$B^0 \to \pi^0\pi^0$ update

M. Dorigo, S. Raiz, D. Tonelli

Trieste

BtoHadrons meeting
Nov 16, 2022

Overview

BF and A_{CP} of $B^0 \to \pi^0\pi^0$ decays: important measurements unique to Belle II.

Status: 189.9 fb^{-1} analysis by Francis shown at ICHEP2022 and targeting PRD submission soon.

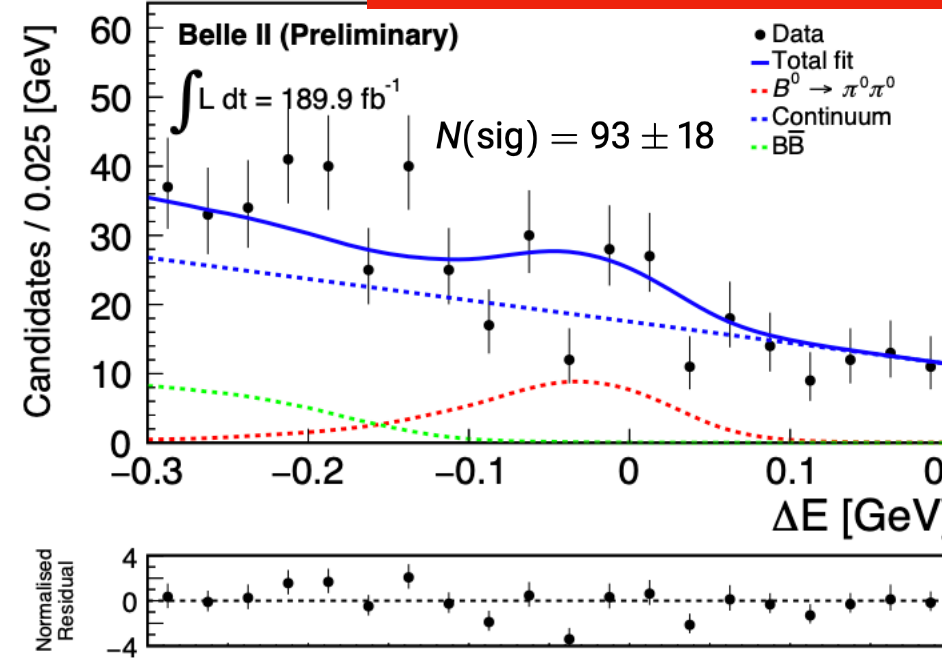
Start from Francis analysis, and try to improve for full LS1 data update.

Last time: photonMVA, CSBDT, specific BDT against ρ's

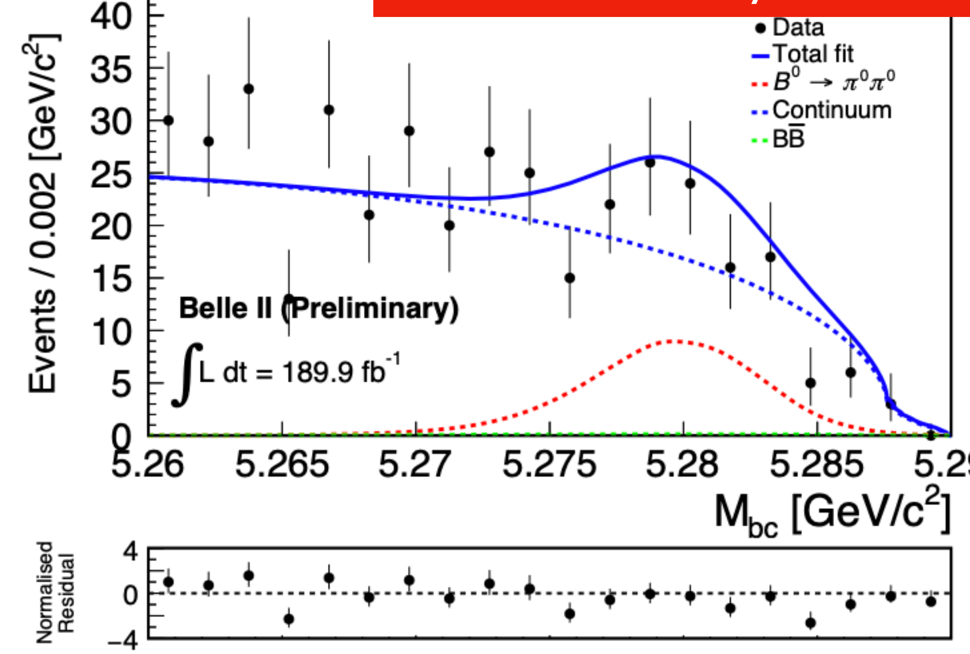
(https://indico.belle2.org/event/7786/contributions/45936/attachments/18498/27519/B02pi0pi0_B2Hadrons_vF.pdf)

Today:
- open points from last time
- simplified fitter and CS selection optimisation
- 2D signal modelling

ICHEP2022 by Francis



ICHEP2022 by Francis



$\mathcal{A}^{CP} = -0.14 \pm 0.46$ (stat.) ± 0.07 (syst.)

$\mathcal{B} = (1.27 \pm 0.25$ (stat.) ± 0.17 (syst.)$) \cdot 10^{-6}$

Samples and selections

Samples

GenericMC: MC15ri

SignalMC: MC15 locally produced (2000000 events)

Data: Proc13 chunk1+chunk2

Off-res data: Proc13 (c1+c2) +Prompt

For data use “all” (no hadron skim).

Base selections

γ: E>0.03 GeV, |clusterTiming|<200, clusterNHits>1.5, 0.30<clusterθ<2.62 (very loose cuts)

π^0: daughterAngle < 0.4, |daughterDiffOfPhi| < 0.4, |cosHelicityAngleMomentum| < 0.99, p > 1.5 GeV/c, 0.115 < InvM < 0.150 GeV/c^2 (very loose cuts)

B^0: -0.3<ΔE<0.2 GeV, $M_{\rm bc}$>5.26 GeV/c^2

Open points from last time

- PhotonMVA: one input variable could have some data/MC discrepancies.

- CSBDT: explore other possible inputs of the BDT.

- Flavor tagger: check if inclusion of Δr and ΔZ in CSBDT sculpts the flavor tagger (change in the FT parameters).

Photon MVA

Open point from last time: photonMVA

Distinguish between signal photons and misreconstructed photons: beam backgrounds, energy releases from other particles…

Combine highly-discriminant cluster- and photon-variables in a MVA.

From last time

Inputs
pt
clusterE1E9
clusterErrorPhi
clusterHighestE
clusterSecondMoment
clusterZernikeMVA
minC2TDist
clusterLAT
clusterNHits
clusterTheta
beamBackgroundSuppression

This is a MVA that is well reproduced in MC, but its main inputs clusterTiming and PulseShapeDiscriminatorMVA are not.

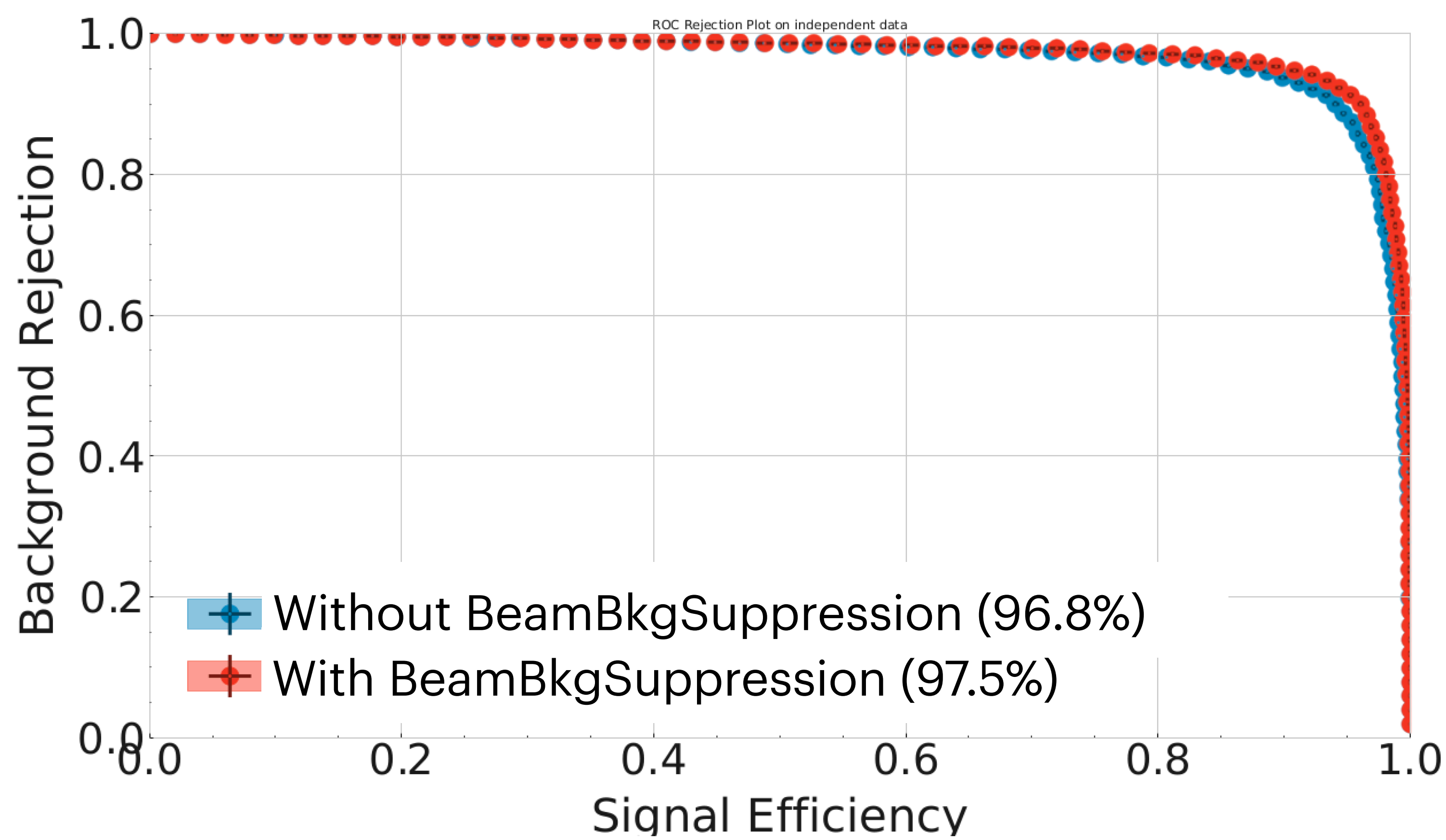
Open point from last time: photonMVA

Distinguish between signal photons and misreconstructed photons: beam backgrounds, energy releases from other particles…

Combine highly-discriminant cluster- and photon-variables in a MVA.

Inputs
pt
clusterE1E9
clusterErrorPhi
clusterHighestE
clusterSecondMoment
clusterZernikeMVA
minC2TDist
clusterLAT
clusterNHits
clusterTheta
~~beamBackgroundSuppression~~

Try removing it



Performance loss is acceptable: remove from inputs.

Photon MVA comparison (after input exclusion)

Simulation

1

Look at photons: reconstruct $B^0 \to \pi^0\pi^0$ in genericMC and apply γ and π^0 selections. Consider as "signal" all real photons, and as "background" all misreconstructed photons. Use MC info to obtain photon signal efficiency and bkg rejection after photonMVA selection. For fixed ε_{sig} (=85%), compare bkg rejection.

Old bkg rejection: 68.5% $\longrightarrow$ My bkg rejection: 84.8%

Simulation

2

Look at B^0 candidates: reconstruct $B^0 \to \pi^0\pi^0$ candidates in genericMC and apply γ and π^0 selections. Consider as "signal" all signal $B^0 \to \pi^0\pi^0$ events, and everything else as "background". Use MC info to obtain signal efficiency and bkg rejection after photonMVA selection. For fixed ε_{sig} (=94.7%), compare bkg rejection.

Old bkg rejection: 15.6% $\longrightarrow$ My bkg rejection: 16.1%

Data

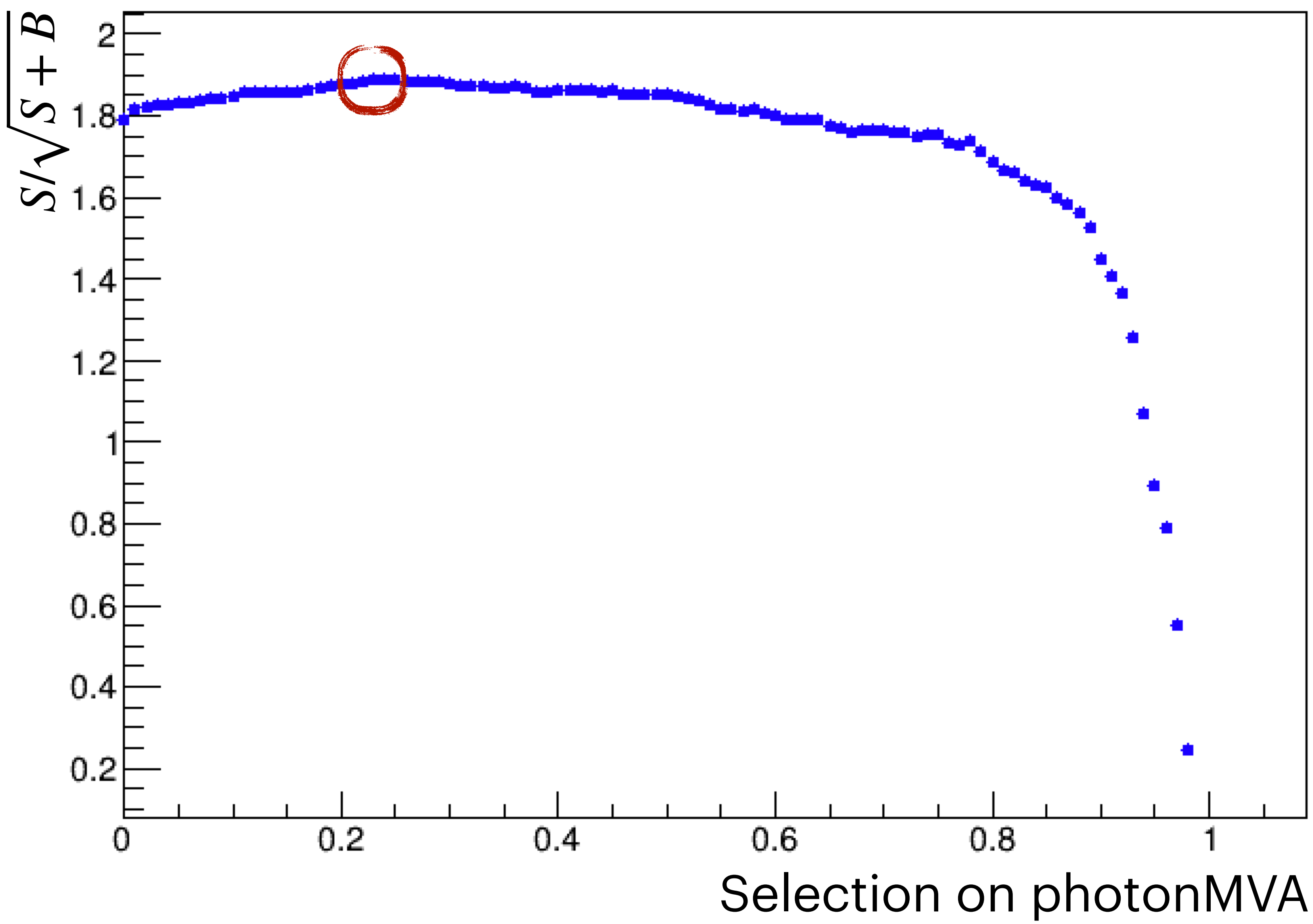
3

Check on data: reconstruct $D^{*+} \to D^0(K^-\pi^+\pi^0)\pi^+$ candidates in data and apply γ and π^0 selections. Reweigh using $p(\pi^0)$. Consider as "signal" all signal $D^{*+} \to D^0\pi^+$ events, and everything else as bkg. Obtain ε_{sig} and bkg rejection as $N_{\text{pass}}/(N_{\text{pass}} + N_{\text{not pass}})$ from fit. For fixed ε_{sig} (=96.6%), compare bkg rejection.

Old bkg rejection: 5.1±0.1% $\longrightarrow$ My bkg rejection: 9.1±0.1%

PhotonMVA: selection optimisation on MC

Use genericMC sample of $B^0 \to \pi^0\pi^0$. Change selection on photonMVA and check $S/\sqrt{S+B}$ in signal region (-0.15<ΔE<0.1 GeV and M_{bc}>5.27 Gev/c^2).



Maximum of significance @photonMVA>0.25

Next step: repeat optimisation on data.

