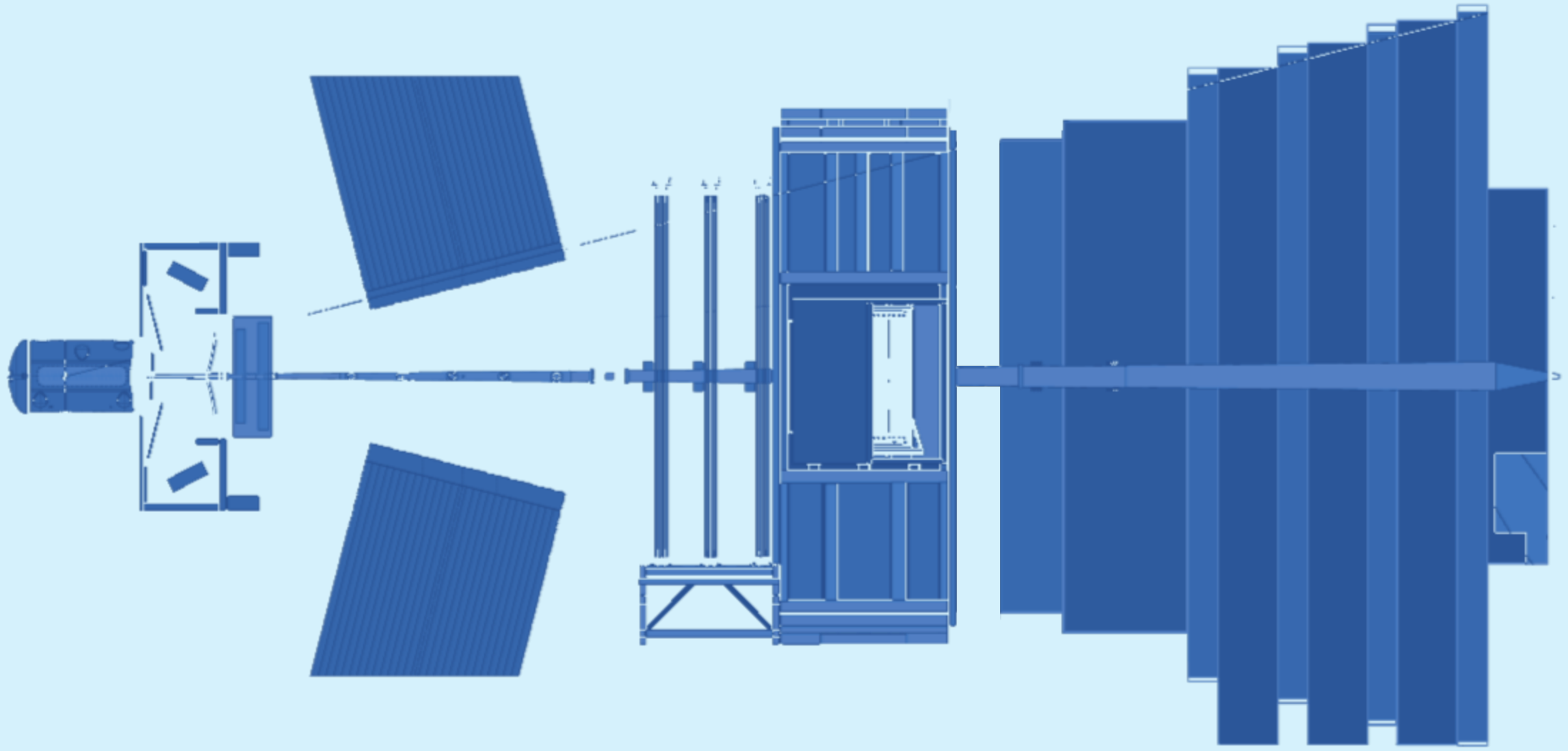
Institut de Ciències del Cosmos
UNIVERSITAT DE BARCELONA
EXCELENCIA MARÍA DE MAEZTU 2020-2023



Performance of Electron Identification in LHCb 2024

By: Pol Vidrier, Carla Marín, Lukas Calefice
2nd COMCHA Workshop

UNIVERSITAT DE BARCELONA

2/10/24

Motivation

ElectronID usage

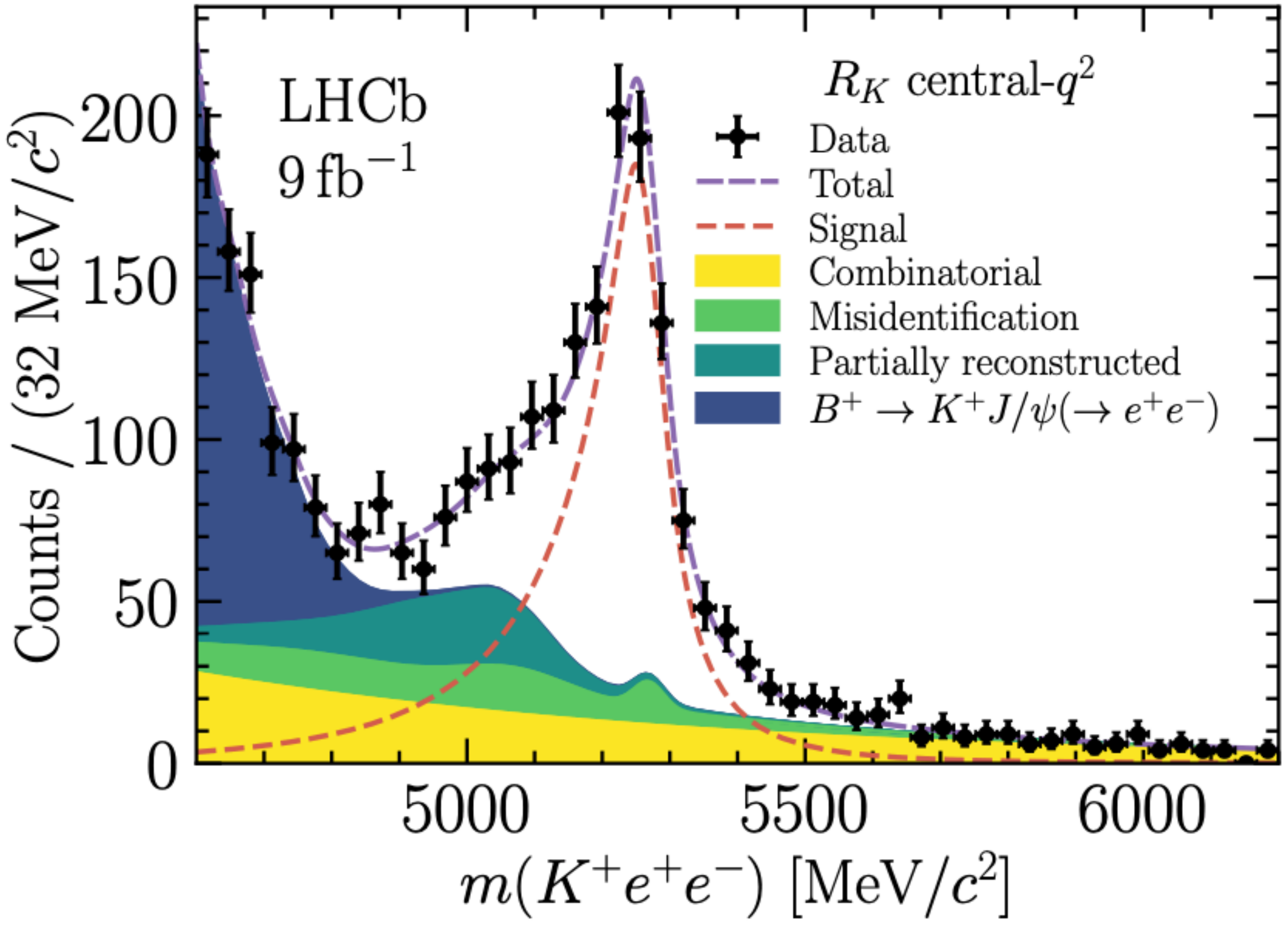
- The correct identification of electrons in LHCb is of vital importance for analyses that involve these particles, like Lepton Flavour Universality tests

Motivation

ElectronID usage

- The correct identification of electrons in LHCb is of vital importance for analyses that involve these particles, like Lepton Flavour Universality tests

The electron misID is significant here!



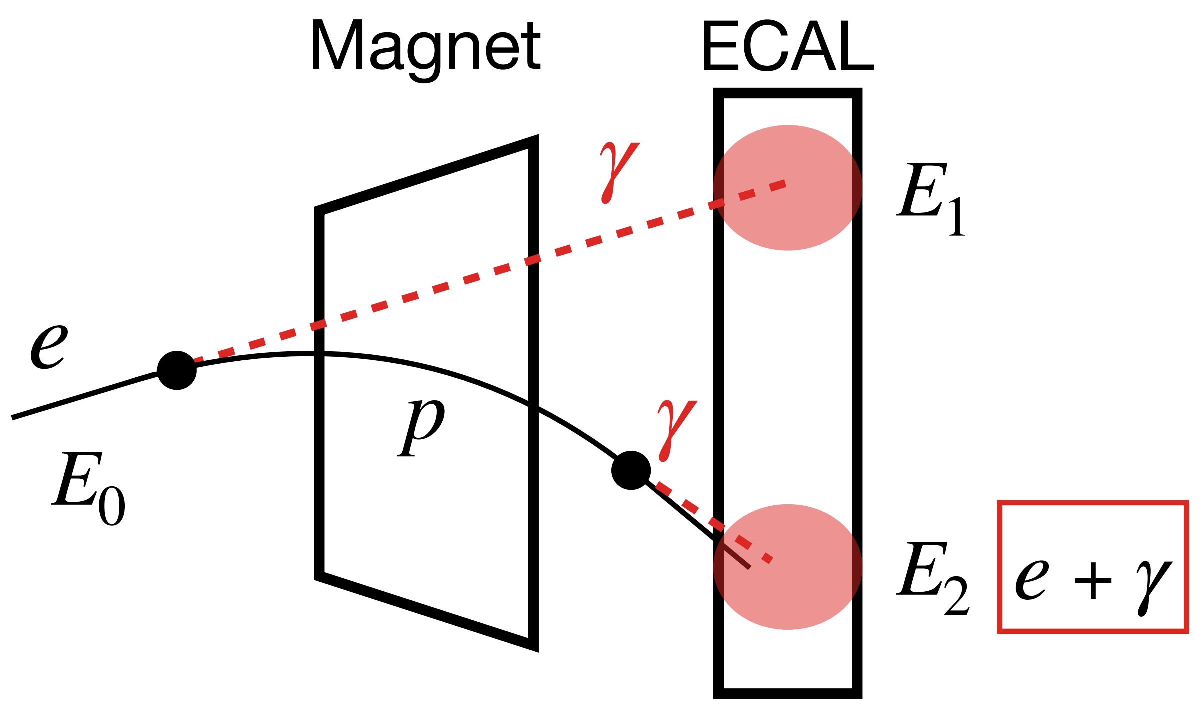
- In this [analysis], the misidentified electrons, in light green, represent an important background that is critical to control and measure accurately. So it is important to compute the efficiency of the electronID to use it in such measurements

Introduction

ElectronID and its difficulties

- ElectronID uses E_2/p, if it is close to 1 $\Longrightarrow$ electron
- The particular magnet-calorimeter setup of LHCb makes that electrons have to be matched to their Bremsstrahlung photons (see [Paloma's presentation]), so we have two types of electrons:
  - Brem photon found
  - Brem photon not found

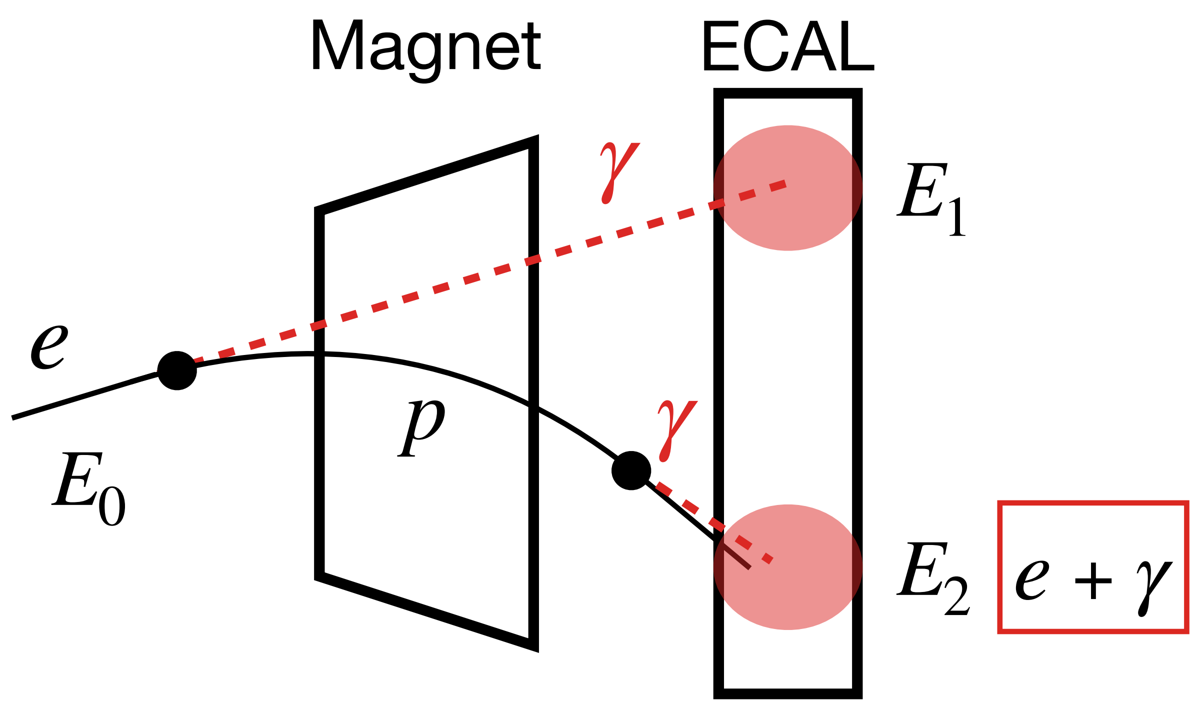
This info is also used in ElectronID!



Introduction

ElectronID and its difficulties

- ElectronID uses E_2/p, if it is close to 1 $\Longrightarrow$ electron



- The particular magnet-calorimeter setup of LHCb makes that electrons have to be matched to their Bremsstrahlung photons (see [Paloma's presentation]), so we have two types of electrons:
  - Brem photon found
  - Brem photon not found

This info is also used in ElectronID!

- The electron identification variables are constructed with the outputs of the electron reconstruction in the trigger system:
  - "PIDe": Δ log-likelihood function ($e-\pi$) using mostly brem, ECAL and RICH information. PIDe>0 $\Longrightarrow$ + likely to be an electron than a pion
  - "ProbNNe": output of Neural Network that also uses tracking information. Probability [0,1]

Particles with ↑ "PID_E" and "PROBNN_E" $\Longrightarrow$ + likely to be an electron

These variables are not reproduced perfectly in simulation so we will use a data-driven approach to evaluate their performance

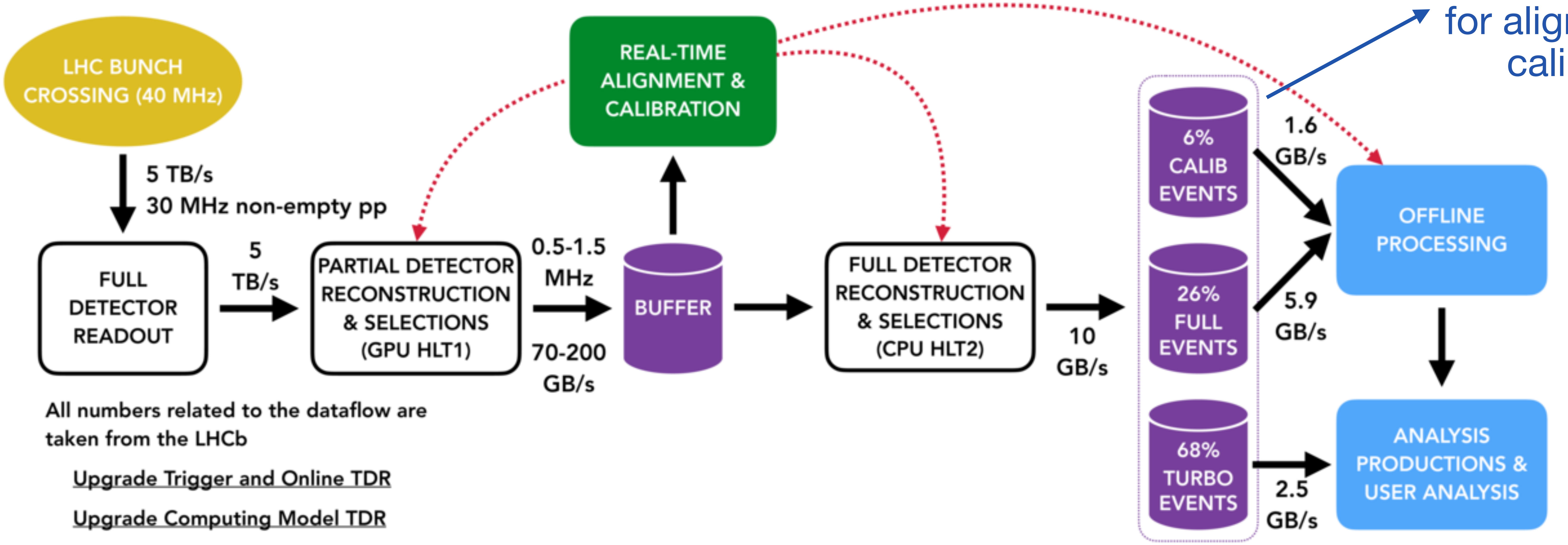
The trigger system

- The detector and the reconstruction have changed significantly for Run 3 (Upgrade I) to accommodate the higher multiplicity, so it is important to validate the electron identification performance with the early data and to provide calibration for analyses using this data. Especially considering that the ECAL itself was not upgraded

