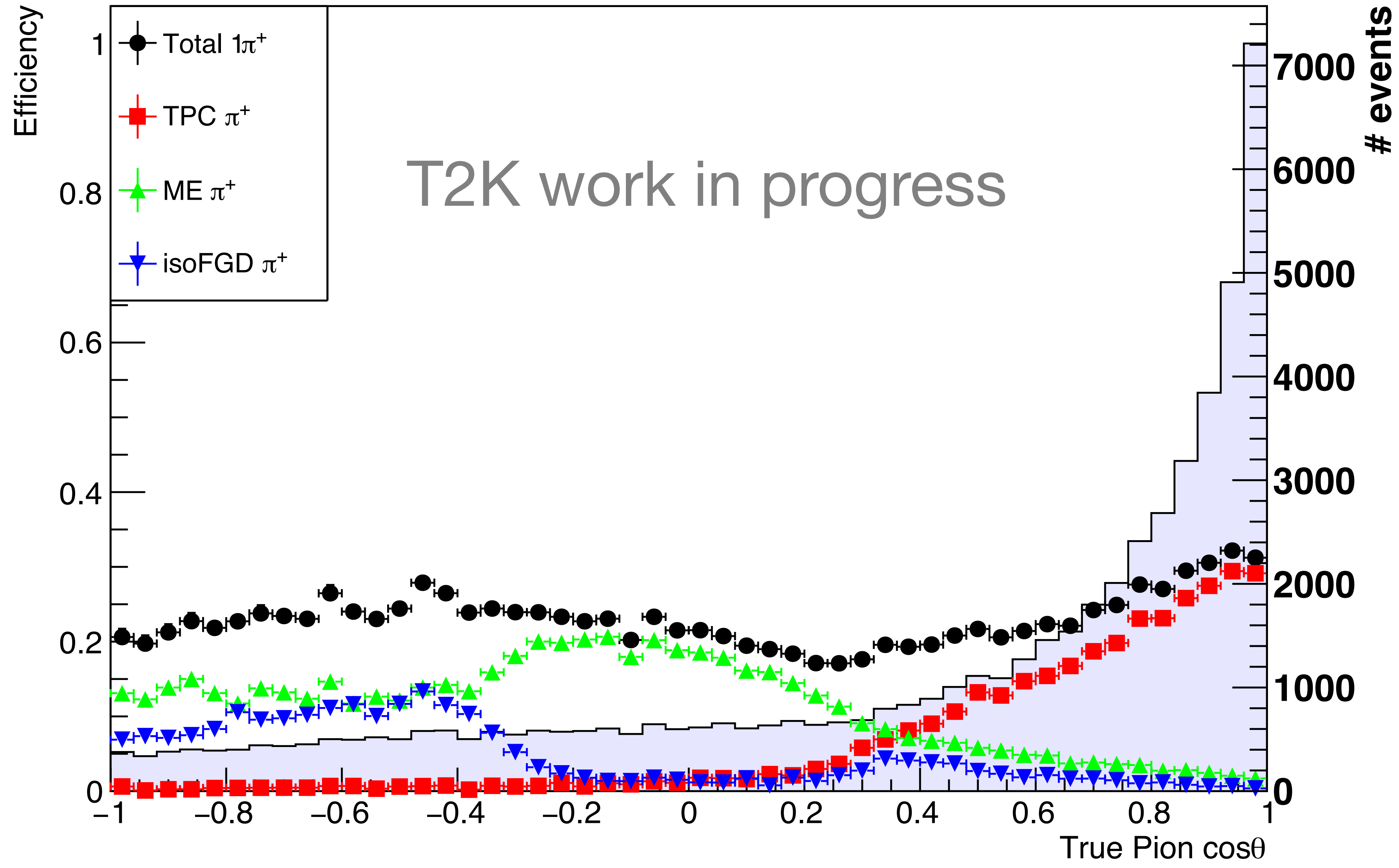
Phase space constraints





Event rates are for MC POT

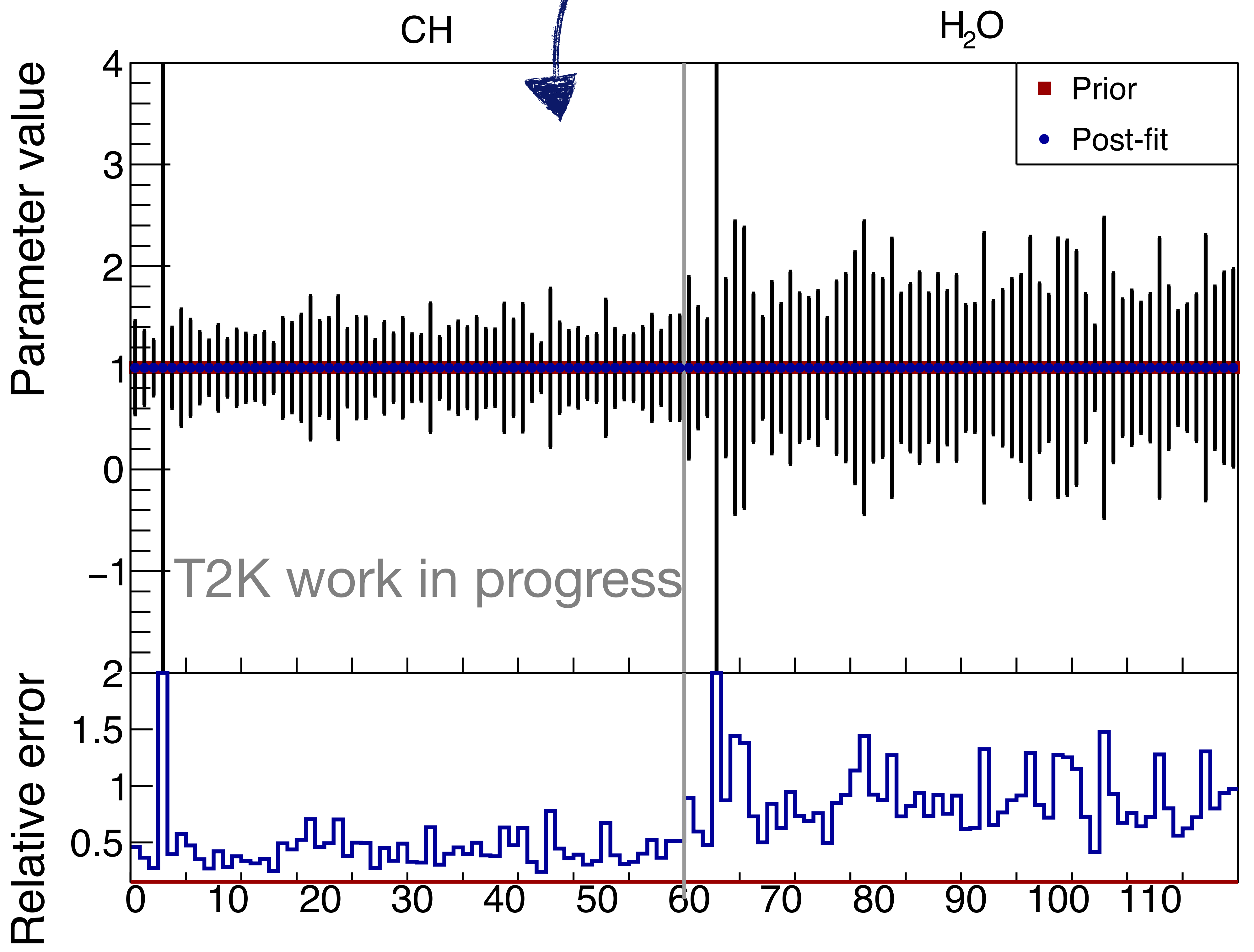




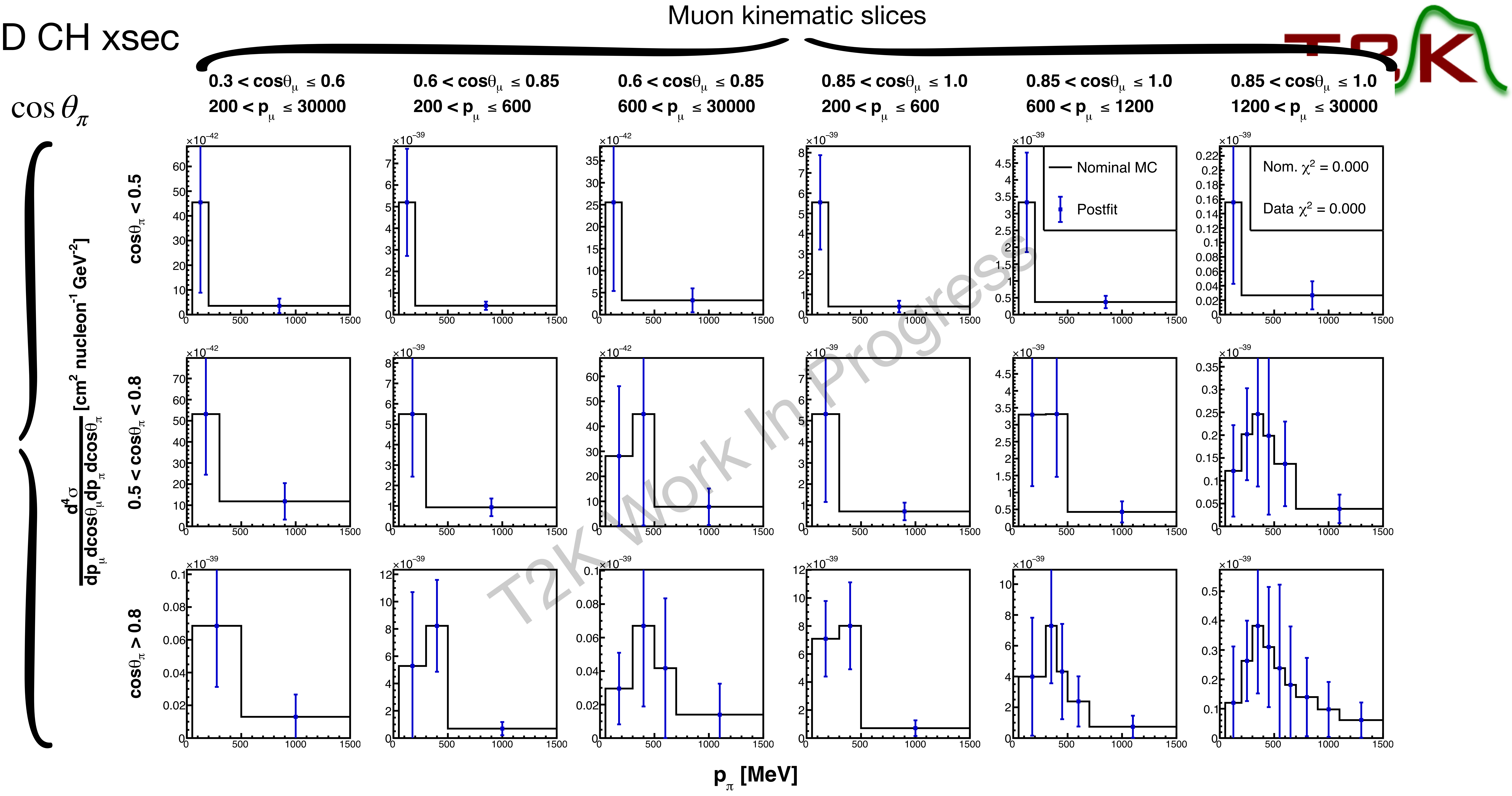
Fine binning scheme in 4D

- Not an easy thing to construct - each 1D distribution implicitly integrates over all other variables, which is what we want to avoid
- Start with 2D efficiency in muon kinematics, place coarse bin edges:
  - $\cos\theta_\mu : 0.3 - 0.6 \rightarrow p_\mu = [200,\ 30000]$
  - $\cos\theta_\mu : 0.6 - 0.85 \rightarrow p_\mu = [200,\ 600,\ 30000]$
  - $\cos\theta_\mu : 0.85 - 1.0 \rightarrow p_\mu = [200,\ 600,\ 1200,\ 30000]$
- In each slice, check 2D efficiency in pion kinematics and place bin edges
- Results in 4 OOPS + 56 in PS bins per target (CH and H_2O) $\rightarrow$ 120 total
- Run the fit, then calculate cross sections (throwing toys from post fit values) with fine binned efficiency corrections and see what we get…