Open point: photonMVA

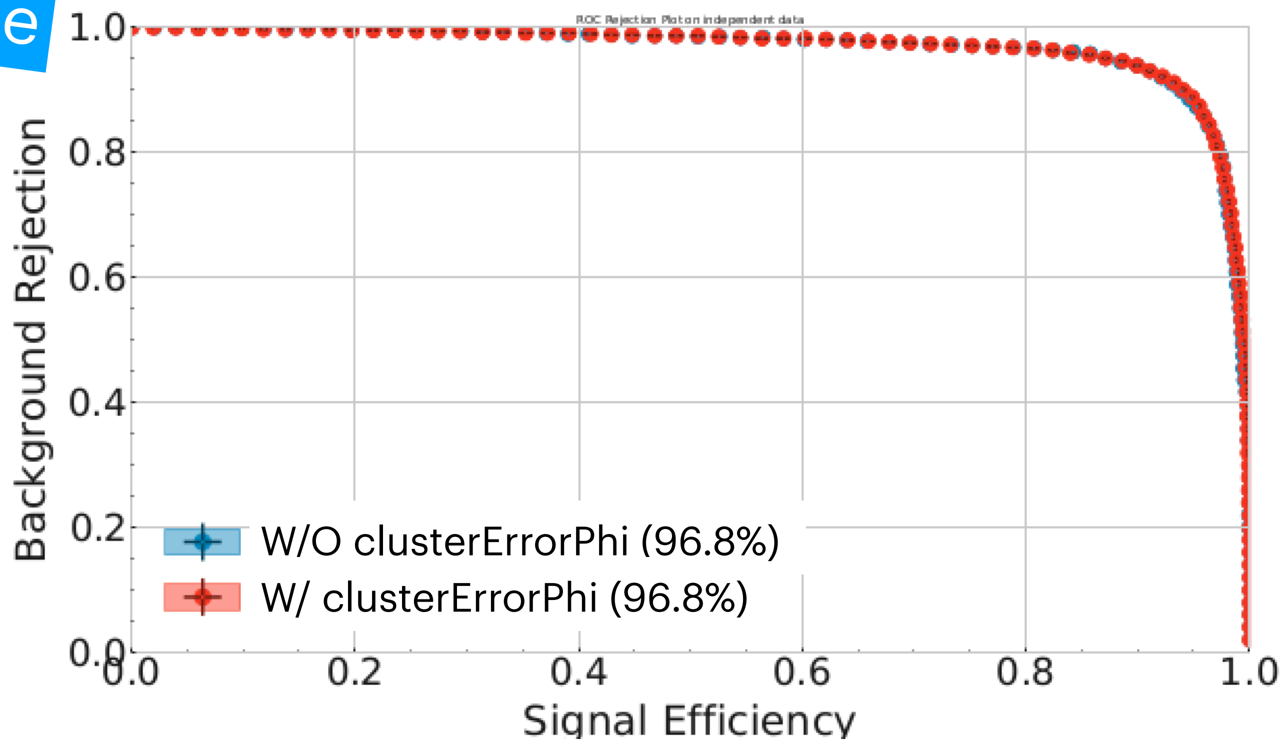
New from last time

Distinguish between signal photons and misreconstructed photons: beam backgrounds, energy releases from other particles…

Combine highly-discriminant cluster- and photon-variables in a MVA.

From last time

Inputs
pt
clusterE1E9
clusterErrorPhi
clusterHighestE
clusterSecondMoment
clusterZernikeMVA
minC2TDist
clusterLAT
clusterNHits
clusterTheta

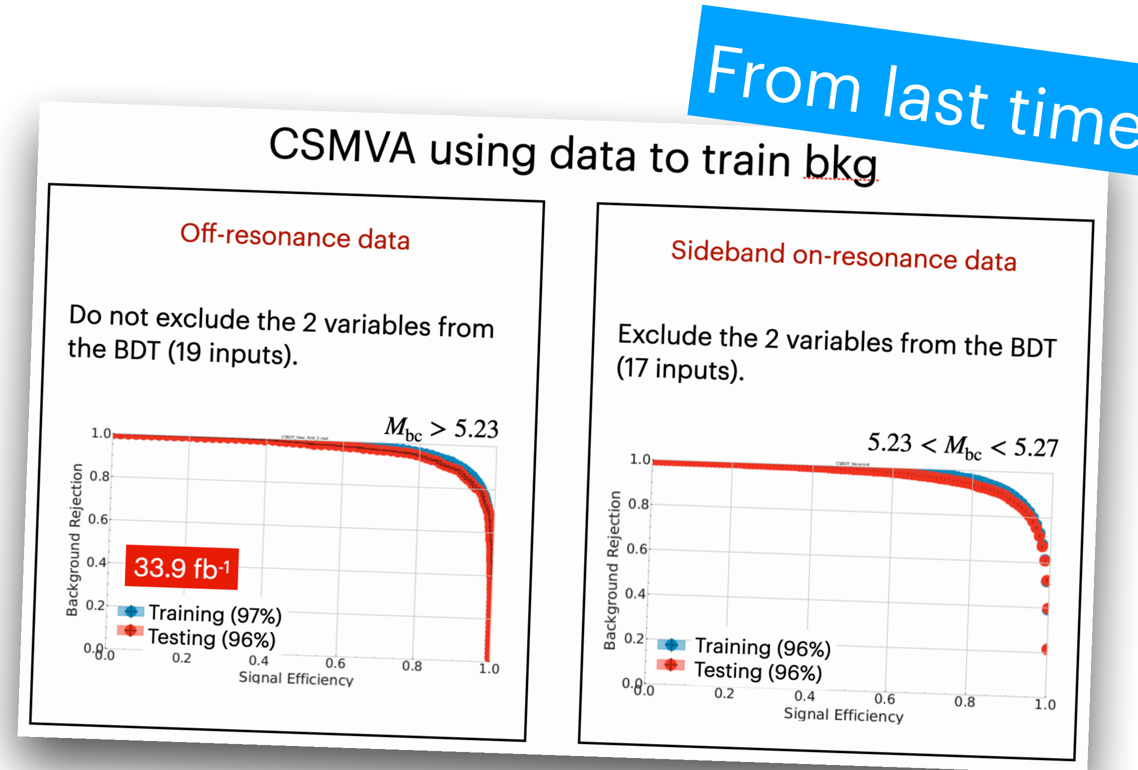


Dummy variable, no effect on MVA.

Remove from inputs.

CSBDT

CSBDT: new possible inputs



Last time: shown CSBDT trained in data.

Today: explore new possible variables with high $B\bar{B}/q\bar{q}$ discriminating power.

1 ROE high-energy tracks ⟷ Probability to find a lepton is higher in $B\bar{B}$ than in $q\bar{q}$ (semileptonic decays).

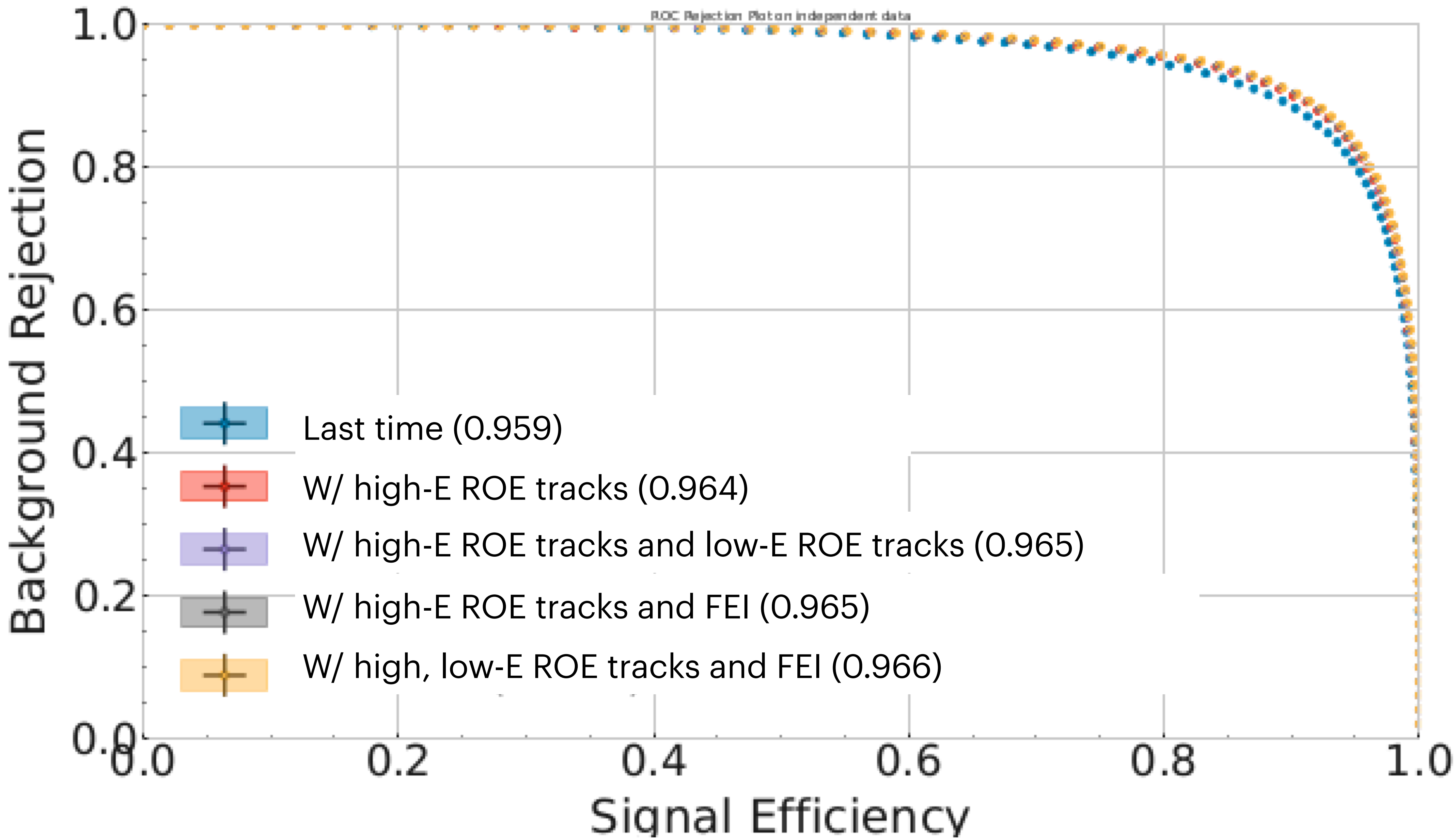
2 ROE low-energy tracks ⟷ Probability to find a soft pion is higher in $B\bar{B}$ than in $q\bar{q}$ ($B \to D\pi$).

3 SL FEI probability ⟷ Low efficiency, but high discrimination.

ROE variables: energy, transverse momentum, PID

Performance with ROCs: MC only

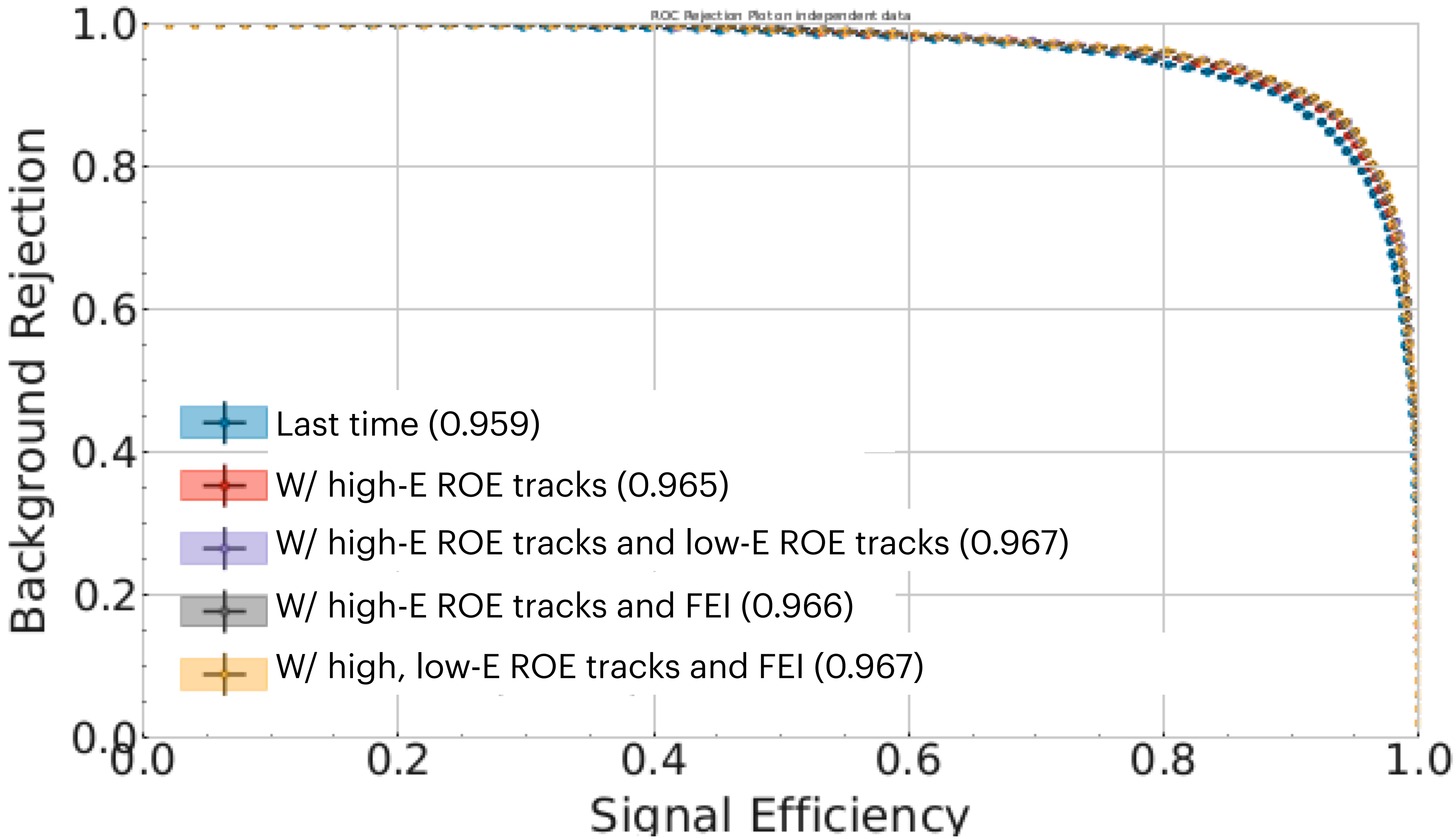
Check performance using MC only:



High-energy tracks variables offer good discrimination.

Off-res data + SignalMC (PID corrected)

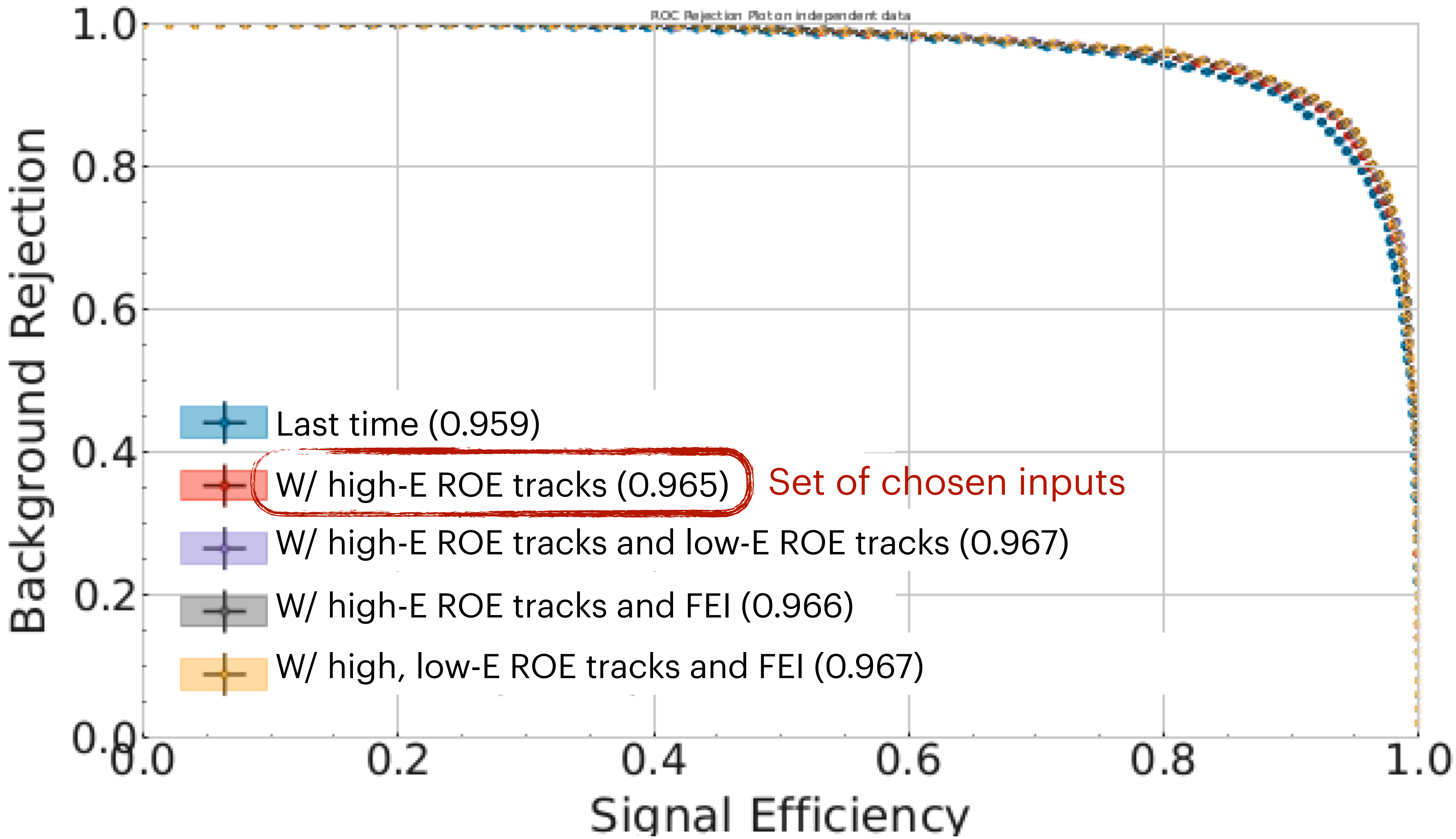
Check performance using signalMC (with corrected PID for high-energy kaons and pions; lepton corrections still not ready)+OffRes data:



Similar performance in data.