The trigger system

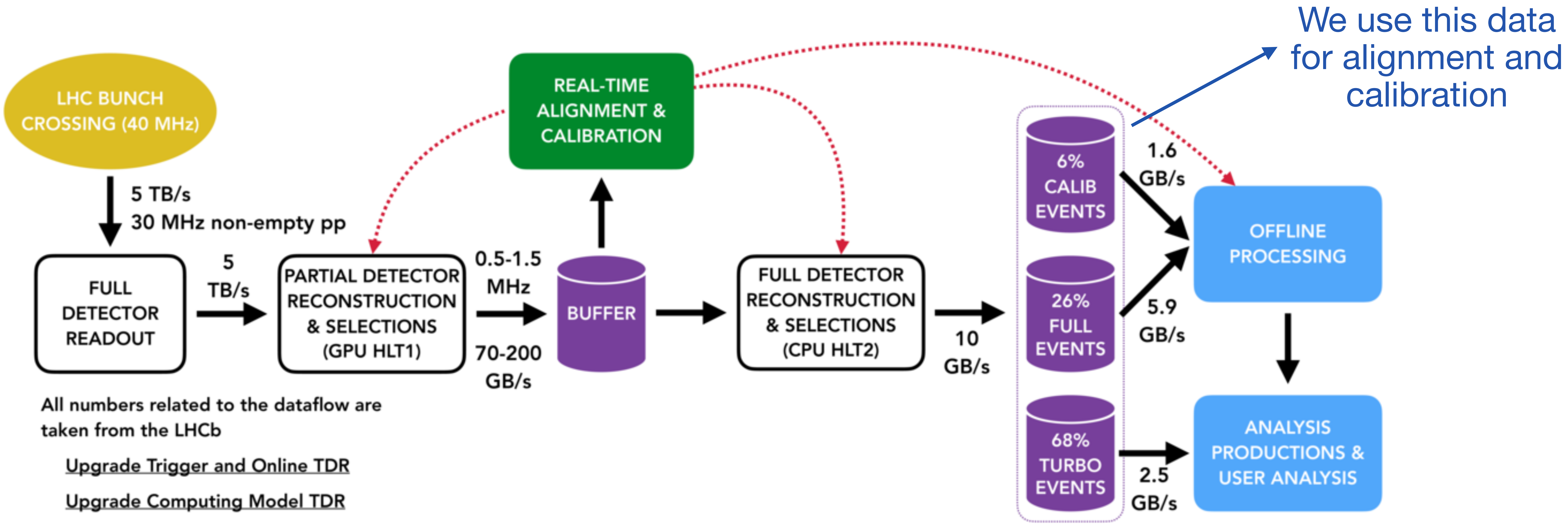
- The detector and the reconstruction have changed significantly for Run 3 (Upgrade I) to accommodate the higher multiplicity, so it is important to validate the electron identification performance with the early data and to provide calibration for analyses using this data. Especially considering that the ECAL itself was not upgraded

We use this data for alignment and calibration



The trigger system

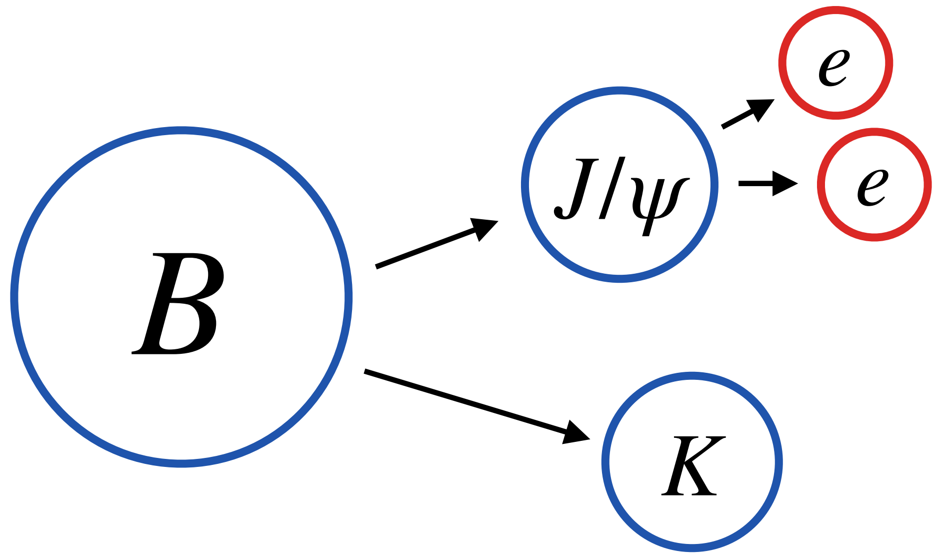
- The detector and the reconstruction have changed significantly for Run 3 (Upgrade I) to accommodate the higher multiplicity, so it is important to validate the electron identification performance with the early data and to provide calibration for analyses using this data. Especially considering that the ECAL itself was not upgraded



- We will present how we obtain the efficiencies of the identification and the misidentification of electrons using both “PIDe” and “ProbNNe” with 2024 data

Data used

- Decay channel: $B^+ \to J/\psi(\to e^+e^-)K^+$, largely studied, high yield and purity $\Longrightarrow$ allows efficiency study in momentum bins



- It is also important to evaluate the electron misidentification (misID). So we study when a pion is confused for an electron using the decay channel $D^{*+} \to D^0(\to K^-\pi^+)\pi^+$

- We use 2024 data ($\sim$2 fb^{-1}) and two MC simulations (for signal and background) with 2024 conditions to develop a selection and model the shape of the mass distributions

Efficiency calculation methodology

- Tag & Probe Method: One electron with PIDe>5 (tag) + the other without (probe), at the HLT2 level. Apply ElectronID>X to the probe $\Longrightarrow$ eff for X

Efficiency calculation methodology

- Tag & Probe Method: One electron with PIDe>5 (tag) + the other without (probe), at the HLT2 level. Apply ElectronID>X to the probe $\Longrightarrow$ eff for X
- Fit & Count Method:

Efficiency calculation methodology

- Tag & Probe Method: One electron with PIDe>5 (tag) + the other without (probe), at the HLT2 level. Apply ElectronID>X to the probe $\Longrightarrow$ eff for X
- Fit & Count Method:
  - PASS: ElectronID>X $\Longrightarrow$ mass fit
  - ALL: no ElectronID cut $\Longrightarrow$ 2nd mass fit

Efficiency calculation methodology

- Tag & Probe Method: One electron with PIDe>5 (tag) + the other without (probe), at the HLT2 level. Apply ElectronID>X to the probe $\Longrightarrow$ eff for X

- Fit & Count Method:
  - PASS: ElectronID>X $\Longrightarrow$ mass fit
  - ALL: no ElectronID cut $\Longrightarrow$ 2nd mass fit

  $\Longrightarrow$ Count yields to obtain the eff for X, using *ROOT TEfficiency* for the correlations

$$Eff = \frac{N^{PASS}_{signals}}{N^{ALL}_{signals}}$$

Efficiency calculation methodology

- Tag & Probe Method: One electron with PIDe>5 (tag) + the other without (probe), at the HLT2 level. Apply ElectronID>X to the probe $\Longrightarrow$ eff for X

- Fit & Count Method:

  - PASS: ElectronID>X $\Longrightarrow$ mass fit
  - ALL: no ElectronID cut $\Longrightarrow$ 2nd mass fit

  $\Longrightarrow$ Count yields to obtain the eff for X, using *ROOT TEfficiency* for the correlations

$$Eff = \frac{N^{PASS}_{signals}}{N^{ALL}_{signals}}$$

- We perform these fits in probe momentum bins

Selection strategy

- We select events using a preselection+BDT and then split into Bremsstrahlung categories:
  - 0brem: no electron with brem energy added back
  - 1brem_tag: the tag electron with brem energy added back
  - 1brem_probe: the probe electron with brem energy added back
  - 2brem: both electrons with brem energy added back

See [Paloma's presentation] for more info on brem recovery

Selection strategy

- We select events using a preselection+BDT and then split into Bremsstrahlung categories:
  - 0brem: no electron with brem energy added back
  - 1brem_tag: the tag electron with brem energy added back
  - 1brem_probe: the probe electron with brem energy added back
  - 2brem: both electrons with brem energy added back

See [Paloma's presentation] for more info on brem recovery

- For the preselection, we use B HLT1 trigger decisions:
  - One or two high-momentum tracks with a displaced vertex
  - We purposely avoid decisions that use ElectronID information

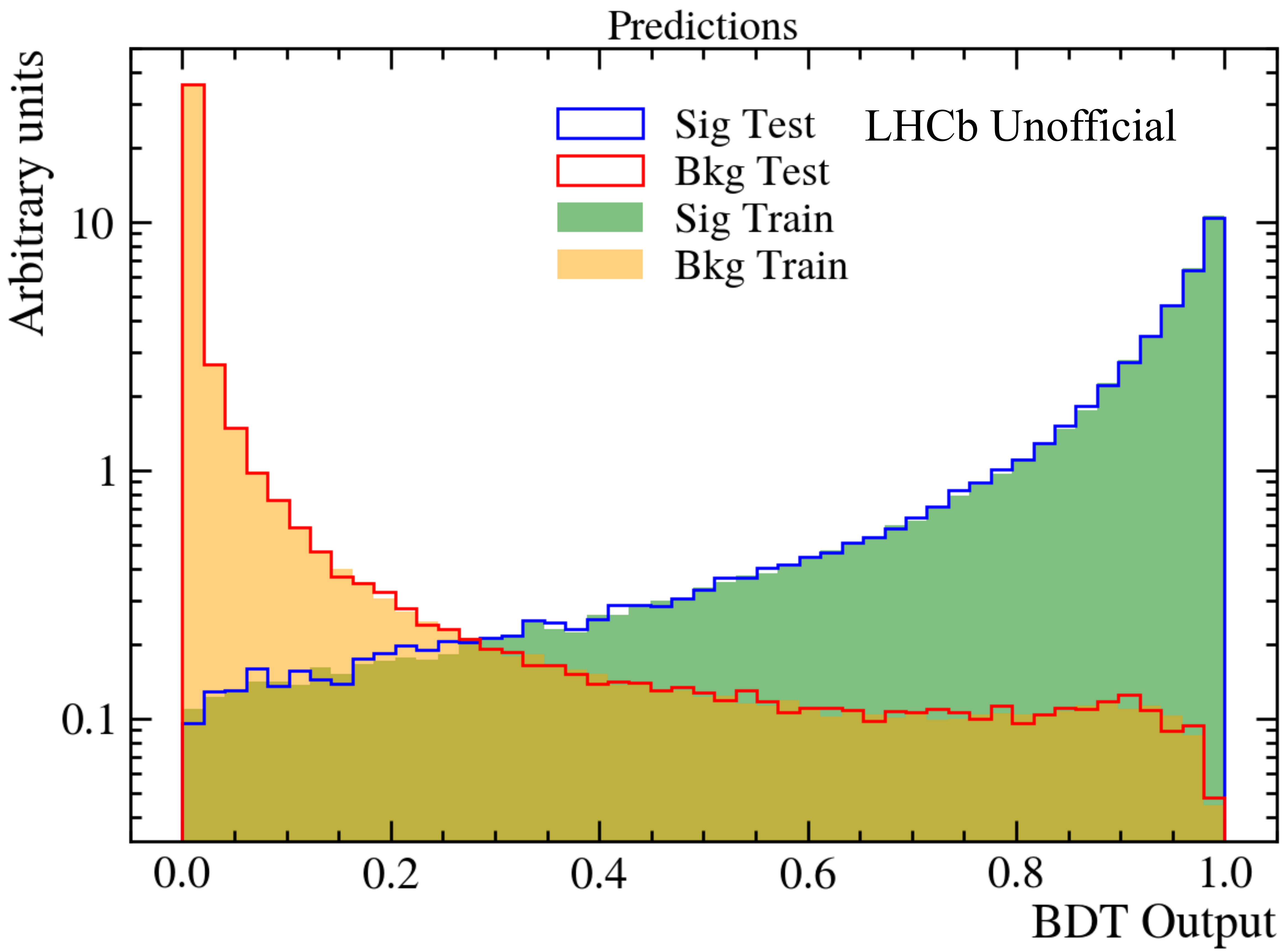
BDT setup

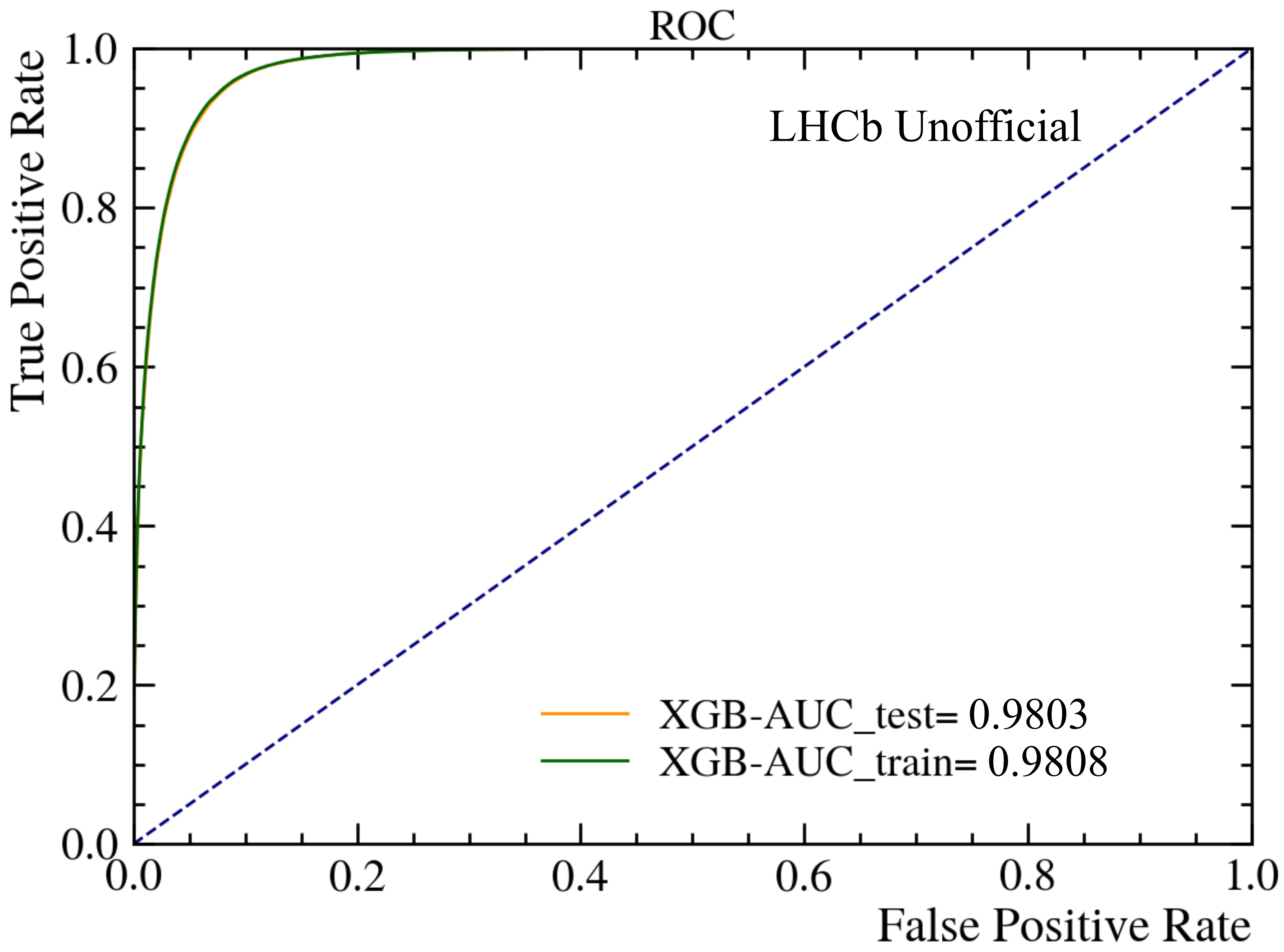
Using [XGBoost]

- We use the MC as a signal proxy and the data in the upper J/ψ mass sideband as a BKG proxy
- We use a 70-30% training-testing split
- Hyper-parameters used are in the Back-Up slides
- Training variables:
  - Kinematic and topological information of B, J/ψ and K selected for their discrimination power
  - We also avoid using ElectronID information to not bias the efficiency measurement

BDT performance

BDT outputs and ROC curve





No overtraining

BDT performance

Figure of Merit

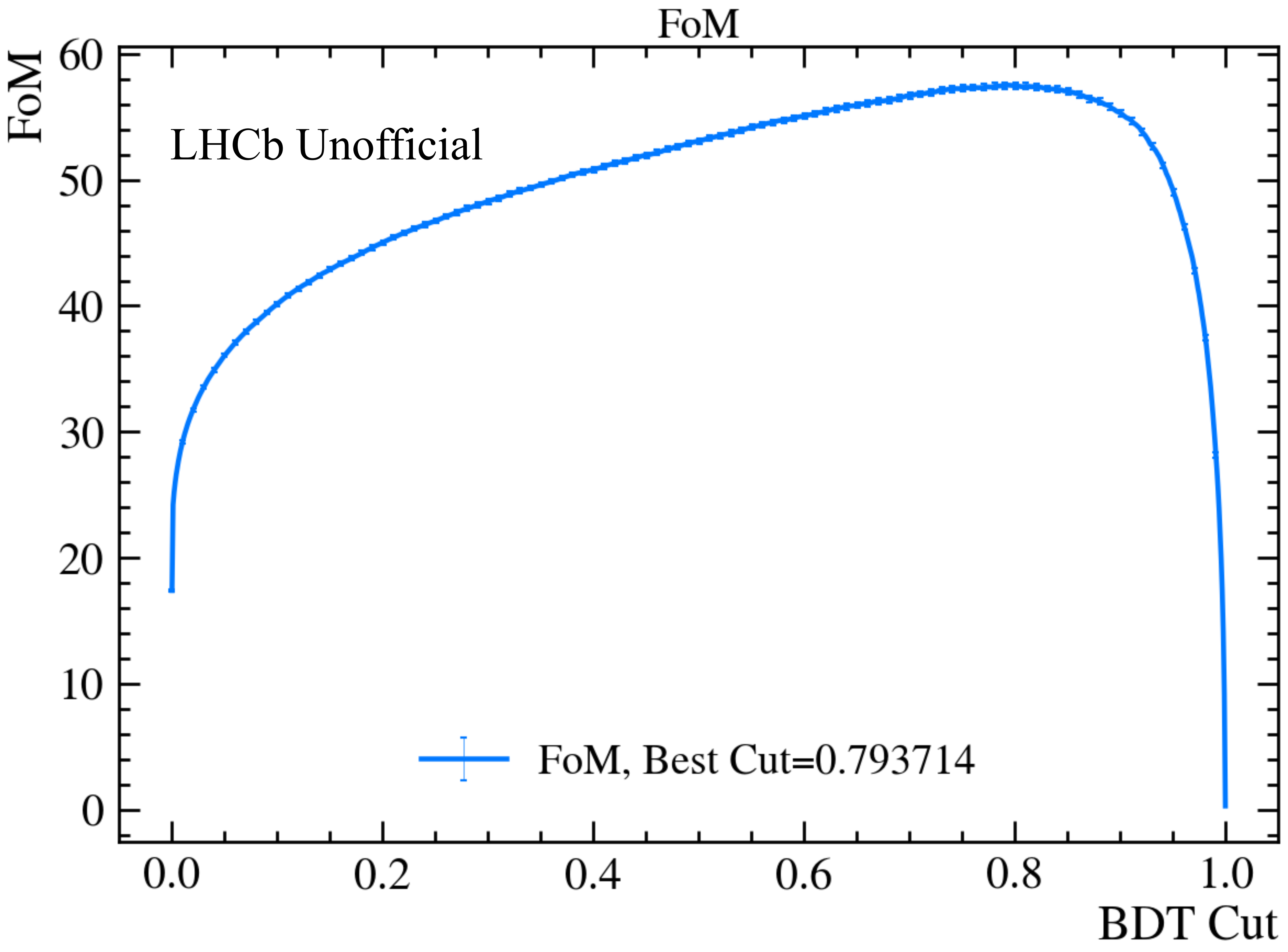
- To get the best cut of the BDT

Significance

$$FoM = \frac{aS}{\sqrt{aS + B}}$$

$a = 0.1$

S and B are the signal and background expected yields



MC efficiency: 73.37%
BKG rejection: 97.92%

Mass fit setup

Using *ROOT RooFit*

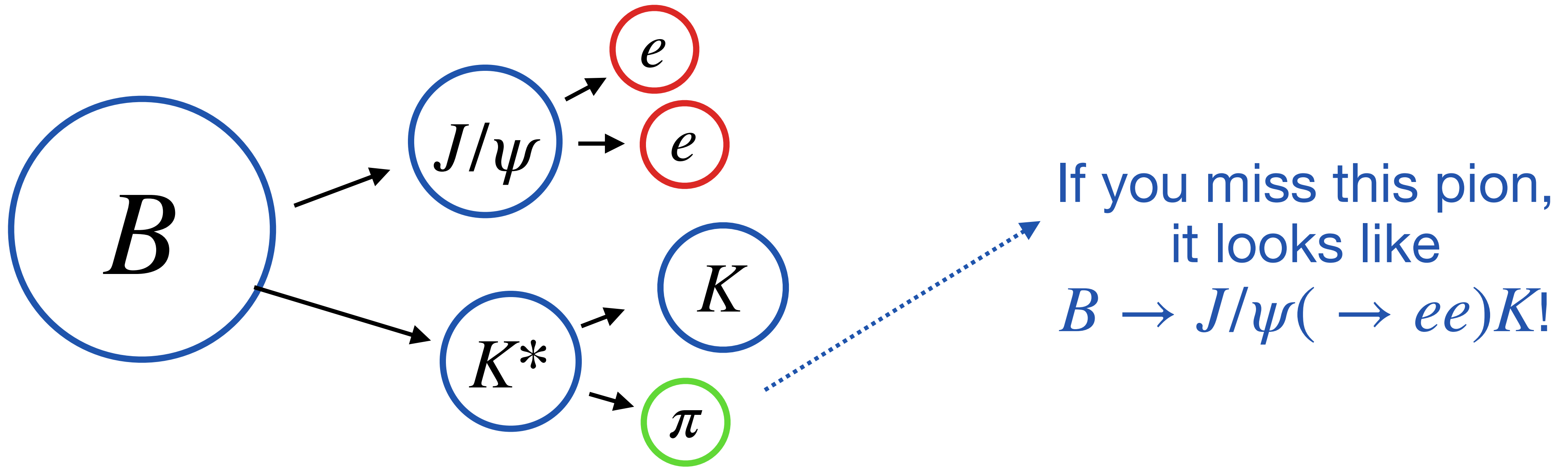
- Variable of choice: B mass with two constraints:
    - Primary Vertex constraint that forces the B reconstruction direction to point to the PV
    - J/ψ mass constraint that $\sim$ mass of the di-electrons = mass of J/ψ

