E_ECL studies for $B^+ \to K^+\nu\bar{\nu}$ against hadronic tag

Claudia Cecchi, Elisa Manoni, Stefano Moneta, Roberta Volpe (INFN & Uni Perugia)

Jacopo Cerasoli, Giulio Dujany, Lucas Martel, Isabelle Ripp-Baudot (IPHC, CNRS)

Joint (S)L / EWP mini-workshop, 05/30/2023

Outline

- Analysis overview and E_ECL usage
- E_ECL studies :
  - definition of the masks
  - first study on the ECL bug
  - photon energy scale corrections
  - thoughts about a correction on the number of extra photons
- Summary and outlook

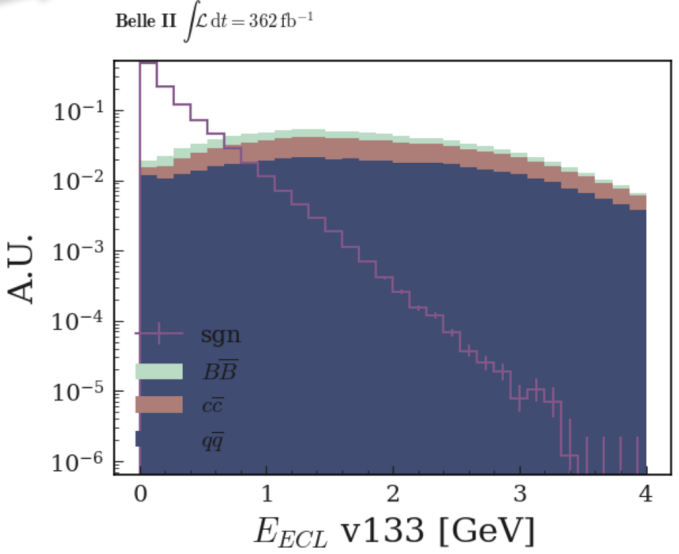
Analysis overview

Search for $B^+ \to K^+\nu\bar{\nu}$ with hadronic FEI tag
(BELLE2-NOTE-PH-2022-045)

Highlights of event reconstruction:

- Reconstruct B_{tag}+K^+ candidate (KID>0.9)
- Request **zero extra** (cleaned tracks, **π^0**, Ks, Λ) in the event
- Train BDT with 14 input variables
- Cut on BDT output, select best candidate according to best Btag (highest FEI signal probability)
- Extract signal strength from fit to BDT output
- Validate analysis in several control regions:
  - wrong Btag-K charge correlation
  - "wrong" KID and piID, lepID sideband
  - J/Psi embedded sample

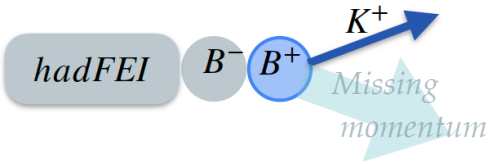
Total extra energy in ECL after preselection
Most discriminant variable in final BDT



Analysis overview

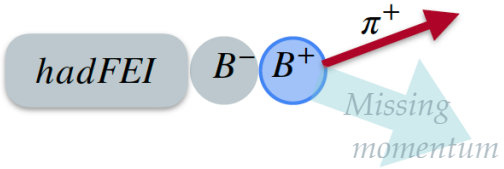
Search for $B^+ \to K^+\nu\bar{\nu}$ with hadronic FEI tag
(BELLE2-NOTE-PH-2022-045)

Definition of sidebands



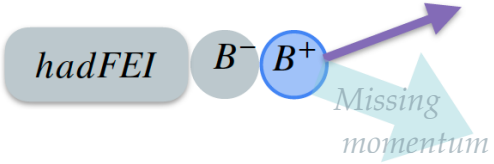
$kaonID > 0.9$

Signal selection



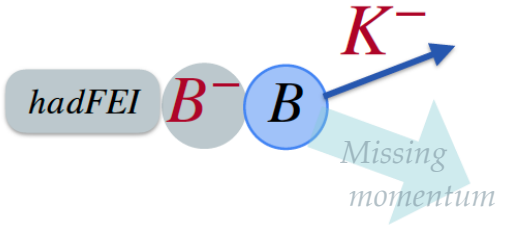
$pionID > 0.9$

Pion sideband



$pionID > 0.5$ and $kaonID > 0.1$

(wrong) Kid sideband

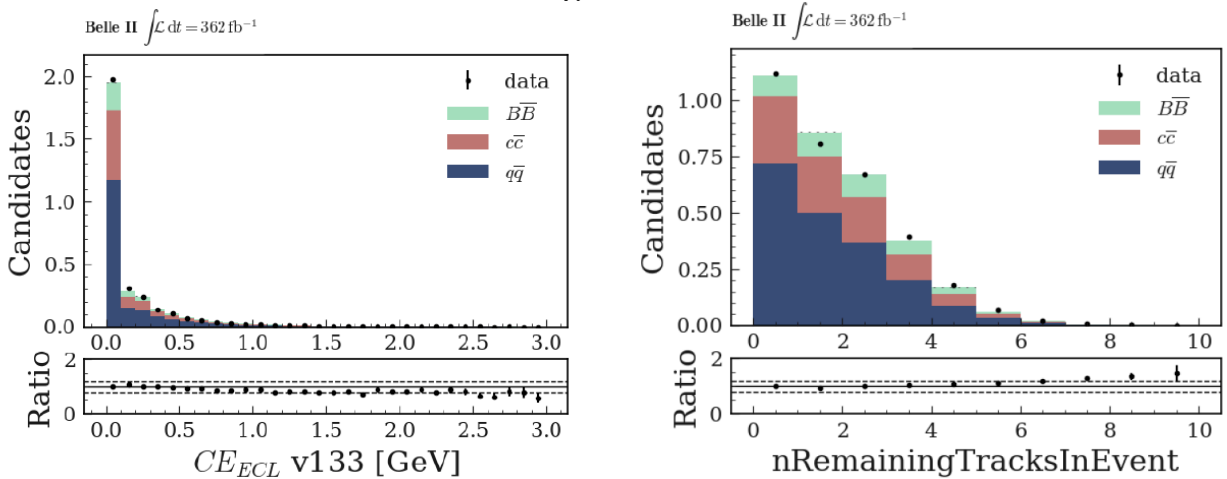


Btag and kaon track have same sign

Wrong charge sideband

Extra energy components

- Charged component: ECL energy associated to extra - remaining tracks, not removed
  N_cleaned_tracks==0 (track-cleaning: dr < 2 cm, |dz| < 4 cm, in CDC accept., nCDCHits>20)



- **Neutral component**: ECL energy associated to extra - photons
- Nominal extra-photon mask:
  - clusterE > (80, 30, 60) MeV in (FWD, BRL, BWD) ECL
  - beamBackgroundSuppression>0.5
  - hadronicSplitOffSuppression>0.3

Definition of v133 (nominal mask)