Off-res data + SignalMC (PID corrected)

Check performance using signalMC (with corrected PID for high-energy kaons and pions; lepton corrections still not ready)+OffRes data:



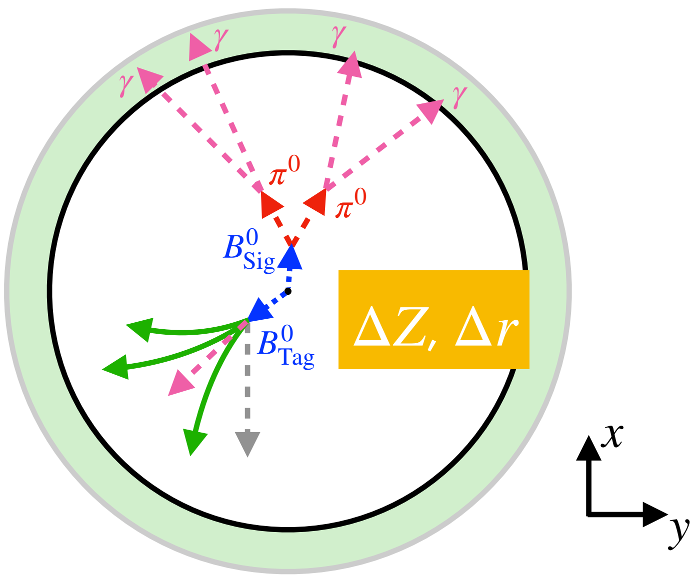
All performances are ~similar, take default+high-E tracks set of inputs.

Open point from last time: flavor tagger sculpting

I need the flavour tagger to distinguish B^0 and $\bar{B}^0$.

My BDT includes B_{Tag} variables Δr and ΔZ (distance of vertex from IP). They also enter in the FT.

Check if these variables sculpt or introduce large correlations in flavour tagger.



Flavor tagging parameters

SignalMC15, light-2210-devonrex

Check flavour tagger parameters obtained in $B^0 \to \pi^0\pi^0$ before and after applying CSBDT selection.

$B^0 \to \pi^0\pi^0$ just after reconstruction

r- Interval	ε_i	$\Delta\varepsilon_i$	$w_i \pm \delta w_i$	$\Delta w_i \pm \delta\Delta w_i$	$\varepsilon_{eff,i} \pm \delta\varepsilon_{eff,i}$	$\Delta\varepsilon_{eff,i} \pm \delta\Delta\varepsilon_{eff,i}$
$0.000 - 0.100$	15.7	-0.07	47.78 ± 0.12	-0.05 ± 0.24	0.0311 ± 0.0033	0.0012 ± 0.0066
$0.100 - 0.250$	15.2	-0.03	41.57 ± 0.12	1.45 ± 0.24	0.4305 ± 0.0122	-0.1488 ± 0.0249
$0.250 - 0.500$	20.0	-0.02	31.03 ± 0.10	-0.00 ± 0.19	2.8817 ± 0.0301	-0.0022 ± 0.0601
$0.500 - 0.625$	11.7	-0.04	21.86 ± 0.11	0.68 ± 0.23	3.7072 ± 0.0314	-0.1922 ± 0.0630
$0.625 - 0.750$	11.1	0.12	15.85 ± 0.10	-0.53 ± 0.21	5.1703 ± 0.0342	0.2178 ± 0.0682
$0.750 - 0.875$	9.2	0.24	9.59 ± 0.09	-0.17 ± 0.18	5.9779 ± 0.0324	0.2058 ± 0.0647
$0.875 - 1.000$	17.1	-0.20	2.05 ± 0.03	0.12 ± 0.06	15.7690 ± 0.0389	-0.2642 ± 0.0778
Total			$\varepsilon_{eff} = \sum_i \varepsilon_i \cdot \langle 1 - 2w_i \rangle^2 = 33.97 \pm 0.08$		$\Delta\varepsilon_{eff} = -0.18 \pm 0.15$	

$B^0 \to \pi^0\pi^0$ after CS selection (>0.7)

r- Interval	ε_i	$\Delta\varepsilon_i$	$w_i \pm \delta w_i$	$\Delta w_i \pm \delta\Delta w_i$	$\varepsilon_{eff,i} \pm \delta\varepsilon_{eff,i}$	$\Delta\varepsilon_{eff,i} \pm \delta\Delta\varepsilon_{eff,i}$
$0.000 - 0.100$	15.2	-0.06	47.73 ± 0.15	0.33 ± 0.29	0.0314 ± 0.0041	-0.0093 ± 0.0083
$0.100 - 0.250$	14.6	-0.09	41.67 ± 0.15	1.53 ± 0.30	0.4039 ± 0.0144	-0.1514 ± 0.0295
$0.250 - 0.500$	19.6	-0.08	30.99 ± 0.12	-0.23 ± 0.24	2.8284 ± 0.0363	0.0576 ± 0.0726
$0.500 - 0.625$	11.7	0.01	21.89 ± 0.14	0.61 ± 0.28	3.7100 ± 0.0384	-0.1566 ± 0.0769
$0.625 - 0.750$	11.1	0.08	15.86 ± 0.13	-0.35 ± 0.25	5.1805 ± 0.0418	0.1454 ± 0.0834
$0.750 - 0.875$	9.3	0.24	9.58 ± 0.11	0.00 ± 0.22	6.0653 ± 0.0398	0.1544 ± 0.0795
$0.875 - 1.000$	18.5	-0.10	1.97 ± 0.04	0.07 ± 0.07	17.1132 ± 0.0489	-0.1375 ± 0.0978
Total			$\varepsilon_{eff} = \sum_i \varepsilon_i \cdot \langle 1 - 2w_i \rangle^2 = 35.33 \pm 0.09$		$\Delta\varepsilon_{eff} = -0.10 \pm 0.19$	

Tagging efficiency is higher, as expected.

Wrong-tags and asymmetries are compatible → no CS visible effect on FlavorTagger.

Flavor tagging parameters

ligh

New from last time

Check flavour tagger parameters obtained in $B^0 \to \pi^0\pi^0$ after applying CSBDT selection and in $B^0 \to D^-(K^+\pi^-\pi^-)\pi^+$ (calibration channel with largest *BF*).

$B^0 \to \pi^0\pi^0$ after CS selection (>0.7)

r- Interval	ε_i	$\Delta\varepsilon_i$	$w_i \pm \delta w_i$	$\Delta w_i \pm \delta\Delta w_i$	$\varepsilon_{eff,i} \pm \delta\varepsilon_{eff,i}$	$\Delta\varepsilon_{eff,i} \pm \delta\Delta\varepsilon_{eff,i}$
$0.000 - 0.100$	15.2	-0.06	47.73 ± 0.15	0.33 ± 0.29	0.0314 ± 0.0041	-0.0093 ± 0.0083
$0.100 - 0.250$	14.6	-0.09	41.67 ± 0.15	1.53 ± 0.30	0.4039 ± 0.0144	-0.1514 ± 0.0295
$0.250 - 0.500$	19.6	-0.08	30.99 ± 0.12	-0.23 ± 0.24	2.8284 ± 0.0363	0.0576 ± 0.0726
$0.500 - 0.625$	11.7	0.01	21.89 ± 0.14	0.61 ± 0.28	3.7100 ± 0.0384	-0.1566 ± 0.0769
$0.625 - 0.750$	11.1	0.08	15.86 ± 0.13	-0.35 ± 0.25	5.1805 ± 0.0418	0.1454 ± 0.0834
$0.750 - 0.875$	9.3	0.24	9.58 ± 0.11	0.00 ± 0.22	6.0653 ± 0.0398	0.1544 ± 0.0795
$0.875 - 1.000$	18.5	-0.10	1.97 ± 0.04	0.07 ± 0.07	17.1132 ± 0.0489	-0.1375 ± 0.0978
Total	$\varepsilon_{eff} = \sum_i \varepsilon_i \cdot \langle 1 - 2w_i \rangle^2 = 35.33 \pm 0.09$					$\Delta\varepsilon_{eff} = -0.10 \pm 0.19$

$B^0 \to D^-(K^+\pi^-\pi^-)\pi^+$

r- Interval	ε_i	$\Delta\varepsilon_i$	$w_i \pm \delta w_i$	$\Delta w_i \pm \delta\Delta w_i$	$\varepsilon_{eff,i} \pm \delta\varepsilon_{eff,i}$	$\Delta\varepsilon_{eff,i} \pm \delta\Delta\varepsilon_{eff,i}$
$0.000 - 0.100$	16.3	0.03	47.39 ± 0.20	-0.43 ± 0.40	0.0445 ± 0.0068	0.0147 ± 0.0137
$0.100 - 0.250$	15.4	-0.10	41.12 ± 0.20	1.61 ± 0.40	0.4847 ± 0.0222	-0.1792 ± 0.0445
$0.250 - 0.500$	20.1	-0.01	31.43 ± 0.17	-0.79 ± 0.33	2.7781 ± 0.0507	0.2355 ± 0.1015
$0.500 - 0.625$	11.8	-0.09	21.36 ± 0.19	0.17 ± 0.38	3.8735 ± 0.0548	-0.0758 ± 0.1096
$0.625 - 0.750$	11.2	0.04	15.27 ± 0.17	-0.91 ± 0.35	5.4134 ± 0.0593	0.3001 ± 0.1187
$0.750 - 0.875$	9.2	0.36	9.04 ± 0.15	-0.95 ± 0.31	6.1702 ± 0.0556	0.5257 ± 0.1112
$0.875 - 1.000$	15.9	-0.23	2.03 ± 0.06	0.05 ± 0.11	14.6735 ± 0.0646	-0.2415 ± 0.1291
Total	$\varepsilon_{eff} = \sum_i \varepsilon_i \cdot \langle 1 - 2w_i \rangle^2 = 33.44 \pm 0.13$					$\Delta\varepsilon_{eff} = 0.58 \pm 0.26$

Tagging efficiency is higher, as expected.

Wrong-tags and asymmetries are compatible → no CS visible effect on FlavorTagger.

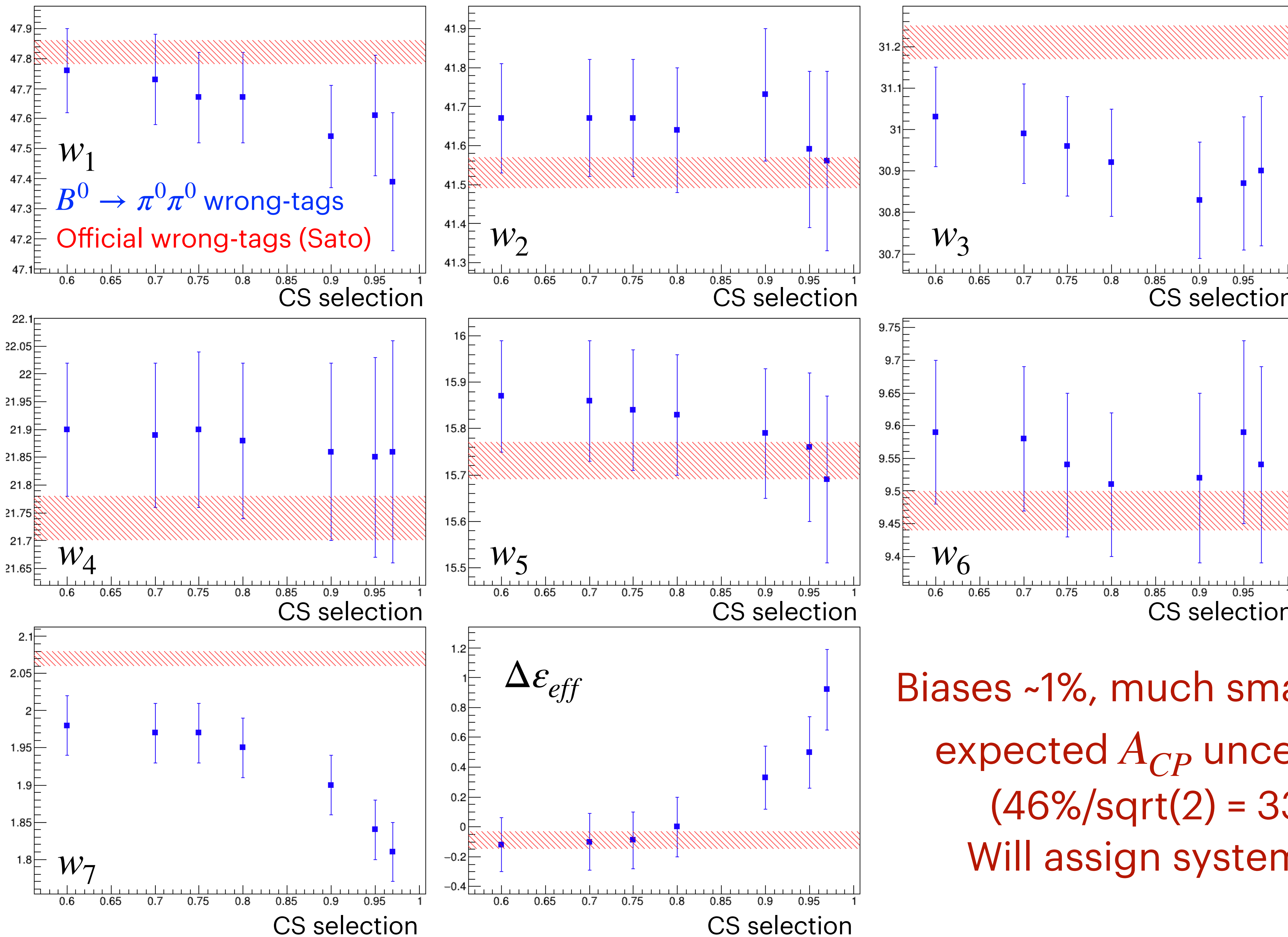
NB: $B^0 \to \pi^0\pi^0$ sample has 2M generated events.

Flavor tagging parameters

ligh

New from last time

Plot wrong tags and $\Delta\varepsilon_{eff}$ in $B^0 \to \pi^0\pi^0$ for various CS selections.



Biases ~1%, much smaller than expected A_{CP} uncertainty (46%/sqrt(2) = 33%).
Will assign systematic.

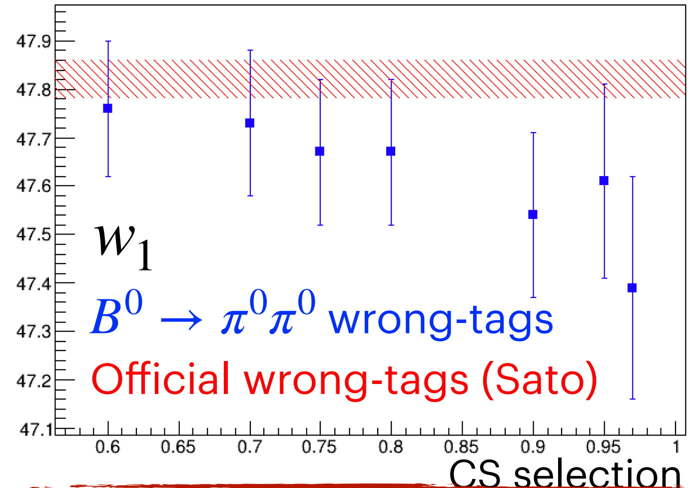
NB: $B^0 \to \pi^0\pi^0$ sample has 2M generated events.

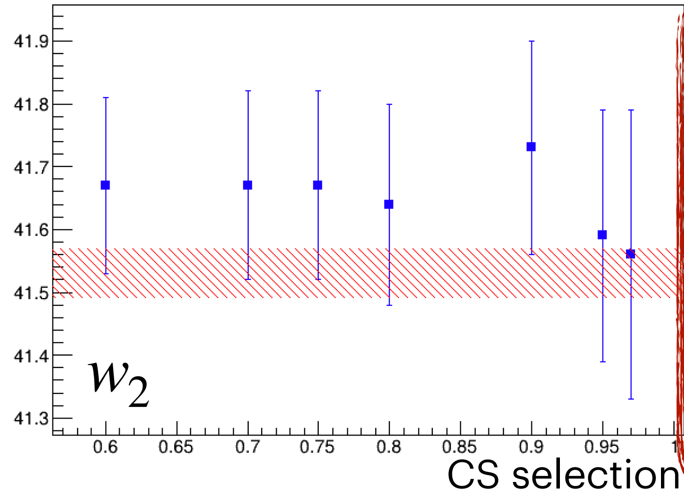
Flavor tagging parameters

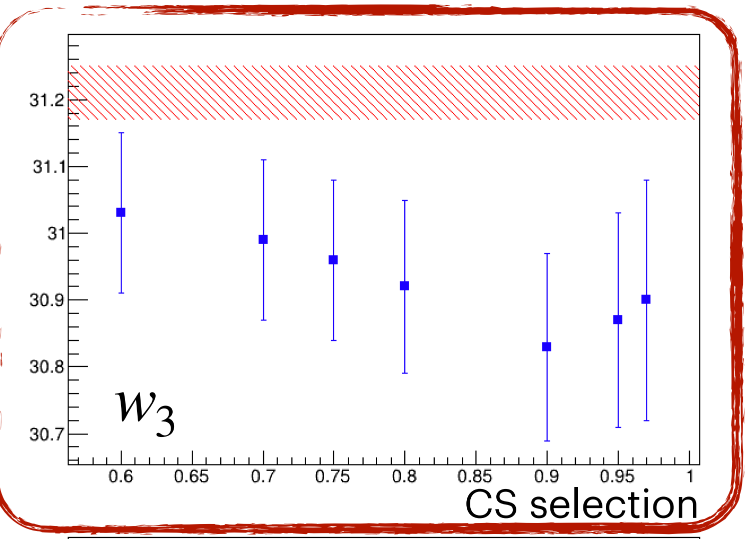
ligh

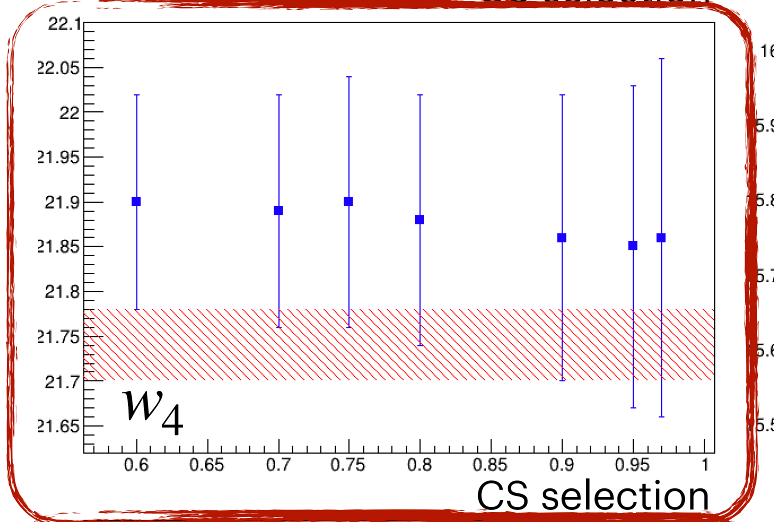
New from last time

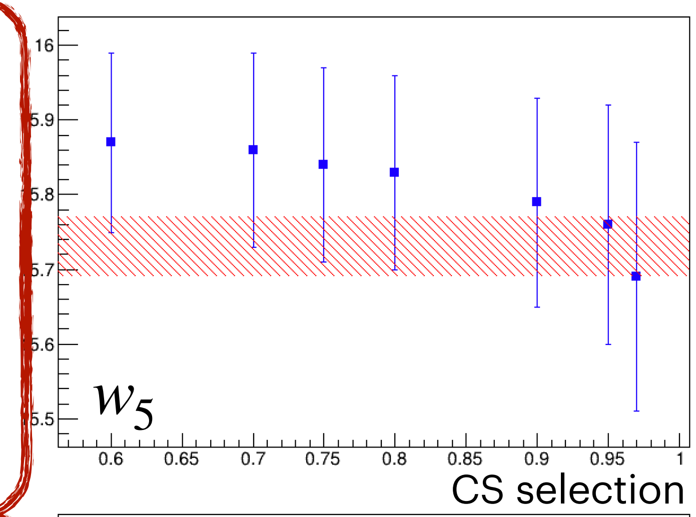
Plot wrong tags and $\Delta\varepsilon_{eff}$ in $B^0 \to \pi^0\pi^0$ for various CS selections.