Mass fit setup

Using *ROOT RooFit*

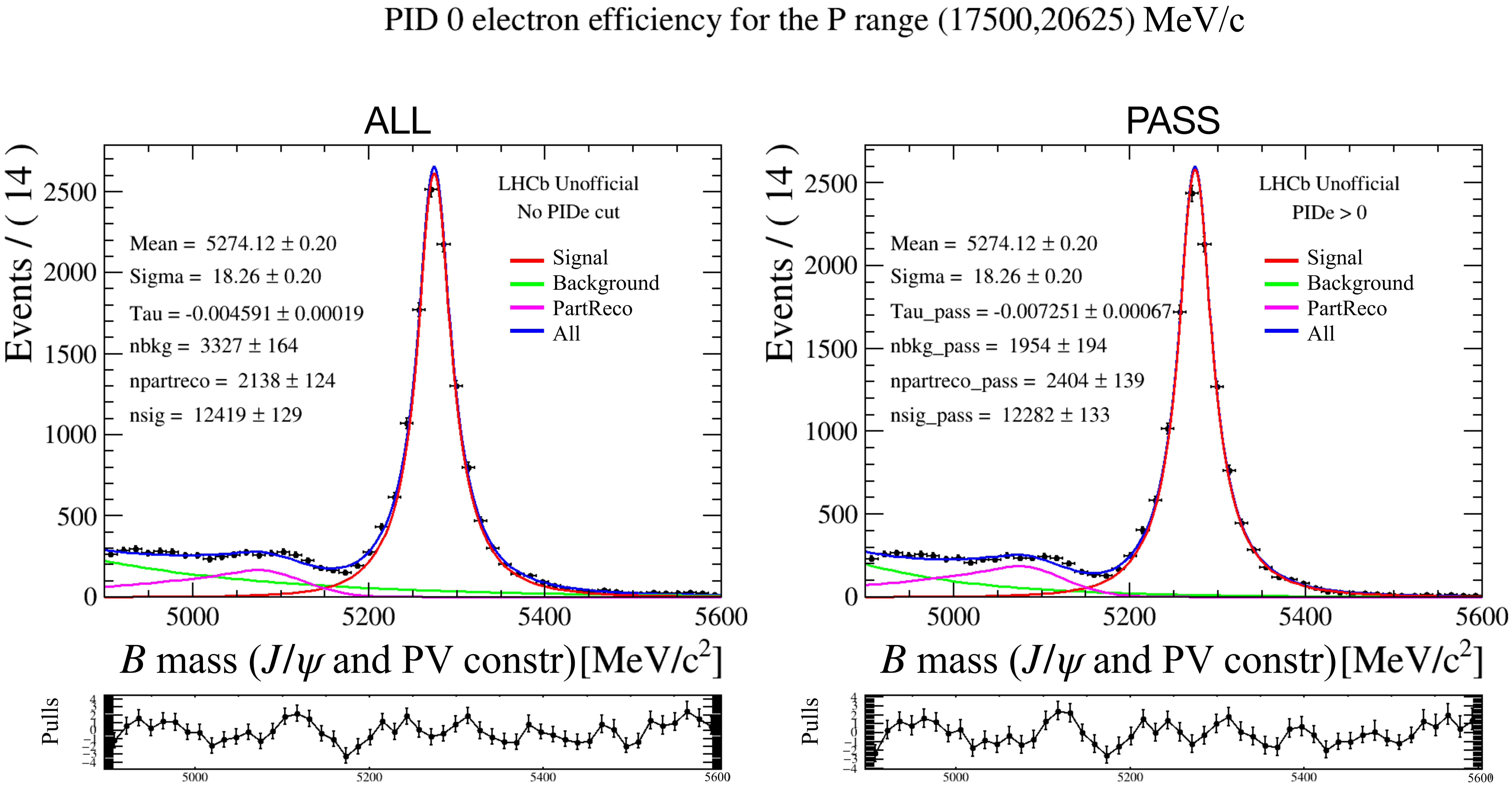
- Variable of choice: B mass with two constraints:
  - Primary Vertex constraint that forces the B reconstruction direction to point to the PV
  - J/ψ mass constraint that $\sim$ mass of the di-electrons = mass of J/ψ
- PDFs used:
  - Signal: Double-sided Crystal Ball with the tails fixed by a fit to the MC and a simultaneous fit to the PASS and ALL sharing the mean and σ
  - Partially Reconstructed Background: Double-sided Crystal Ball fixed by a fit to the MC
  - Combinatorial Background: Exponential

This is when you miss a particle in the reconstruction. We see $B^0 \to J/\psi(\to e^+e^-)K^{*0}$ $\Longrightarrow$ small peak on the left side of the real B mass value