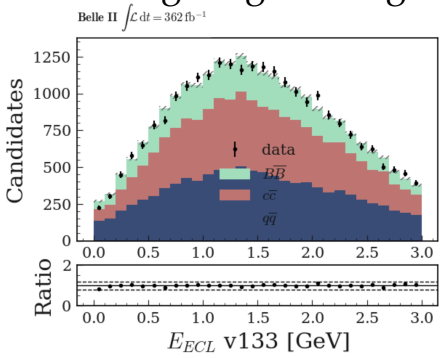
Extra photon mask choice

after pre-selection

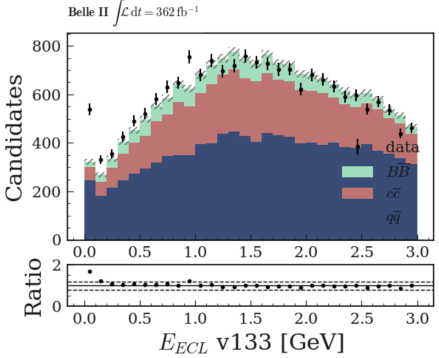
Several extra photon masks tested in the past:

- using different energy threshold
- use timing/minC2T distance to suppress machine bg/hadronic split-off instead of MVA
- Choice of nominal definition based on data/MC agreement in sidebands: **good overall data/MC agreement in 3 out of 4 control samples, both before and after BDT cut**

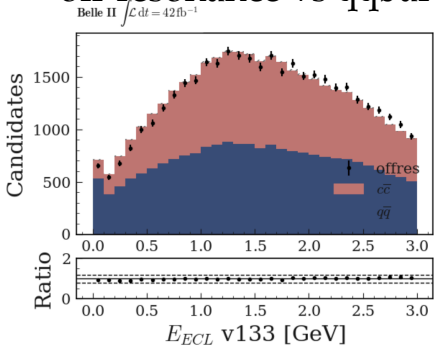
wrong Btag-K charge

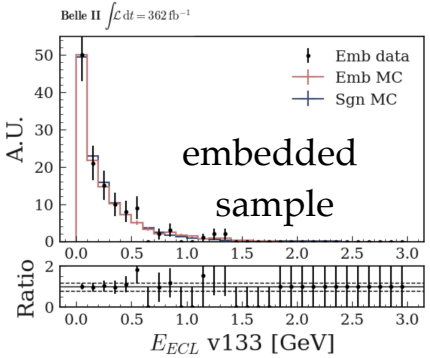


inverted signal K ID (similar behavior in pi sideband)



off-resonance vs qqbar



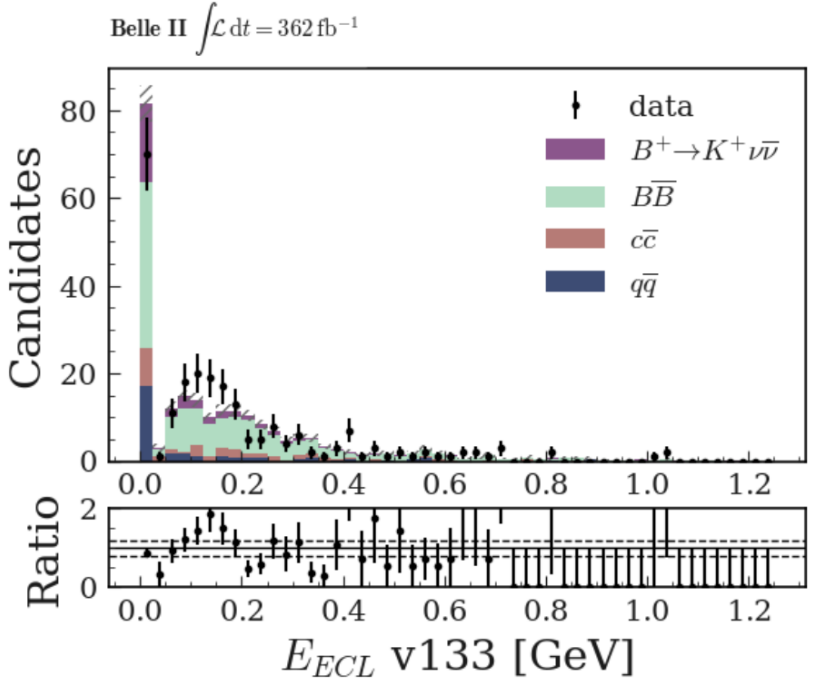


E_{ECL} distribution in signal region after unblinding

E_{ECL} distribution in signal region after final fit

- discrepancy between fitted signal+background and data around 150MeV
- correlation with other variables checked → no hint of mis-modeling due to specific regions/topology

This (mainly) triggered the need for further studies on E_ECL

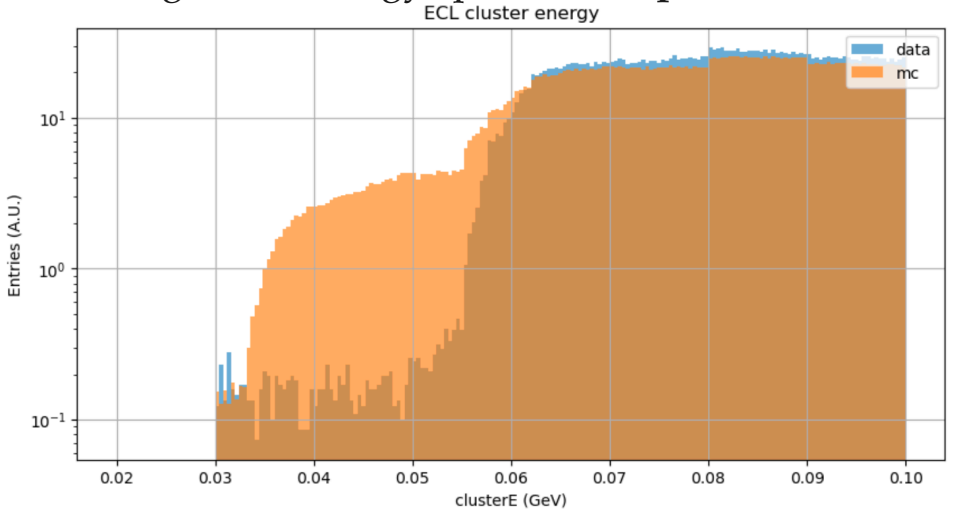


Low energy problem

Disagreement below 60 MeV:

- due to a cut on the ratio timing/timing_error, applied at reconstruction level, for clusters with E<50 MeV

ROEgamma energy spectrum in **pi-sideband**



$E_{cluster}[GeV]$

→ Investigate alternative gammaROE masks with **higher energy threshold** and removal of beamBackgroundSuppression and hadronicSplitOffSuppression BDTs