Template parameter values for an Asimov fit

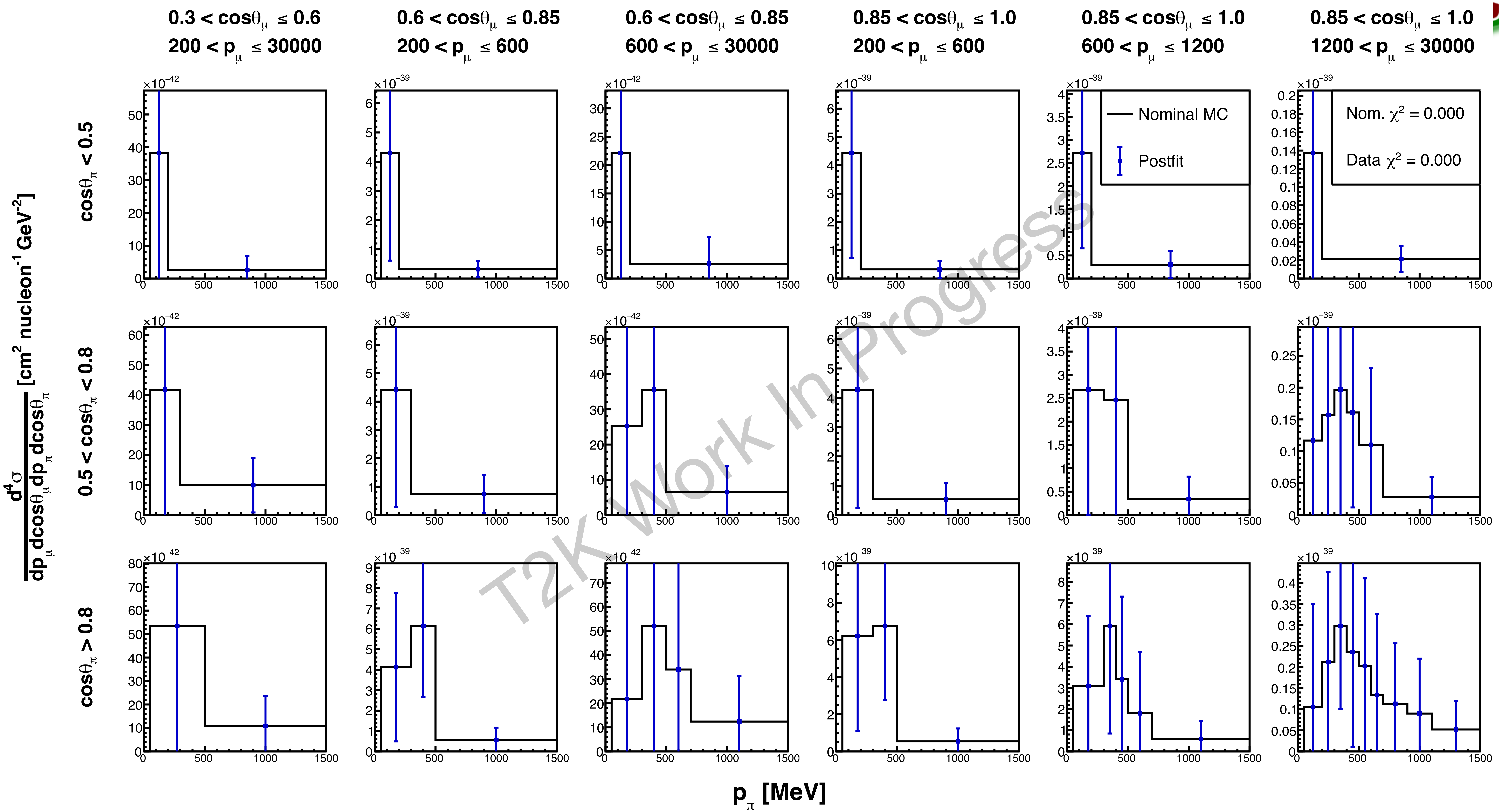


4D CH xsec



Muon kinematic slices
$\cos\theta_\pi$
0.3 < cosθ_μ ≤ 0.6
200 < p_μ ≤ 30000
0.6 < cosθ_μ ≤ 0.85
200 < p_μ ≤ 600
0.6 < cosθ_μ ≤ 0.85
600 < p_μ ≤ 30000
0.85 < cosθ_μ ≤ 1.0
200 < p_μ ≤ 600
0.85 < cosθ_μ ≤ 1.0
600 < p_μ ≤ 1200
0.85 < cosθ_μ ≤ 1.0
1200 < p_μ ≤ 30000
cosθ_π < 0.5
0.5 < cosθ_π < 0.8
cosθ_π > 0.8
d⁴σ / dp_μ dcosθ_μ dp_π dcosθ_π [cm² nucleon⁻¹ GeV⁻²]
Nominal MC
Postfit
Nom. χ² = 0.000
Data χ² = 0.000
T2K Work In Progress
p_π [MeV]

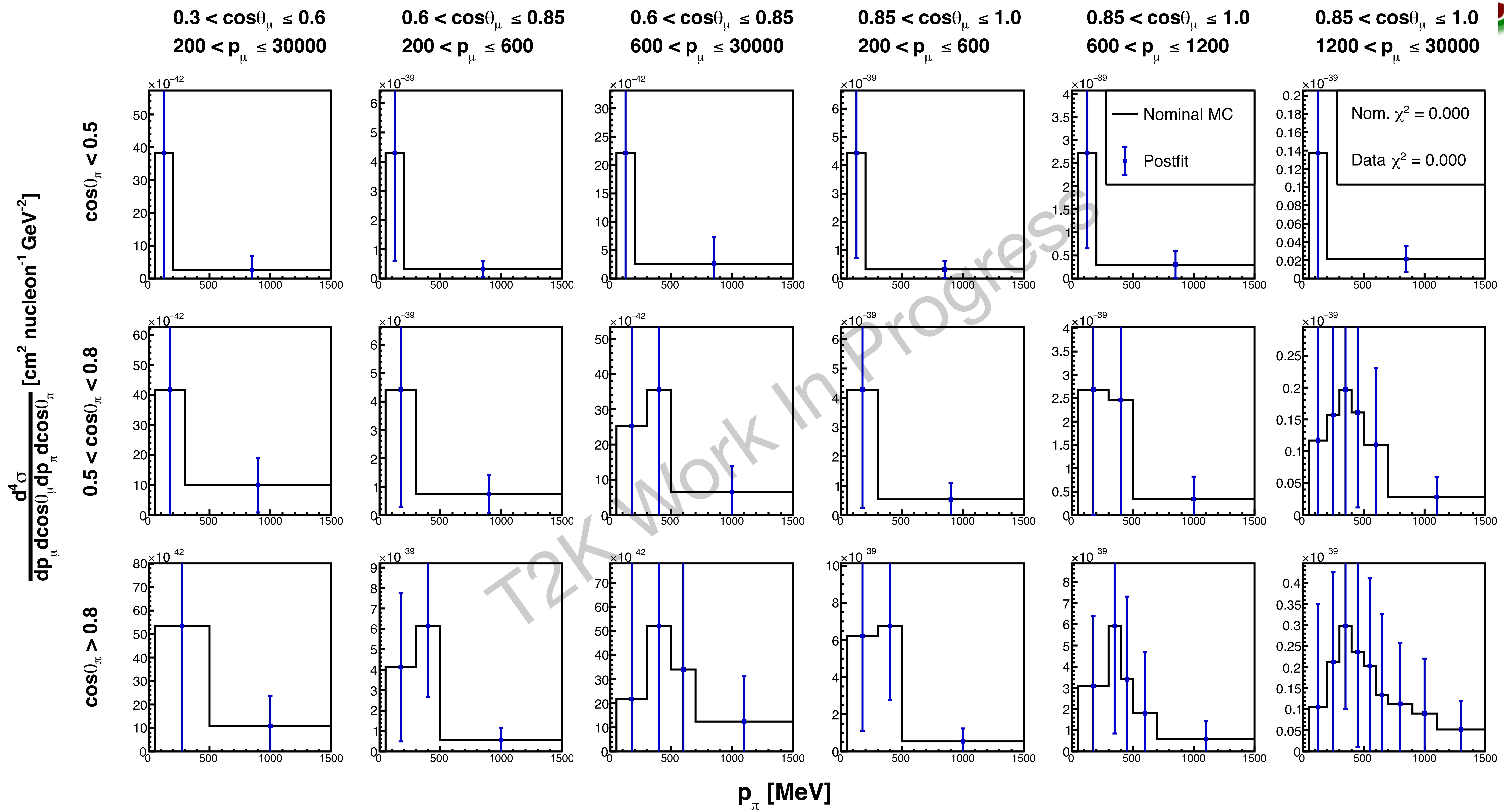
4D H_2O xsec



0.3 < cosθμ ≤ 0.6
200 < pμ ≤ 30000
0.6 < cosθμ ≤ 0.85
200 < pμ ≤ 600
0.6 < cosθμ ≤ 0.85
600 < pμ ≤ 30000
0.85 < cosθμ ≤ 1.0
200 < pμ ≤ 600
0.85 < cosθμ ≤ 1.0
600 < pμ ≤ 1200
0.85 < cosθμ ≤ 1.0
1200 < pμ ≤ 30000
cosθπ < 0.5
0.5 < cosθπ < 0.8
cosθπ > 0.8
d⁴σ / dpμ dcosθμ dpπ dcosθπ [cm² nucleon⁻¹ GeV⁻²]
Nominal MC
Postfit
Nom. χ² = 0.000
Data χ² = 0.000
T2K Work In Progress
pπ [MeV]

4D H_2O xsec

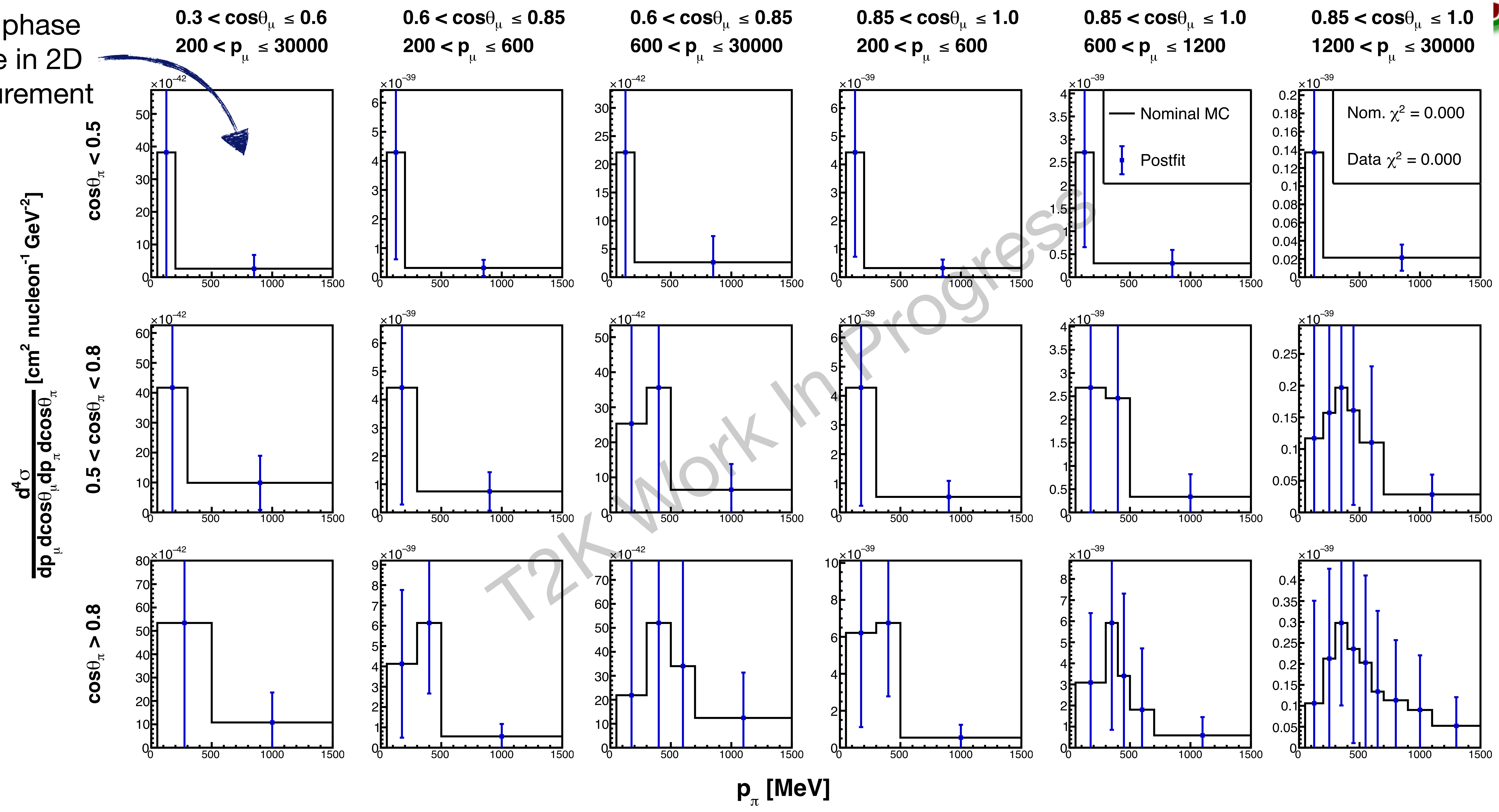
Integrate over muon kinematics to get a 2D result in pion kinematics