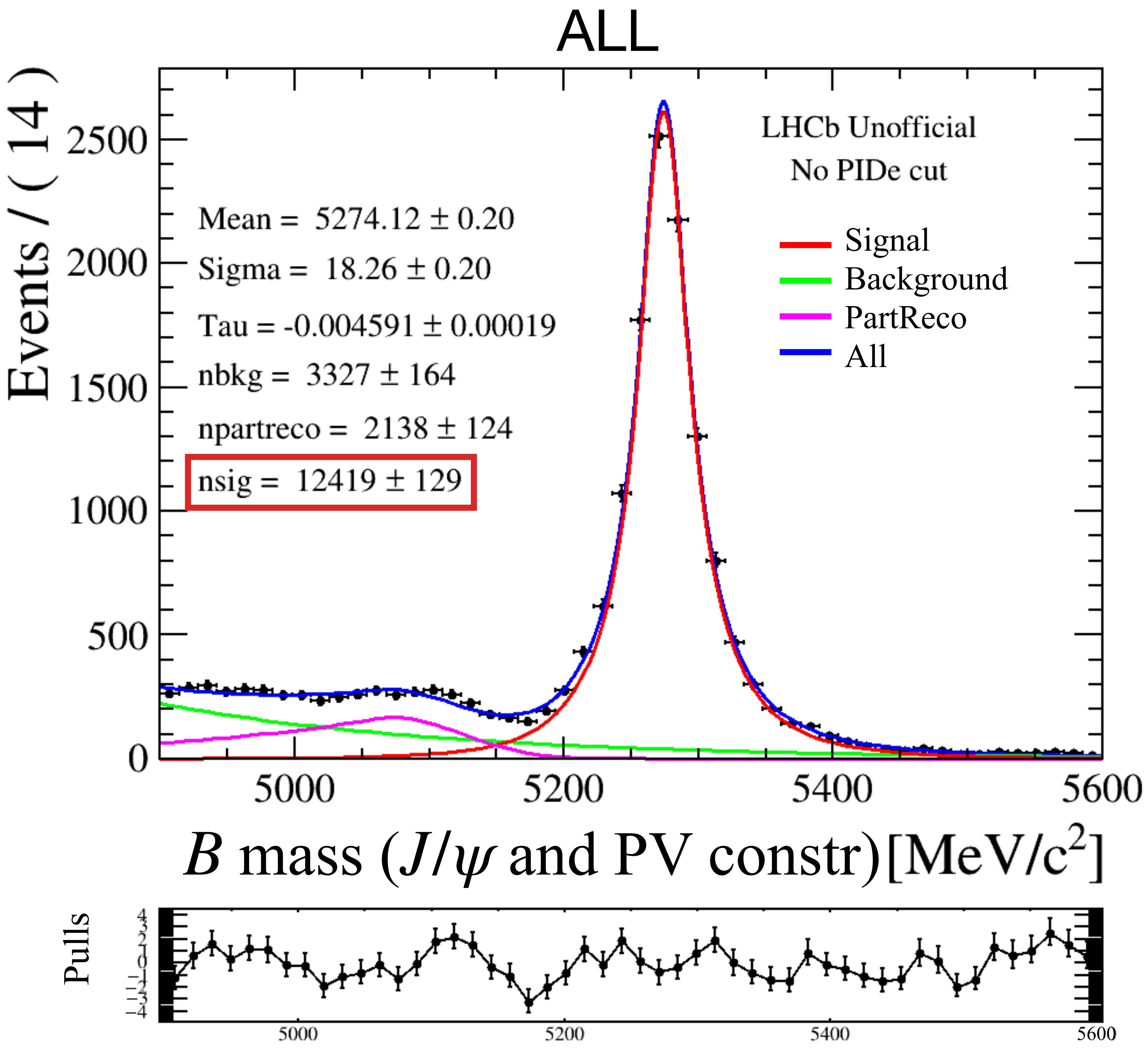
PIDe fit example

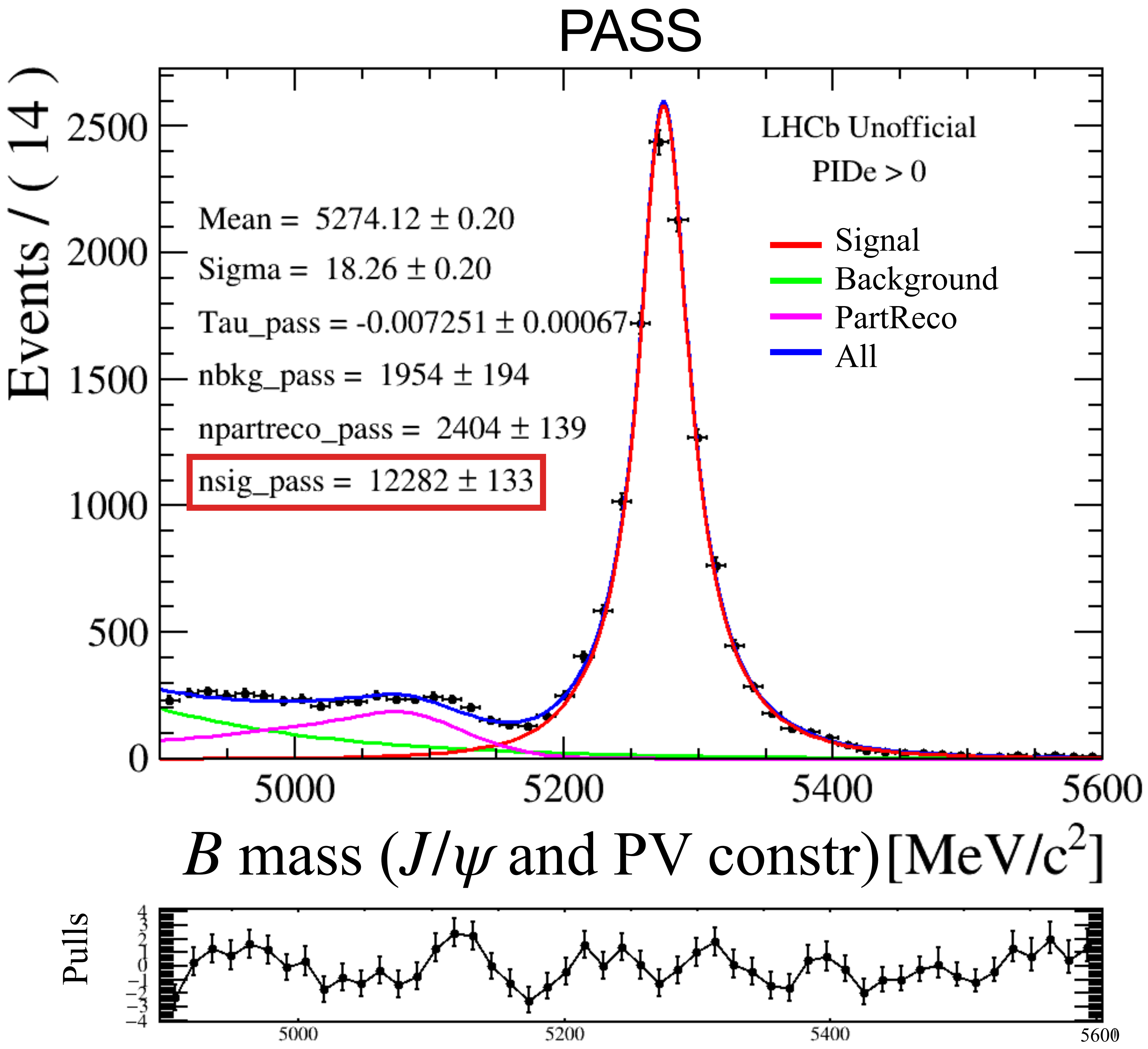
2brem category



2brem category

PID 0 electron efficiency for the P range (17500,20625) MeV/c



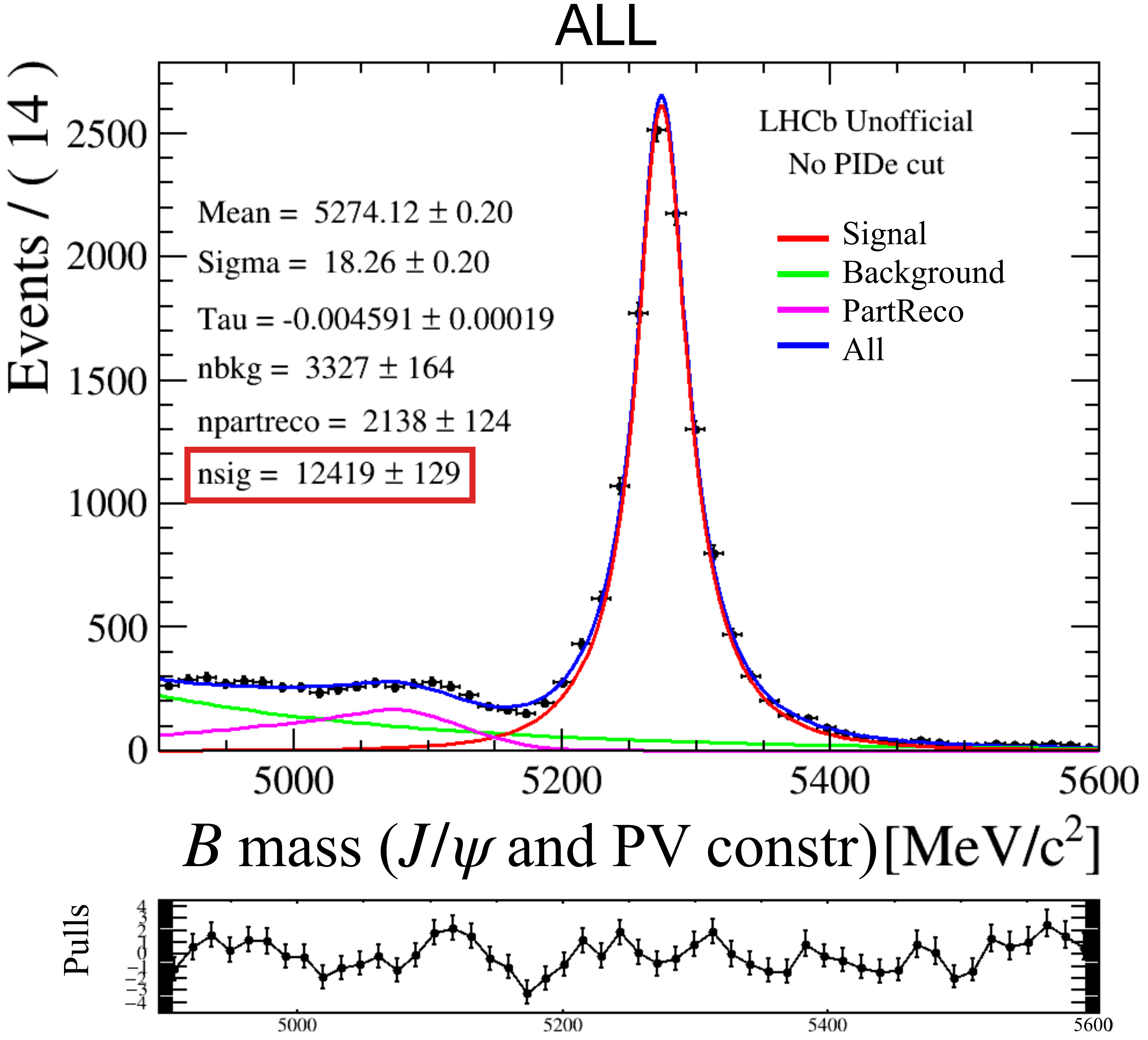


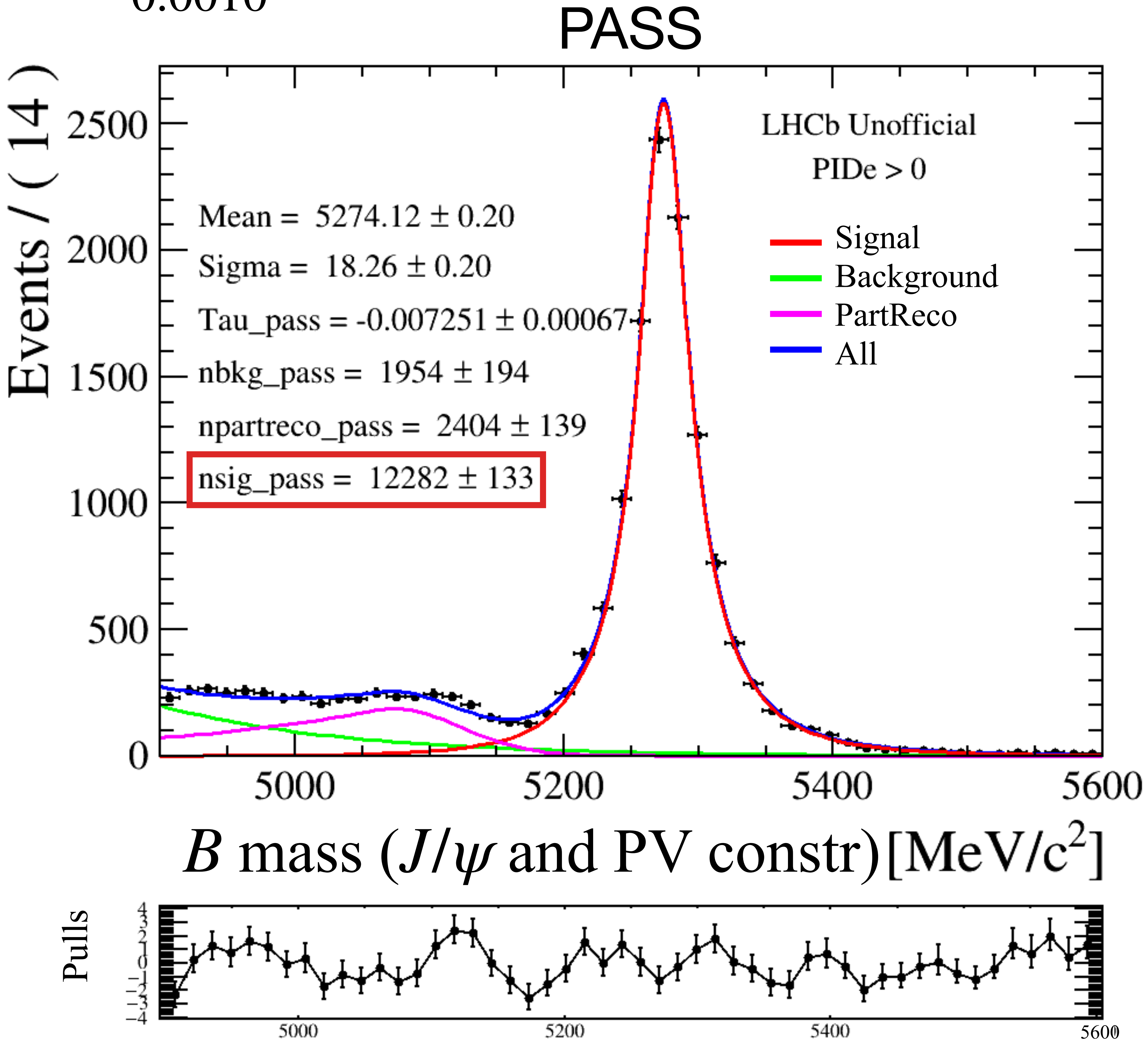
PIDe fit example

2brem category

PID 0 electron efficiency for the P range (17500,20625) MeV/c

$$Eff = 0.9890^{+0.0009}_{-0.0010}$$





MisID

- 2024 data of $D^{*+} \to D^0(\to K^-\pi^+)\pi^+$ where the π^+ from the D^0 (probe) is misID as an electron

- Mass fit setup, using the D^0 mass variable:
  - Signal: Gaussian with a simultaneous fit to the PASS and ALL sharing the mean and σ
  - Background: Exponential

- MisID probably overestimated

Selection used (inspired by studies of pion identification performance):

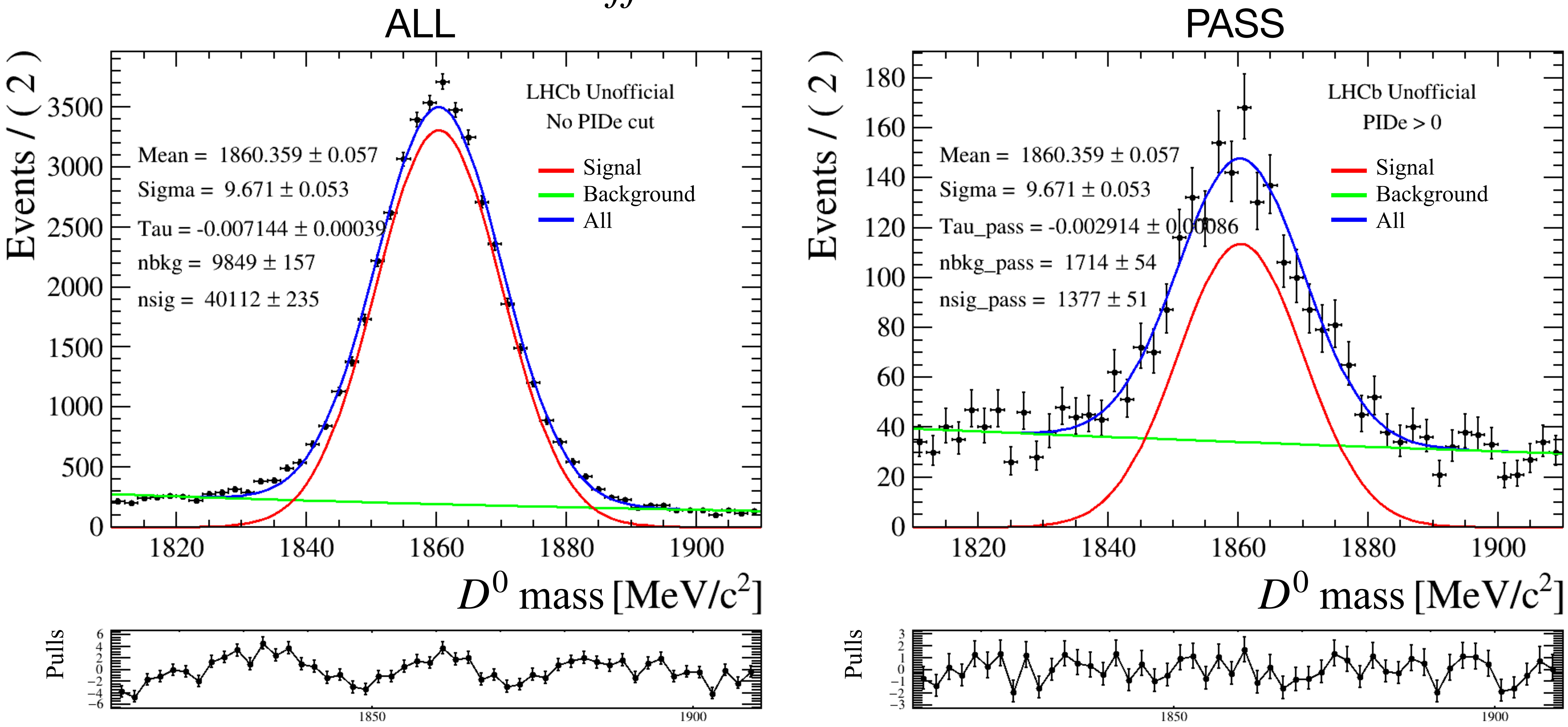
- $(D^{*+}$ - $D^0)$ mass window
- Remove $D^0 \to KK$