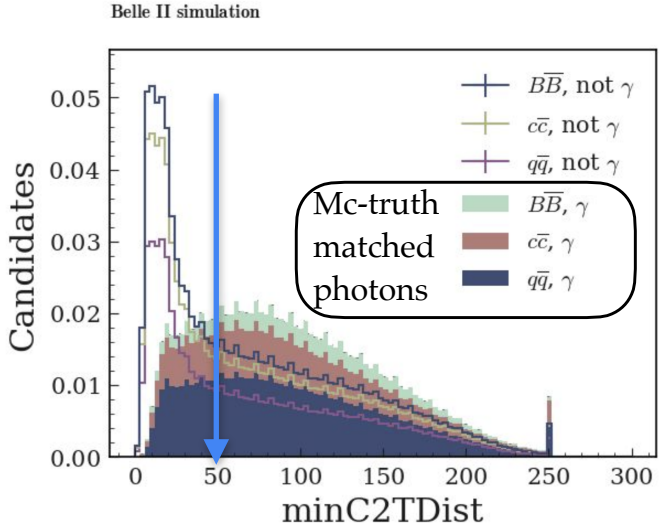
Example of tested masks

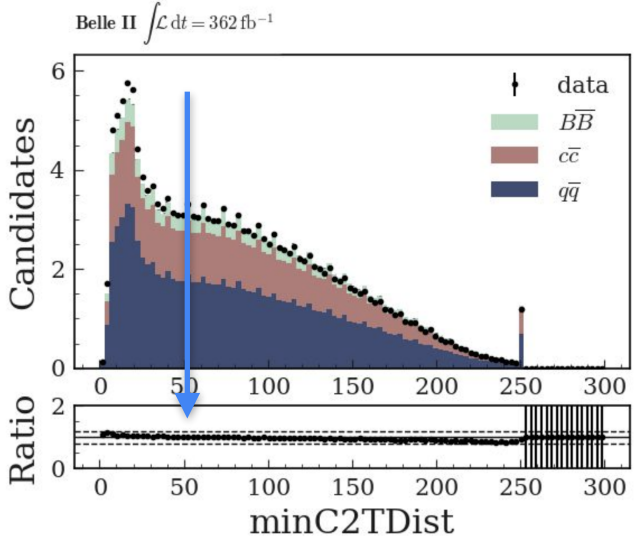
Other masks

	Selection	Comments
v133	E>(80, **30**, 60) MeV, BDTs	Nominal mask
vENE_MC2TD	**E>60 MeV**, θ_{clu} in CDC accept., **minC2Tdist>50 cm**	Removing BDTs and increasing energy threshold

Pion sideband after preselection and:

- E>60 MeV
- θ in CDC accept.

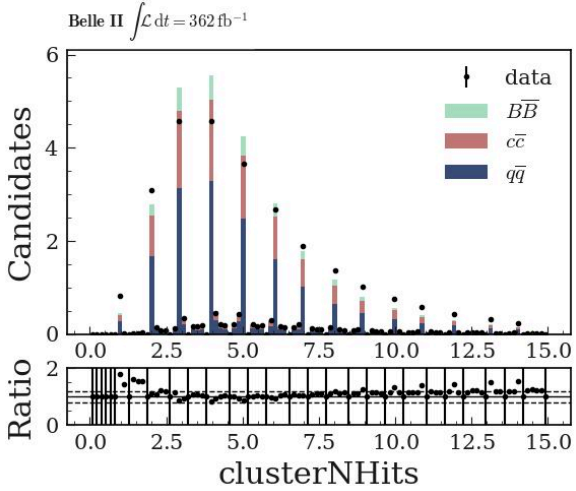


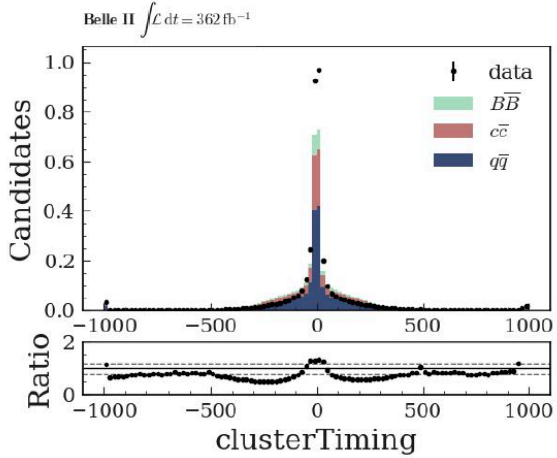


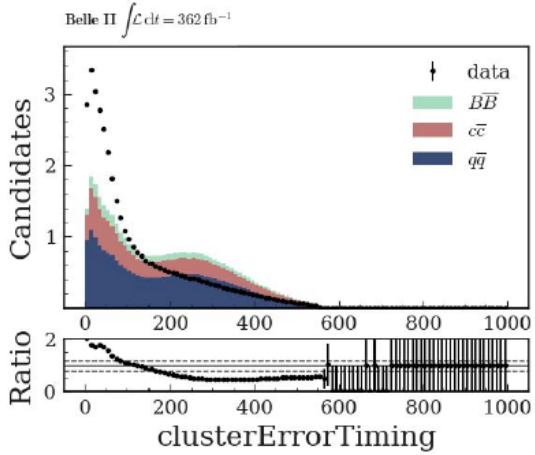
Other variables checked

- Checked also number of ELC hits associated to cluster, recommended in some photon list → poor data/MC agreement

- Timing infos not usable for the same reason



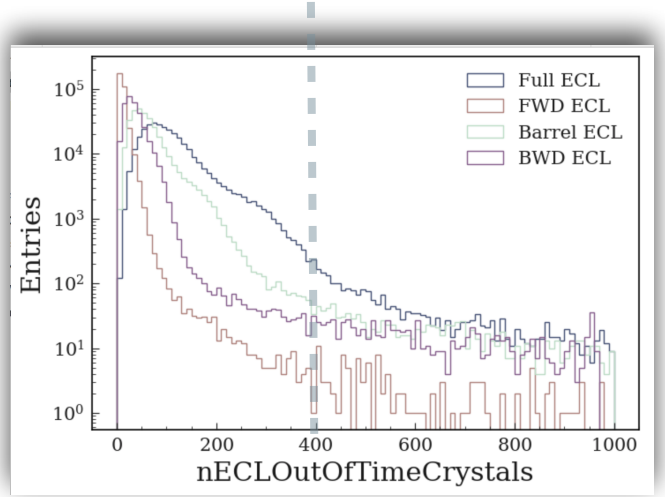


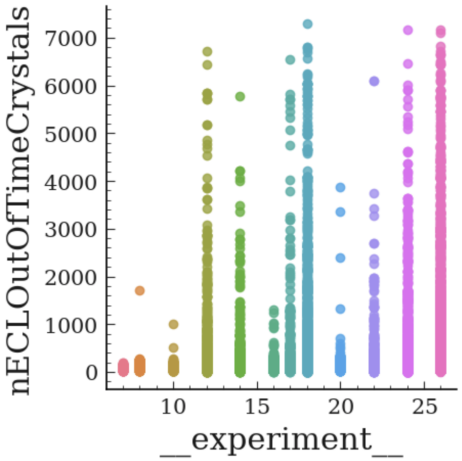


Pion sideband before any photon selection mask

First studies on ECL bug in high energy conditions (I)

- Effective clustering inefficiency for high-energy photons in case of high background conditions (Chris' slides)
- Recommended to study nECLOutOfTimeCrystals event-level analysis variable for selected events,
- problem visible for nECLOutOfTimeCrystals>400
  - plots below for pion sideband
  - fraction of events with nECLOutOfTimeCrystals>400: ~ 1% (similarly in signal region)