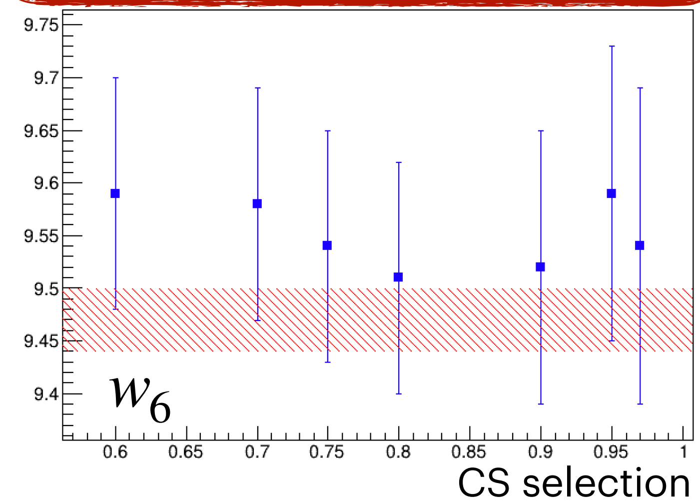


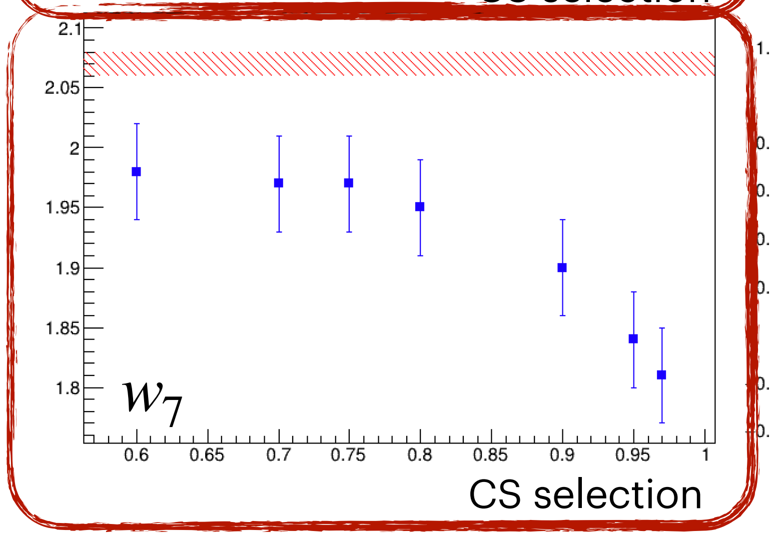


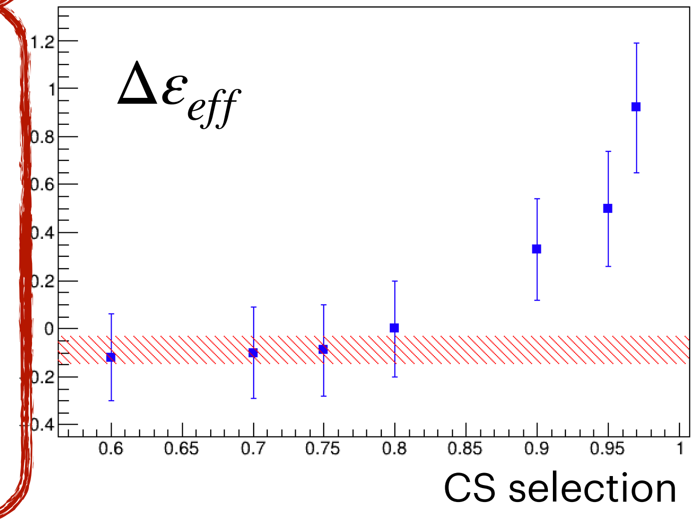












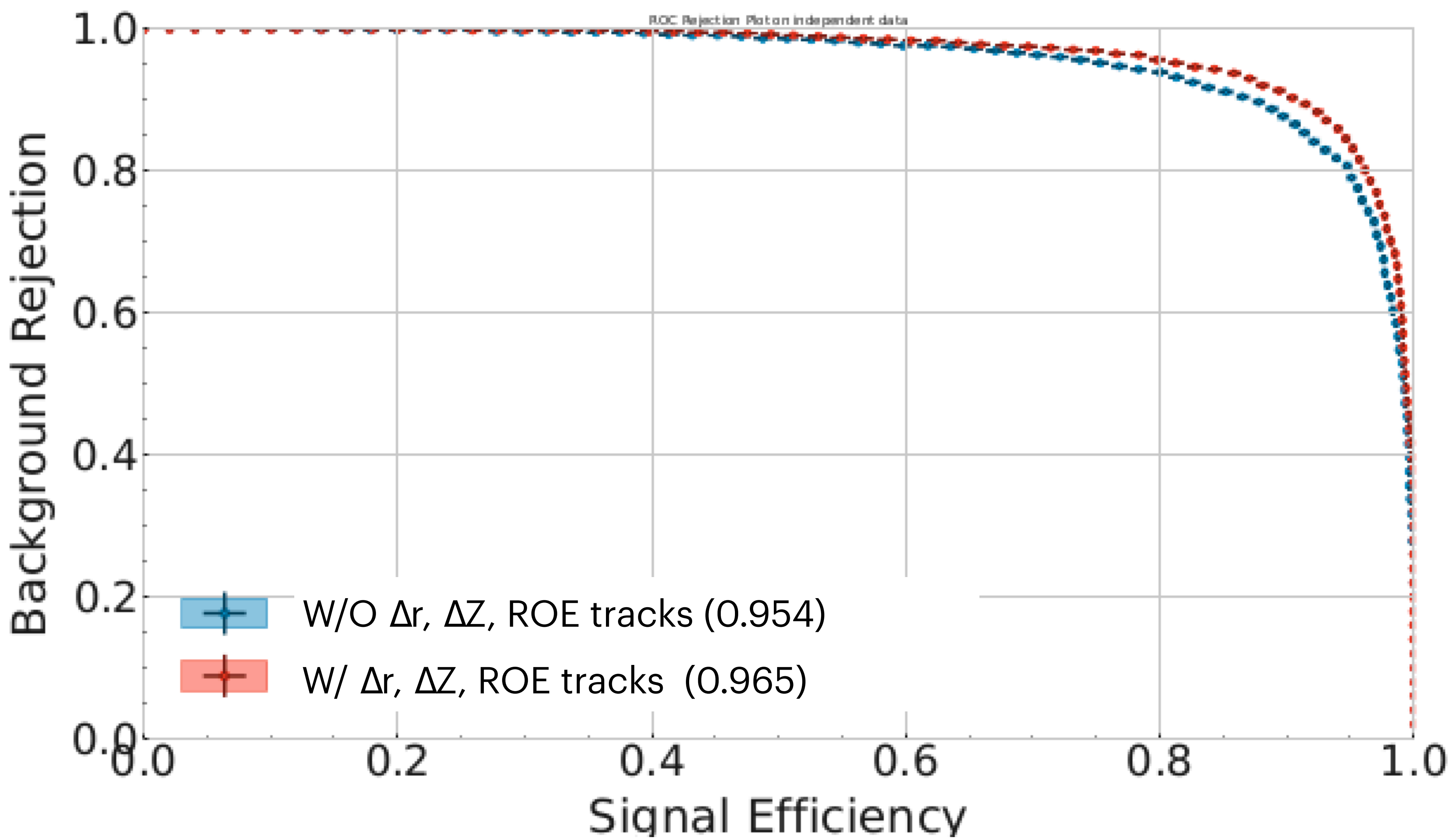
Values are systematically higher or lower than official value.
But, uncertainties in my case (<1000 signal events) will be much larger.

Flavor tagging parameters

ligh

New from last time

Systematic will be ~1%. What about removing flavour tagger variables? How much do we lose in sensitivity?



1.1% loss in AUC if FT variables are removed

At the current precision, using Δr and ΔZ as inputs does not bias the result after CS selection.

Optimise CS selection by minimising *BF* uncertainty.

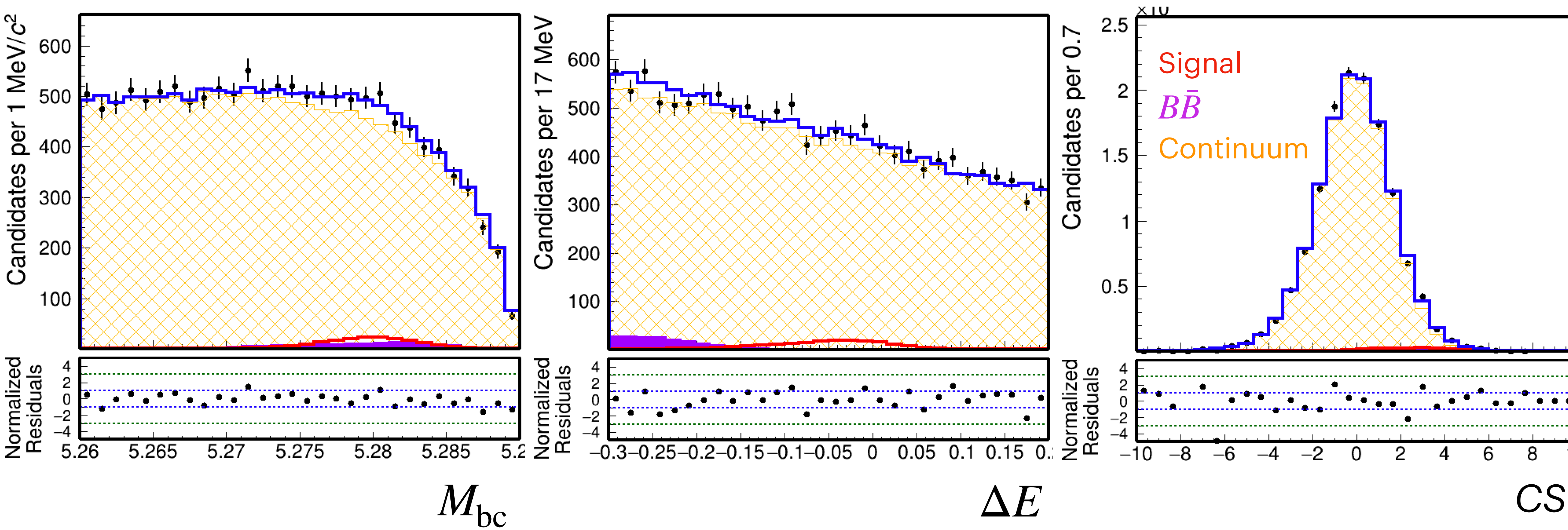
→ Need a fitter

qr-integrated fit of MC sample

Simplified version of the fit: pdf is factorised, no flavor tagger.

Fit realistic $B^0 \to \pi^0\pi^0$ sample of 365 fb^{-1} (qr-integrated) with CS>0.6:

- Non-extended, unbinned ML fit. Components: signal, $B\bar{B}$, continuum.
- Fix all parameters from MC. Only free parameters are *BF* and $B\bar{B}$ yield.



$$pdf = pdf_1(M_{\rm bc}) \times pdf_2(\Delta E) \times pdf_3(CS')$$

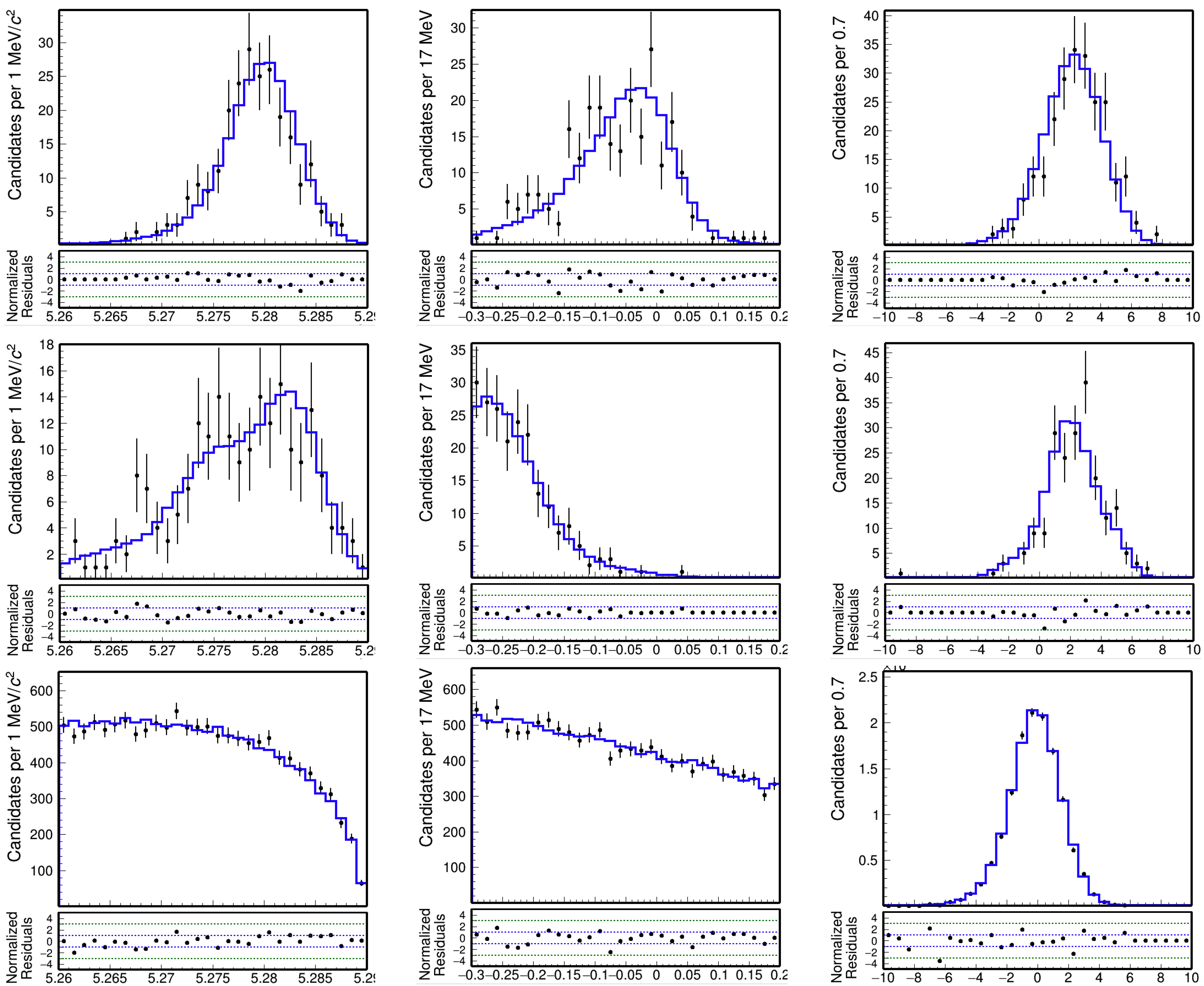
Looks good.