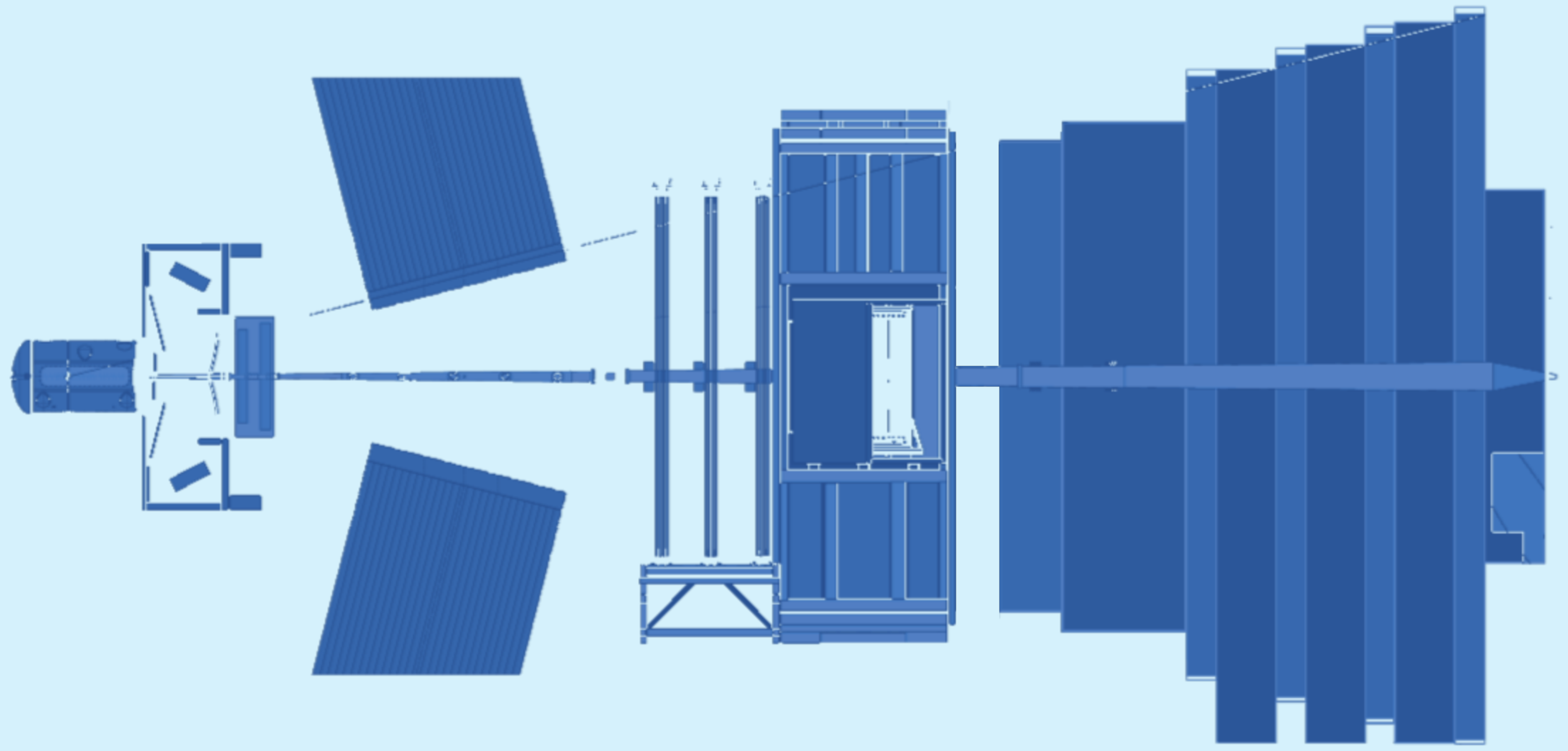
PIDe misID fit example

2brem category

PID 0 electron efficiency for the P range (17500,20625) MeV/c

$$Eff = 0.0343 \pm 0.0009$$

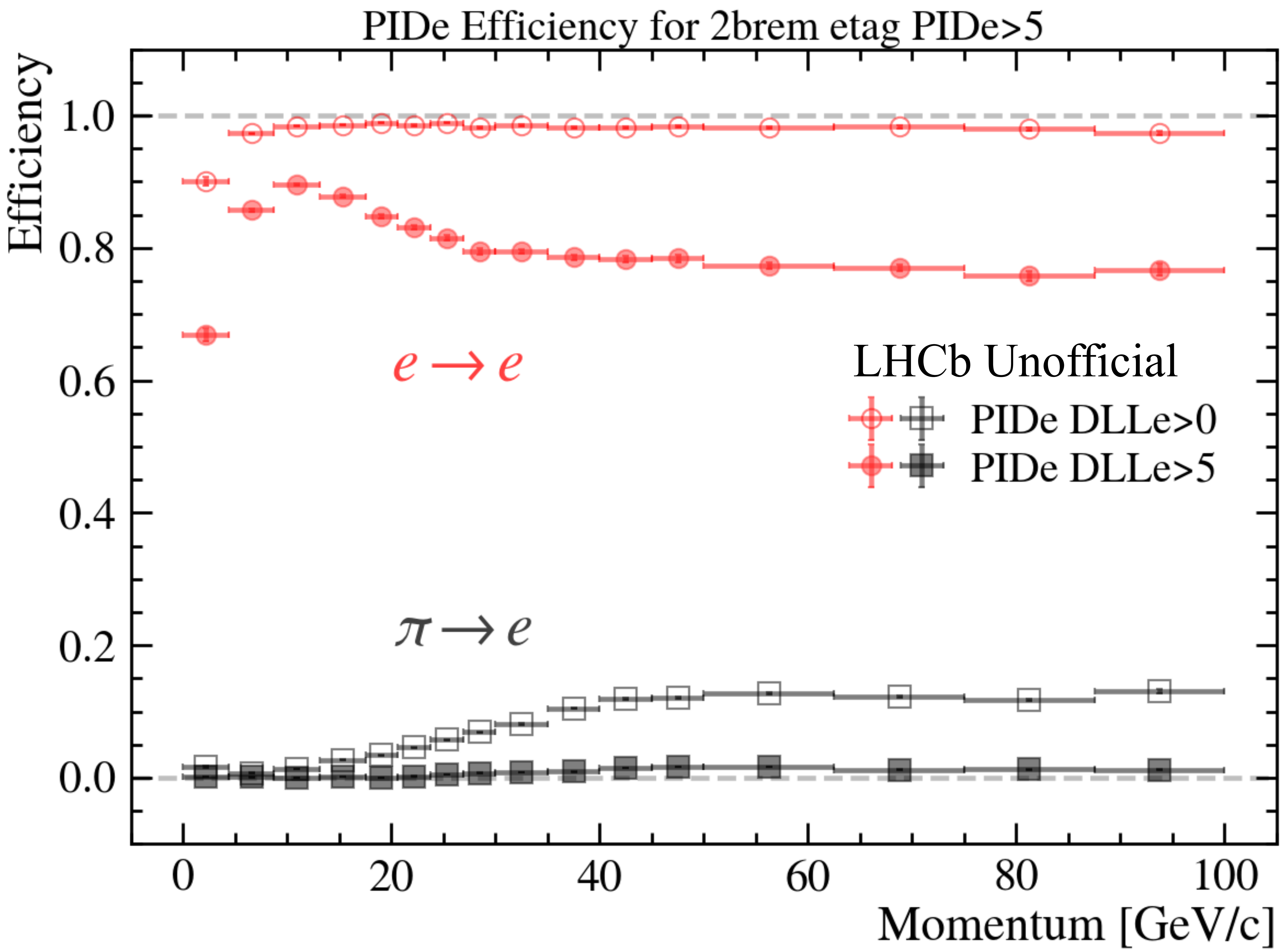




PIDe and ProbNNe 2024 performance results

PIDe 2024 Performance

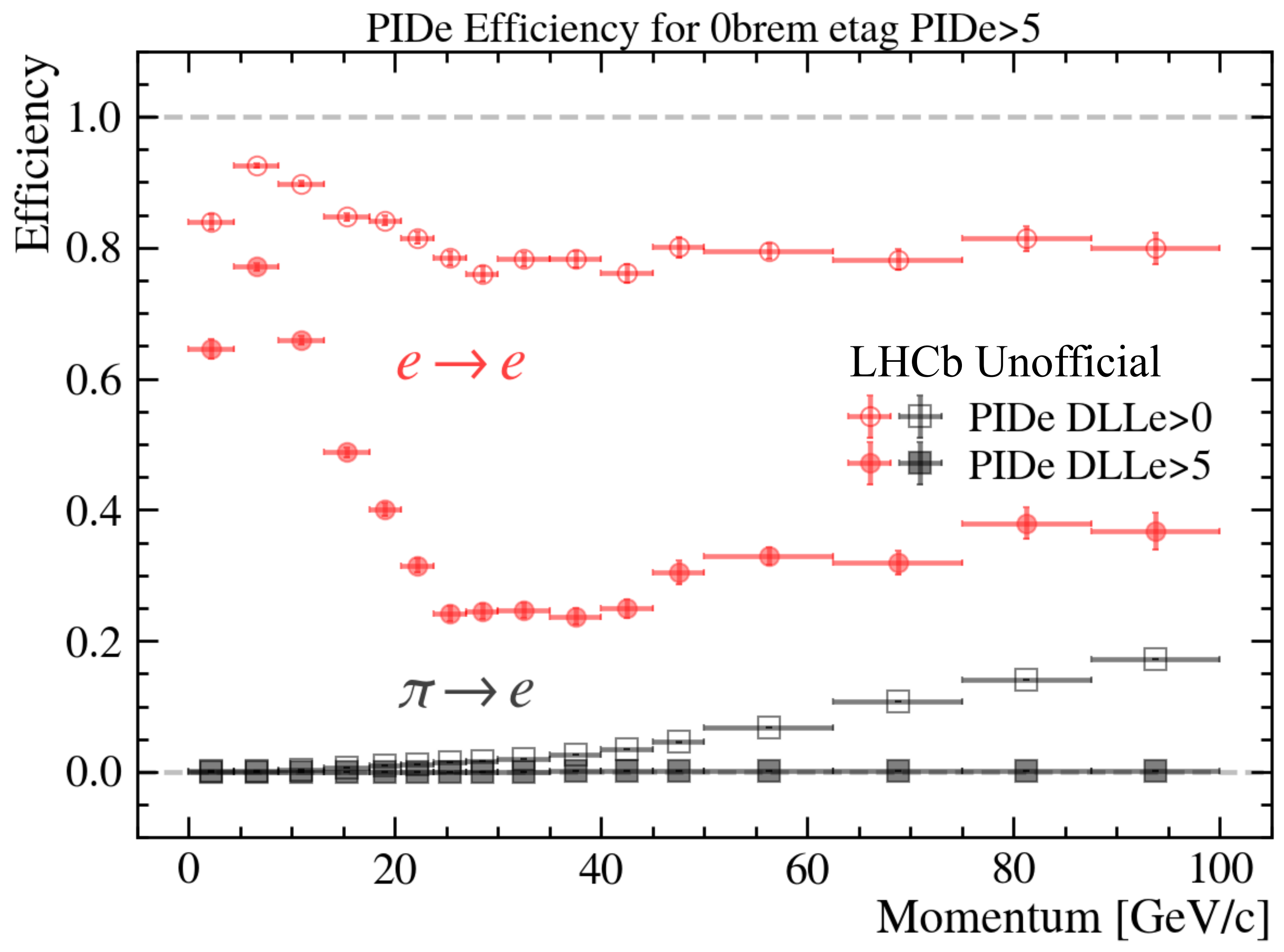
In probe momentum bins, 2brem



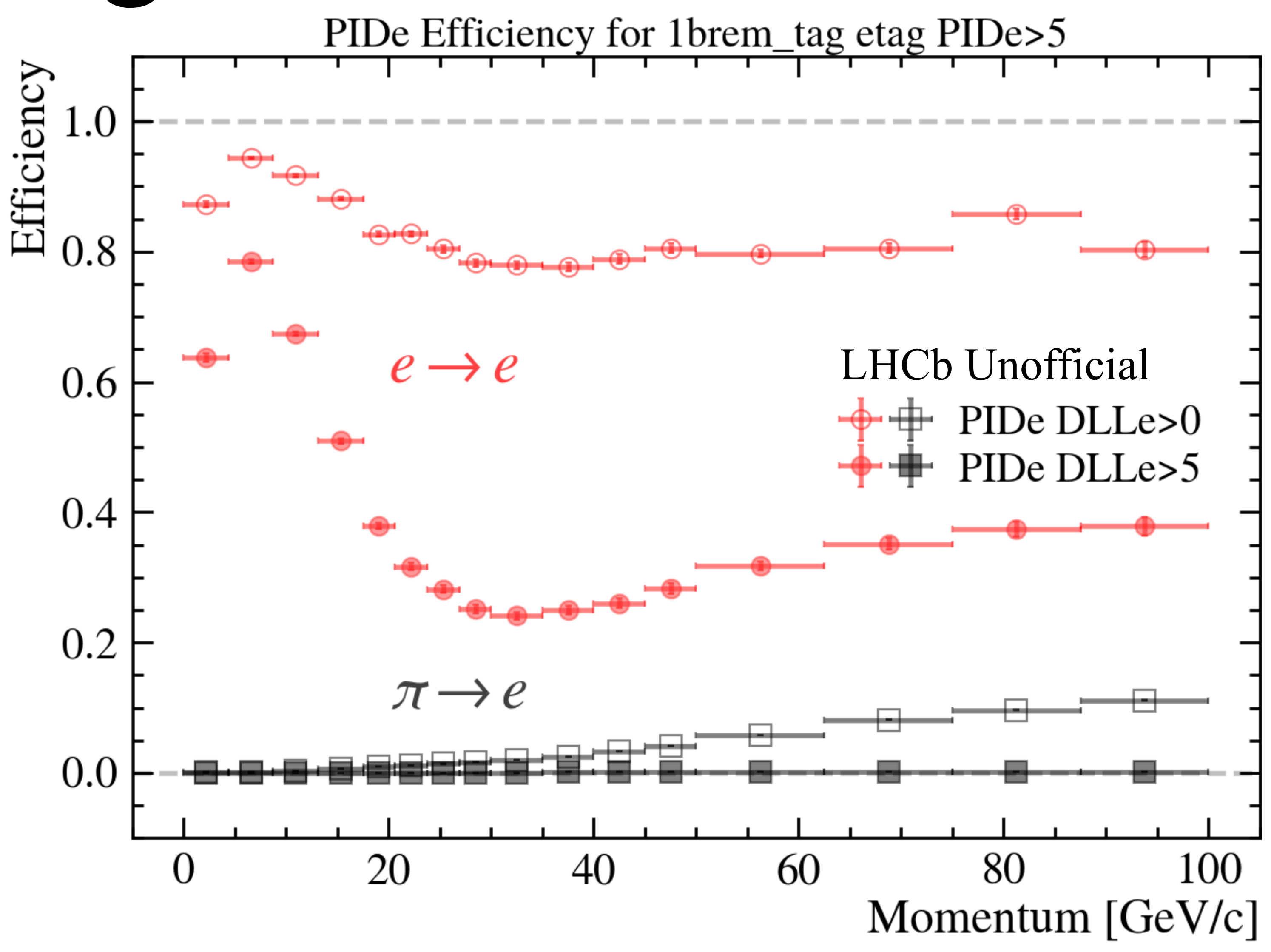
PIDe 2024 Performance

In probe momentum bins, all brem categories

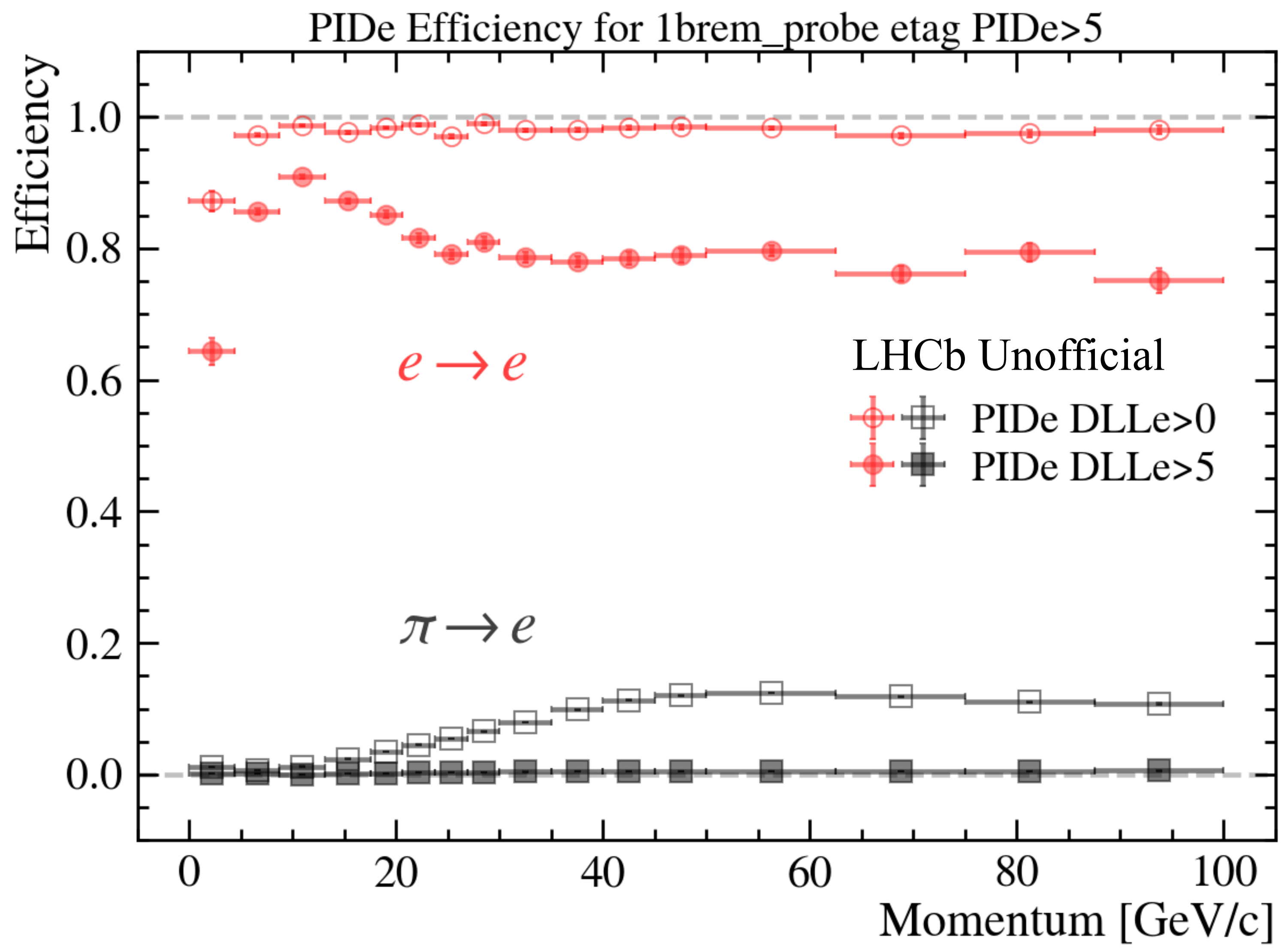
0brem



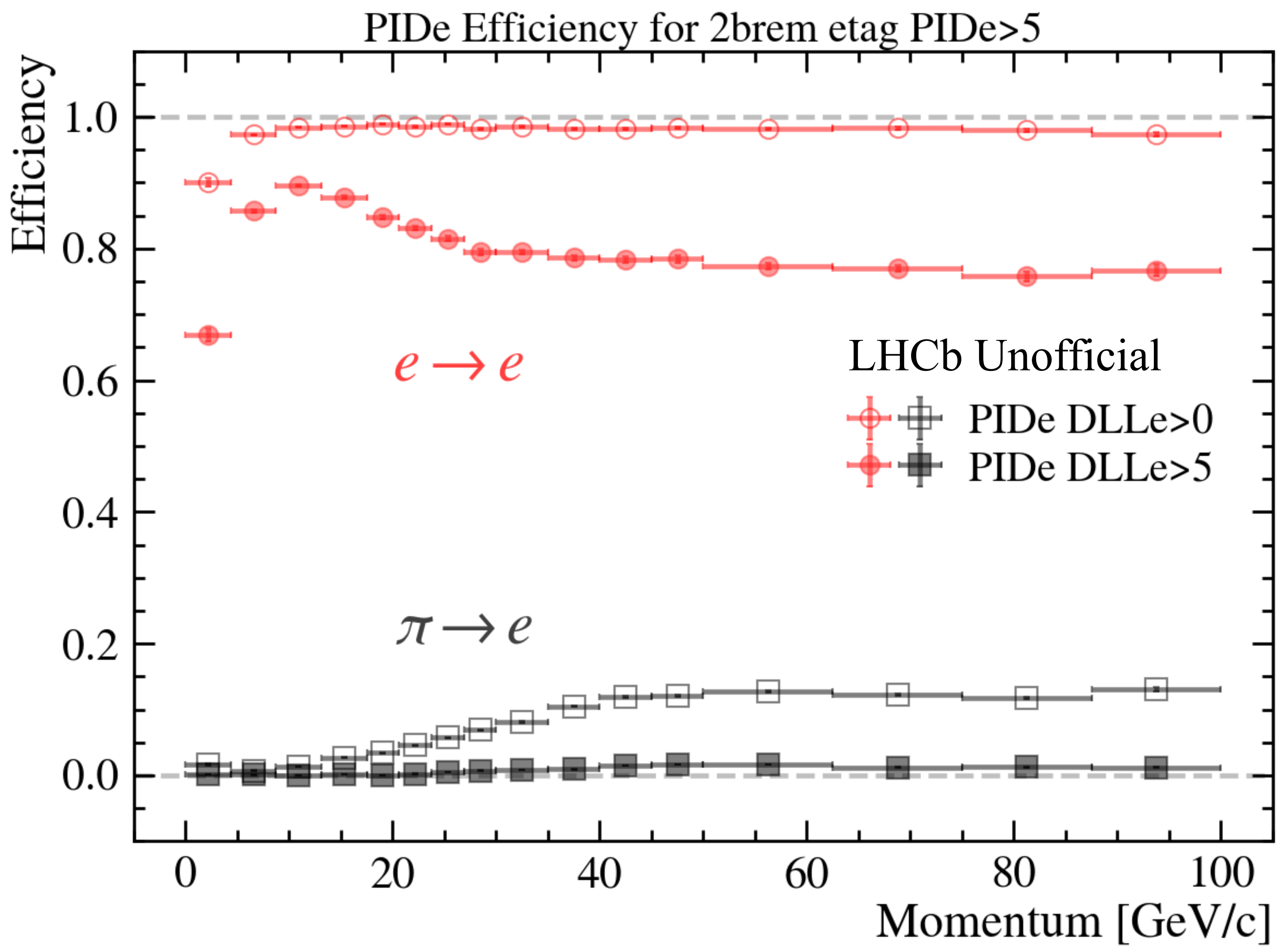
1brem_tag



1brem_probe



2brem

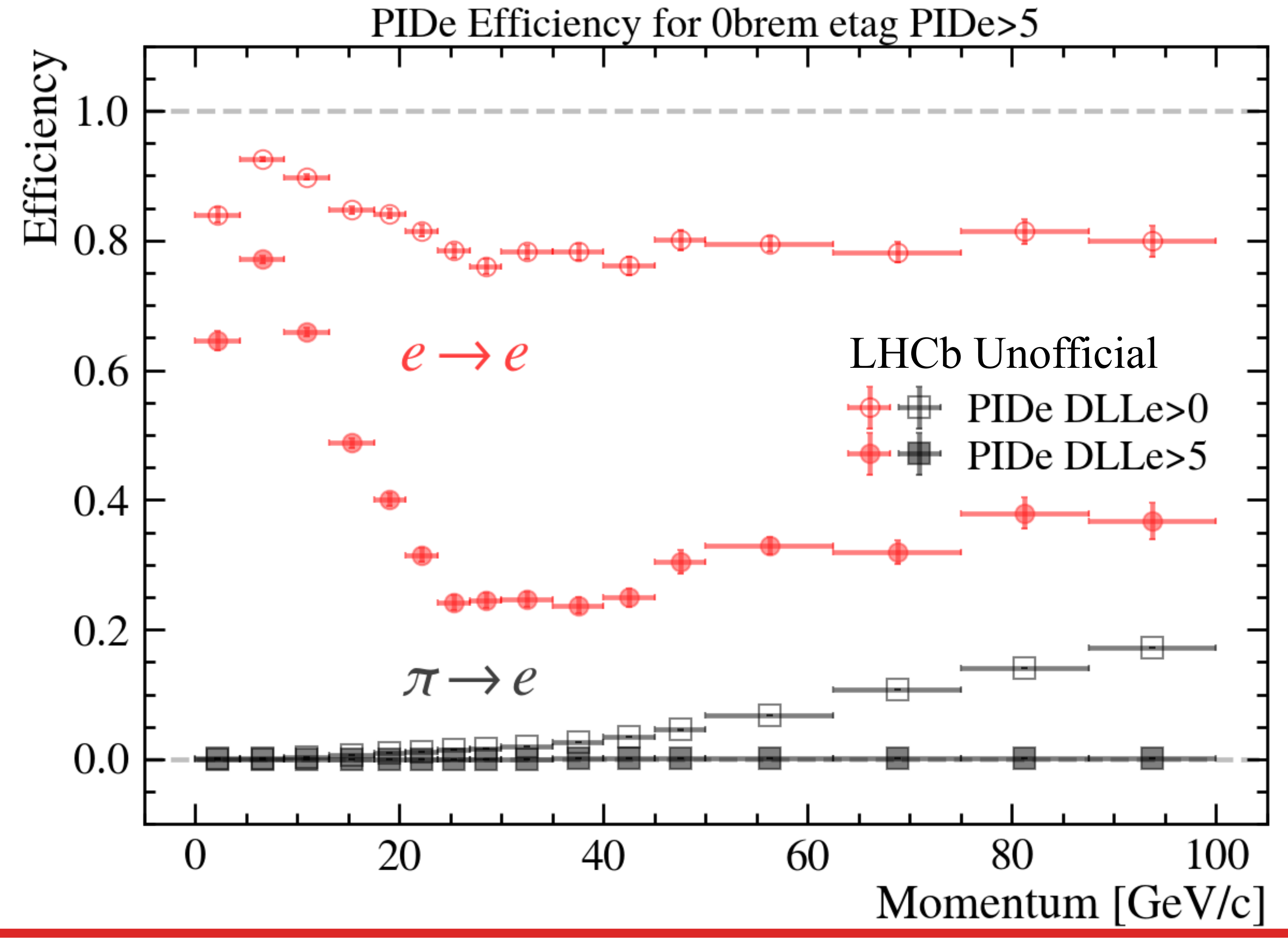


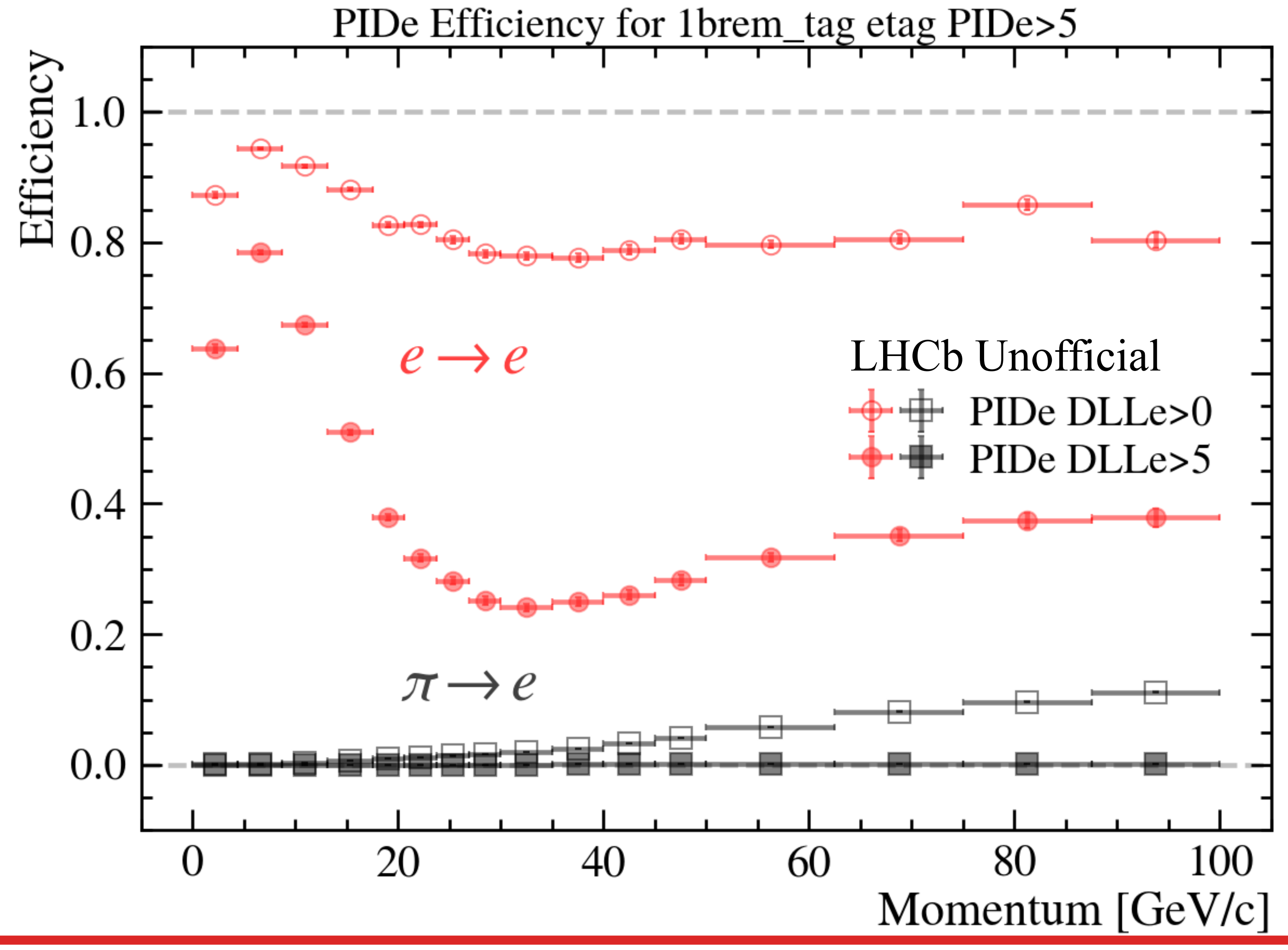
PIDe 2024 Performance

In probe momentum bins, all brem categories

No brem on probe

0brem

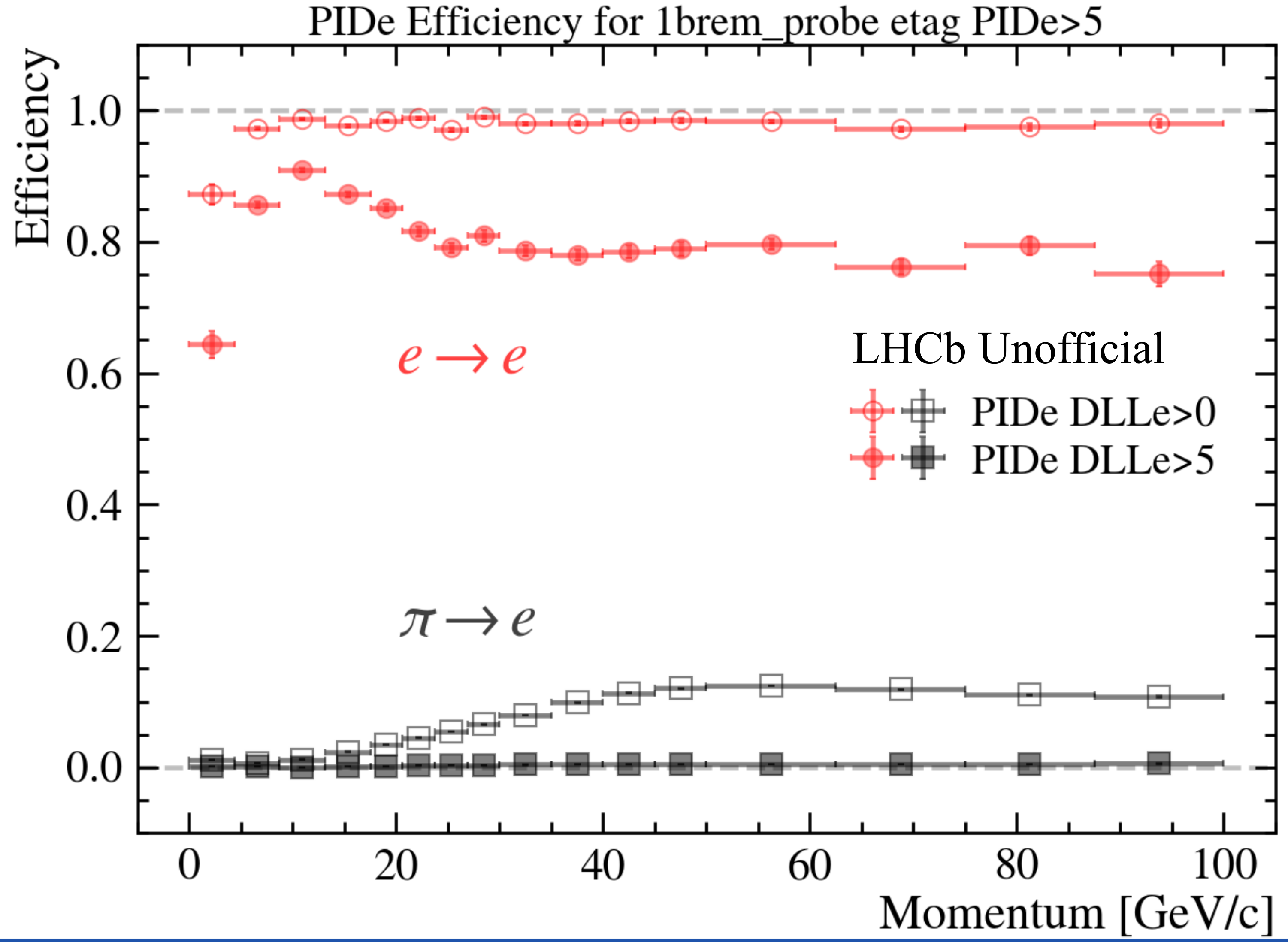


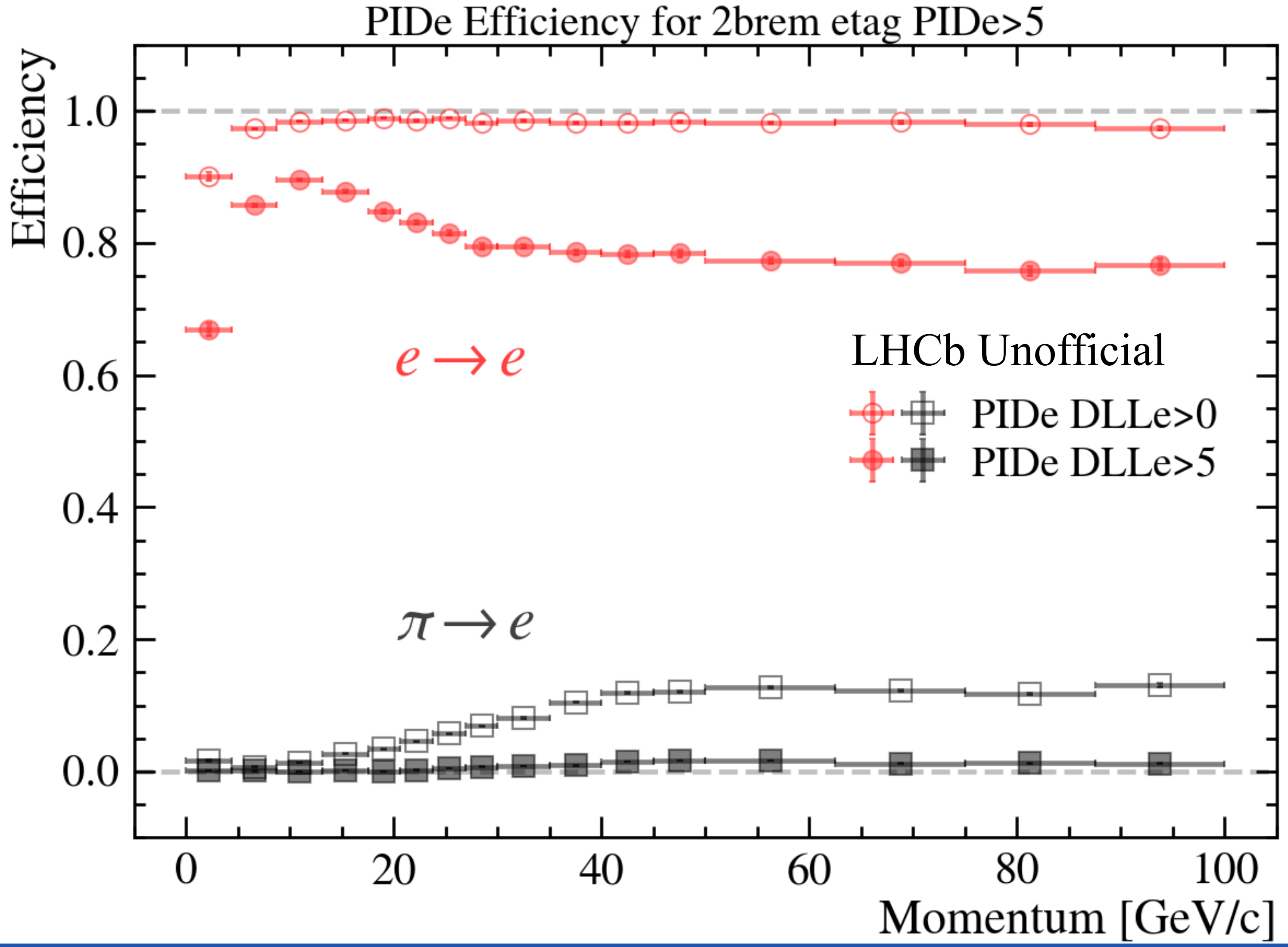


1brem_tag

Brem on probe

1brem_probe

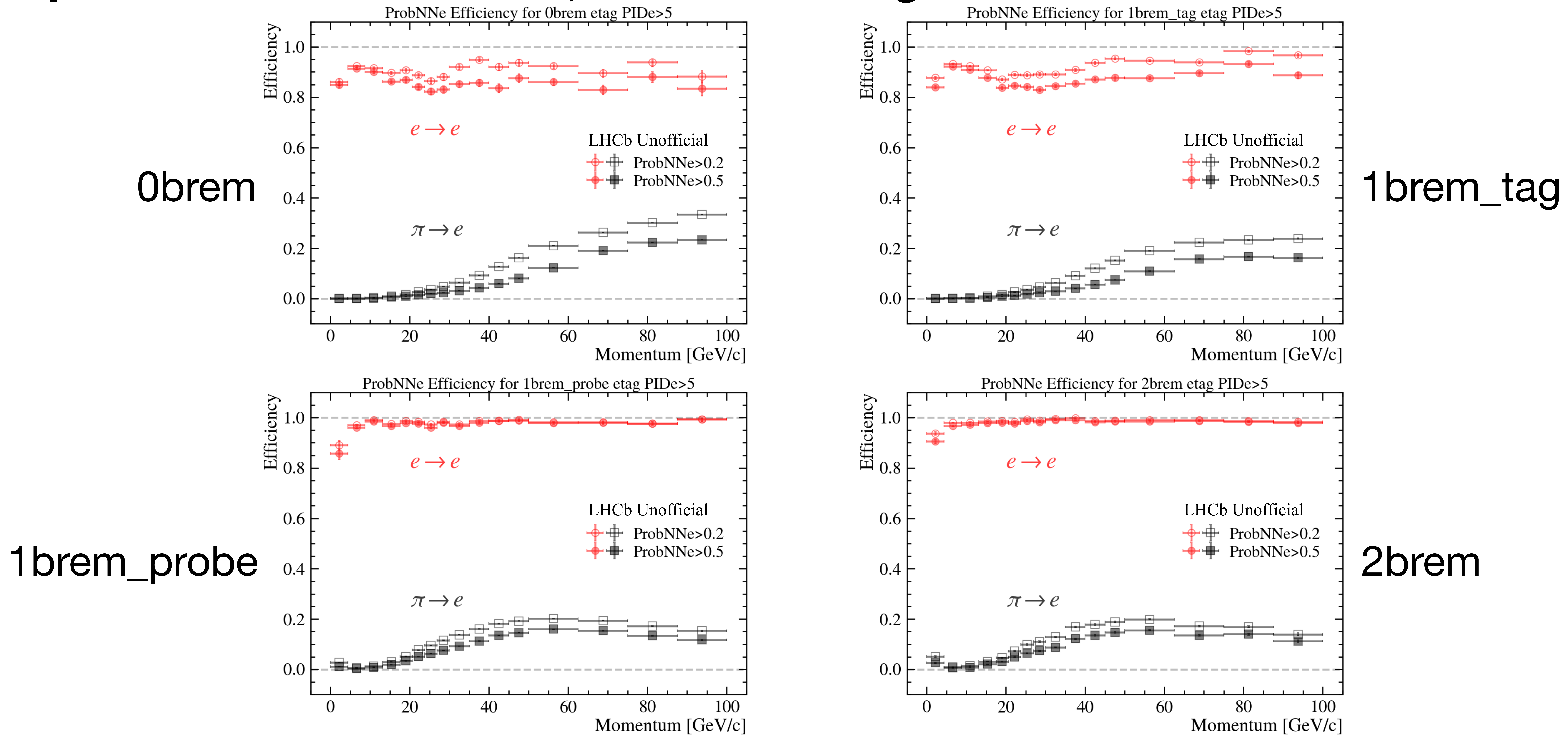




2brem

ProbNNe 2024 Performance

In probe momentum bins, all brem categories



Summary and conclusions

- We use the Tag & Probe and the Fit & Count methods to compute the efficiency of both the ID and misID of electrons in LHCb Upgrade I

- For signal selection, both linear cuts and a BDT are used and they offer a great reduction of the combinatorial background

- The final efficiency is computed with the yields of a simultaneous mass fit to the PASS and ALL samples

Summary and conclusions

- We use the Tag & Probe and the Fit & Count methods to compute the efficiency of both the ID and misID of electrons in LHCb Upgrade I