4D H_2O xsec

Integrate over muon kinematics to get a 2D result in pion kinematics

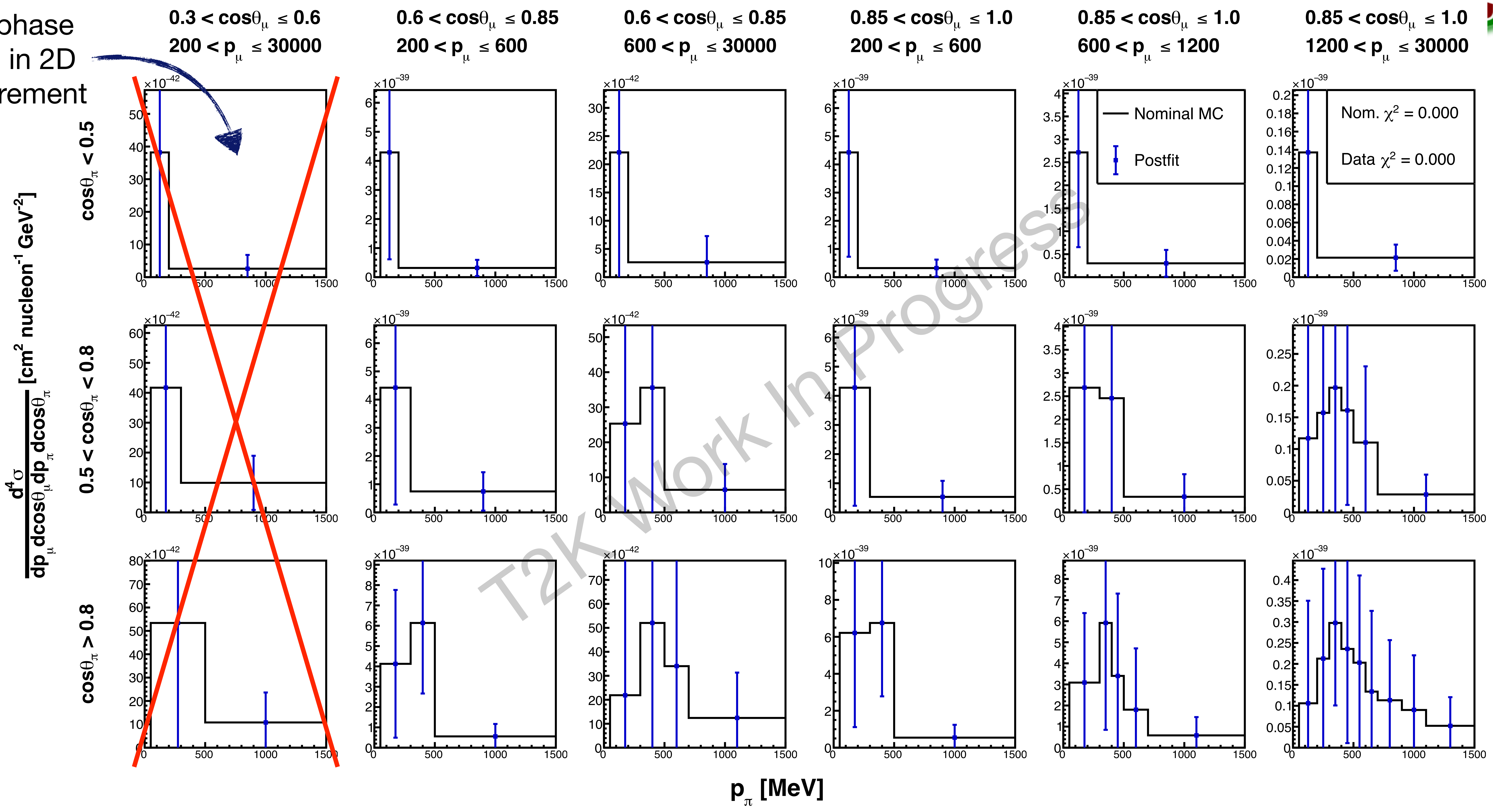
Limit phase space in 2D measurement



4D H_2O xsec

Integrate over muon kinematics to get a 2D result in pion kinematics

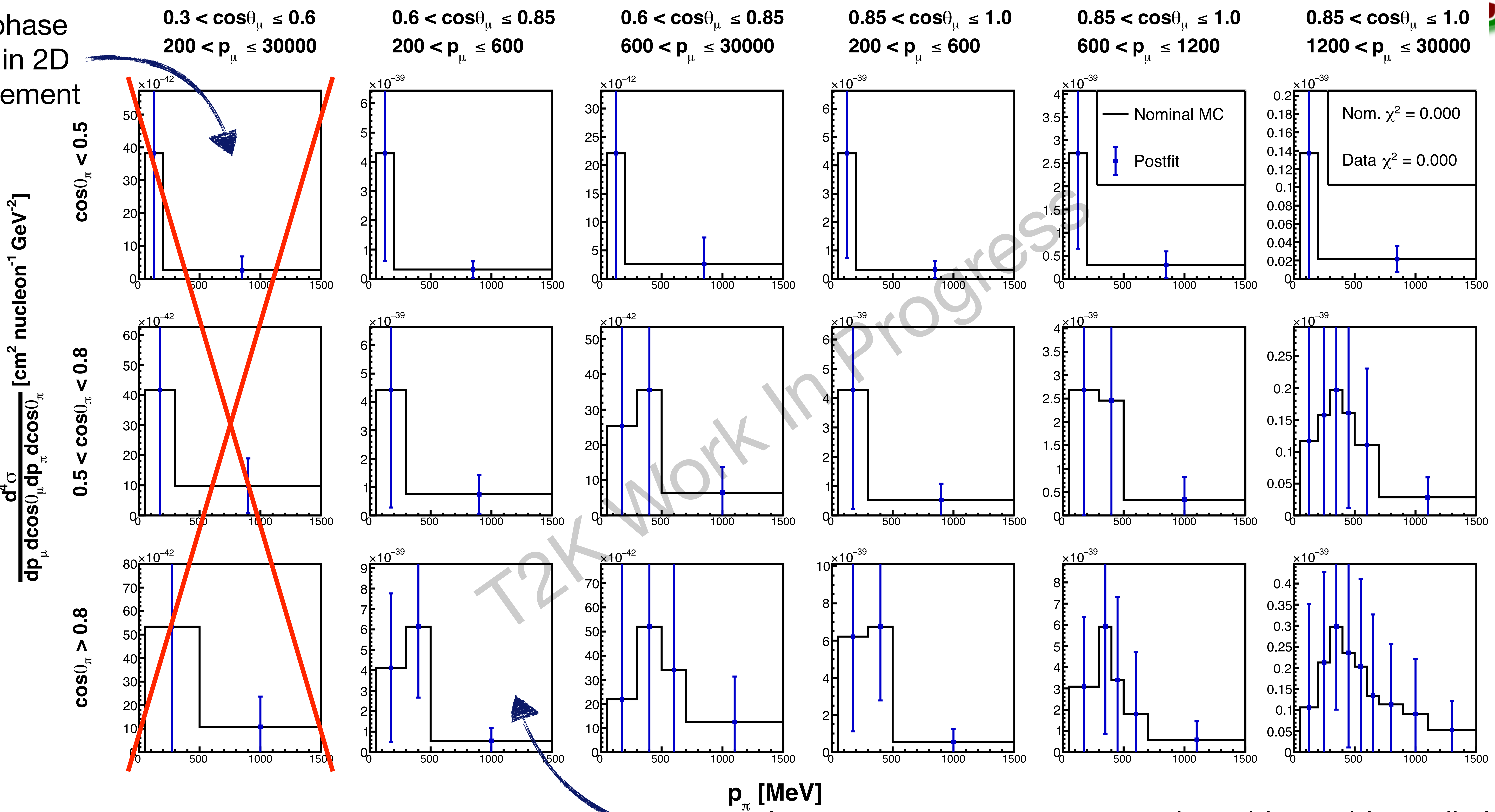
Limit phase space in 2D measurement



4D H_2O xsec

Integrate over muon kinematics to get a 2D result in pion kinematics

Limit phase space in 2D measurement

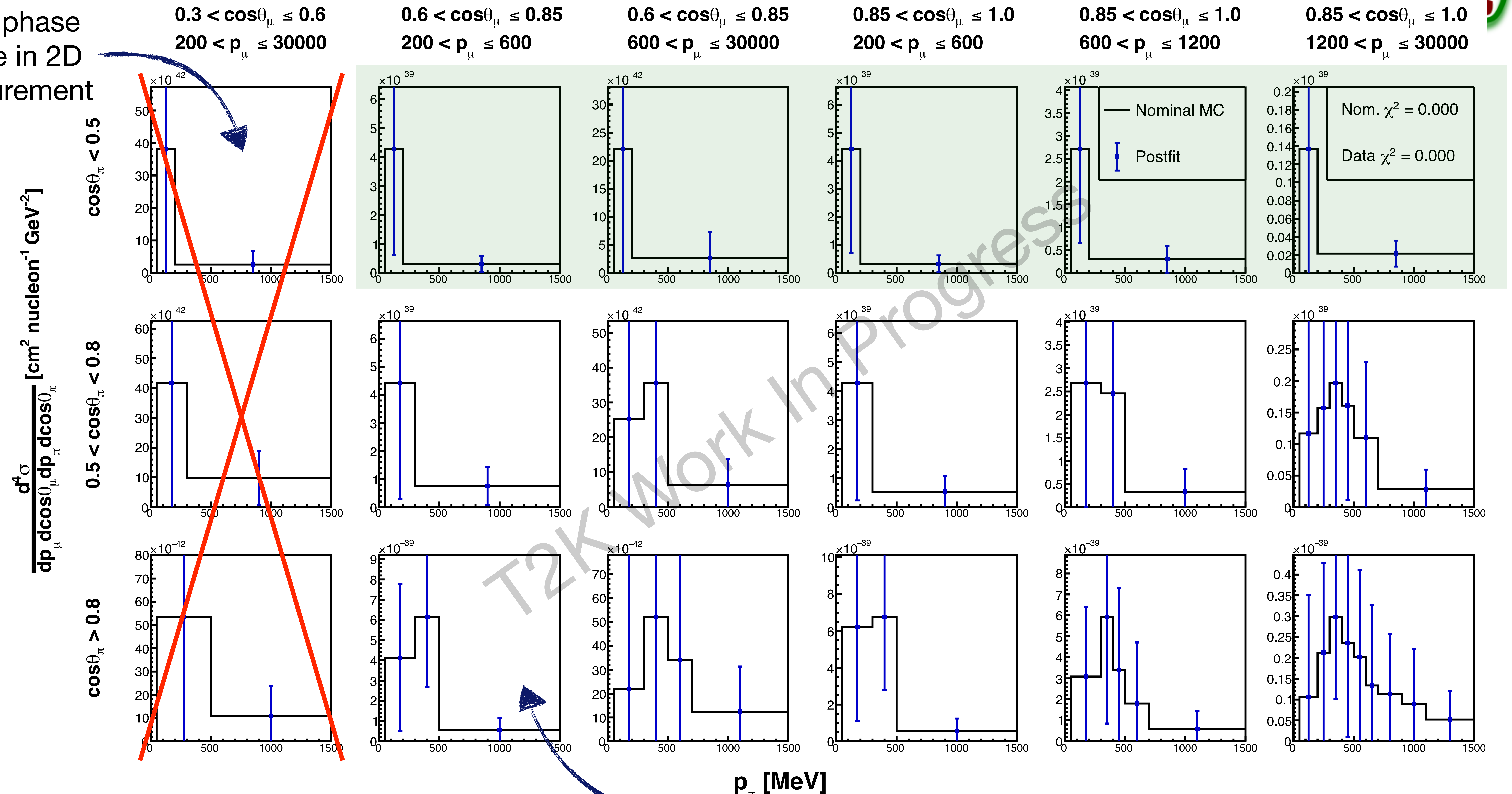


Integrate over muon space, and combine p_π bins to limiting case

4D H_2O xsec

Integrate over muon kinematics to get a 2D result in pion kinematics

Limit phase space in 2D measurement

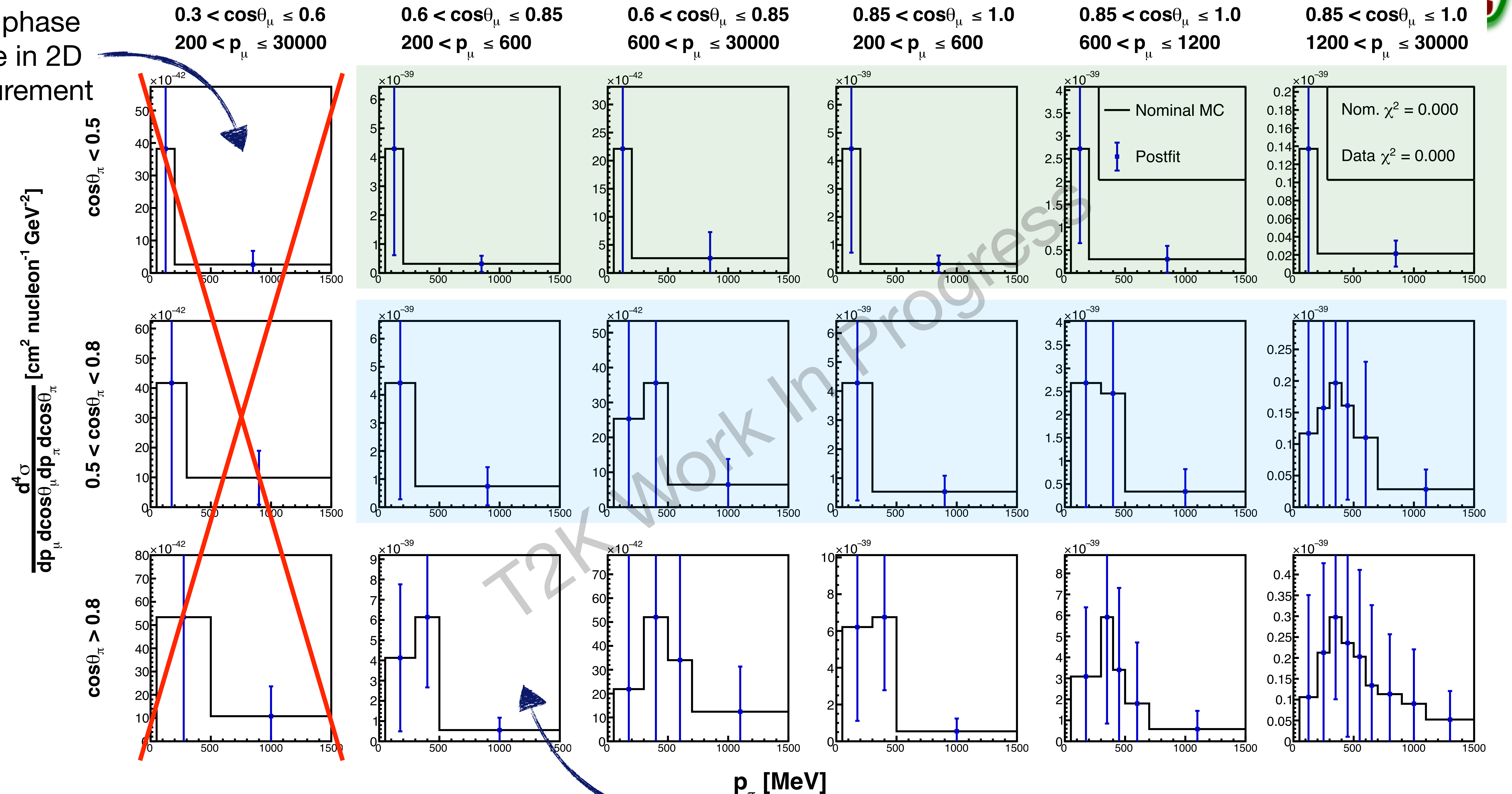


Integrate over muon space, and combine p_π bins to limiting case

4D H_2O xsec

Integrate over muon kinematics to get a 2D result in pion kinematics