Fit projections

Signal

BBbar

Cont

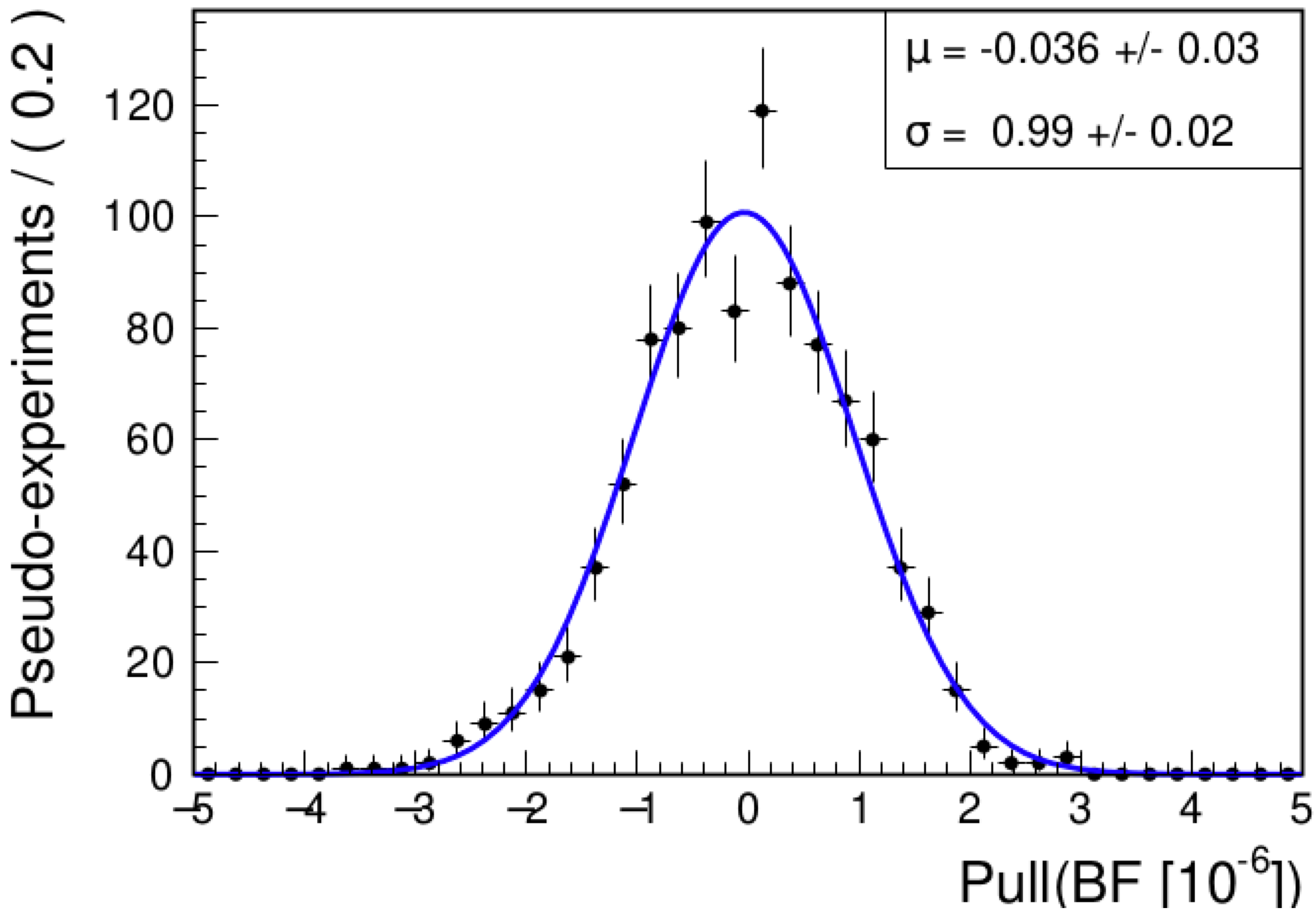


$M_{\rm bc}$ ΔE CS'

Looks good.

Check pulls

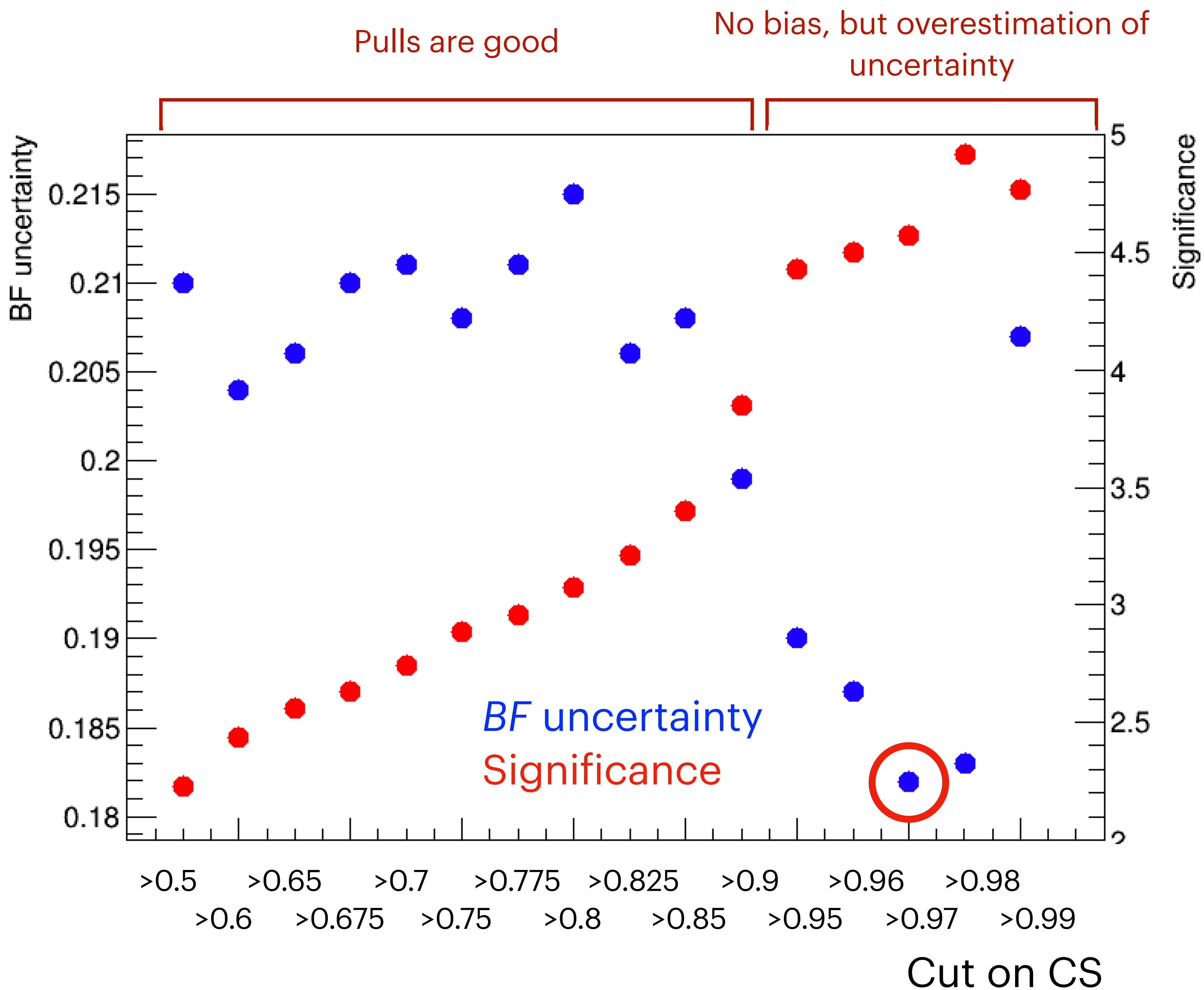
Generate 1000 toys from pdfs and fit them to search for possible biases:



Toys drawn from pdf look fine.

CS selection optimisation

Vary CS selection and minimise *BF* uncertainty. Generate and fit 1000 toys (from pdf) for each one, and compare σ_{BF} and significance.



Selection optimisation using $B^+ \to K^+\pi^0$

New from last time

PhotonMVA

- Optimise photonMVA selection on data.

(wait for the photon energy corrections)

CSBDT

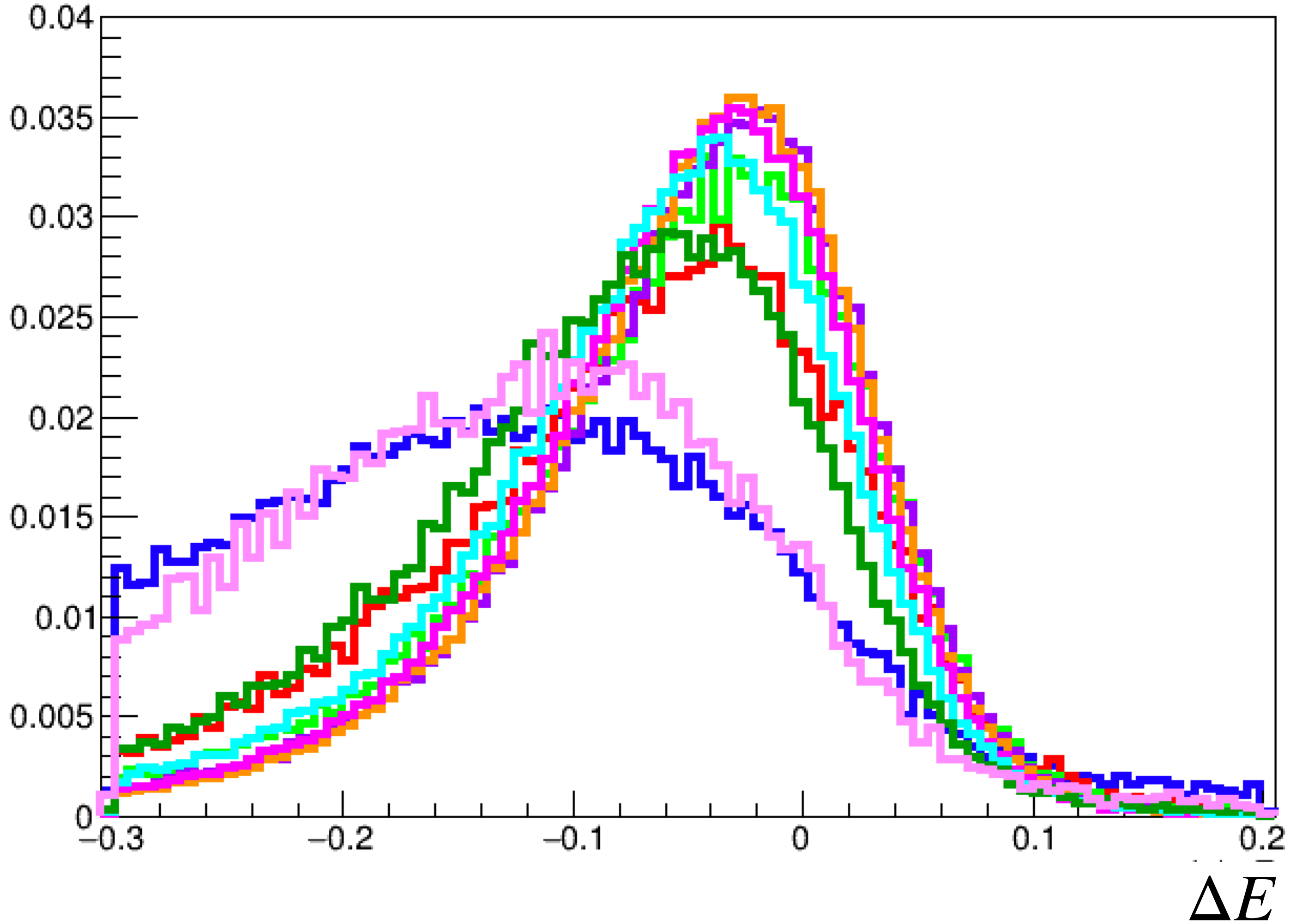
- Optimise CS selection on MC.
- If results are similar to what I obtain in $B^0 \to \pi^0\pi^0$ MC: optimise CS selection of $B^0 \to \pi^0\pi^0$ channel using $B^+ \to K^+\pi^0$ data.

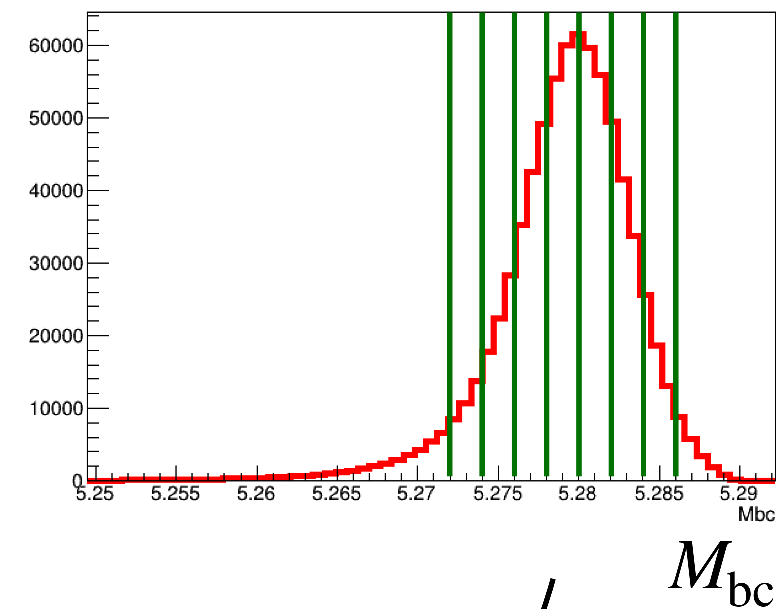
Now I'm reconstructing $B^+ \to K^+\pi^0$ in MC, then I'll pass to data

From factorised pdfs to 2D signal modelling

Signal dependencies: ΔE—M_{bc}

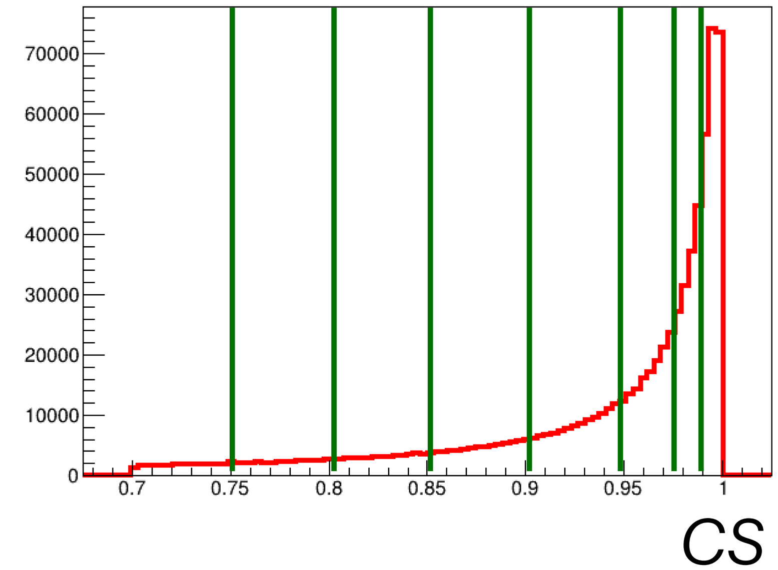
Plot ΔE in slices of M_{bc} in realistic signalMC (@CS>0.7):