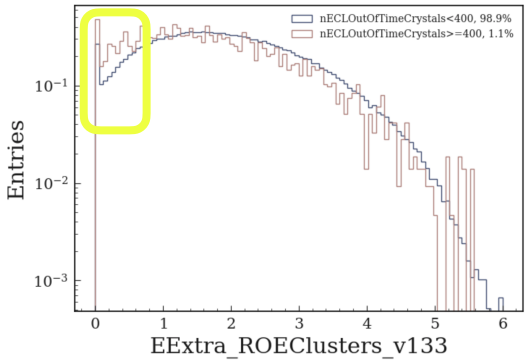


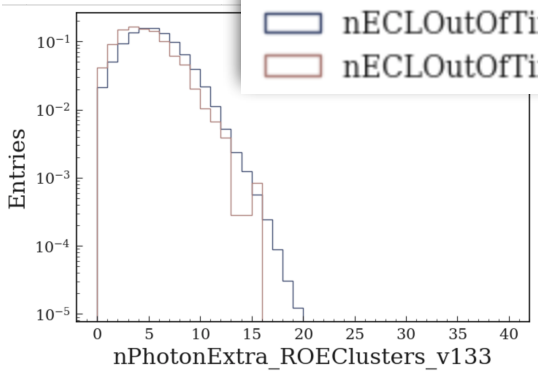
- Higher numbers of out-of-time crystals in exp 12,18,26
- mainly in Barrel and BwdEC

First studies on ECL bug in high energy conditions (II)

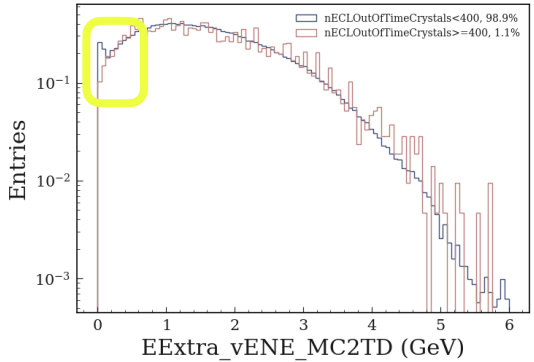
Compare nominal ROE mask (V133) with vENE_MC2TD

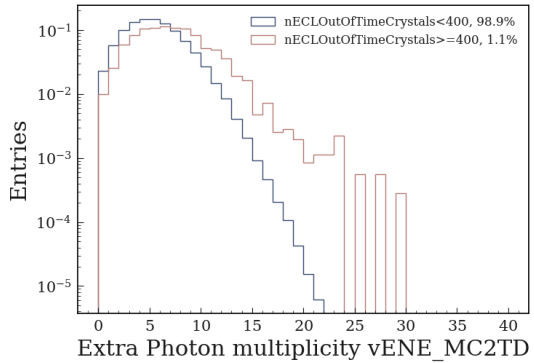
v133





vENE_MC2TD

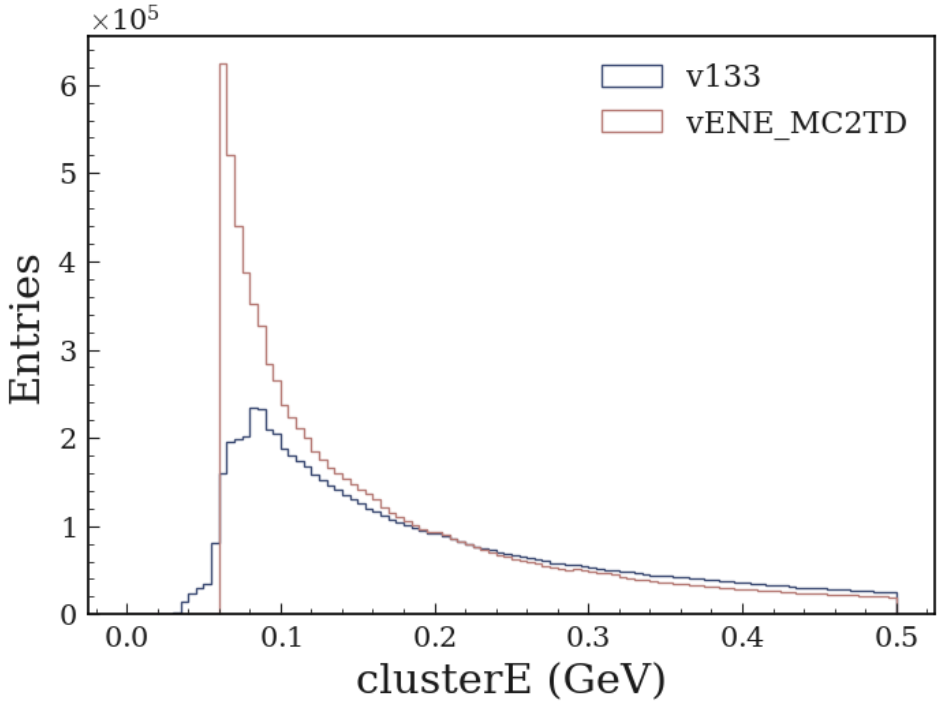




- In v133: events with high numbers of out-of-time crystals has lower photon multiplicity and peak at lower EExtra
- In vENE_MC2TD, the other way around

First studies on ECL bug in high energy conditions (III)

Compare nominal ROE mask (V133) with vENE_MC2TD

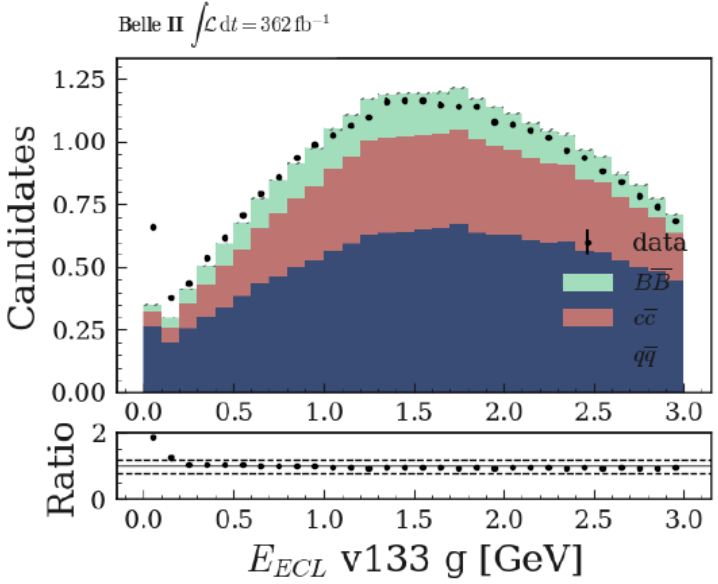


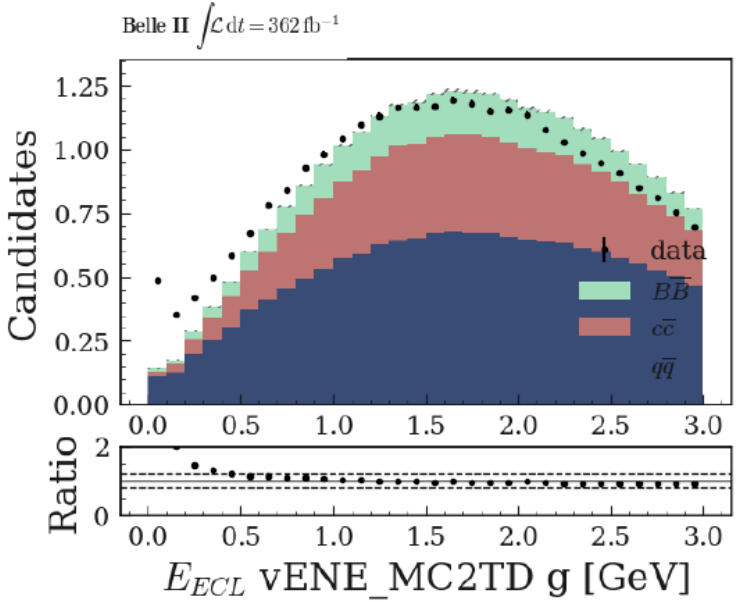
- **v133**: photon requirements tighter than in vENE_MC2TD (mainly due to BDTs in v133)
- **vENE_MC2TD** probably "apparently" less affected by ECL bug (the events affected by the bug are less signal like because of the presence of other clusters)