Limit phase space in 2D measurement

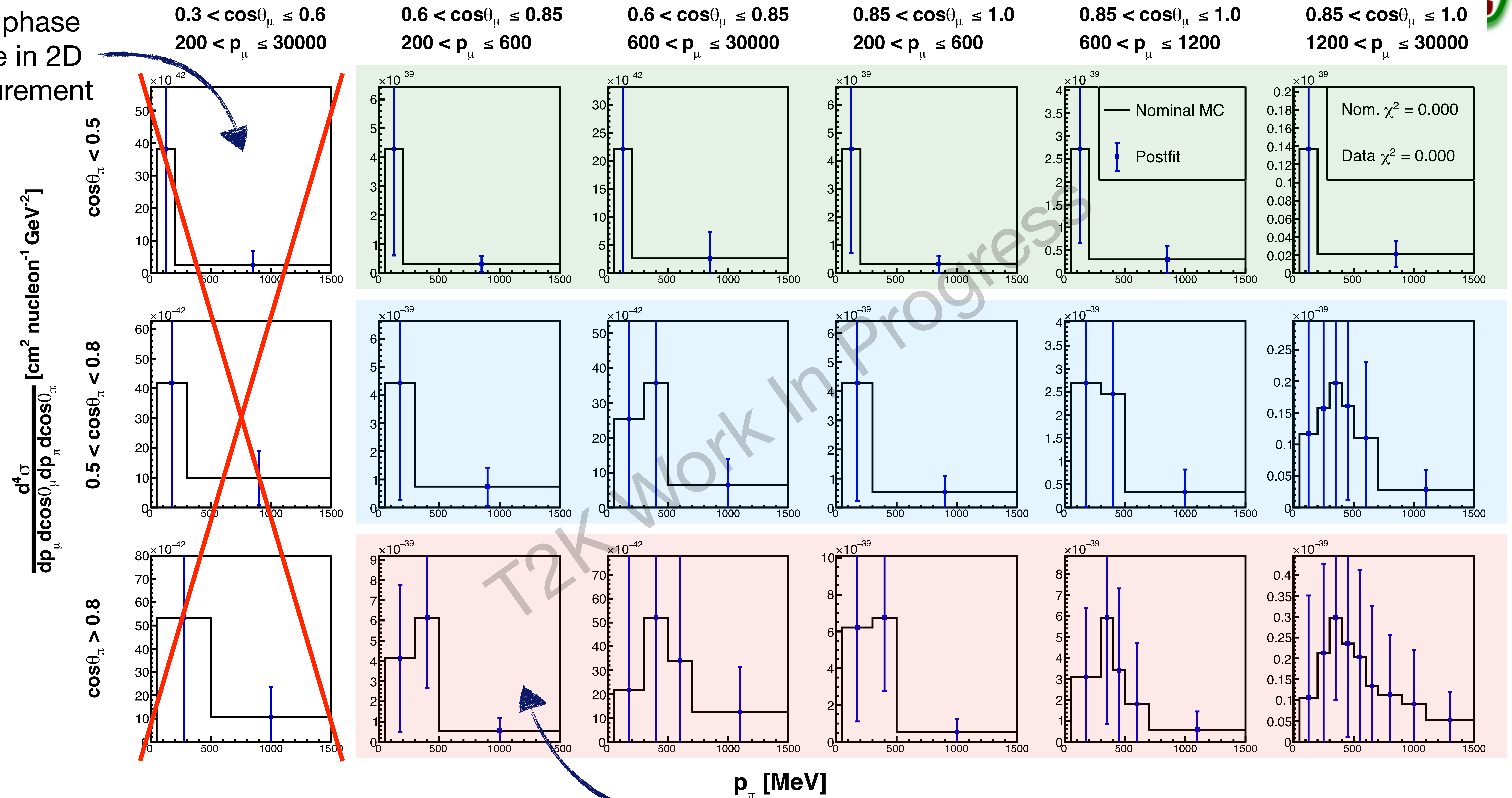


Integrate over muon space, and combine p_π bins to limiting case

4D H_2O xsec

Integrate over muon kinematics to get a 2D result in pion kinematics

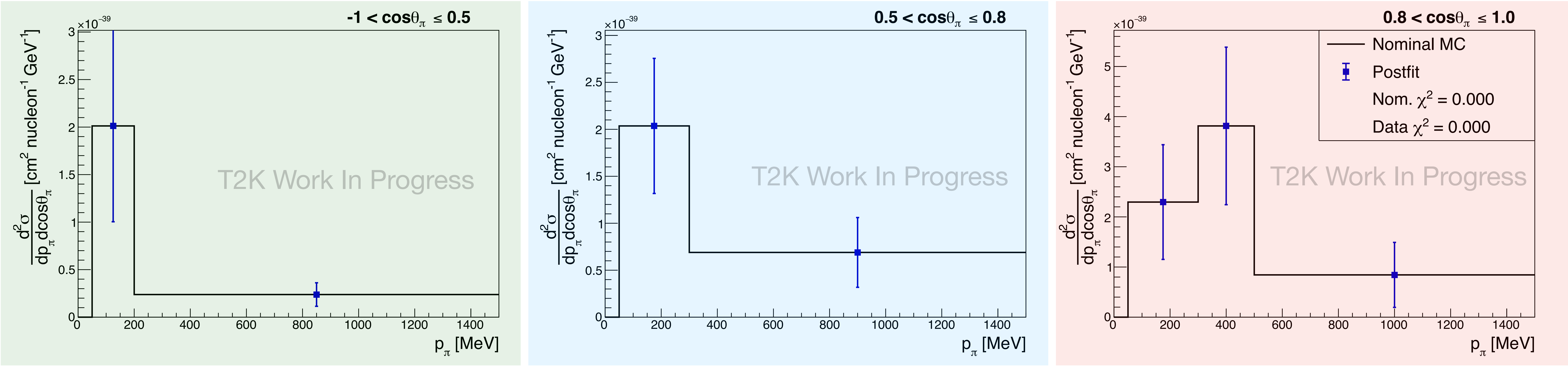
Limit phase space in 2D measurement



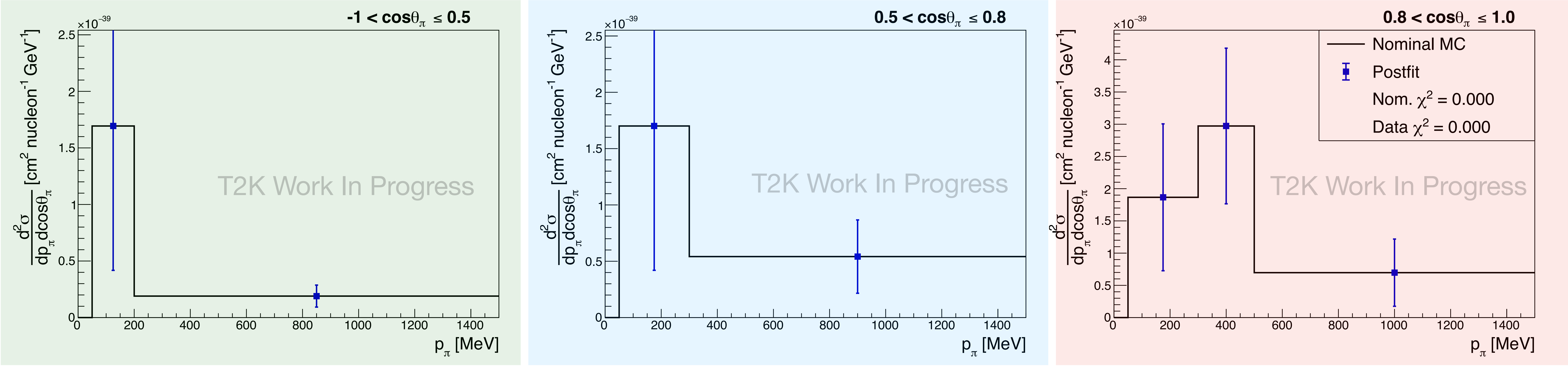
Integrate over muon space, and combine p_π bins to limiting case

2D collapsed xsec

CH

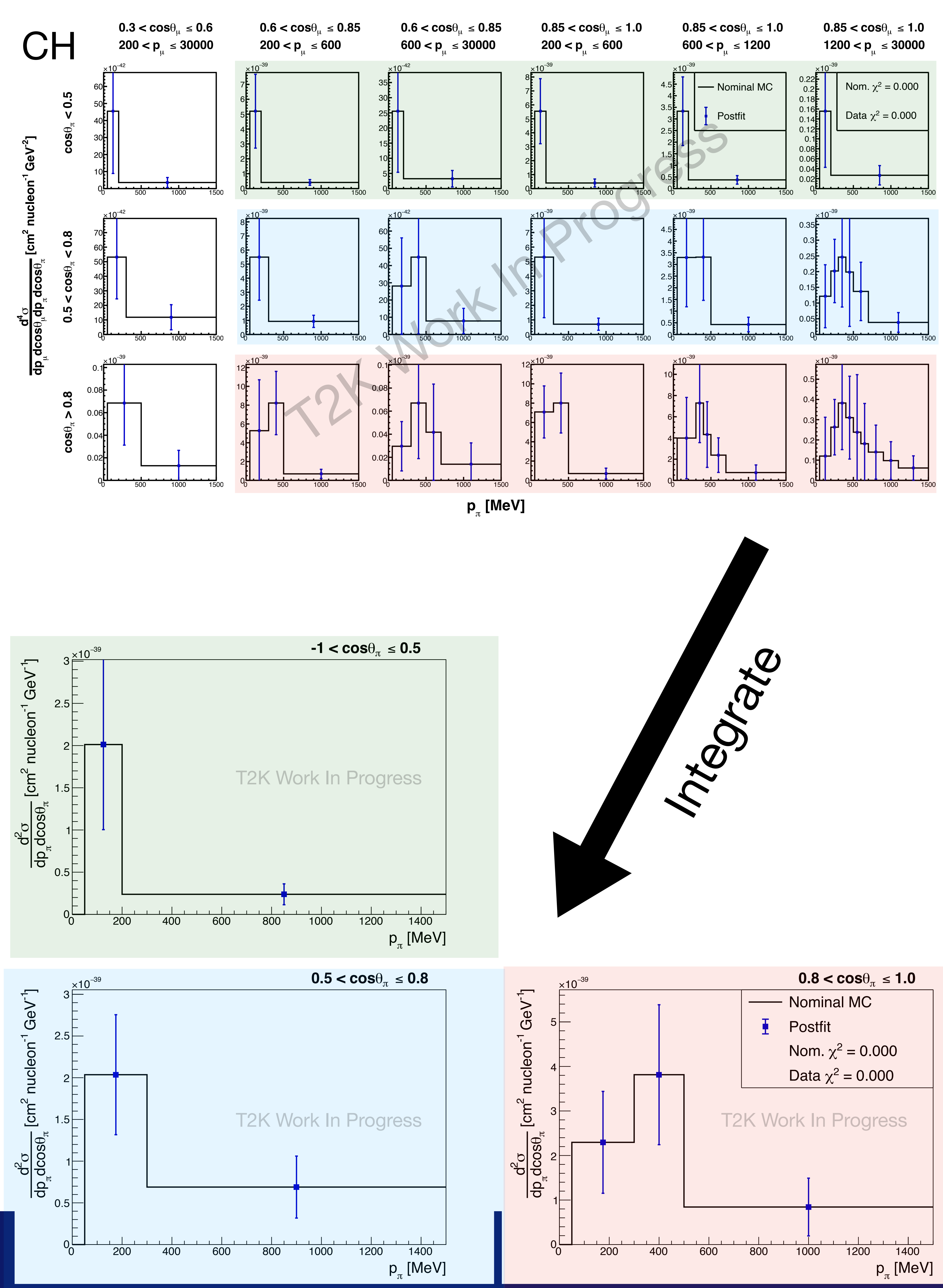


H_2O



2D collapsed xsec

- Extracting in a fine 4D scheme allows us to perform efficiency corrections in a model independent way
  - Also provides a 4D result we can feasibly use (albeit with high stat. error)
- Collapsing to 2D in pion kinematics *after* extraction maintains the model independence, but gives reduction in stat. error
- Can collapse down to any number of dimensions less than 4, binning scheme permitting