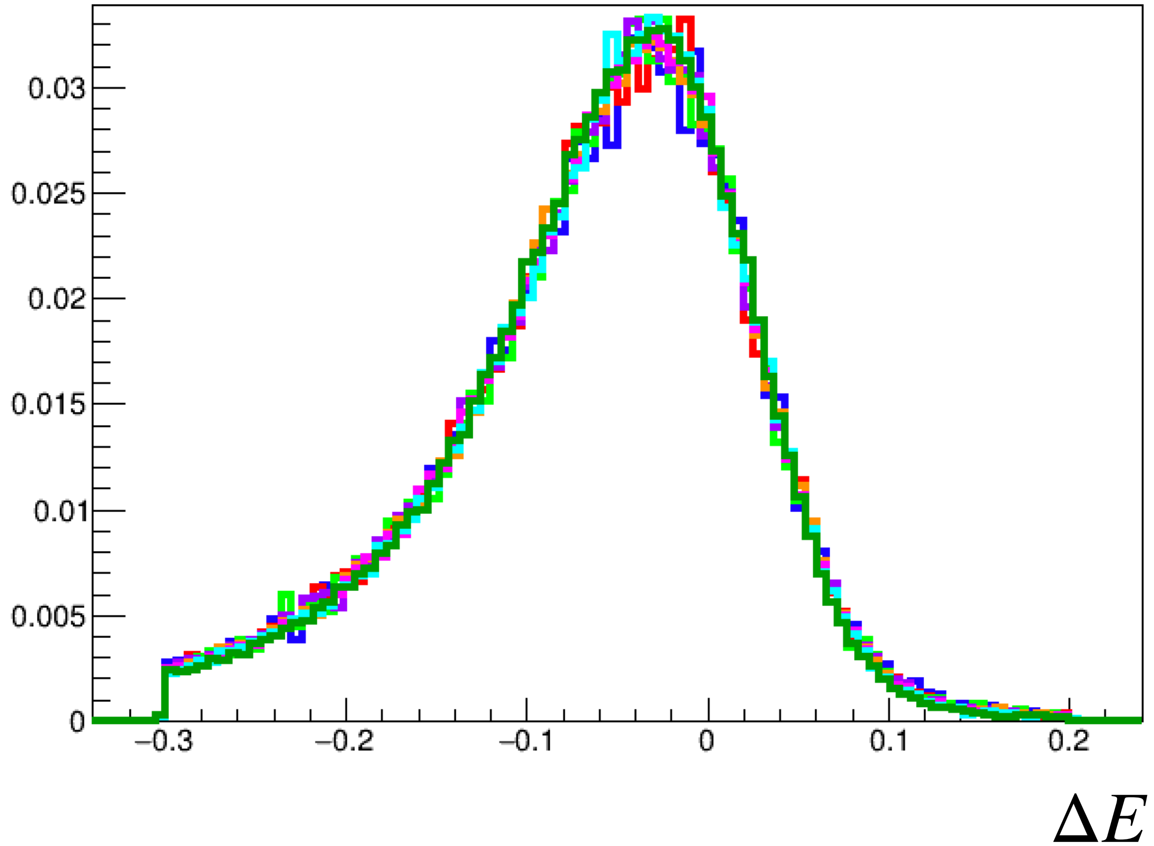


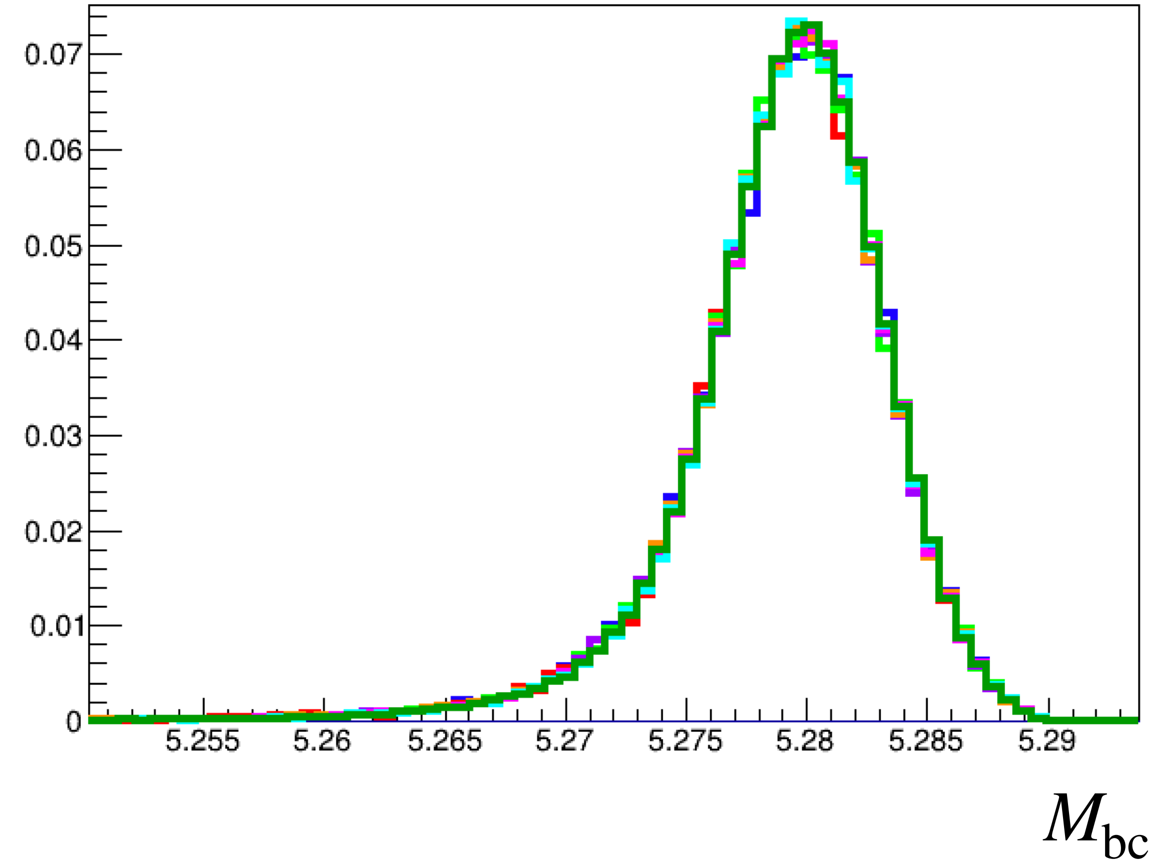
Large ΔE—M_{bc} dependence.

Signal dependencies: ΔE vs CS, M_{bc} vs CS



Plot ΔE and M_{bc} in slices of *CS* in realistic signalMC (@CS>0.7):

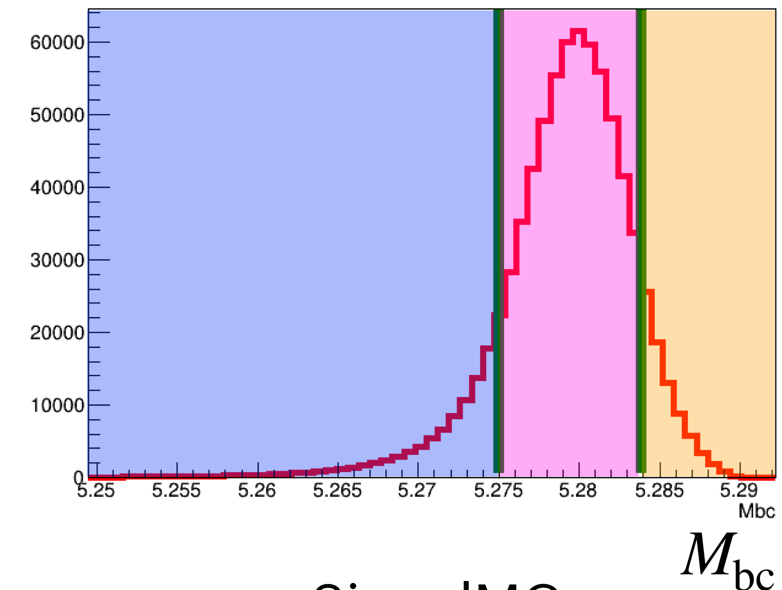


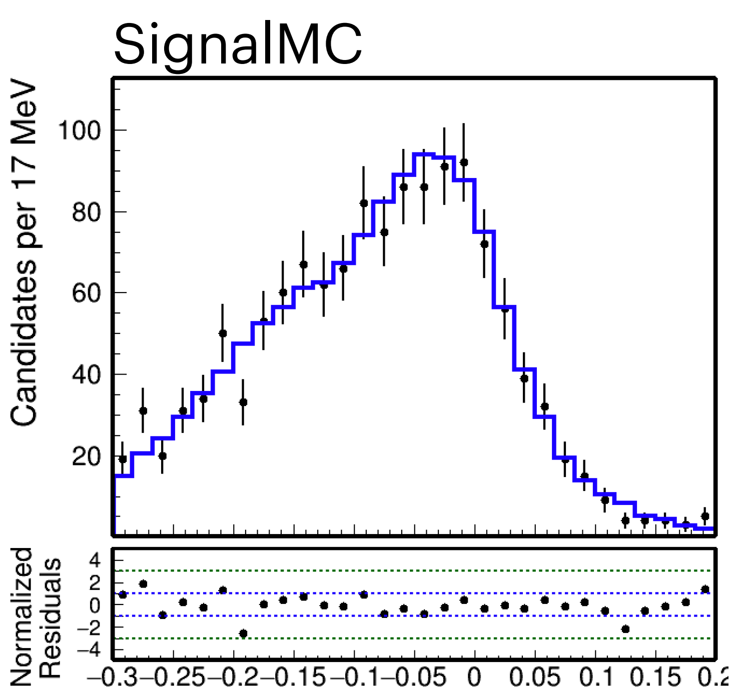


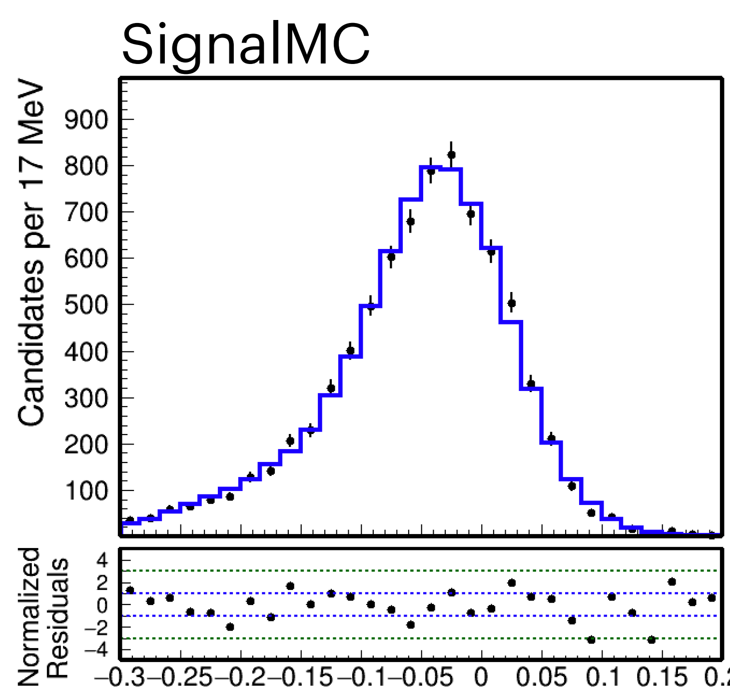
No dependencies in ΔE vs *CS* or M_{bc} vs *CS*.

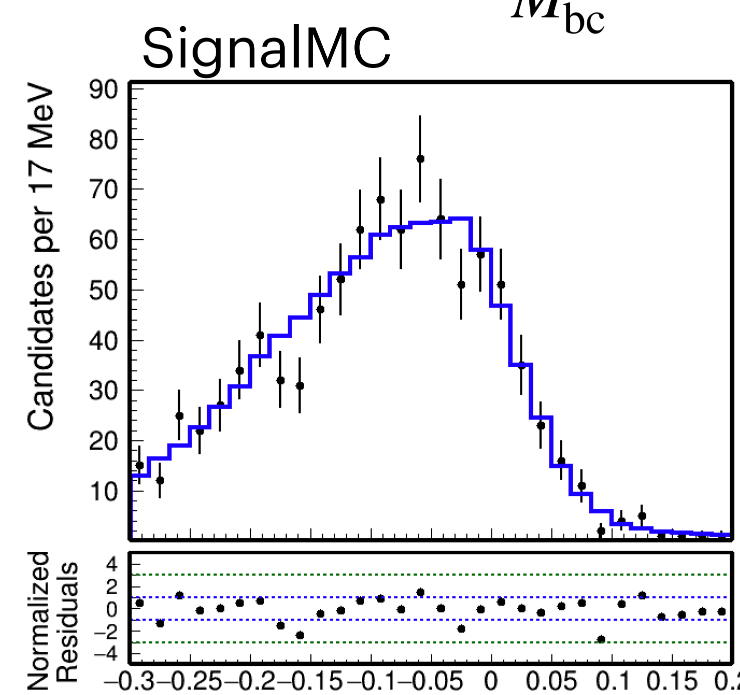
Conditional ΔE—$M_{\rm bc}$ function

Write 2D conditional signal function $f(\Delta E\,|\,M_{\rm bc})$:
different ΔE model for each $M_{\rm bc}$ bin.









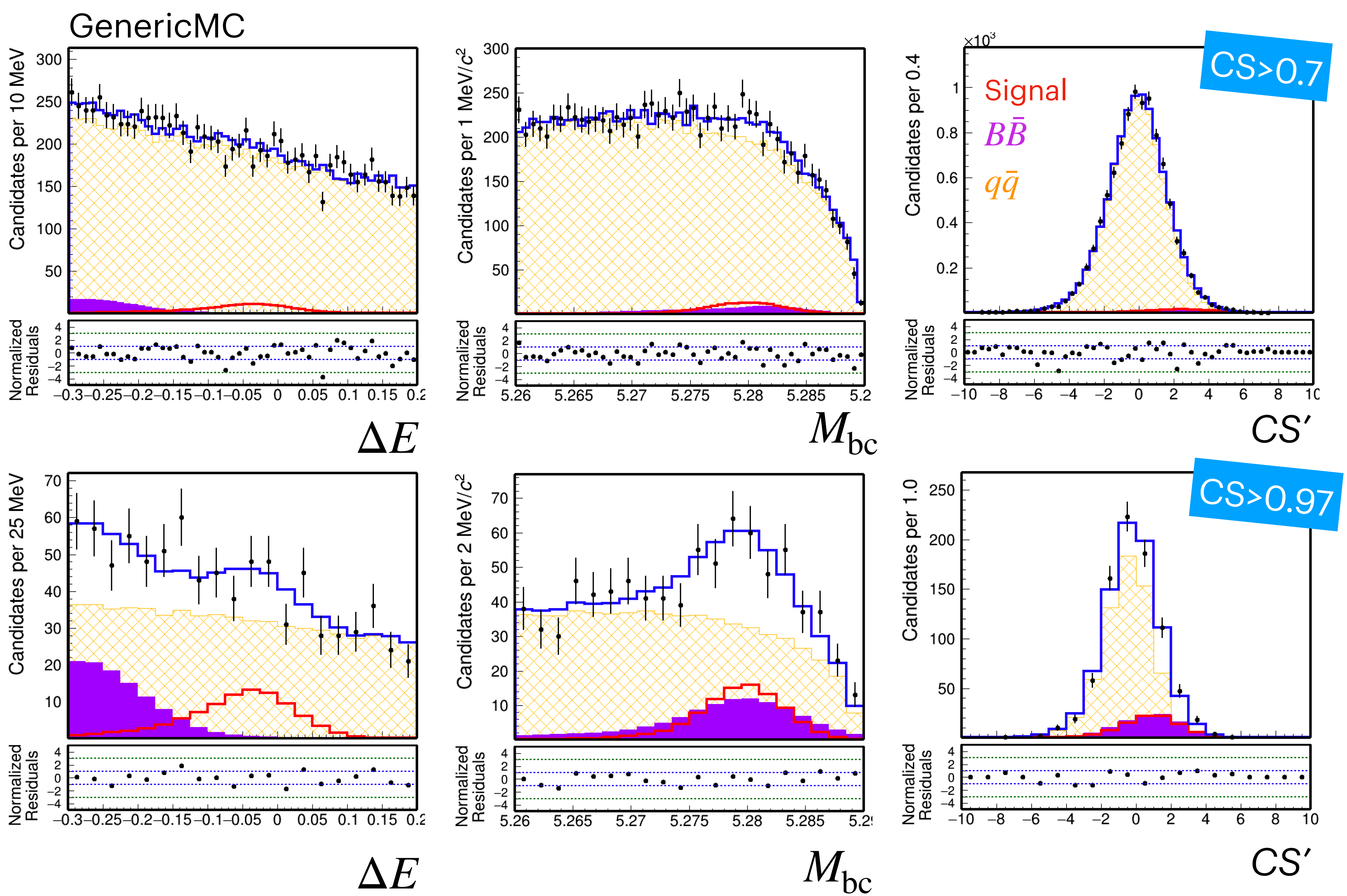
$pdf(\Delta E, M_{\rm bc}) = f(\Delta E, {\rm par}_1)\cdot g(M_{\rm bc})$ if $M_{\rm bc}$ value is in bin 1

$f(\Delta E, {\rm par}_2)\cdot g(M_{\rm bc})$ if $M_{\rm bc}$ value is in bin 2

$f(\Delta E, {\rm par}_3)\cdot g(M_{\rm bc})$ if $M_{\rm bc}$ value is in bin 3

Projections look good.