Plot of E_ECL before the corrections with different masks

Several masks have similar poor agreement in data/MC comparison at low energy

"nominal mask"





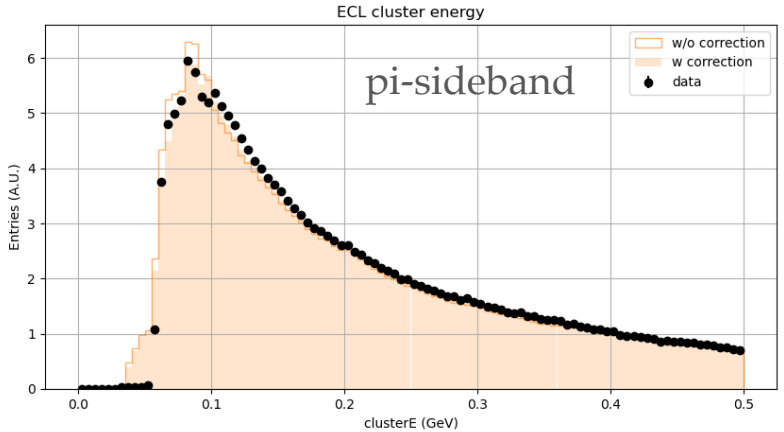
Try to correct the photon energy, both for real and fake/bkg photons

Photon energy scale correction(I)

- No impact from official **photon energy bias correction derived from neutral group, affecting mainly real photons** (because derived from a sample dominated by real photons)
- Attempt to derive energy corrections for clusters not associated to photons
  - Compare ROE gamma energy distribution for data and MC
  - Apply an overall energy scale
  - Perform Kolmogorov-Smirnov test for different correction to chose the one which maximize the agreement

$$E^{n}_{\mathrm{ROE}}(f_h) = \sum_i E^{\gamma}_i + f_h \sum_j E^{n}_j$$

- $i \in$ ECL clusters matched to photons.
- $j \in$ ECL clusters not matched to photons.
- $f_h \equiv$ scale factor quantifying accuracy of energy calibration.



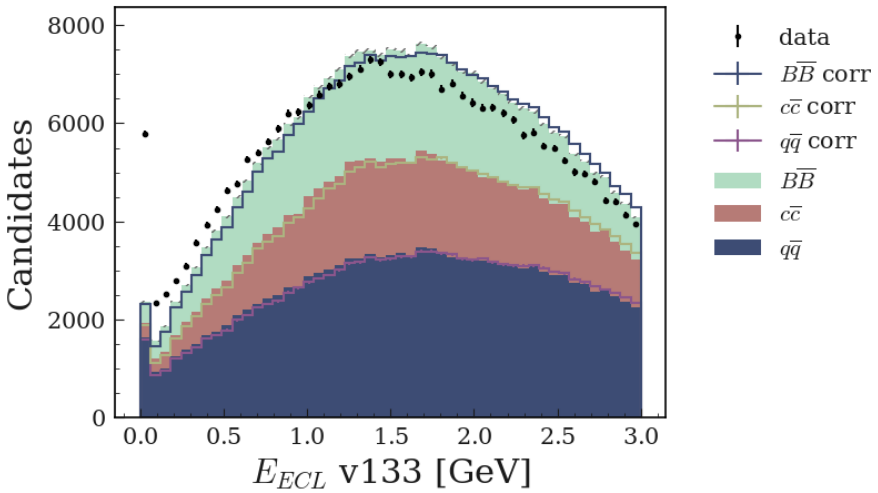
Same exercise performed for several sidebands and several masks

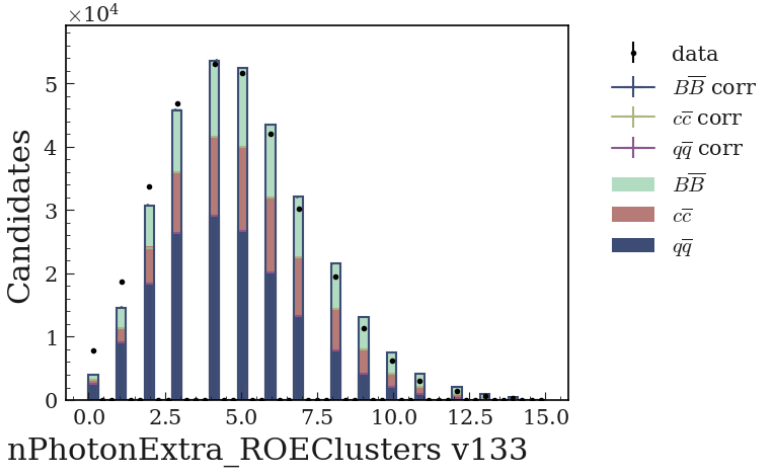
Photon energy scale correction(II)

Results for Other masks

pi-sideband:
- fraction of unmatched ROEgamma ~ 15%
- best correction: +17%

Propagating to E_ECL





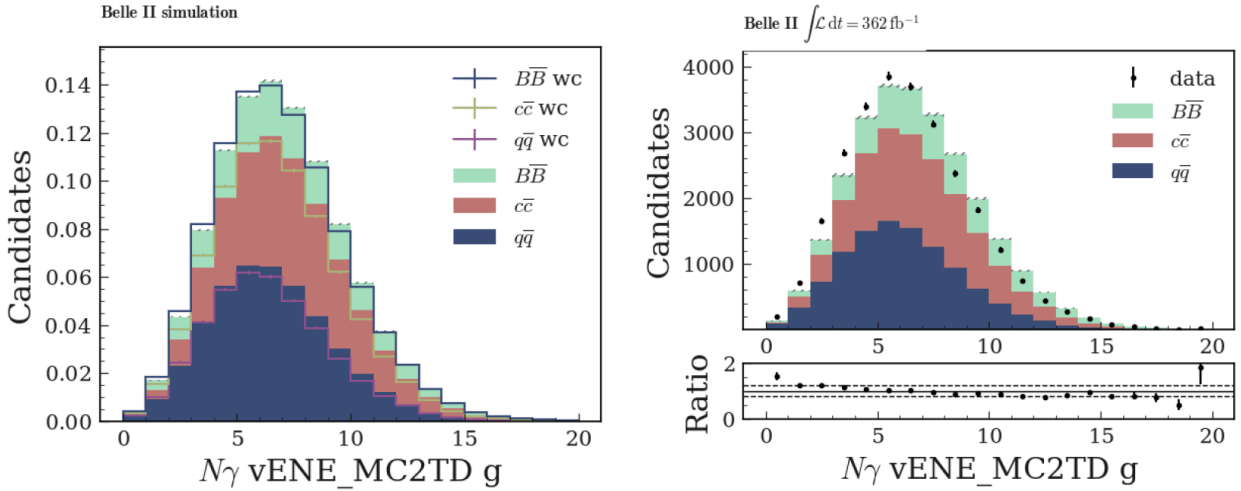
The disagreement is still present, also considering different ROEgamma masks and different sideband