- For signal selection, both linear cuts and a BDT are used and they offer a great reduction of the combinatorial background
- The final efficiency is computed with the yields of a simultaneous mass fit to the PASS and ALL samples

- Both PIDe and ProbNNe are working fine for 2024
- For future work, we want to polish these results and expand them to all data-taking periods of 2024 so analysts will be able to use them to calibrate the efficiency of their analysis

Thank you for your attention!

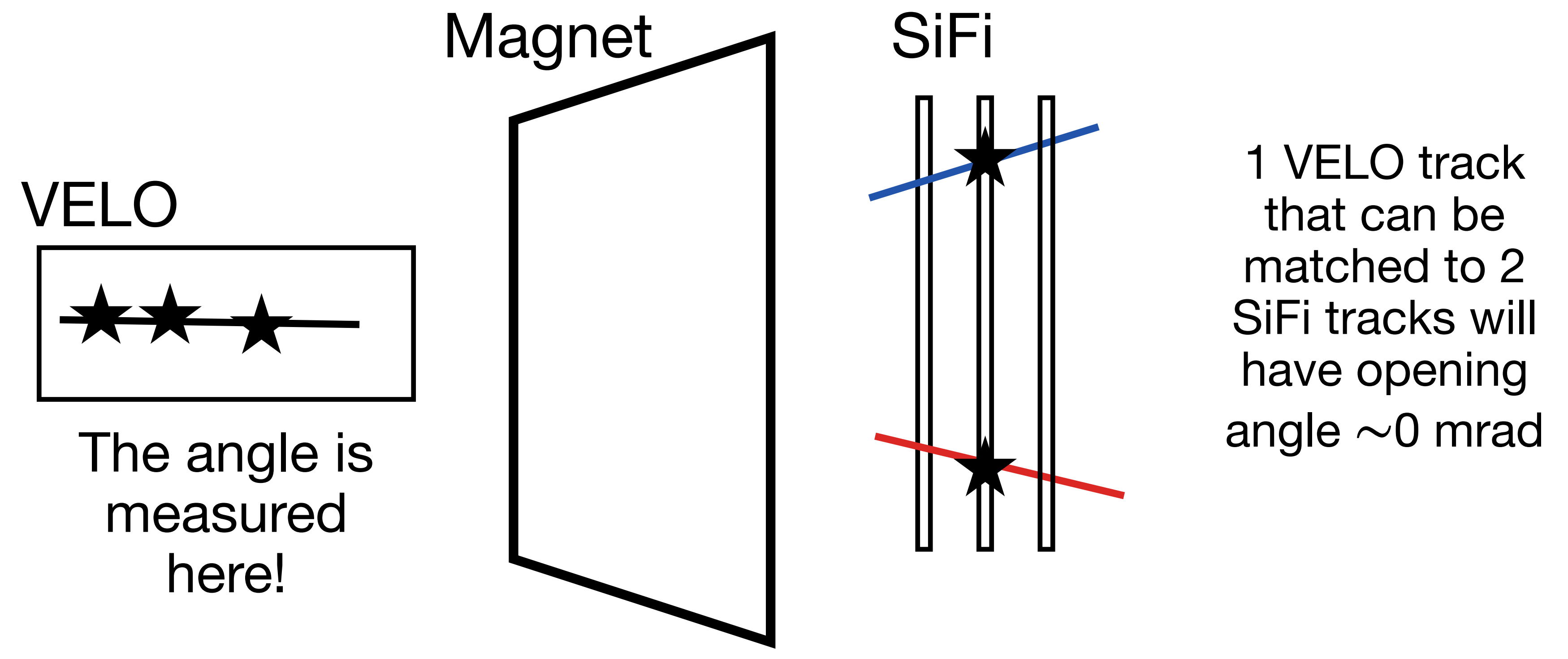
Back-Up

Preselection

Linear cuts

- Common cuts for MC and Data:
  - $\theta(e^-, e^+) > 0.5$ mrad
  - $\theta(e^-, K) > 0.5$ mrad
  - $\theta(e^+, K) > 0.5$ mrad

To avoid clone tracks



- B HLT1 trigger decisions:
  - One or two high-momentum tracks with a displaced vertex
  - We purposely avoid decisions that use ElectronID information

- Cuts to MC only:
  - Match reconstructed B candidate to the true signal
- Cuts to Data, so we can use it as a BKG proxy for the BDT training:
  - J/ψ Mass > 3200 MeV

BDT hyper-parameters

Same as in 2023

Parameter	Best value
Learning rate	0.2
Minimum loss	0.25
Maximum depth	2
Minimum child weight	0
Maximum delta step	0
Subsample	0.75
Number of trees	275

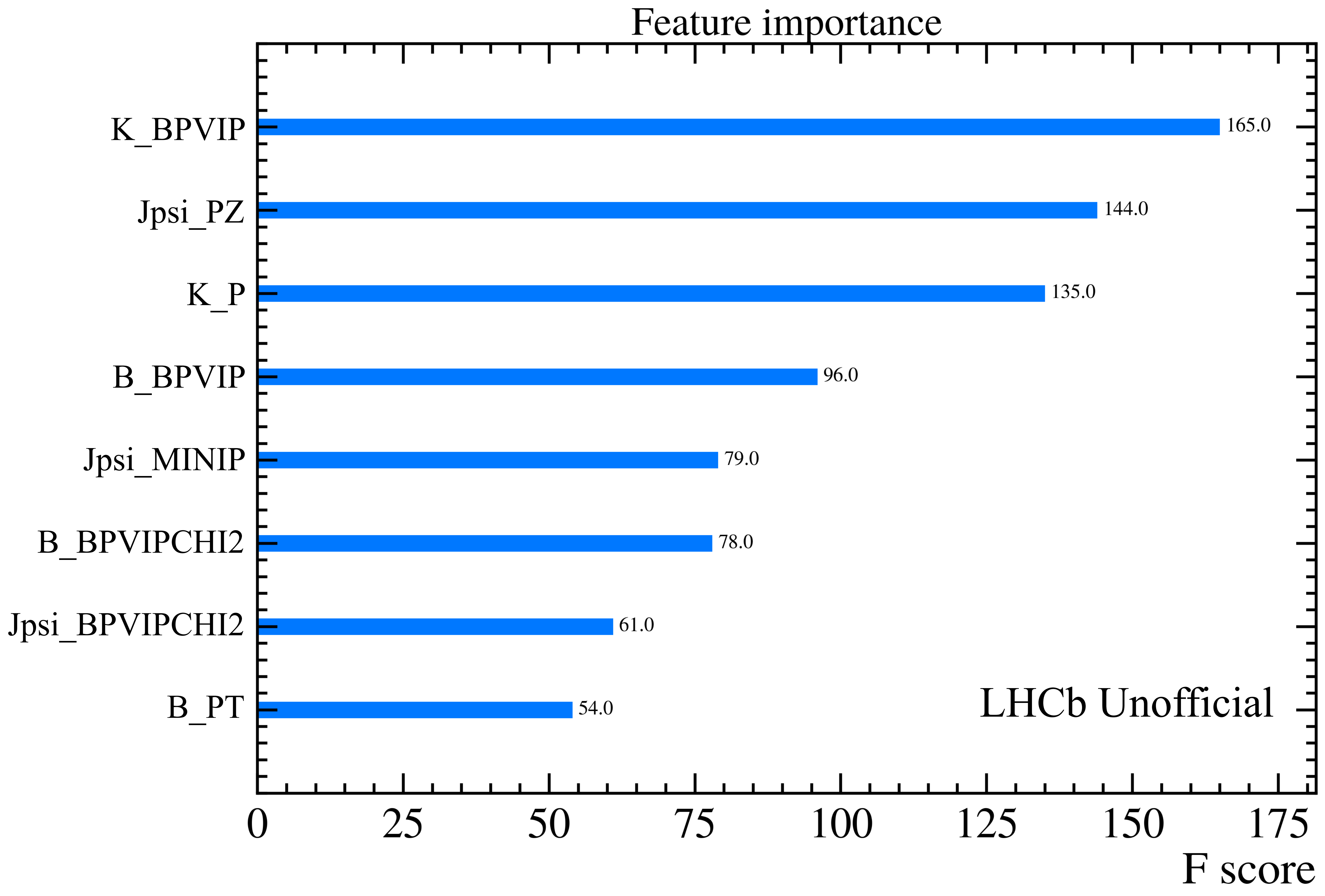
- The subsample sets the subsample ratio of the training instances. This means that if one chooses 0.75, the algorithm randomly uses only 75% of the training data for the first tree and then when it goes over the next tree, it takes only 75% of that. This is done to prevent overtraining
- And lastly there is the number of trees. The fewer, the less chance of overtraining but also worse classification