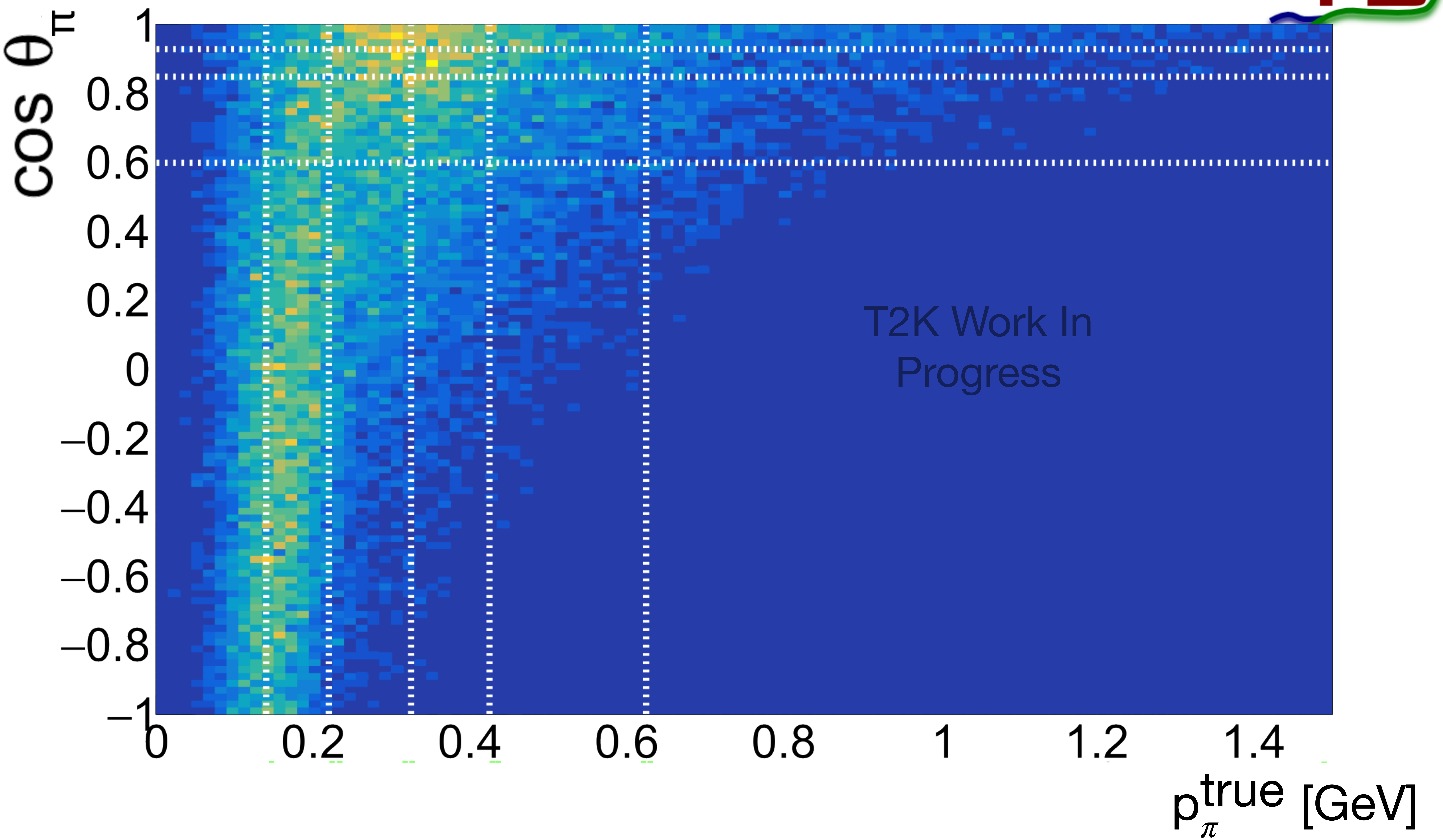


But wait, there's more!

- Fitting and performing efficiency corrections in fine binning before integrating to the desired result is clearly advantageous in terms of error coverage and avoiding model bias
- But it can also sometimes give us more information than we would otherwise get
  - When integrating wider bins, we keep track of the correlations between them
- Efficiency corrections in $N > 1$D performed using a single set of toy throws
- Separately integrate down to a single variable (p_π), and then a different one ($\cos\theta_\pi$)
- We have results in 2 variables, plus the correlations between those 2!

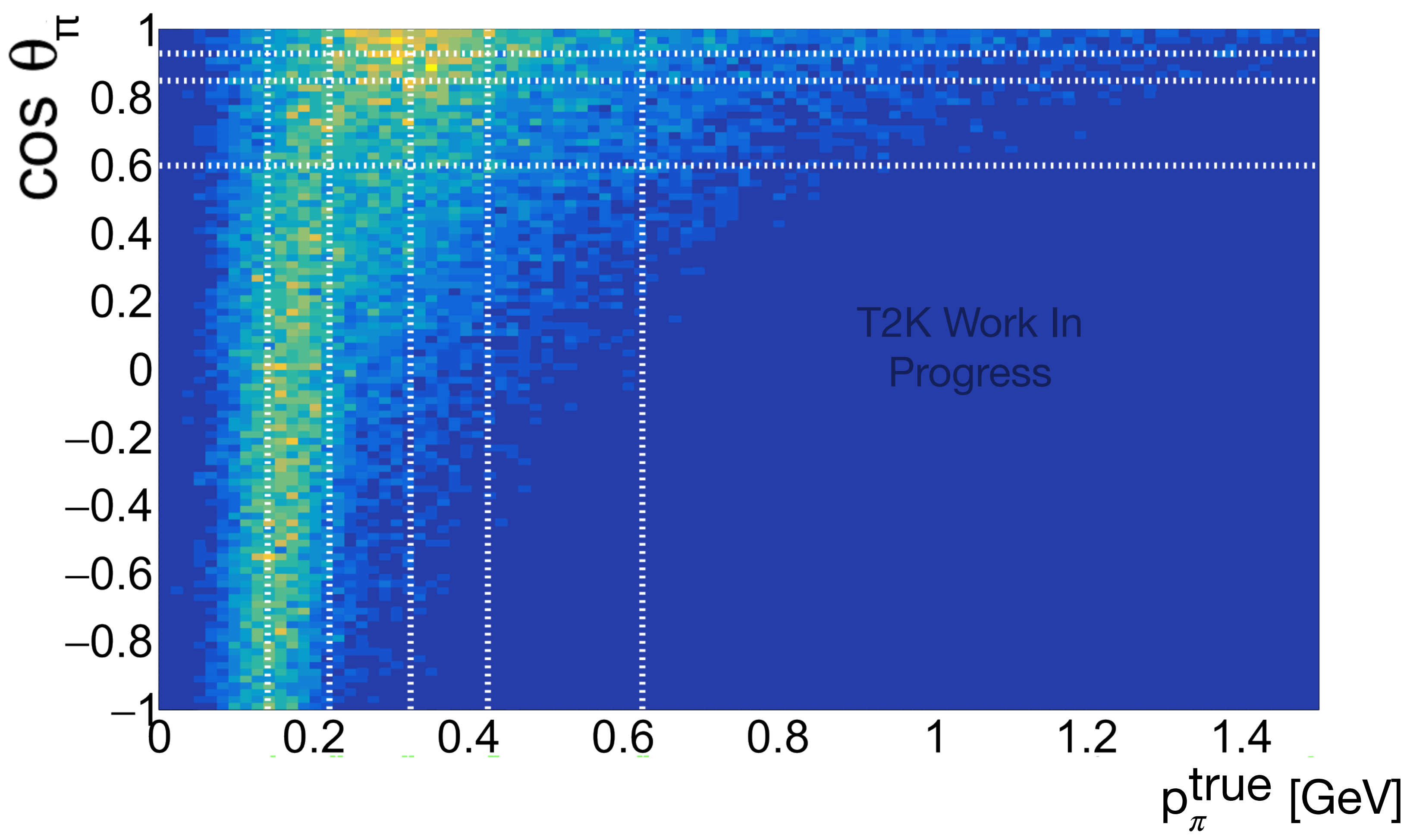
$\bar{\nu}_\mu CC1\pi^-$ cross section

Work by Liam O'Sullivan (JGU)



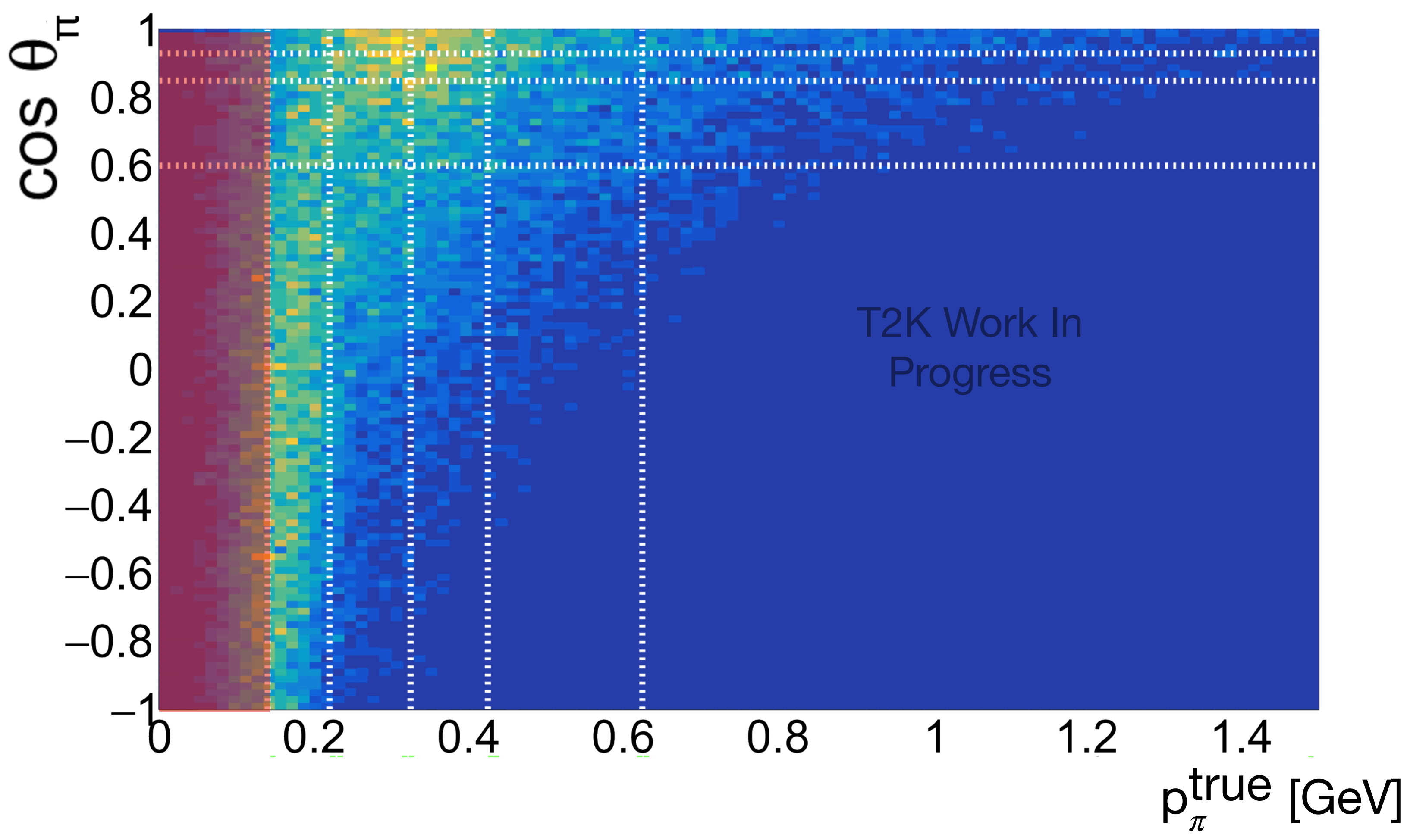
$\bar{\nu}_\mu CC1\pi^-$ cross section

Work by Liam O'Sullivan (JGU)