Fit to 365fb^{-1} MC sample using 2D signal model

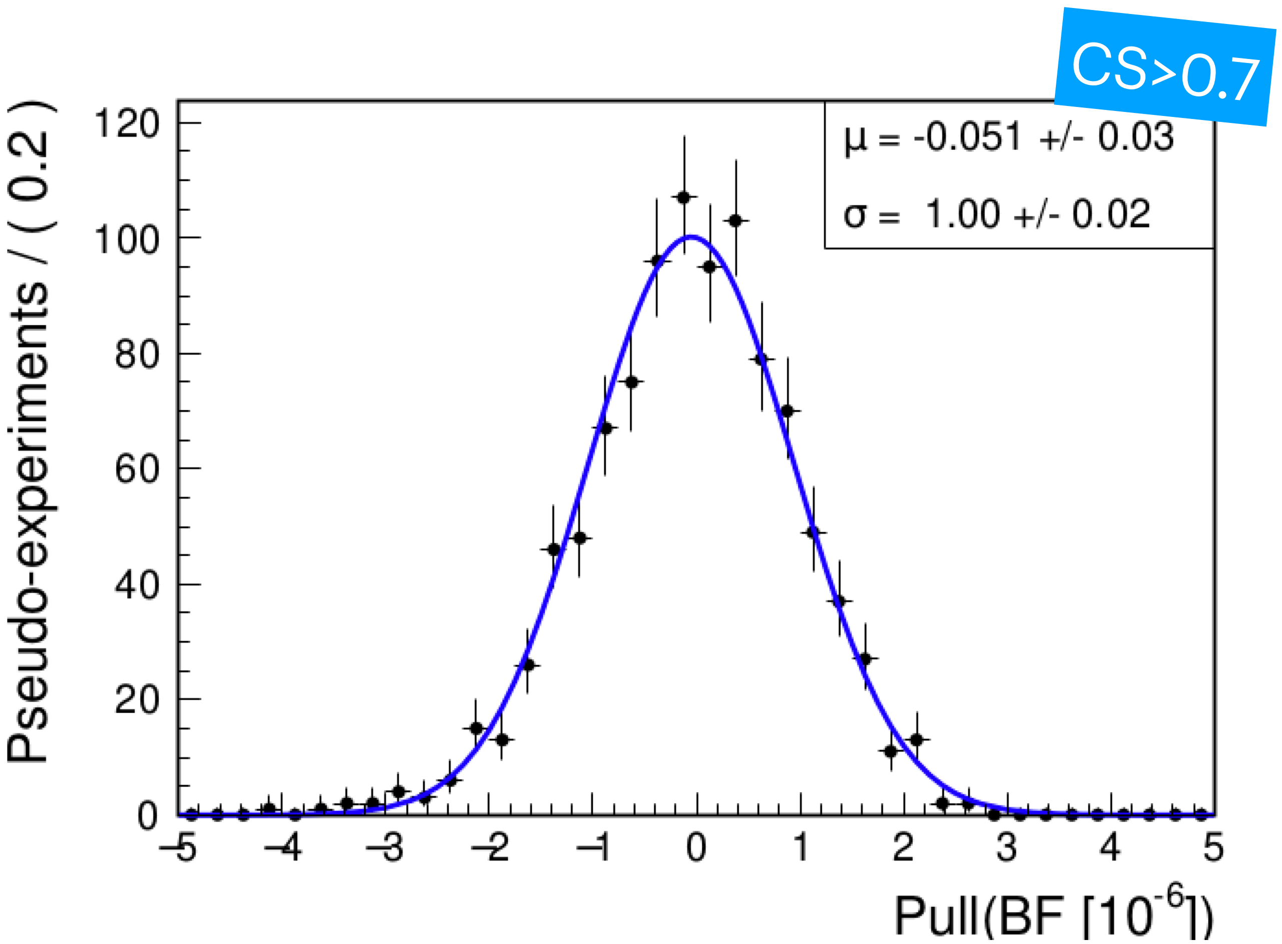


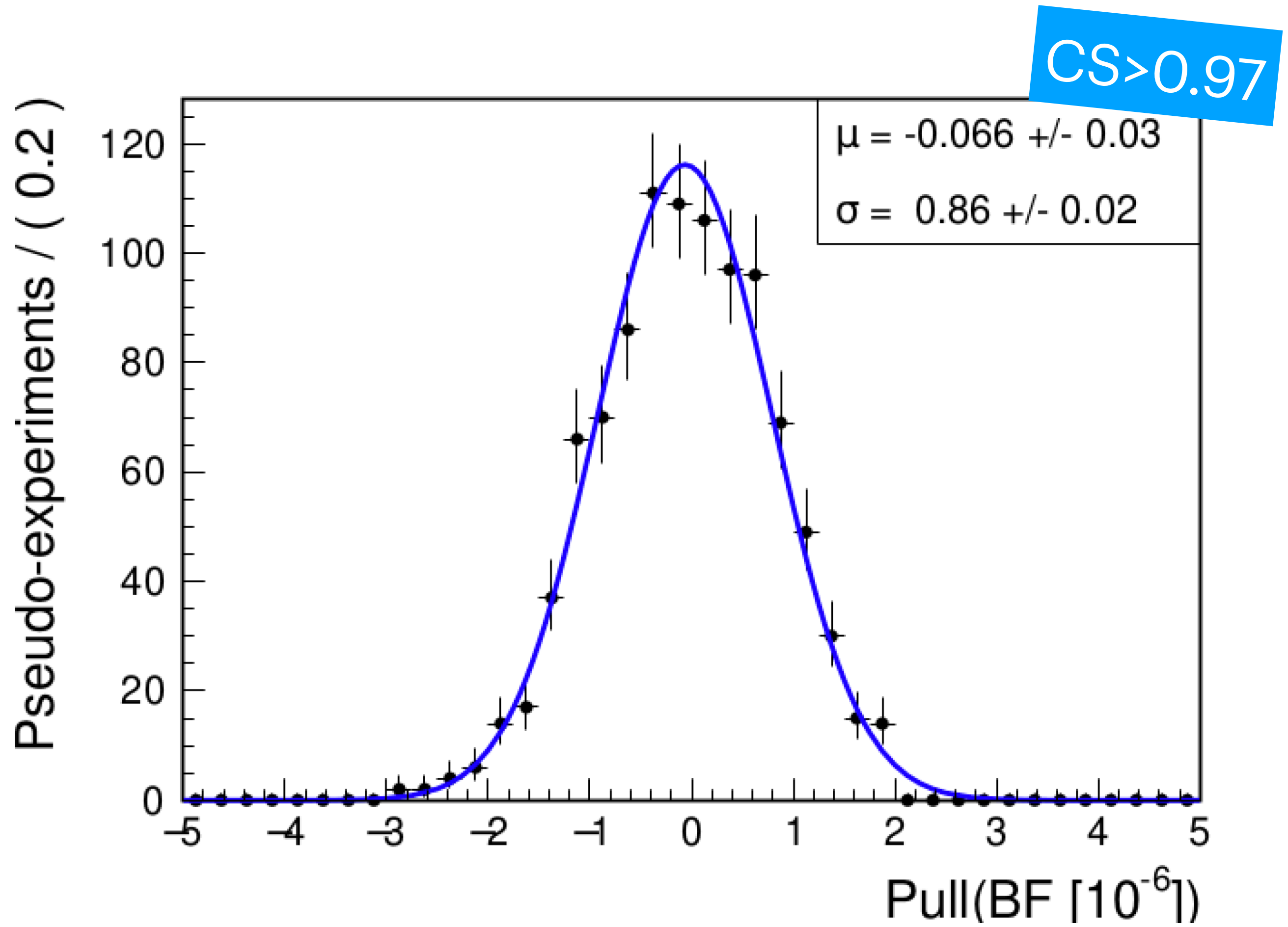
$$pdf = pdf_1(\Delta E, M_{\mathrm{bc}}) \times pdf_2(CS')$$

Projections look good.

Check pulls using 2D signal model

Generate 1000 toys from pdfs and fit them:





Pulls look fine.

Next step: pass from ideal toys drawn from pdf to realistic toys (bootstrapped from MC — must pay attention to bootstrap bias).

Summary

Goal: LS1 update of $B^0 \to \pi^0\pi^0$ analysis.

Today:

- closed open points from last time: final set of photonMVA inputs, selection optimisation on MC, final set of CSBDT inputs, check on flavor tagger parameters;

- simplified fitter (factorised likelihood, no flavor tagger) and CS selection optimisation (for *BF* measurement);

- 2D signal model using conditional function.

Next steps:

- optimise photonMVA selection on data;

- finalise CS optimisation on data control sample;

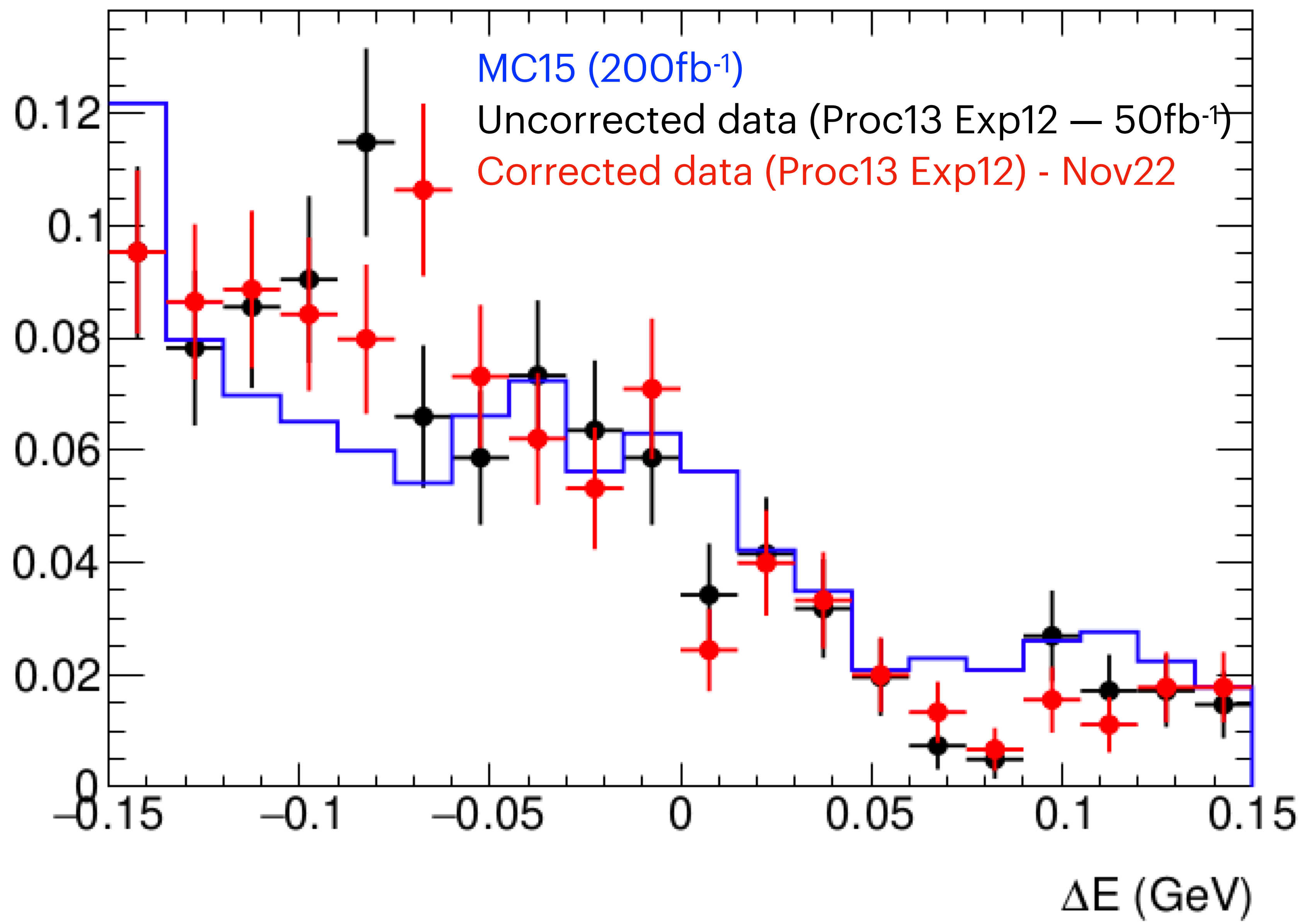
- check pulls on realistic toys (consider dependences between variables).

Backup

Photon-energy corrections

New from last time

Check $B^0 \to D^0(K^-\pi^+\pi^0)\pi^0$ before/after corrections wrt MC.

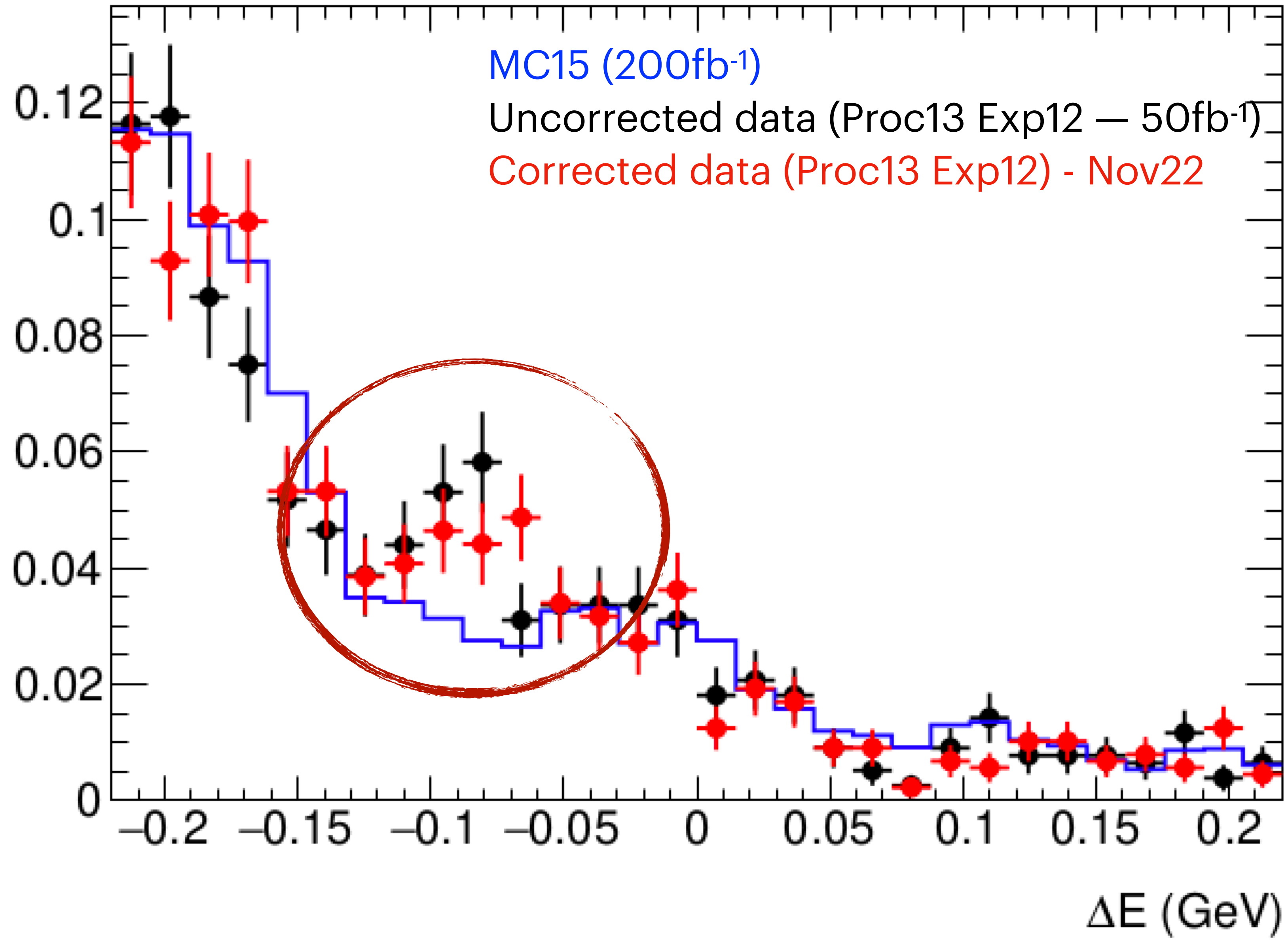


Statistics is still too small. Large shift seems to be present (maybe data $B\bar{B}$ goes out of the range?)

Photon-energy corrections

New from last time

Check $B^0 \to D^0(K^-\pi^+\pi^0)\pi^0$ before/after corrections wrt MC.



Statistics is still too small. Large shift seems to be present.
When all payloads will be ready (all-in-one) repeat with full stats.

Others

New from last time

Yesterday: started working with Benigno on fitter generalisation (n-bins). Seems feasible in short time.

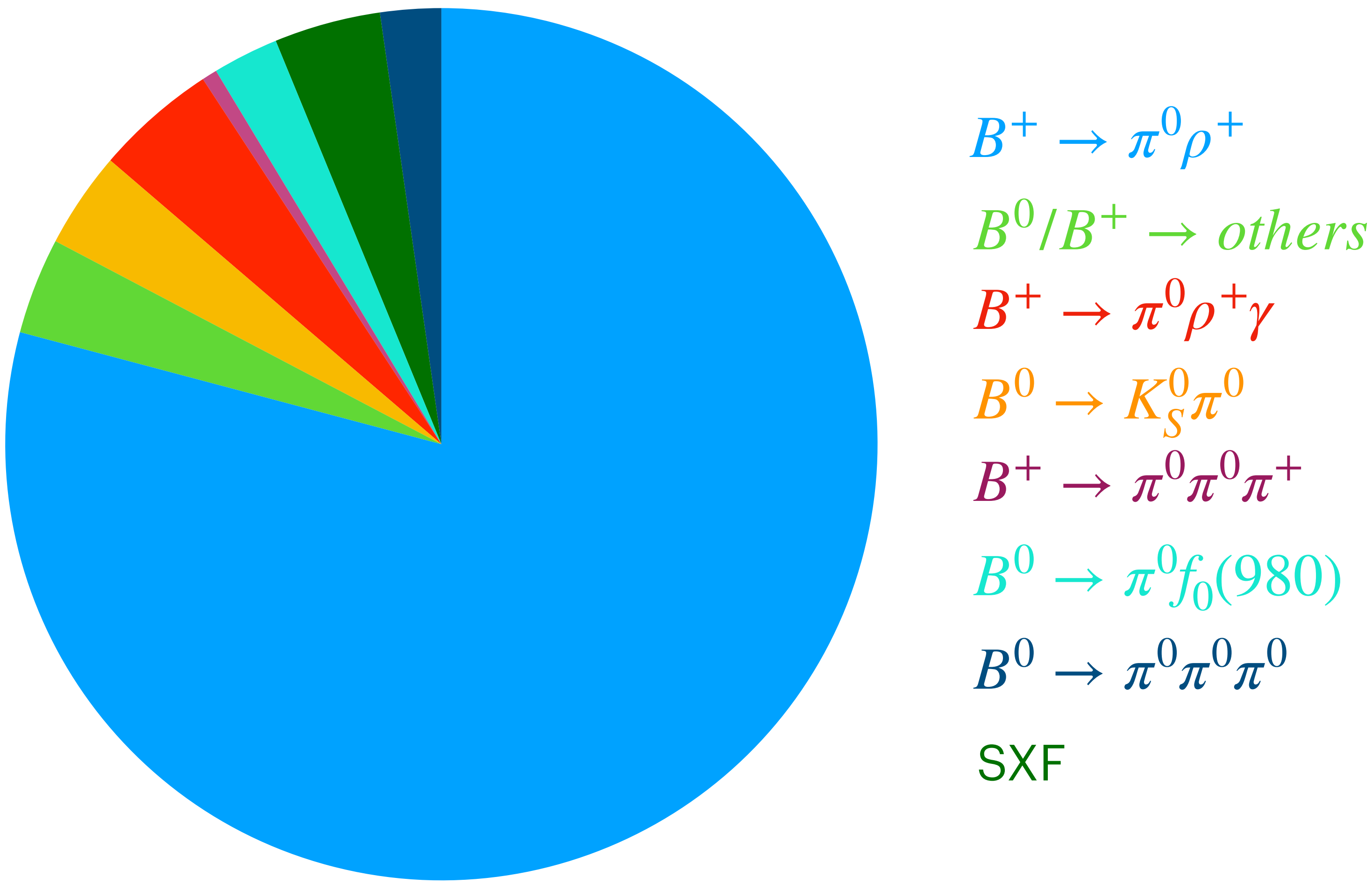
Using different functions for each bin is more tricky and requires more time (but in case there's an "ugly" shortcut).

Topoana

New from last time

I tried using Topoana on $B\bar{B}$ to check if there's something useful (after CS>0.7).
Also, modelling $B\bar{B}$ using the peaking bkg tool could be part of its validation.