- The learning rate, which after each time we go from one tree to the other, one can get the weights of the features, and this parameter shrinks these weights
- The minimum loss, which is the minimum loss reduction required to make another partition on a leaf node of the tree
- The maximum depth, which is the maximum depth of a tree, thus related to how many nodes can a tree have. Increasing this value will make the model more complex and more likely to have overtraining but also more capable of classifying difficult cases
- The minimum child weight sets the minimum of the sum of the weights needed in a the subset after a node. If the sum of the weights is lower than this minimum, then the partition will stop
- The maximum delta step sets the maximum weight allowed for each tree so it does not become infinitely large

BDT variables and their importance

Variables used for training:

- B minimum ip
- J/ψ minimum ip
- K BPV ip
- B BPV ip
- J/ψ BPV $ip\ \chi^2$
- B BPV $ip\ \chi^2$
- $K\ p$
- $J/\psi\ \ p_Z$
- $B\ p_T$



Sample's statistics after the BDT

- The BDT is trained and applied to the data before splitting into brem categories and it has:
  - MC efficiency: 73.37%
  - BKG rejection: 97.92%

PIDe+ProbNNe24 with tag PIDe>5	Nº of events in sample
0brem	192361
brem on tag	594238
brem on probe	93945
2brem	352372

Double-sided Crystal Ball formula

$$f(x;\alpha_L,n_L,\alpha_R,n_R,\mu,\sigma)=\begin{cases} A_L\left(B_L-\frac{x-\mu}{\sigma}\right)^{-n_L} & \text{for } \frac{x-\mu}{\sigma}<-\alpha_L \\ \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) & \text{for } -\alpha_L\le\frac{x-\mu}{\sigma}\le\alpha_R \\ A_R\left(B_R+\frac{x-\mu}{\sigma}\right)^{-n_R} & \text{for } \frac{x-\mu}{\sigma}>\alpha_R \end{cases}$$

$$A_i=\left(\frac{n_i}{|\alpha_i|}\right)^{n_i}\exp\left(-\frac{|\alpha_i|^2}{2}\right)$$

Where $i=L,R$

$$B_i=\frac{n_i}{|\alpha_i|}-|\alpha_i|$$

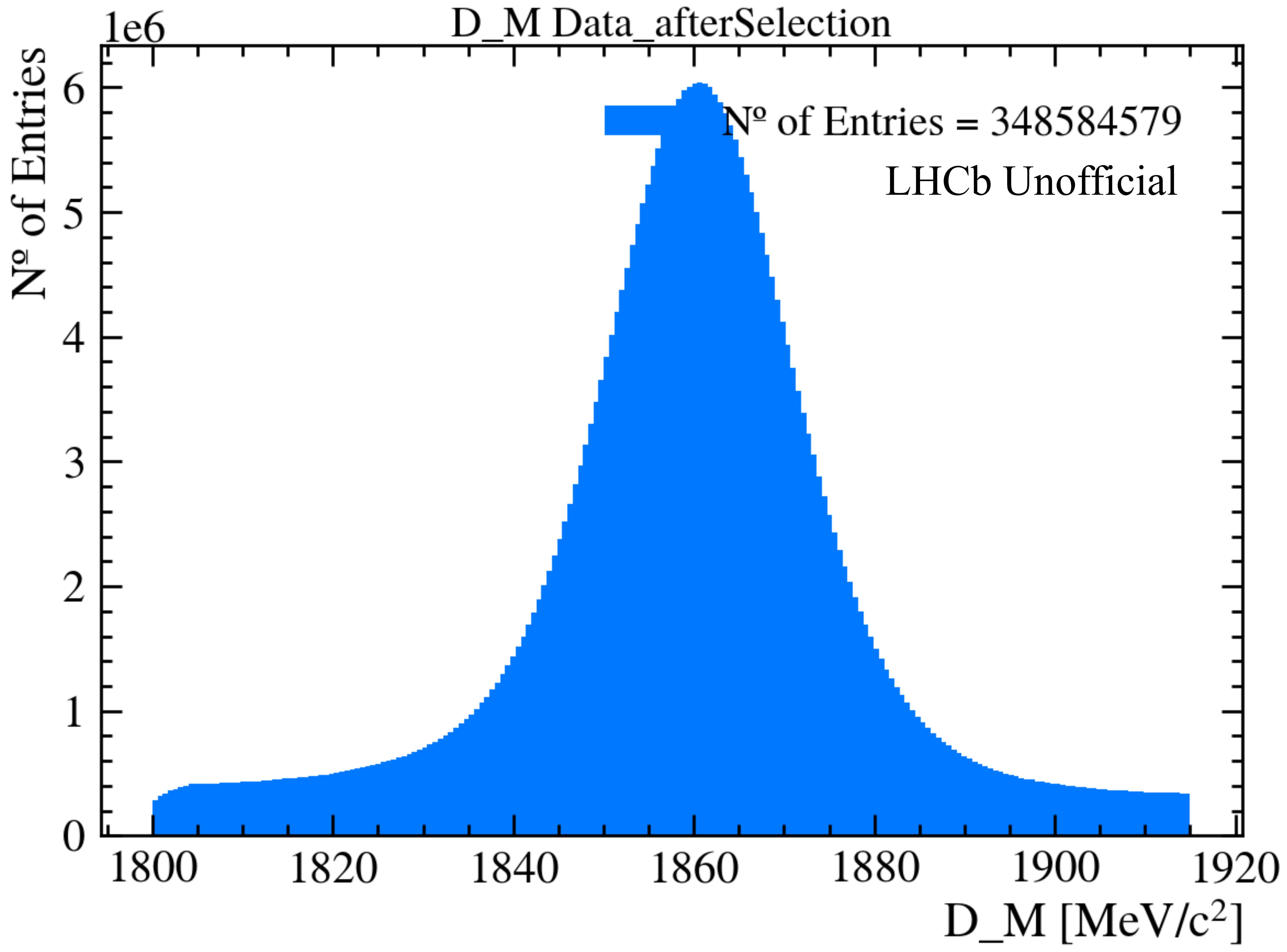
Selection for misID

- Inspired by studies of pion identification performance:
  - (D^{*+} Mass - D^0 Mass) > 141 MeV/c^2
  - (D^{*+} Mass - D^0 Mass) < 152 MeV/c^2
  - ($D^0_{hypo(KK)}$ Mass < (1864.84 - 25.0) MeV/c^2 || $D^0_{hypo(KK)}$ Mass > (1864.84 + 25.0) MeV/c^2)

D^0 mass plot and sample's statistics after the selection

- After the cuts, before splitting into brem categories

PIDe+ProbNNe24 with tag PIDe>5	Nº of events in sample
0brem	314192507
brem on tag	22291863
brem on probe	10570329
2brem	1529880



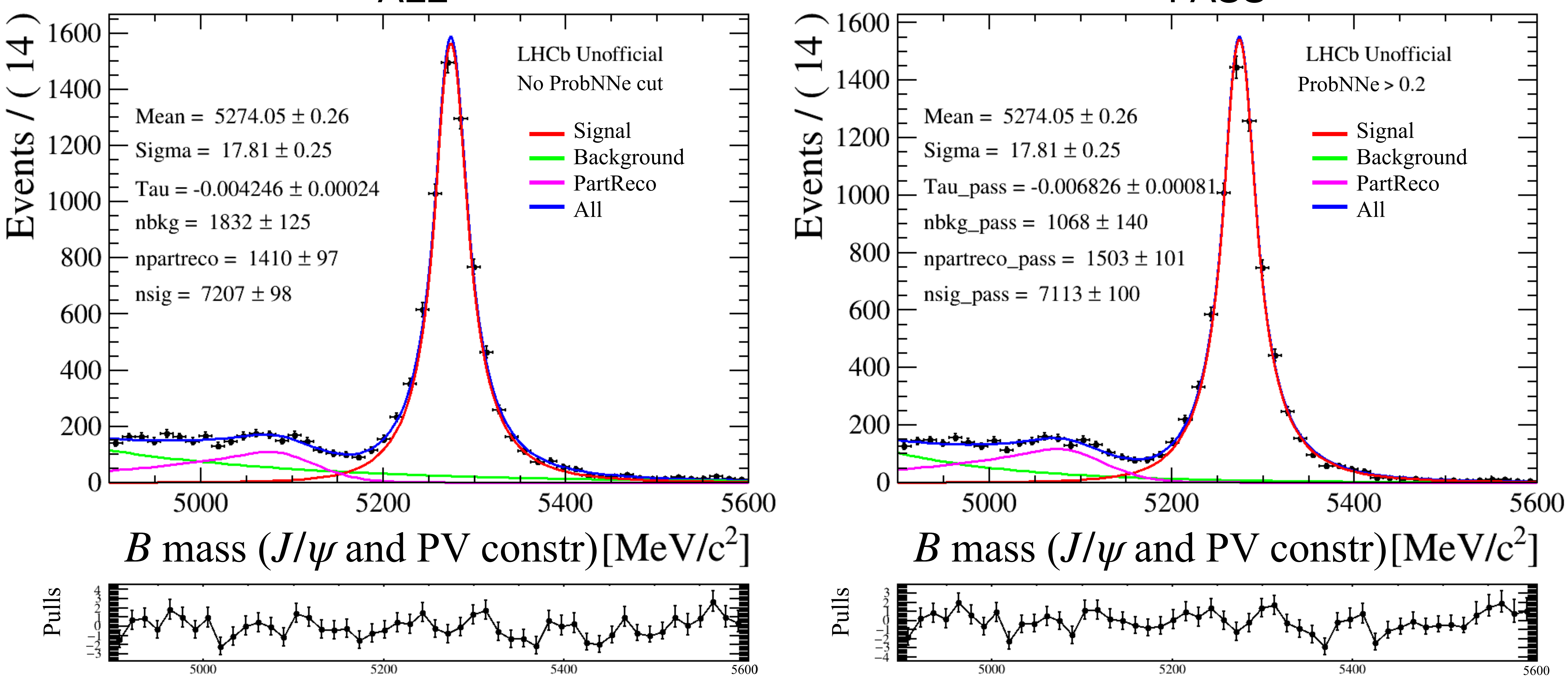
ProbNNe fit example

2brem

ProbNNe 0.2 electron efficiency for the P range (17500,20625) MeV/c

$$Eff = 0.9869^{+0.0013}_{-0.0015}$$

ALL | PASS



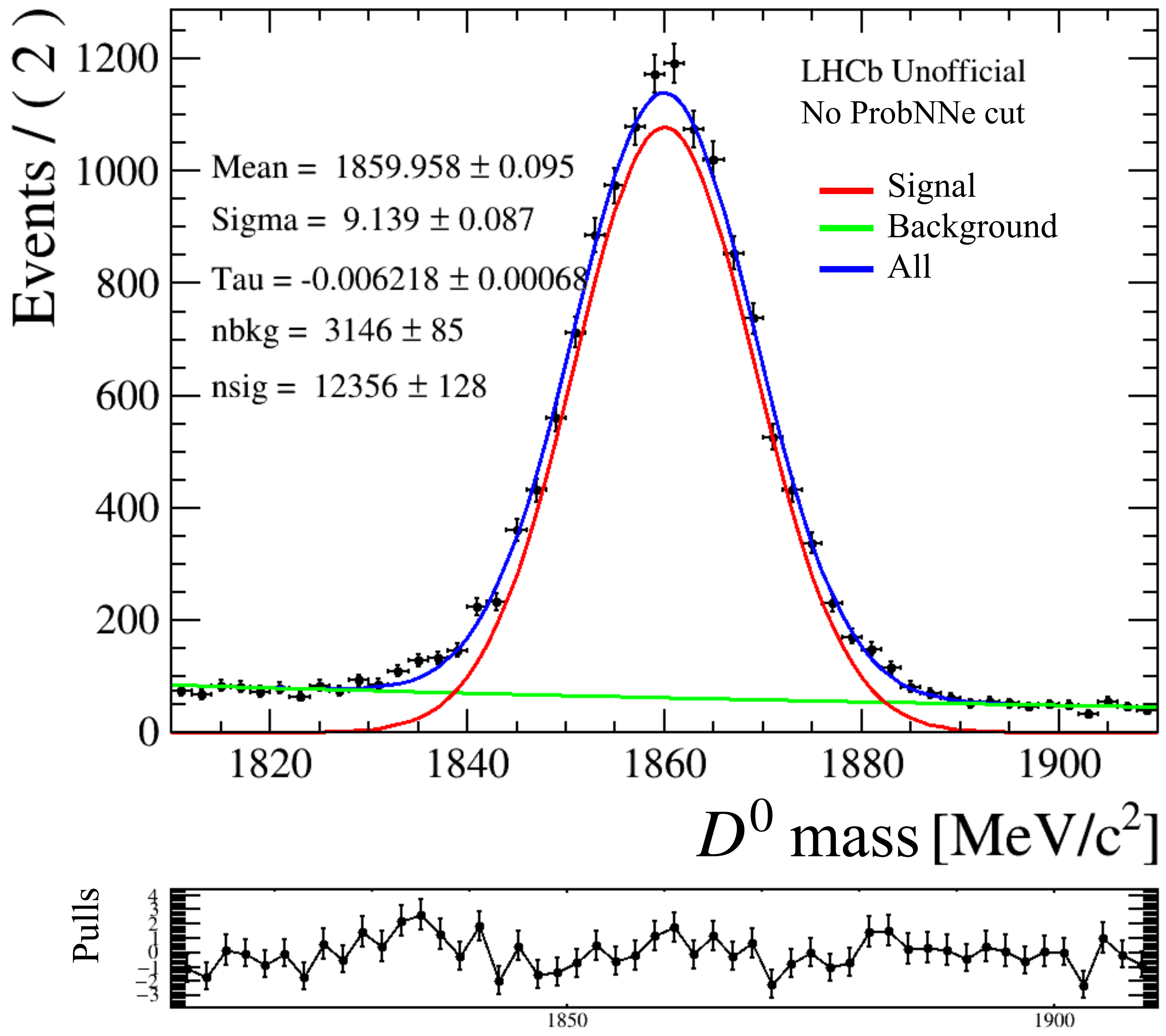
ProbNNe misID fit example

2brem

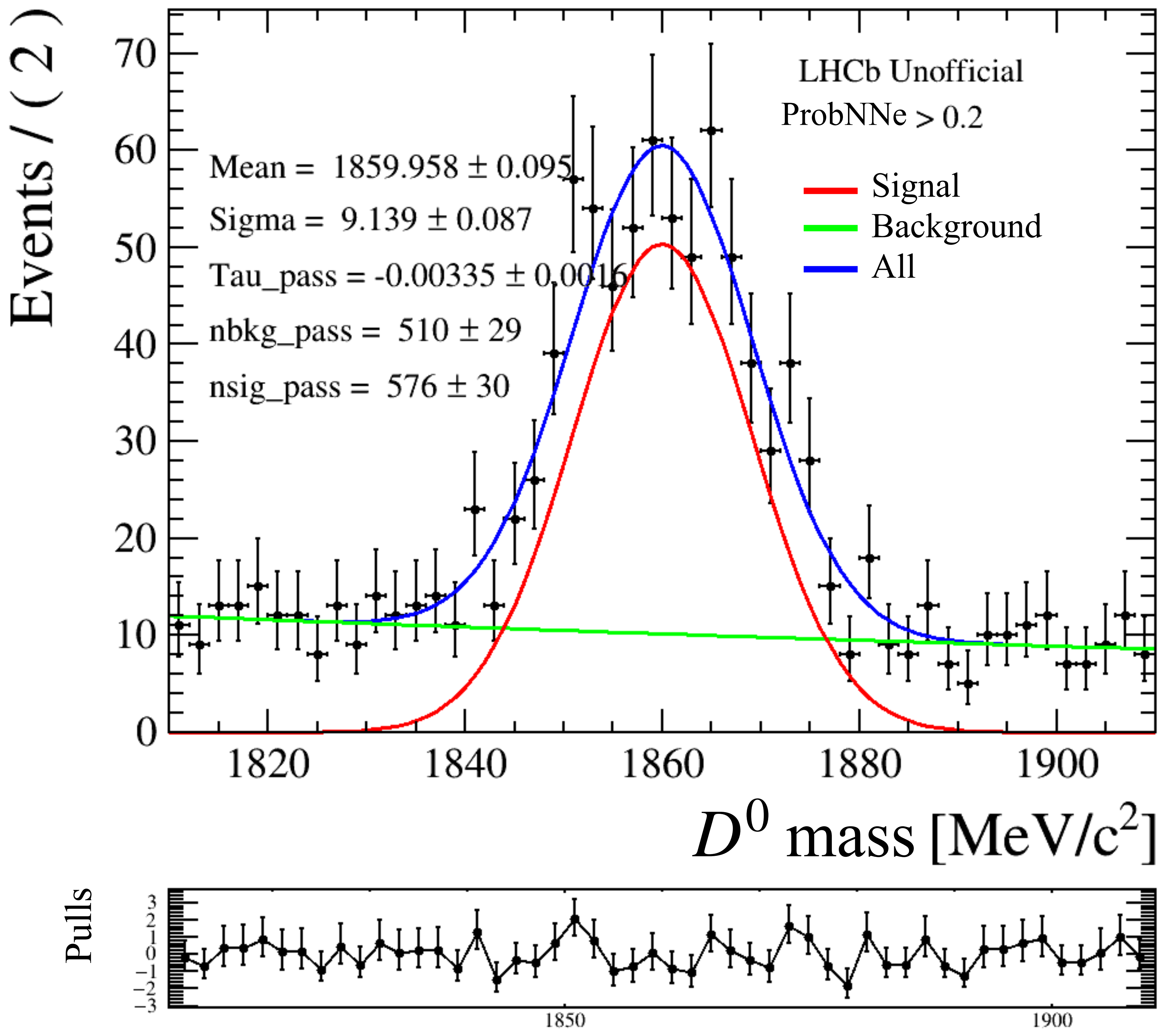
ProbNNe 0.2 electron efficiency for the P range (17500,20625) MeV/c