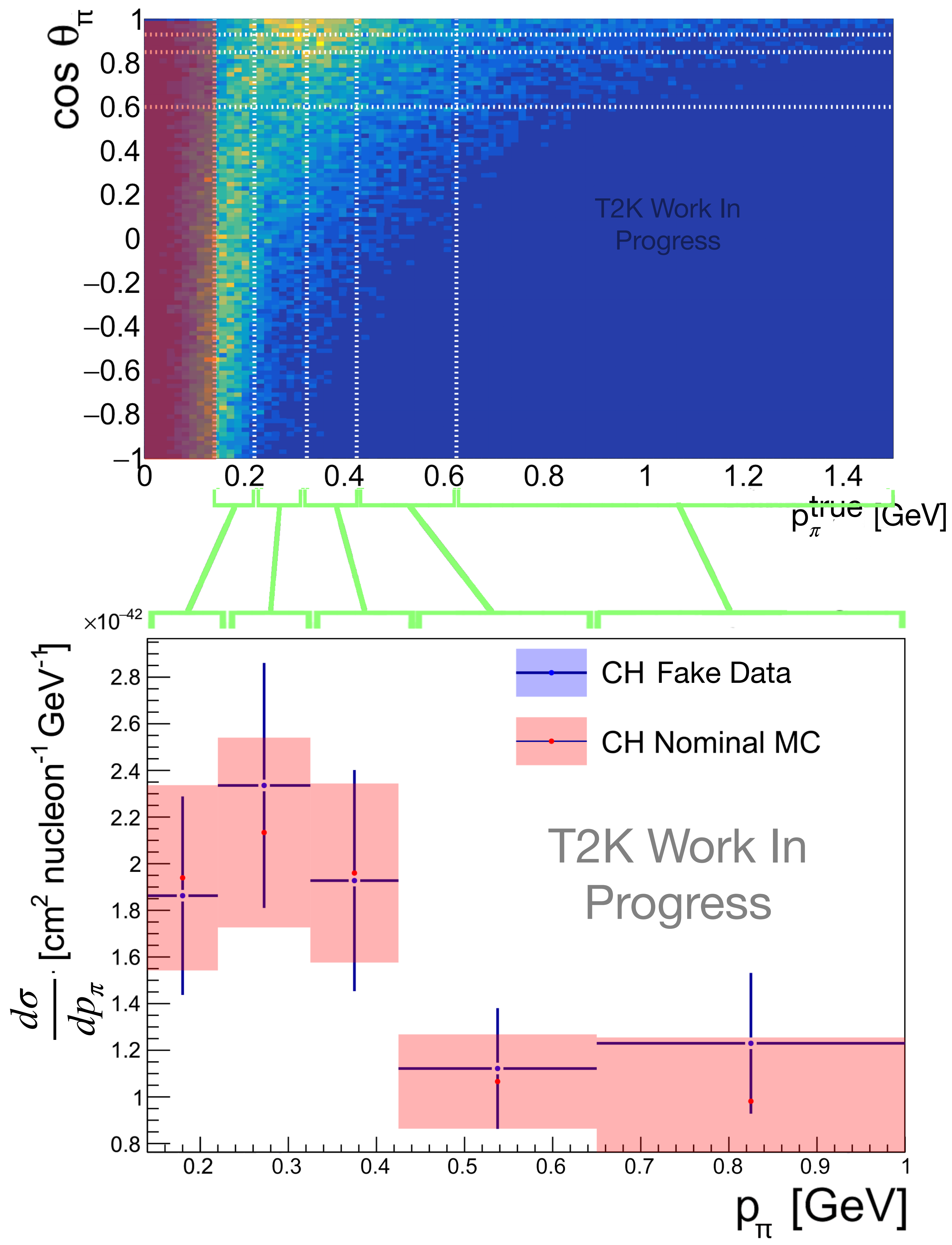
$\bar{\nu}_\mu CC1\pi^-$ cross section

Work by Liam O'Sullivan (JGU)



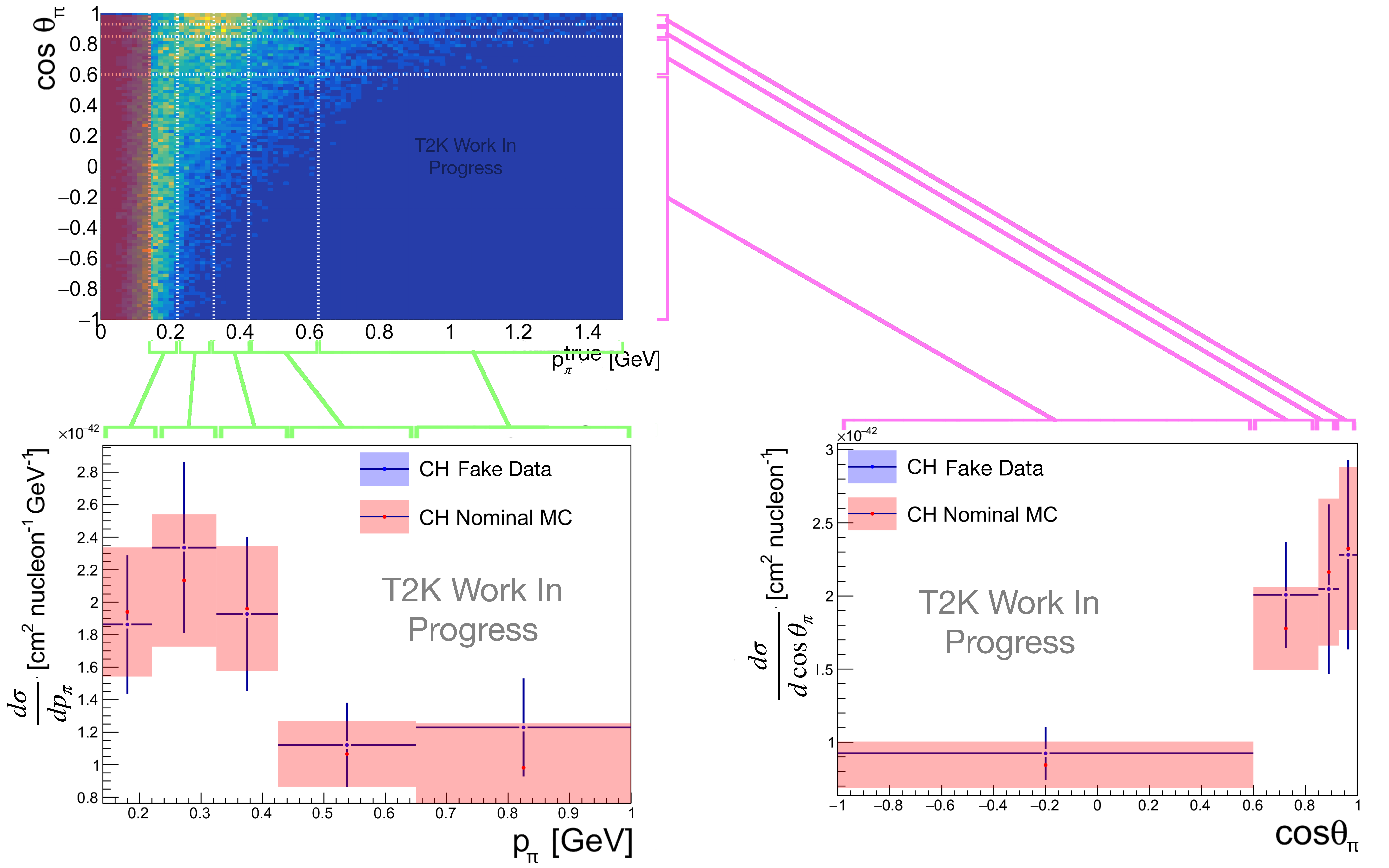
$\bar{\nu}_\mu CC1\pi^-$ cross section

Work by Liam O'Sullivan (JGU)



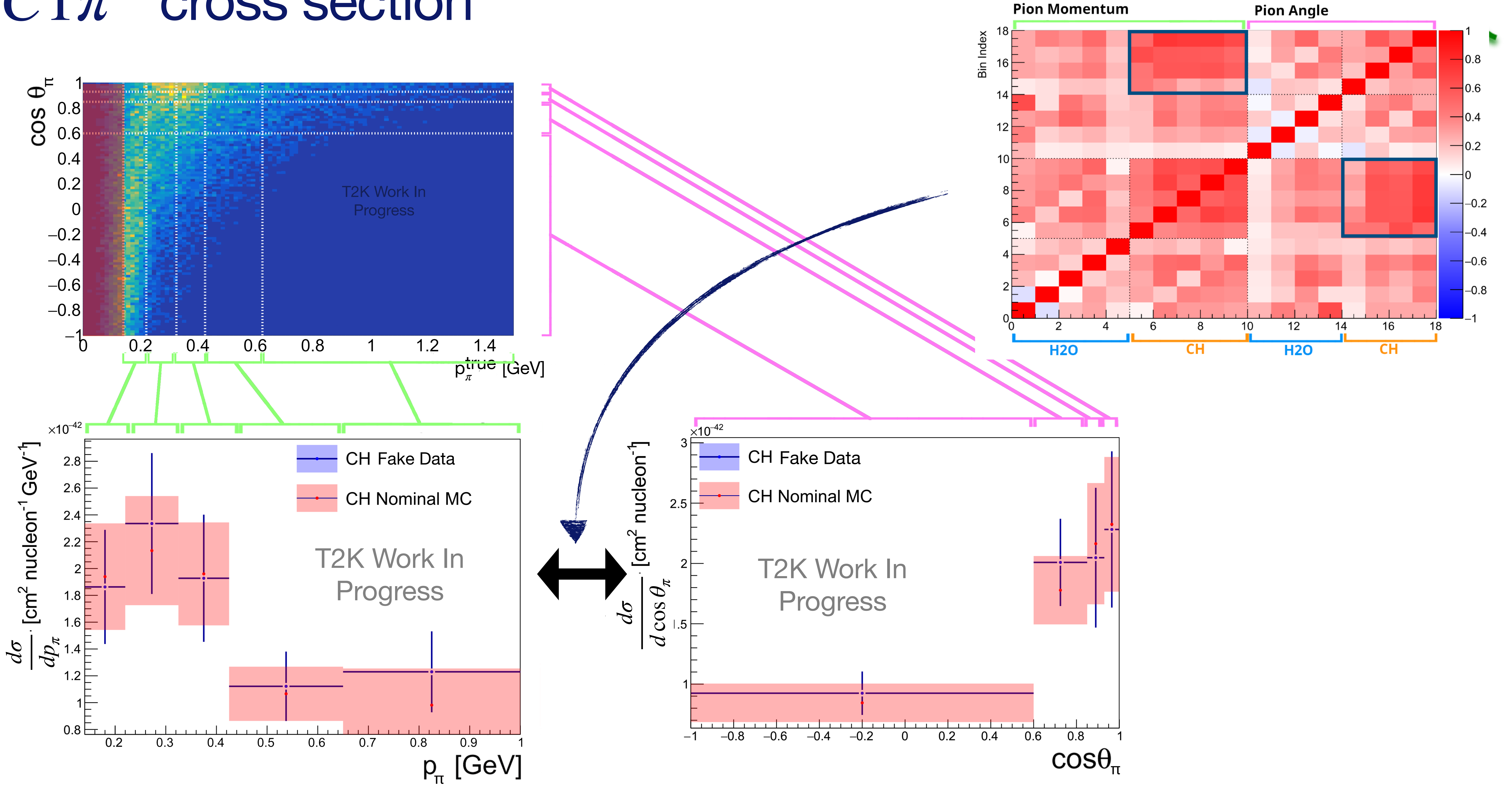
$\bar{\nu}_\mu CC1\pi^-$ cross section

Work by Liam O'Sullivan (JGU)



$\bar{\nu}_\mu CC1\pi^-$ cross section

Work by Liam O'Sullivan (JGU)



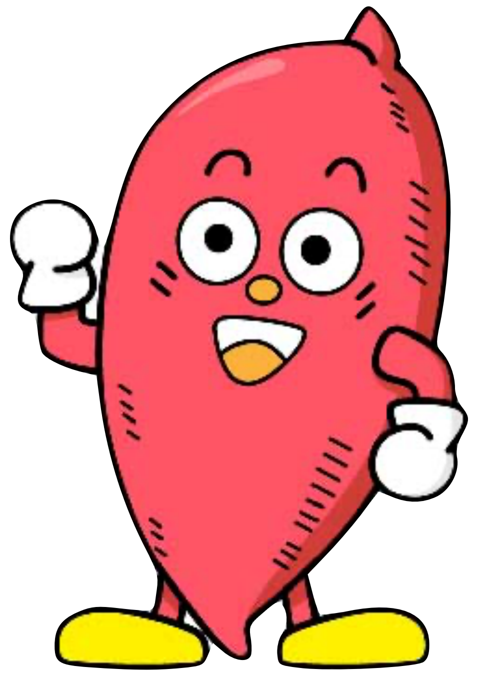
Things to be mindful of

- Still not a perfect approach
- Low statistics in the full 4D binning scheme can cause us to run into additional issues
- Have to be careful about bins with 0 events - in data this is okay, but 0 predicted events will cause issues
  - Aim for ~10 events as a minimum in a bin as a general rule
- Potential to obtain negative best fit results (when template parameters are negative)
  - Non-physical, but as long as the error bar covers 0 this is numerically okay
  - Combining bins after can deal with this
- Requires fitting with a large number of parameters, which can lead to technical issues in running the fit
  - My analysis has a total 757 parameters
  - More parameters = slower fits - have to optimise as best we can
  - Can cause numerical instability in calculating covariance matrix