$B\bar{B}$ composition in genericMC



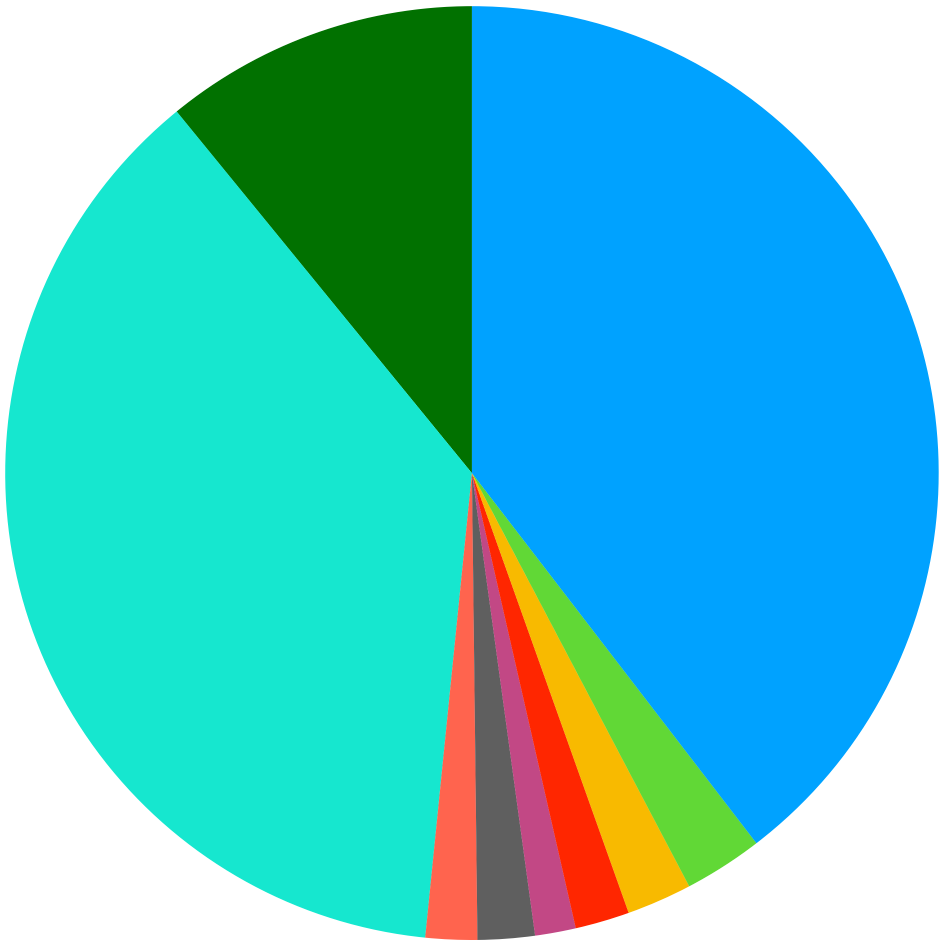
Francis approximation (use $B^+ \to \pi^0\rho^+$ and $B^0 \to K_S^0\pi^0$ only) seems not valid.

Topoana

New from last time

I tried using Topoana on $B\bar{B}$ to check if there's something useful (after CS>0.7).
Also, modelling $B\bar{B}$ using the peaking bkg tool could be part of its validation.

$B\bar{B}$ composition in genericMC (sum of B and $\bar{B}$)



$B^+ \to \pi^0\rho^+$

$B^+ \to others\ (<15\ \mathrm{each})$ (hundreds of decays)

$B^0 \to others\ (<15\ \mathrm{each})$ (hundreds of decays)

$B^+ \to e^+\nu_e\bar{D}^{*0}$

$B^+ \to \pi^0\rho^+\gamma$

$B^0 \to \pi^0\pi^0$

$B^+ \to \mu^+\nu_\mu\bar{D}^{*0}$

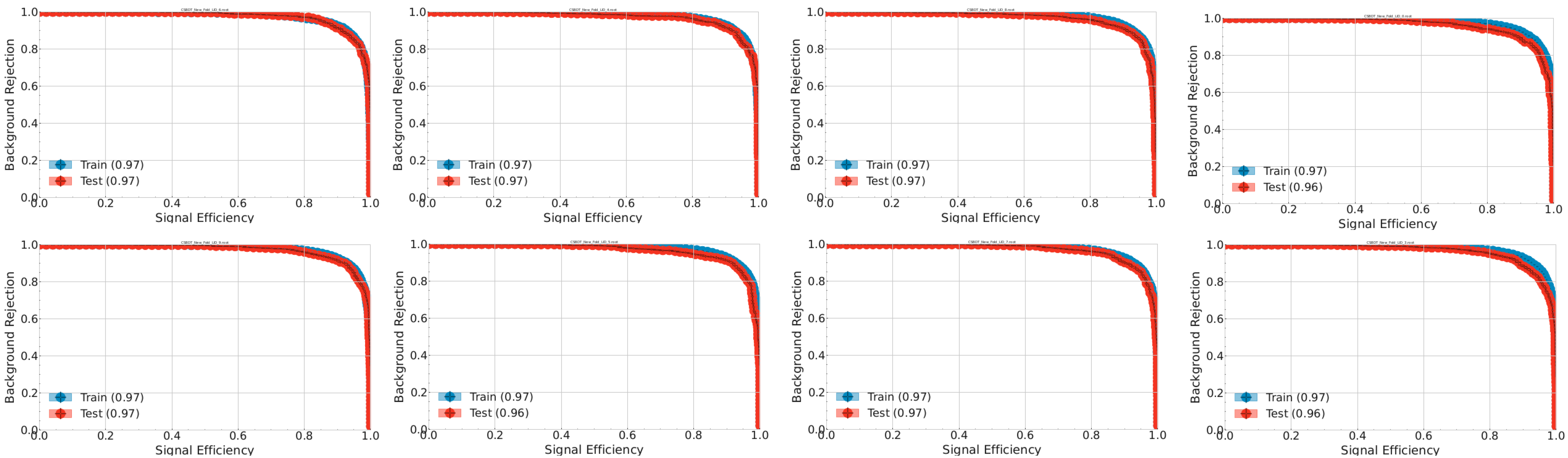
$B^0 \to K_S^0\pi^0$

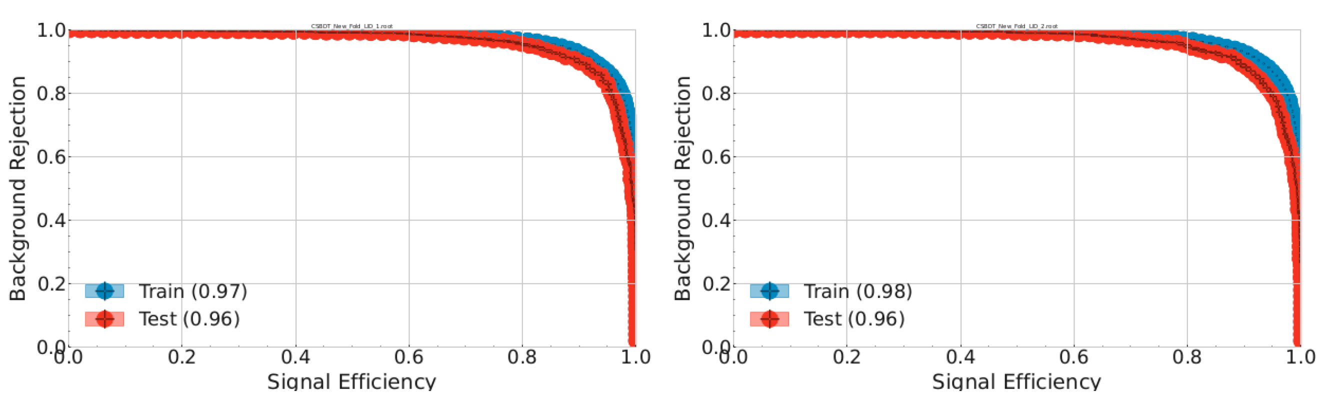
$B^+ \to e^+\nu_e\bar{D}^0$

Francis approximation (use $B^+ \to \pi^0\rho^+$ and $B^0 \to K_S^0\pi^0$ only) seems not valid.
Model $B\bar{B}$ using genericMC (no peaking bkg tool). Too many single decays

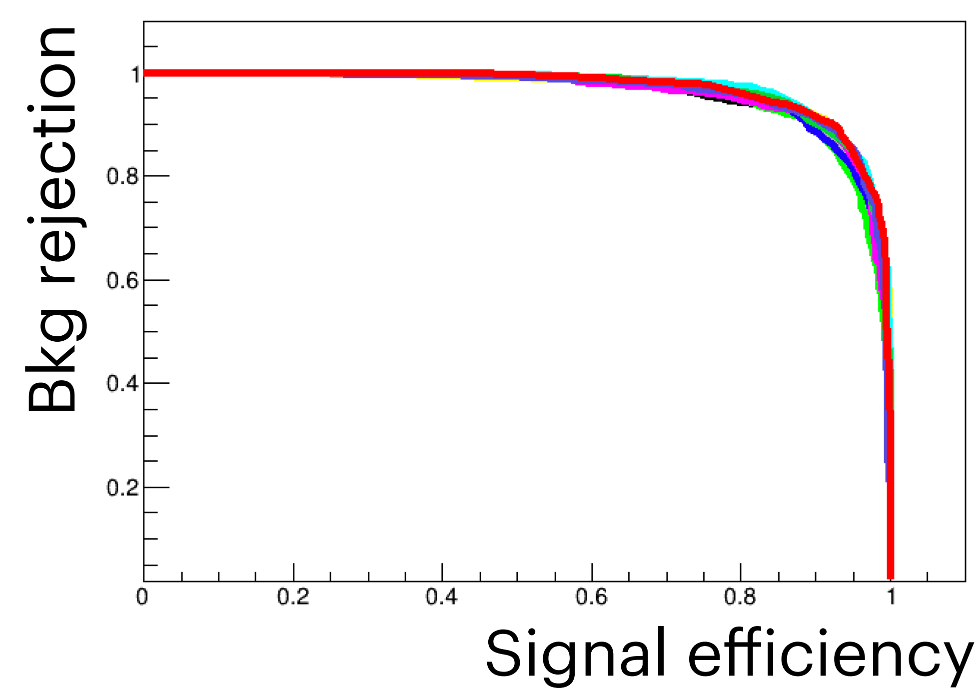
CSBDT: K-fold validation

Go back to default CSBDT, and use all off-res data (possible final configuration). Perform k-fold cross evaluation.



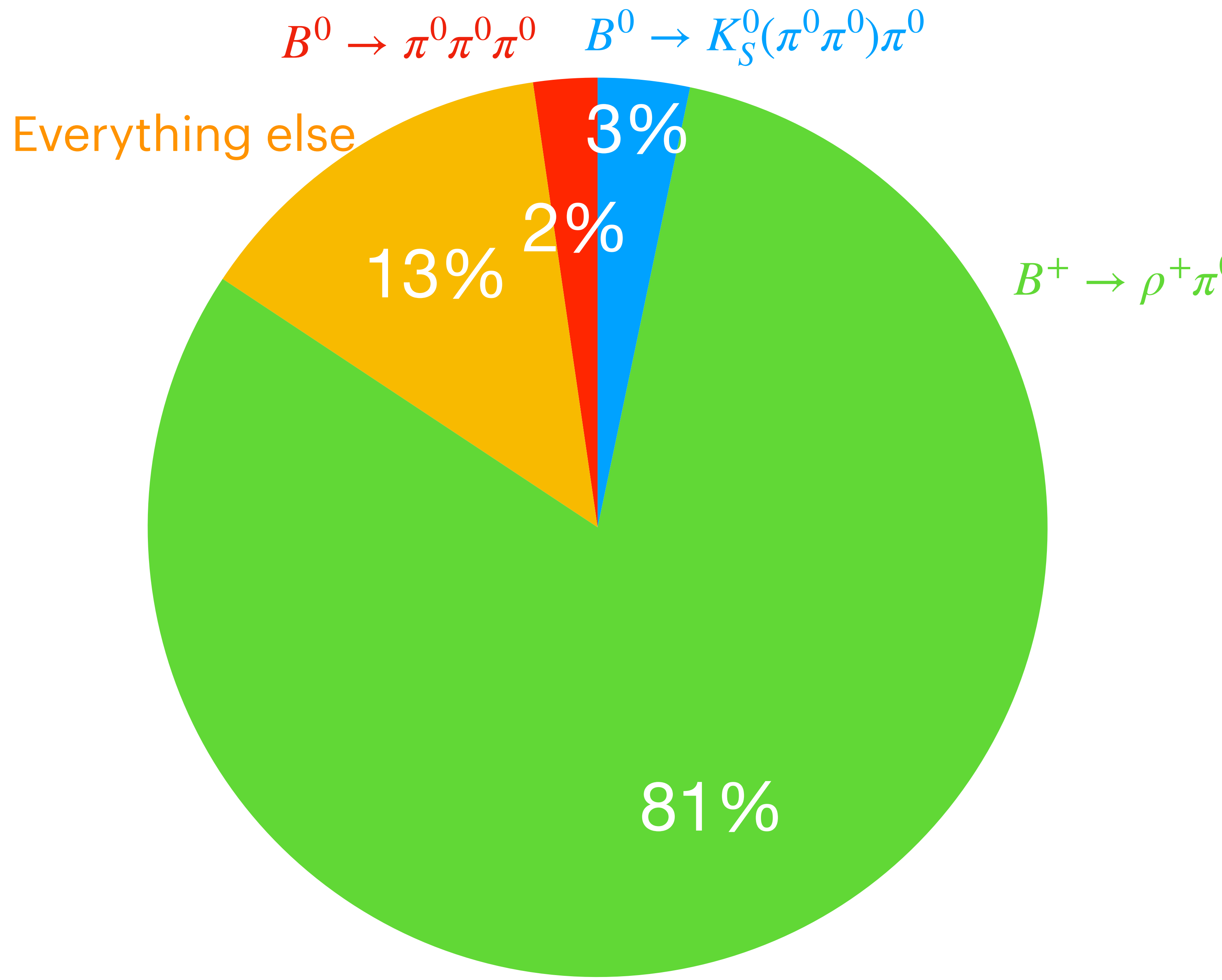


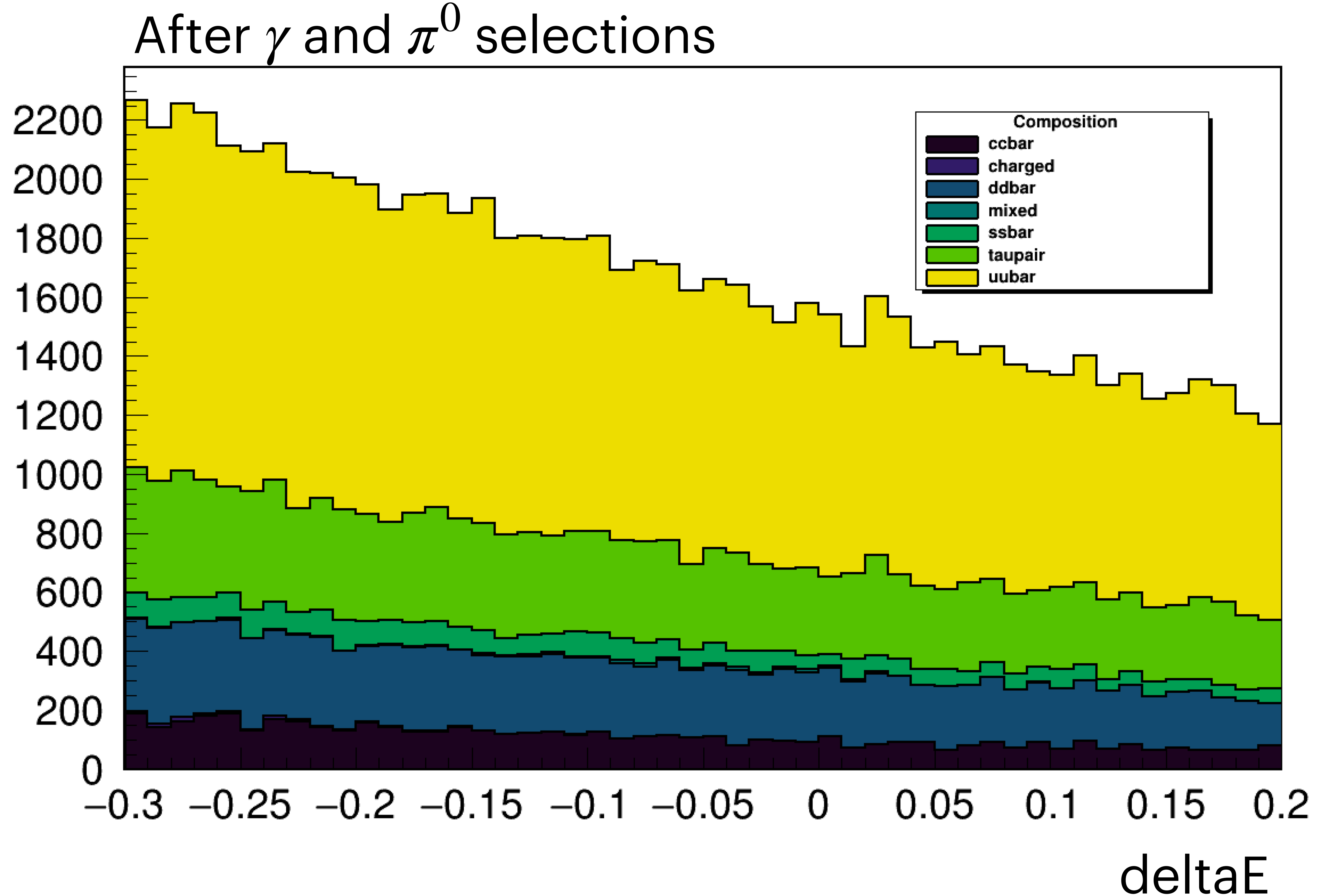
No bias observed. BDT is stable.
Use all data for the training.



BBbar composition for modelling

BBbar composition (CS>0.7):




After γ and π^0 selections
Composition
ccbar
charged
ddbar
mixed
ssbar
taupair
uubar
0
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
−0.3
−0.25
−0.2
−0.15
−0.1
−0.05
0
0.05
0.1
0.15
0.2
deltaE