- Overall energy scale not sufficient to fix data/MC disagreement
- This correction should be energy dependent and at least in 3 theta regions

Extra photon multiplicity correction

- In addition to correct the photon energy, we can correct also the photon efficiency, e.g. from extra photon cluster multiplicity
- Need to choose sideband region in which ROE cluster multiplicity better resemble the signal region one and we have a good data/MC agreement

Wrong charge sideband is a good candidate:



Summary and outlook

- After finding some data/MC discrepancies we performed several checks on E_ECL variable
- We tested several masks for the neutral cluster selection
- First study on the effect of the ECL reco bug on the analysis identifies the mask which is "less apparently affected"
- Tried to apply an overall correction to the energy scale of "fake/bkg" photons, but it is not effective (likely there is an important dependence on energy)

- Two possible developments:
  - Use an energy (and ECL region) dependent correction
  - Add a correction on the number of photons

Work in progress

Extra slides

ECL bug

From Chris' slides

- We have discovered a feature of the ECL reconstruction code that can cause high-energy photons to be not reconstructed in high background events.

The source of the problem

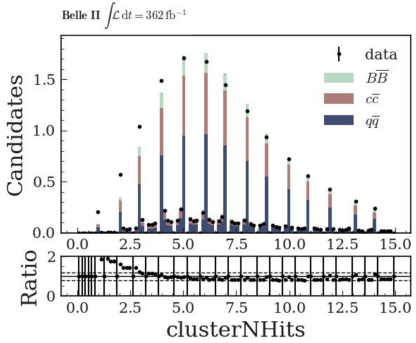
- In fact, not every local maximum forms an ECLShower. If there are >10 in a connected region, we form only one ECLShower.
- This was a design choice. The bug is that we intended to keep the highest energy local maximum, but instead kept the first (i.e. random) one.
- According to full-luminosity simulated backgrounds at the time, we would never get >10 local maxima in a connected region.

Tested masks

pi sideband after presel

	Selection	Comments
1.	E>(80, 30, 60) MeV, BDTs	Nominal mask
2.	E>100 MeV, BDTs	Increasing energy threshold
3.	E>60 MeV, gamma_theta in CDC accept., minC2Tdist>50 cm	Removing BDTs and increasing energy threshold
4.	E>100 MeV, gamma_theta in CDC accept., minC2Tdist>50 cm	

- Checked also number of ELC hits associated to cluster, recommended in some photon list → poor data/MC agreement
- Timing infos not usable for the same reason



“Real” photon energy corrections

- The payload for photon energy bias corrections can be applied only to a list of particles (see here)
- Neutral Extra energy (NEExtra) can be corrected while EExtra cannot
- Decided to get the corrections from **NEExtra_corr - NEExtra**

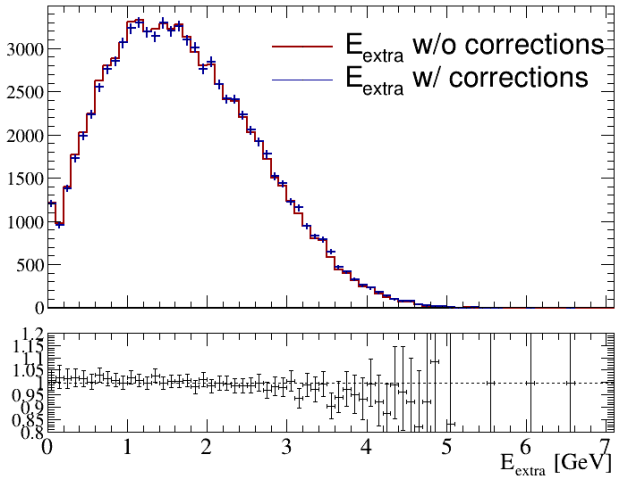
No impact from official photon energy bias correction derived from neutral group

- derived from ee->mumugamma sample, reconstructed clusters are mostly associated to real photons
- <%level change in number of selected events at pre-selection stage, almost no impact on the shape

considered as a check, no further syst applied

EExtra_corr = EExtra + NEExtra_corr - NEExtra

After preselection + BCS



The effect is very small

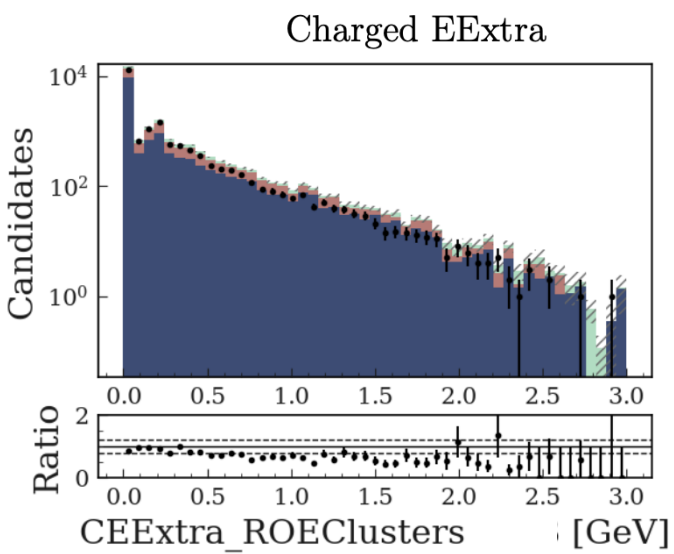
Summary of corrections derived from piID sidebands

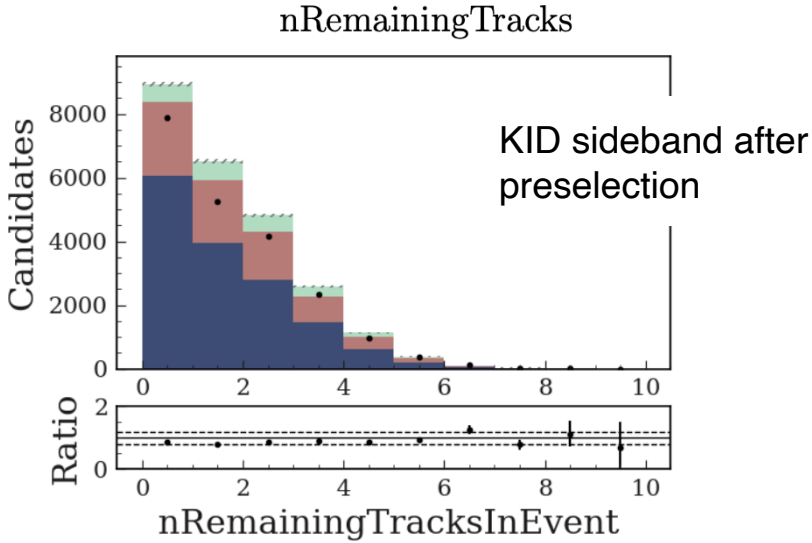
	selection	fraction of unmatched photons in piID sideband	correction factor from piID sideband
v133	E>(80, 30, 60) MeV, BDTs	**15.9%**	**+17%**
vITA	E>100 MeV, CDC acc	**37.4%**	**+2%**
v233	E>100 MeV, BDTs	**14.7%**	**-14%**