$$Eff = 0.0467^{+0.0020}_{-0.0019}$$

ALL



PASS

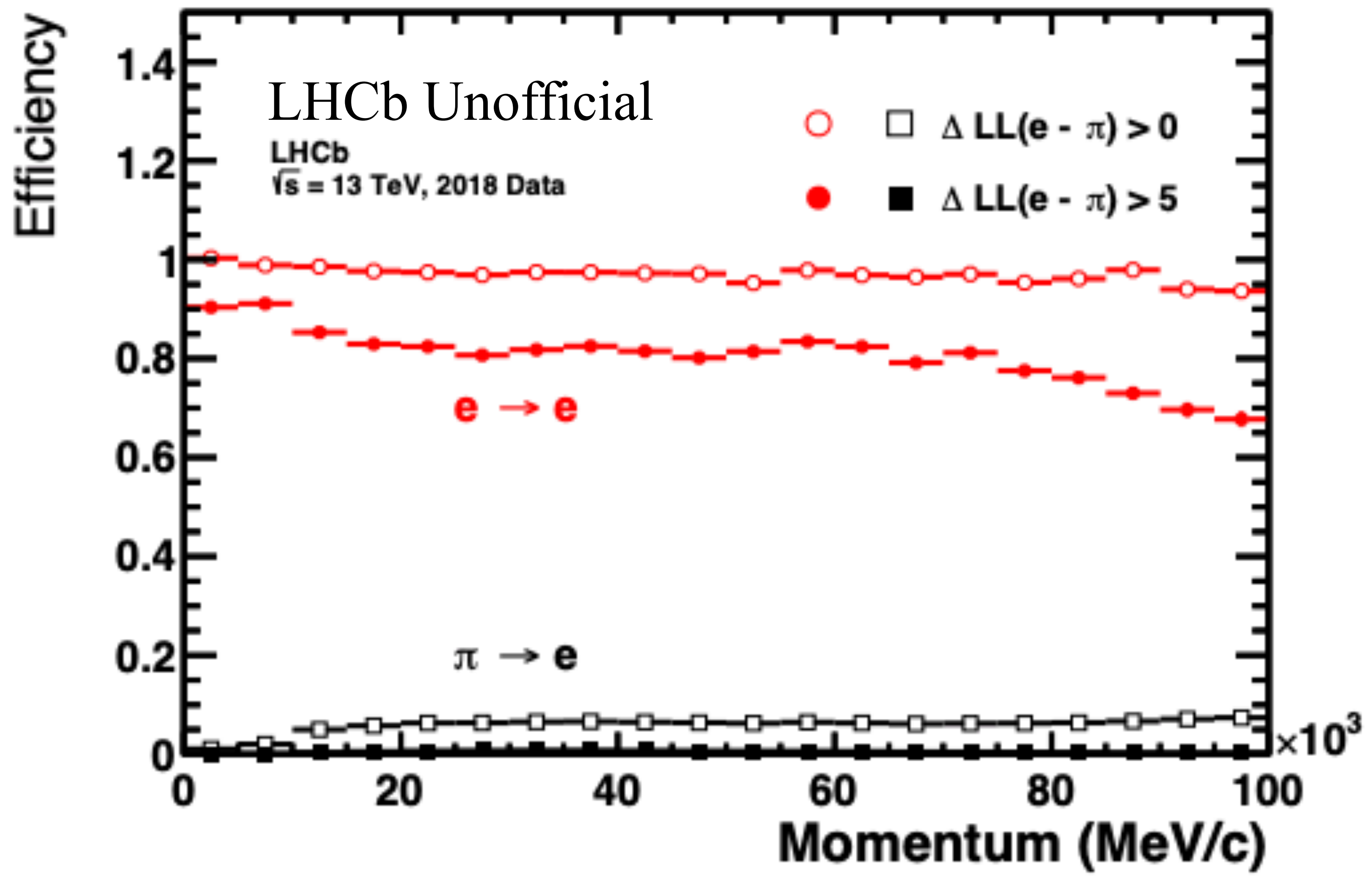


PIDe 2024 Performance

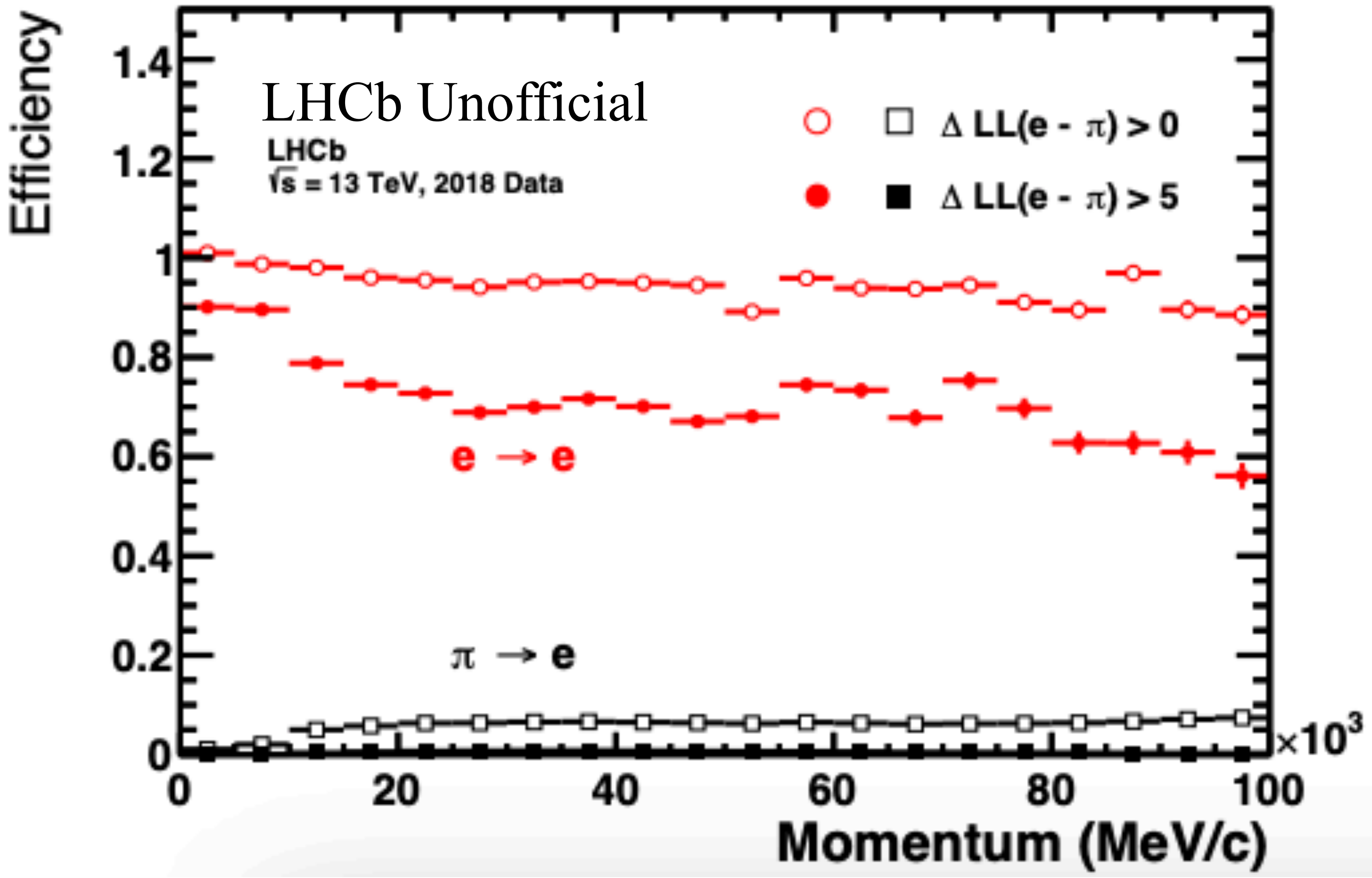
Integrated results

PIDe24 with tag PIDe>5	Probe PIDe>0		Probe PIDe>5	
	$e \to e$	$\pi \to e$	$e \to e$	$\pi \to e$
0brem	0.853 ± 0.002	0.027970 ± 0.000011	0.489 ± 0.003	0.000635 ± 0.000002
brem on tag	0.8616 ± 0.0010	0.02083 ± 0.00004	0.4929 ± 0.0015	0.000549 ± 0.000006
brem on probe	0.9835 ± 0.0006	0.06056 ± 0.00009	0.8350 ± 0.0018	0.00290 ± 0.00002
2brem	0.9855 ± 0.0003	0.0665 ± 0.0003	0.8311 ± 0.0010	0.00646 ± 0.00010

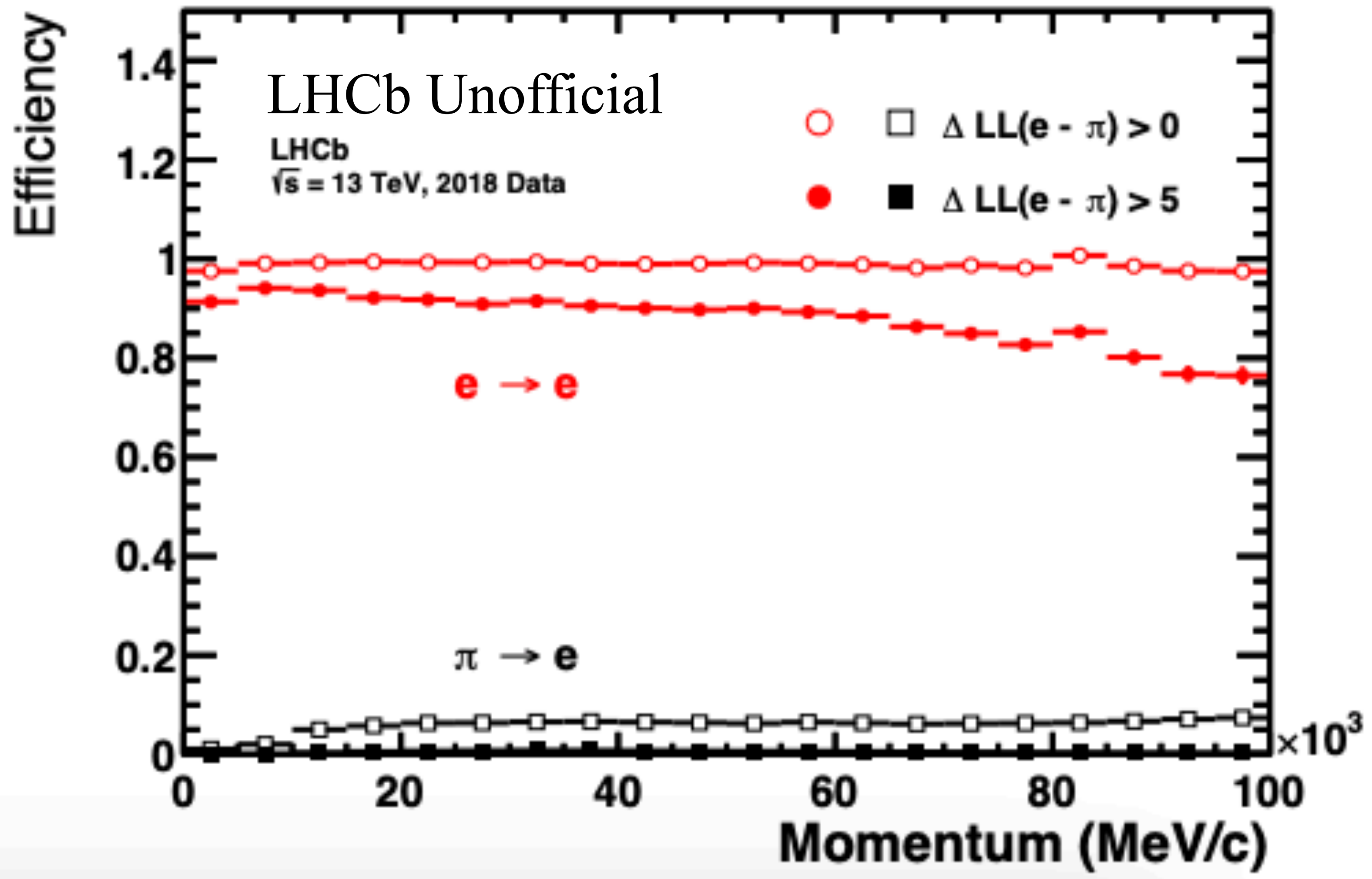
PIDe Run 2 results



no brem photons

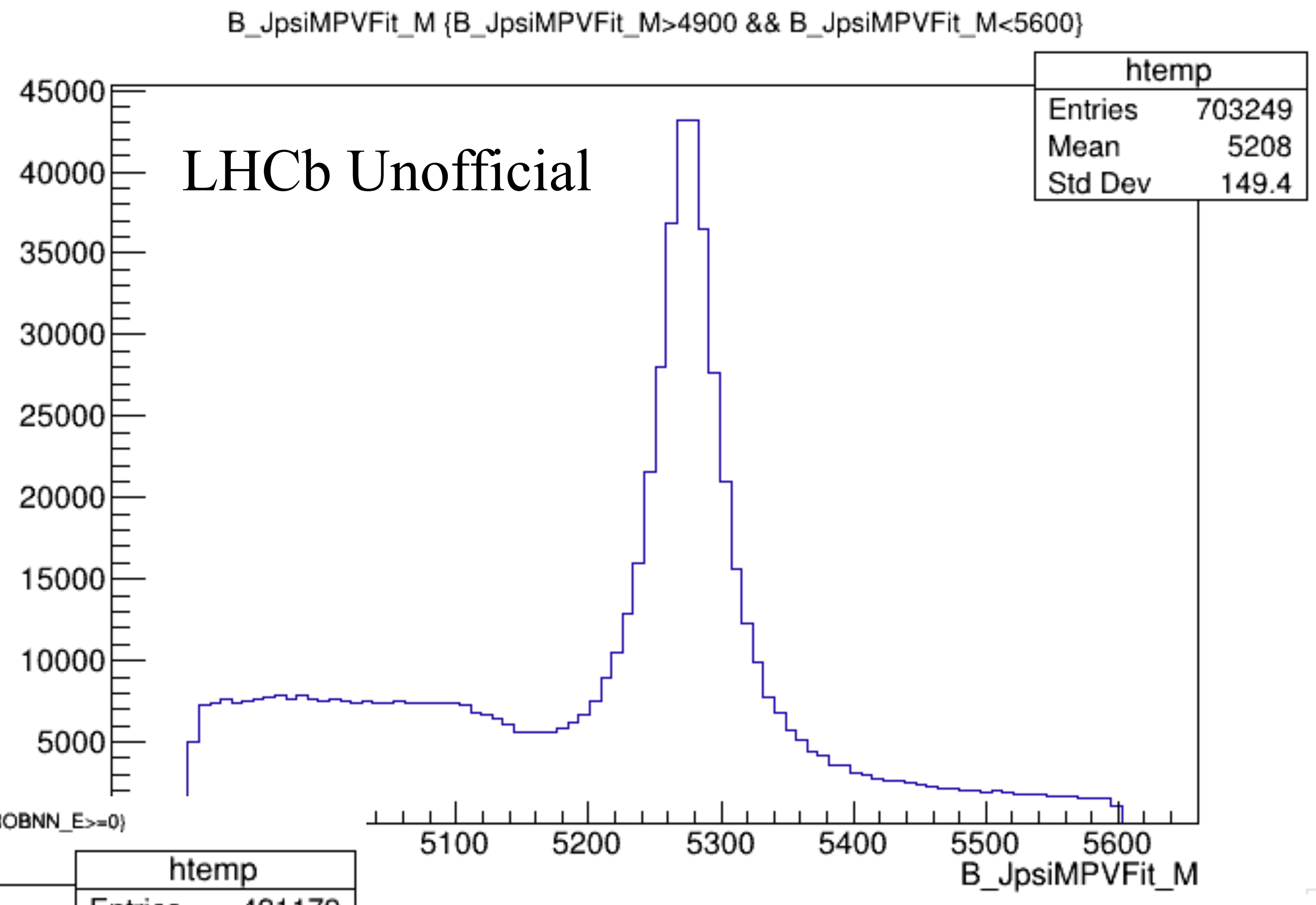


has brem photons

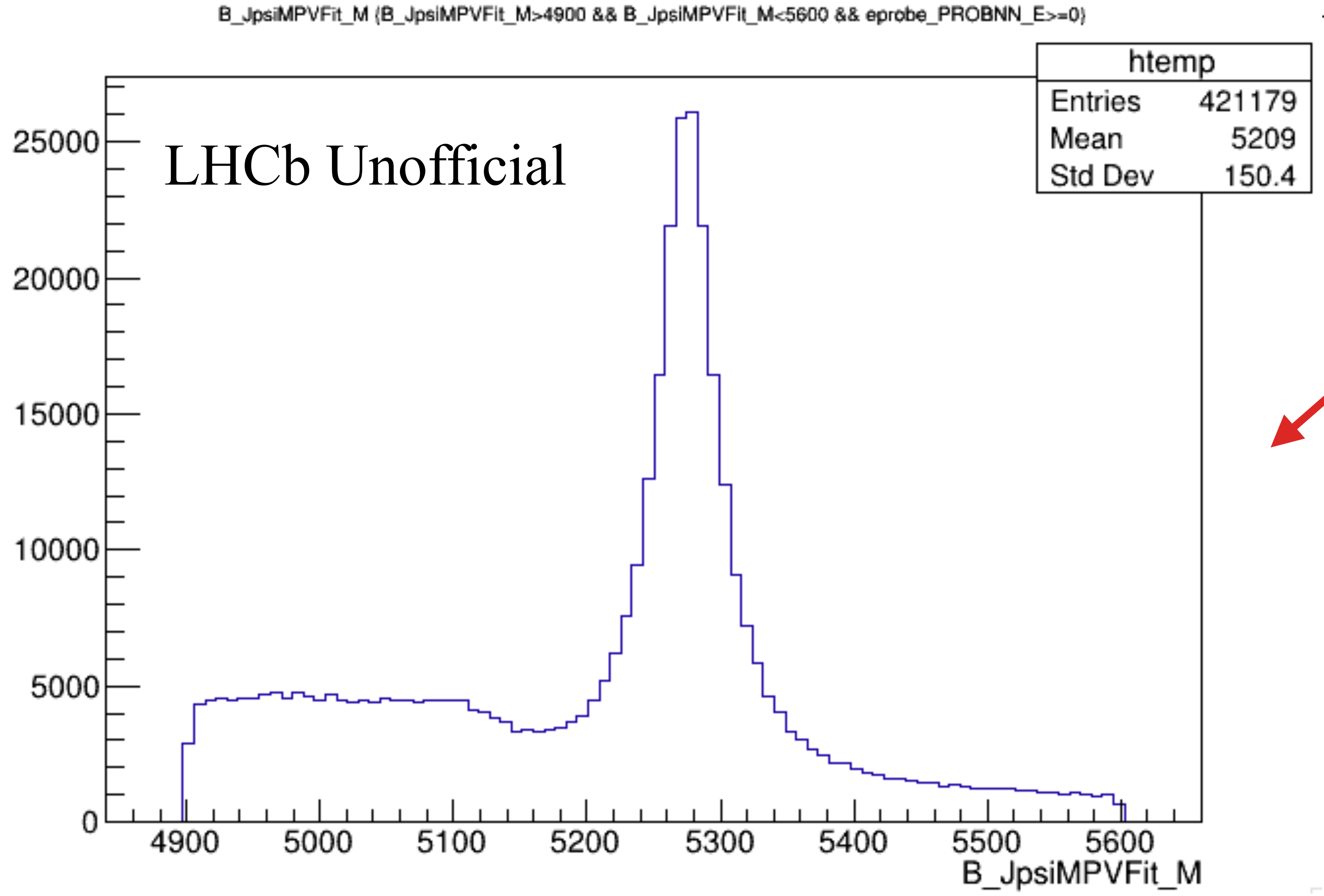


NaN caveat in ProbNNe

- For the data taken before the MayMD, the ProbNNe entries are all NaNs, so we have to remove these fills. Those represent ~40% of our sample
- So the fills used are:
  - MagDown: 9618 - 9708
  - MagUp: 9653 - 9691
- And for the misID:
  - Fills: 9618 - 9708



Removing the NaNs



ProbNNe 2024 Performance

Integrated results

ProbNNe24 with tag PIDe>5	Probe ProbNNe>0.2		Probe ProbNNe>0.5	
	$e \rightarrow e$	$\pi \rightarrow e$	$e \rightarrow e$	$\pi \rightarrow e$
0brem	0.915 ± 0.002	0.07995 ± 0.00004	0.878 ± 0.002	0.04670 ± 0.00003
brem on tag	0.9140 ± 0.0011	0.06381 ± 0.00011	0.8807 ± 0.0013	0.03521 ± 0.00009
brem on probe	0.9851 ± 0.0008	0.1001 ± 0.0002	0.9788 ± 0.0009	0.0718 ± 0.0002
2brem	0.9881 ± 0.0004	0.1022 ± 0.0007	0.9805 ± 0.0005	0.0746 ± 0.0006

ProbNNe Run 2 results