See Laura's talk on new tools ft. GUNDAM

Conclusions

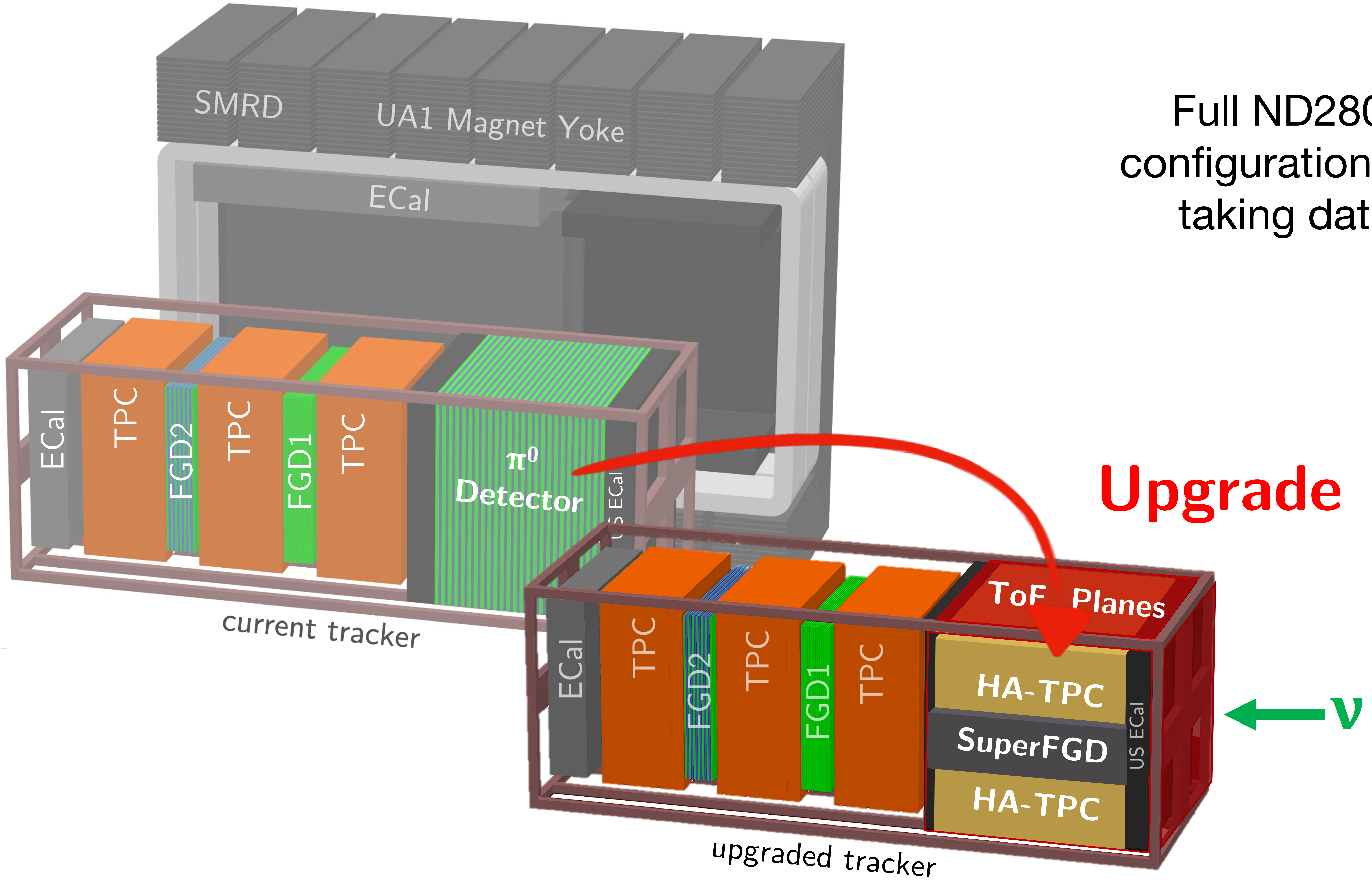
- The way we perform efficiency corrections in cross section analyses has to be carefully considered
  - Particularly when performing analyses in multiple variables
- Apply phase space constraints, limiting to regions we know we can measure
- Perform fit and extract cross section using a fine binning scheme, then integrate down to wider bins/lower dimensions to reduce statistical error
- If the same sets of toys are used, we also get correlations between measurements at lower dimensions




T2K
February ‘23

Backup

ND280 Upgrade



Full ND280Upgrade configuration due to start taking data in 2024

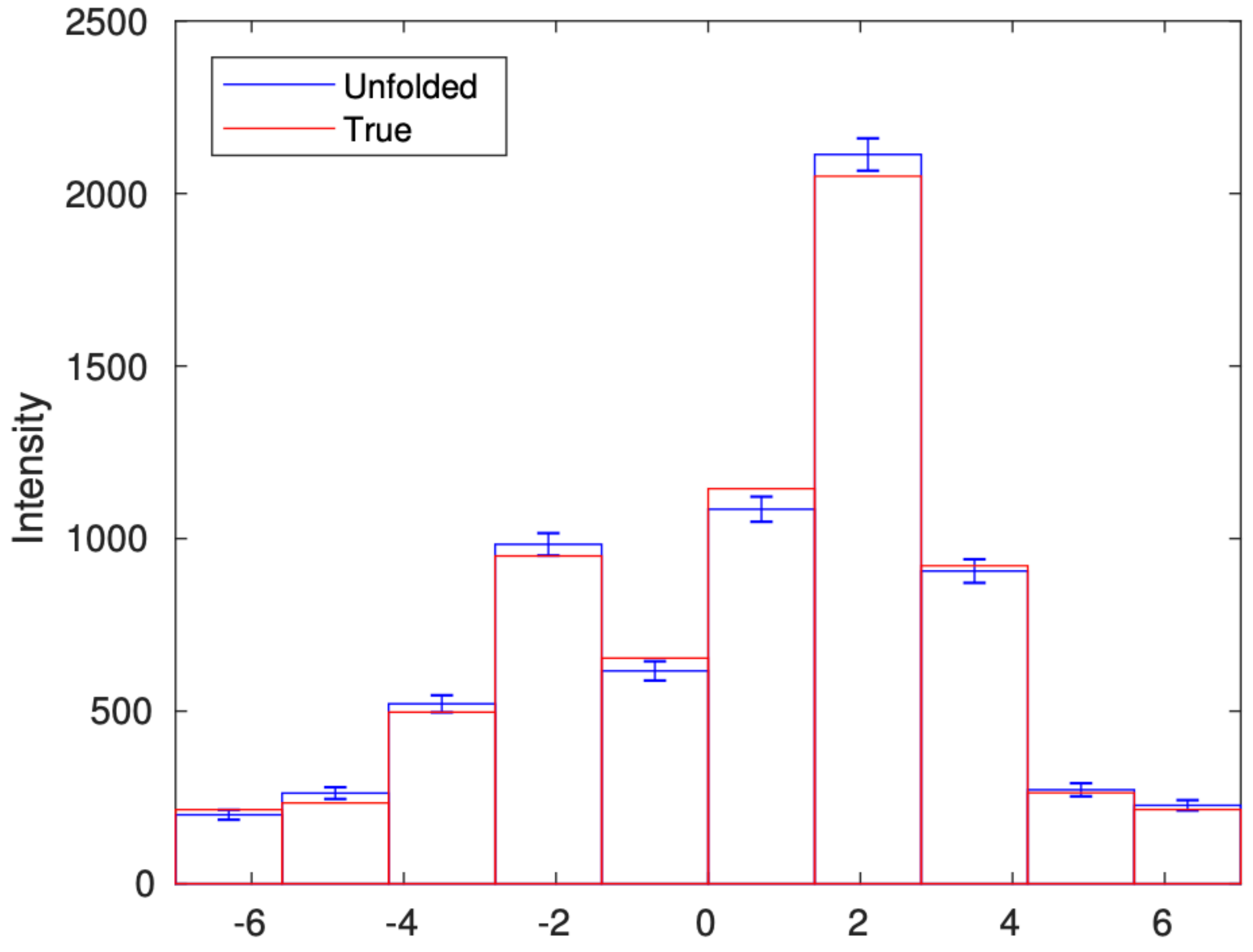
Wide bin unfolding

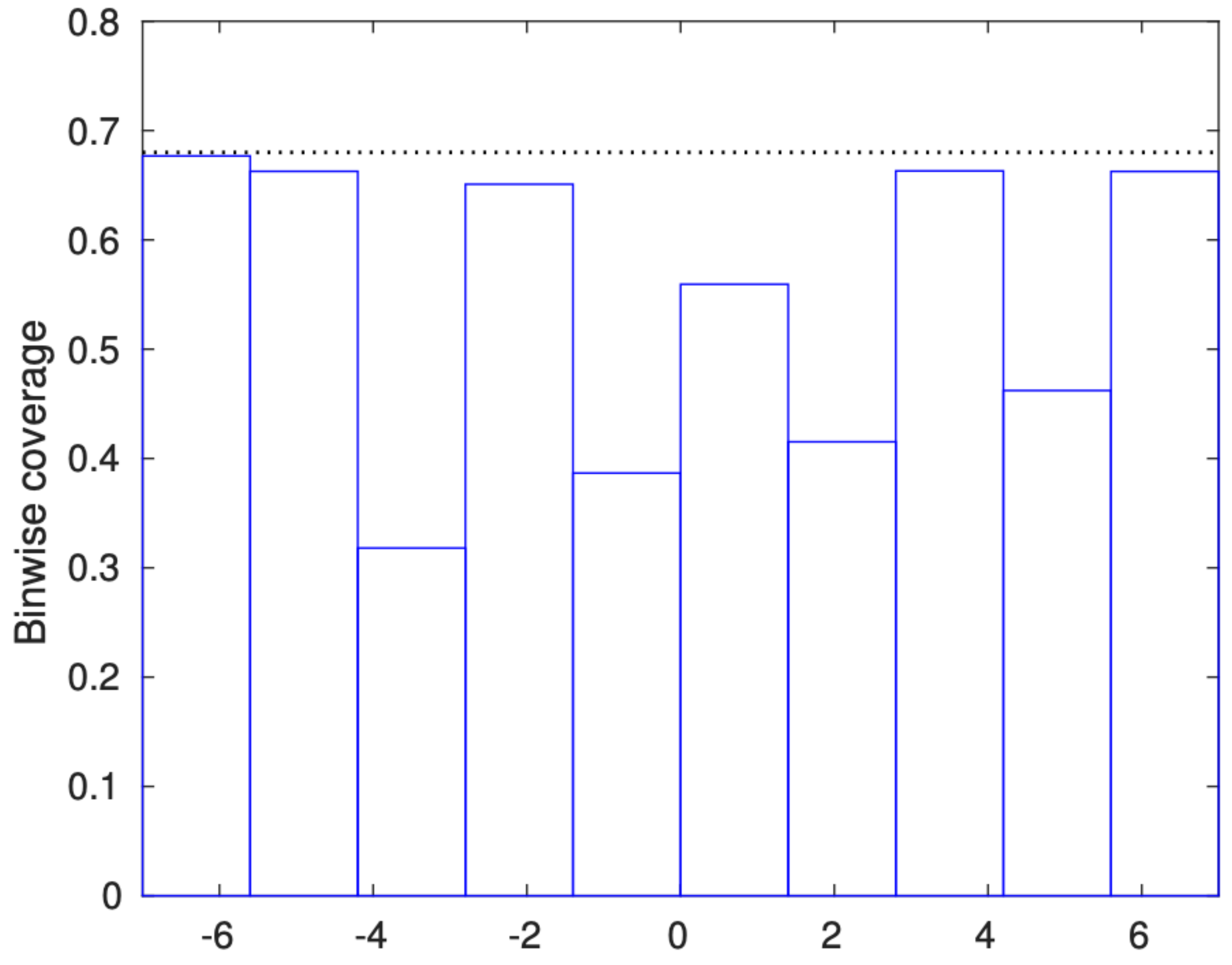
- One functional we should be able to recover without explicit regularization is the integral of f over a *wide* unfolded bin:

$$H_j[f] = \int_{T_j} f(t)\,\mathrm{d}t, \quad \text{width of } T_j \text{ large}$$

- But one cannot simply arbitrarily increase the particle-level bin size in the conventional approaches, since this increases the MC dependence of $\boldsymbol{K}$
- To circumvent this, *it is possible to first unfold with fine bins and then aggregate into wide bins*
- Let's see how this works!
  - Simulation setup: $\hat{\boldsymbol{\lambda}} = \boldsymbol{K}^{\dagger}\boldsymbol{y}$, convolution kernel $\mathcal{N}(0, 0.35^2)$, slightly different f^{MC}, otherwise as before