- only with v233 we get negative energy corrections
- no EExtra definition is very well modeled, even after corrections
- a simple energy scaling does not solve the issue

Neutral component of EECL

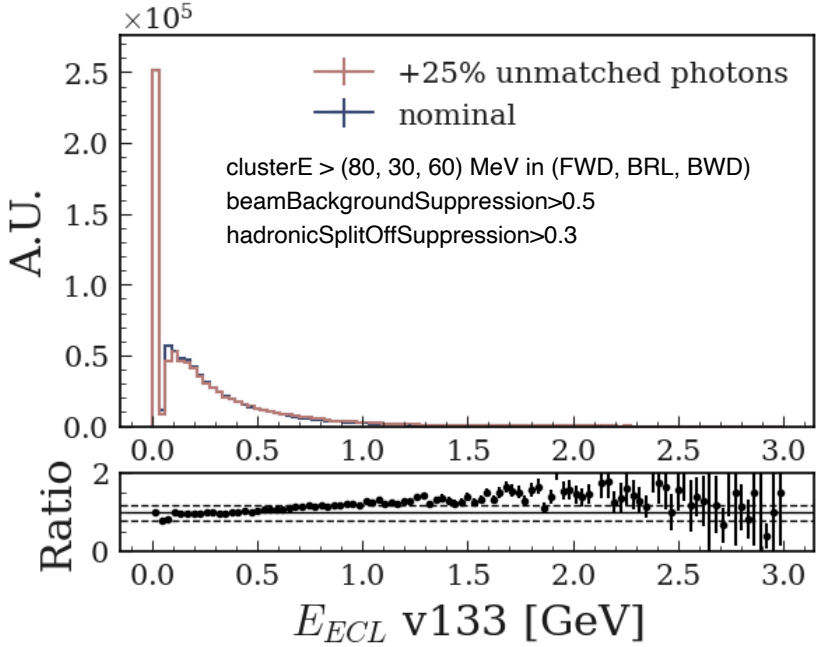
Energy associated to extra - remaining tracks, not passing track cleaning requirements (dr < 2 cm, |dz| < 4 cm, in CDC accept., nCDCHits>20)

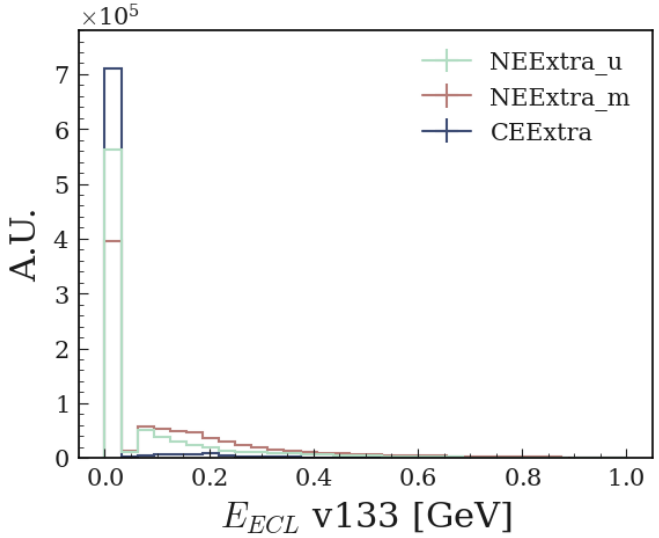




Signal MC studies

- Nominal mask
- Signal MC events passing reconstruction + preselection (zero extra cleaned tracks, π^0, Ks, Λ in the event)
  - fraction of unmatched photons over total number of extra photons ~30%
- Correction produces some effect the region around $E_{ECL}>0$
- Need to check data/MC agreement to be conclusive (ntuple production with all necessary info on-going)



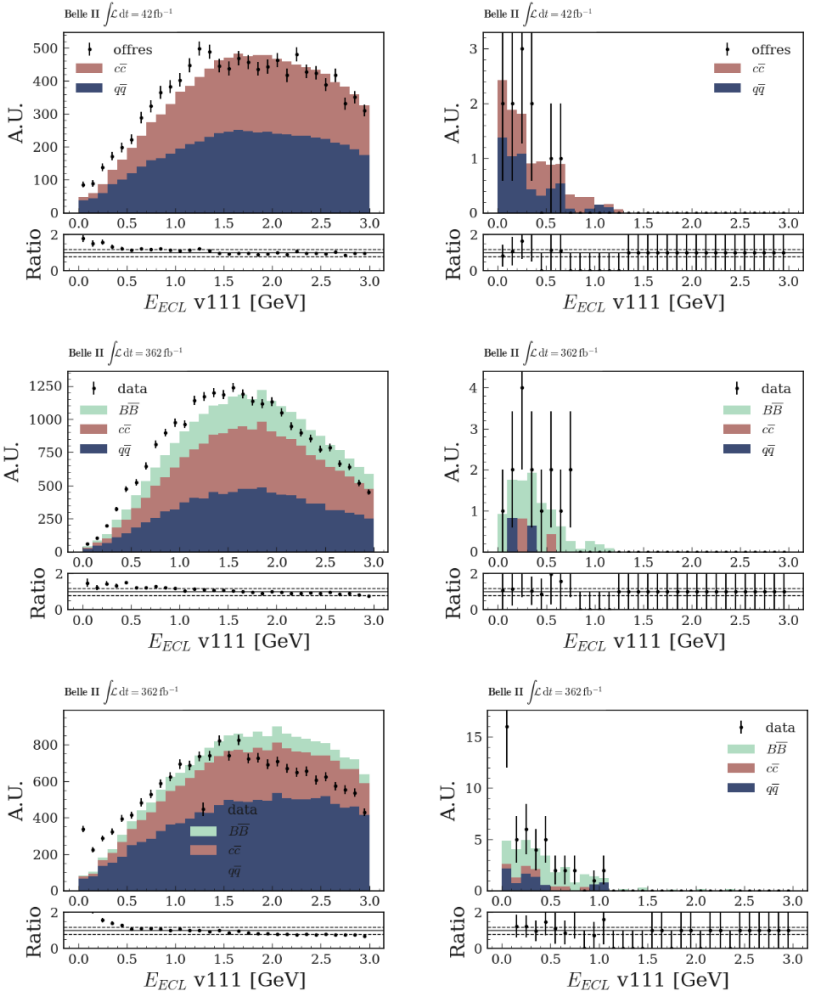


Eextra V111

"[[clusterReg == 1 and clusterE>0.080] or [clusterReg == 2 and clusterE>0.030] or [clusterReg == 3 and clusterE>0.060]]"

"abs(formula(clusterTiming / clusterErrorTiming))<2.0 and abs(clusterTiming) < 200"

"minC2TDist>20"