Wide bins, standard approach, perturbed MC

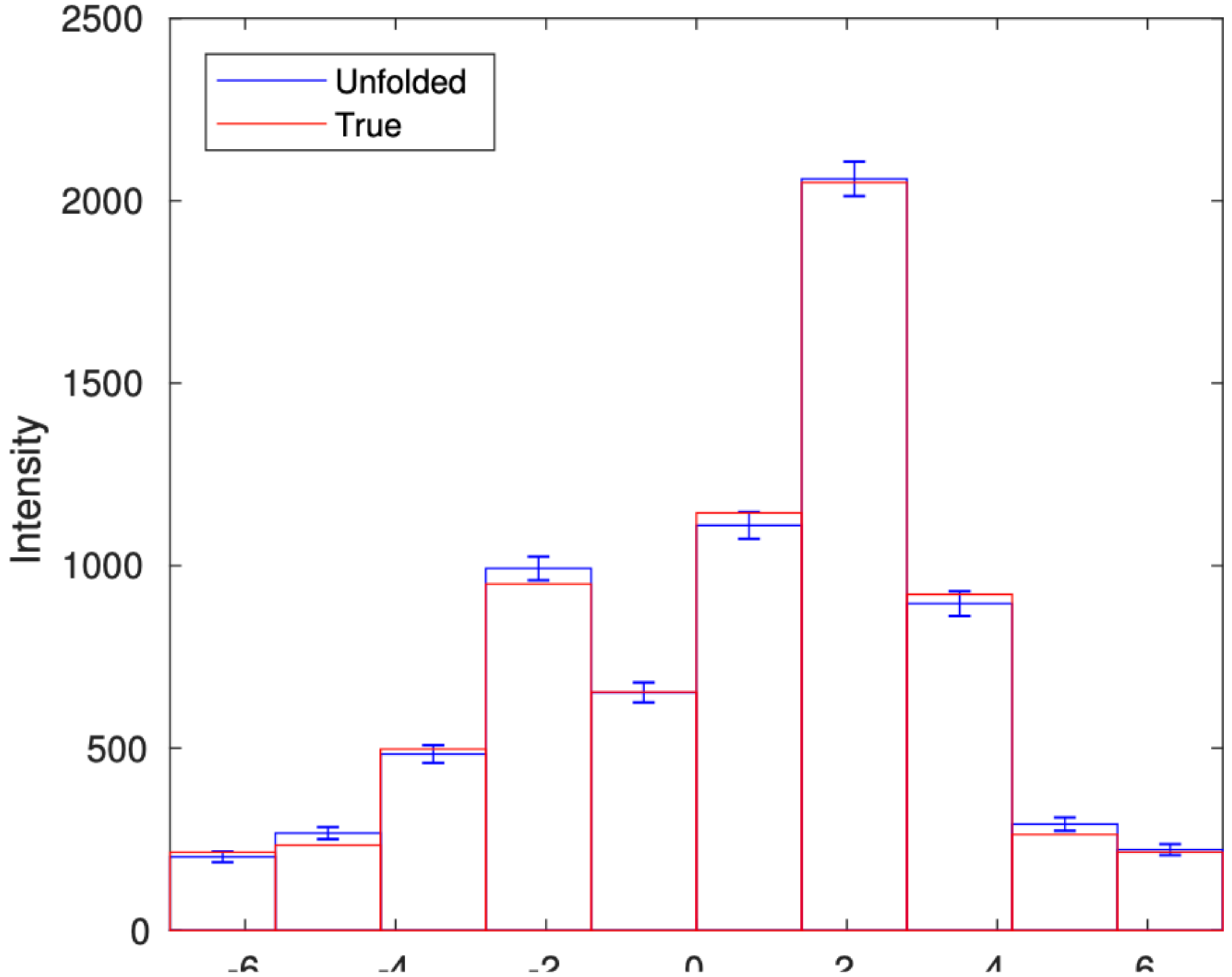


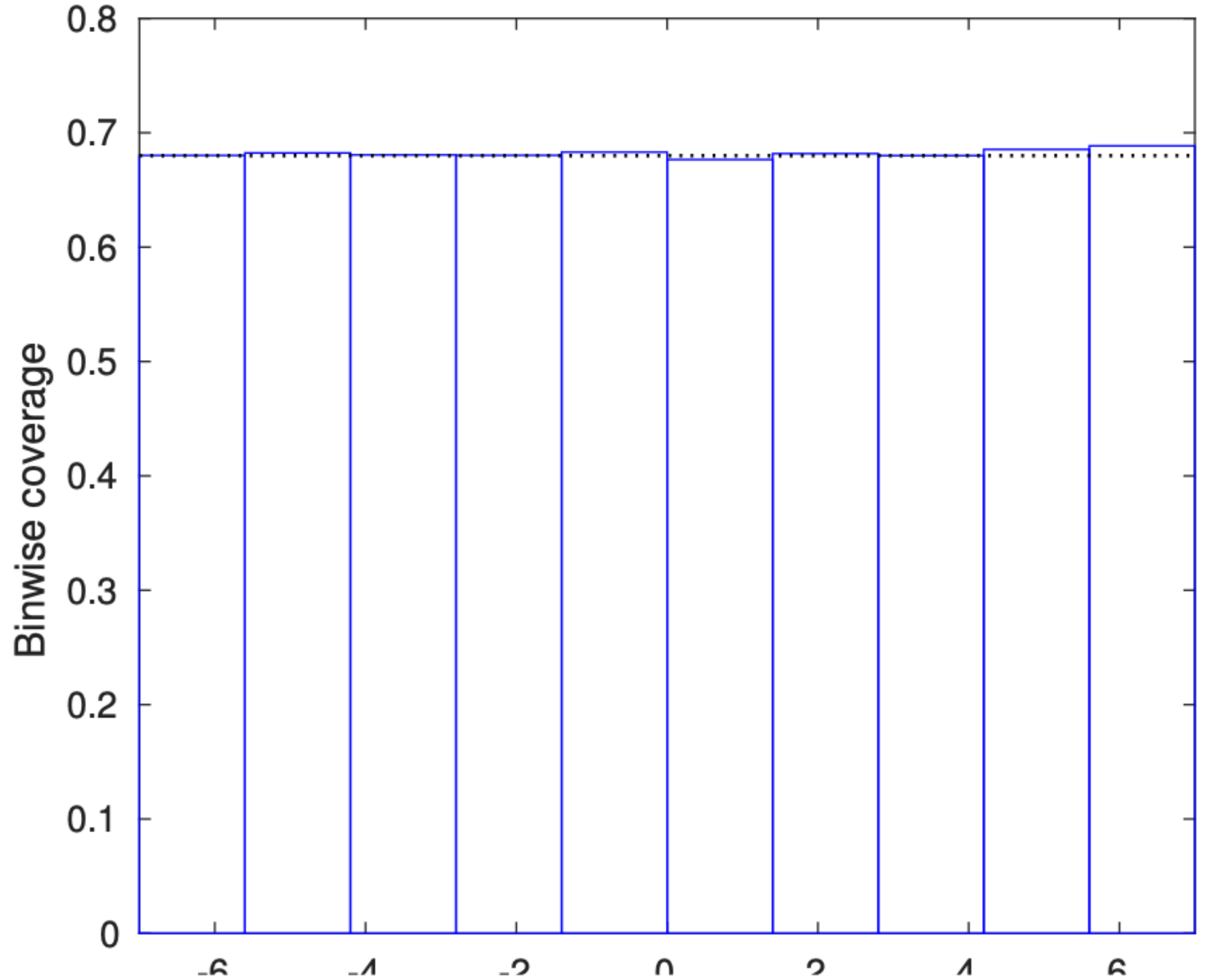


The response matrix $K_{i,j} = \frac{\int_{S_i} \int_{T_j} k(s,t) f^{\mathrm{MC}}(t)\,dt\,ds}{\int_{T_j} f^{\mathrm{MC}}(t)\,dt}$ depends on f^{MC}

$\Rightarrow$ Undercoverage if $f^{\mathrm{MC}} \neq f$

Wide bins, standard approach, correct MC

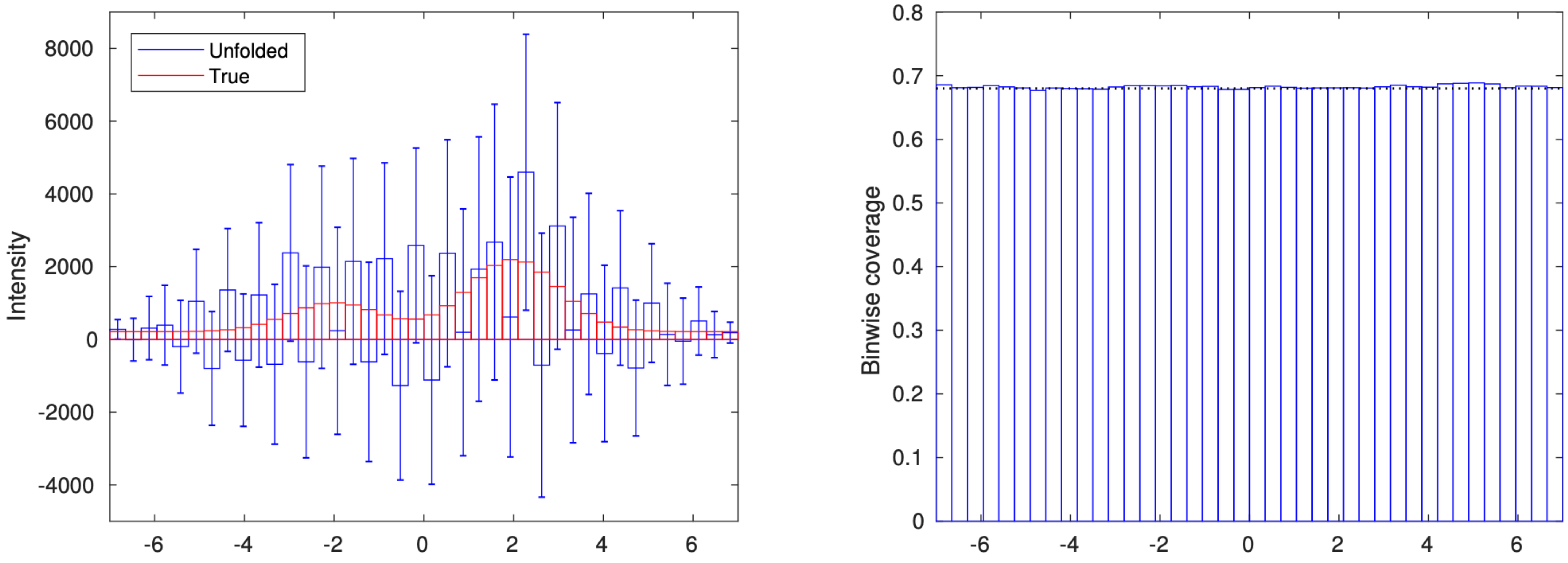




If $f^{\mathrm{MC}} = f$, coverage is correct

$\Rightarrow$ But this situation is unrealistic because f of course is unknown

Fine bins, standard approach, perturbed MC

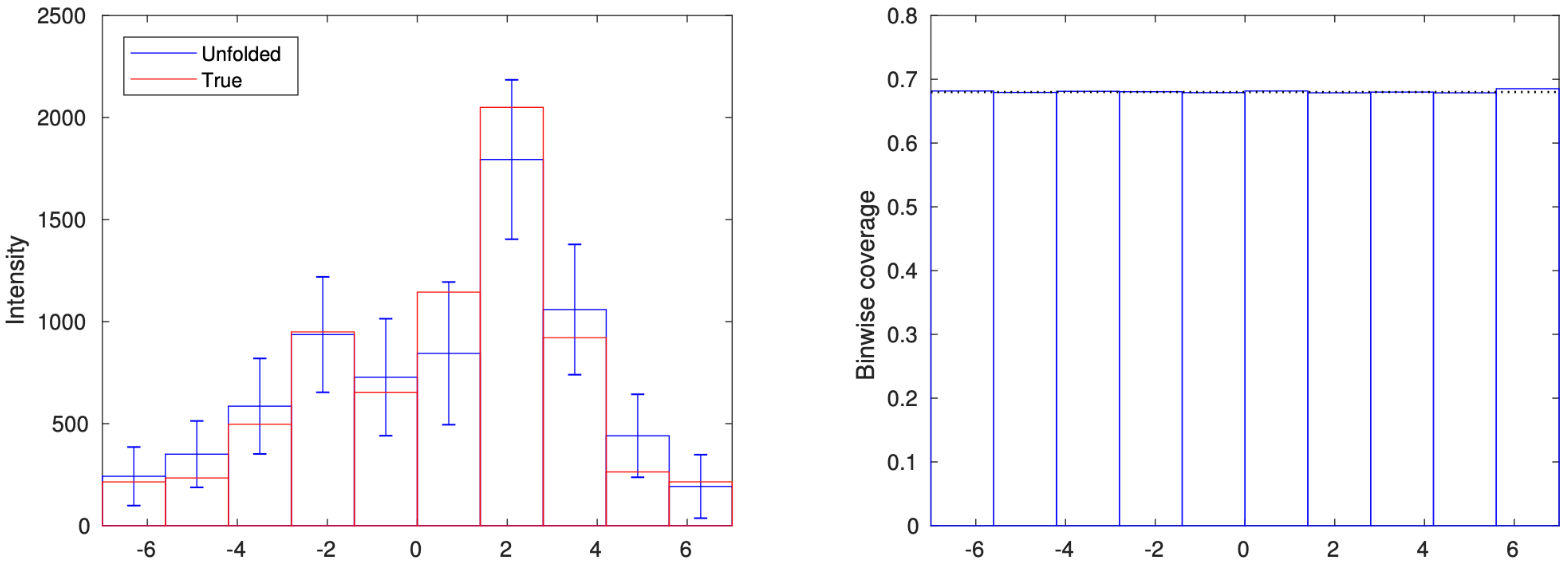


With narrow bins, less dependence on f^{MC} so coverage is correct, but the intervals are very wide[1]

$\Rightarrow$ Let's aggregate these into wide bins, keeping track of the correlations

[1] More unfolded realizations given in the backup.

Wide bins via fine bins, perturbed MC



Wide bins via fine bins gives both correct coverage and intervals with reasonable length[2]

[2]More unfolded realizations given in the backup.