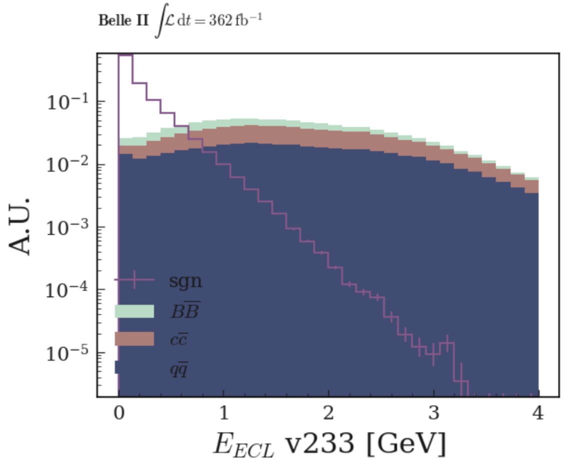


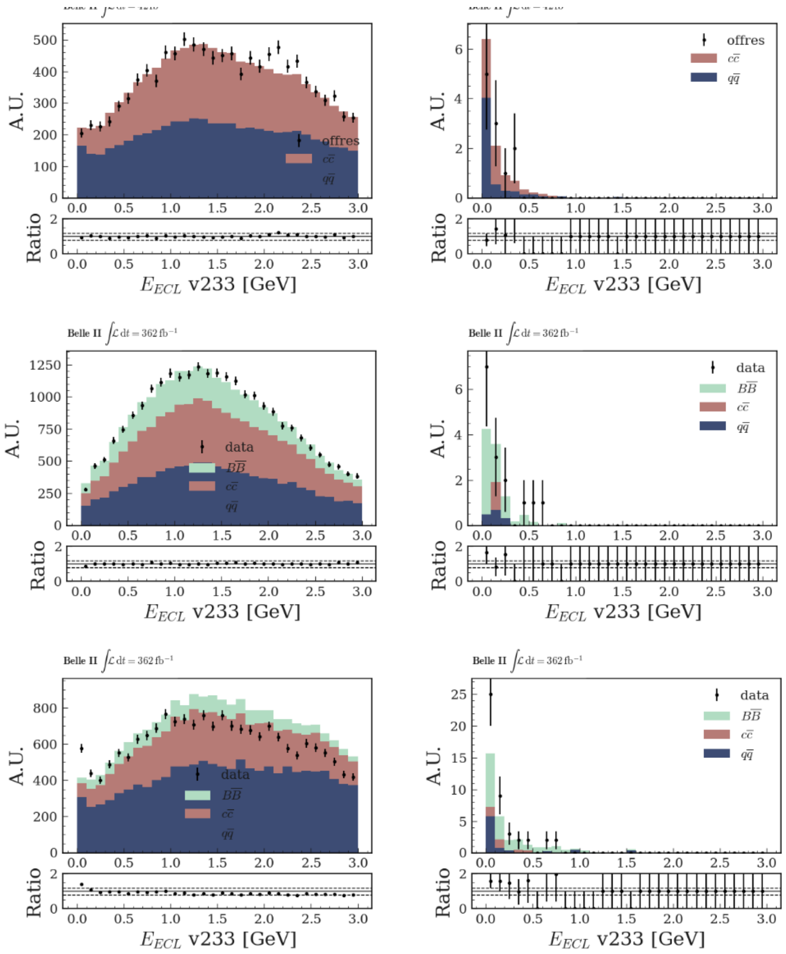
Eextra V233

"clusterE>0.1"

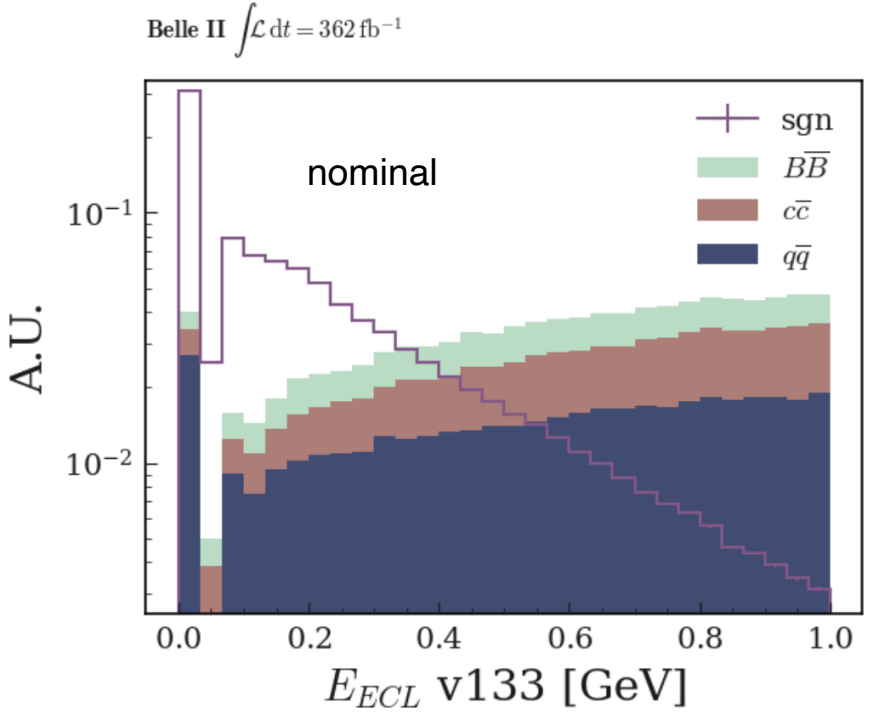
"abs(formula(clusterTiming / clusterErrorTiming))<2.0 and abs(clusterTiming) < 200"

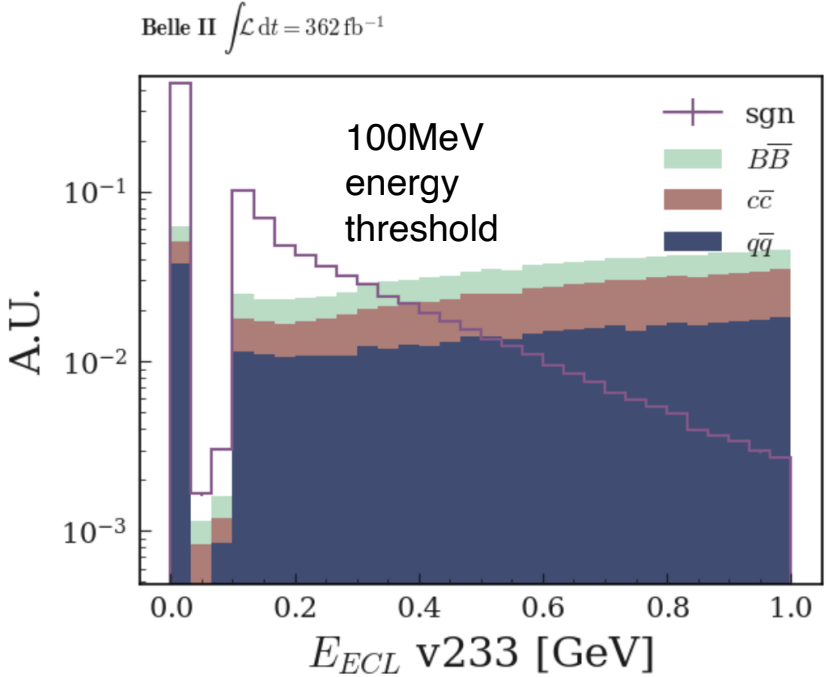
"minC2TDist>20"



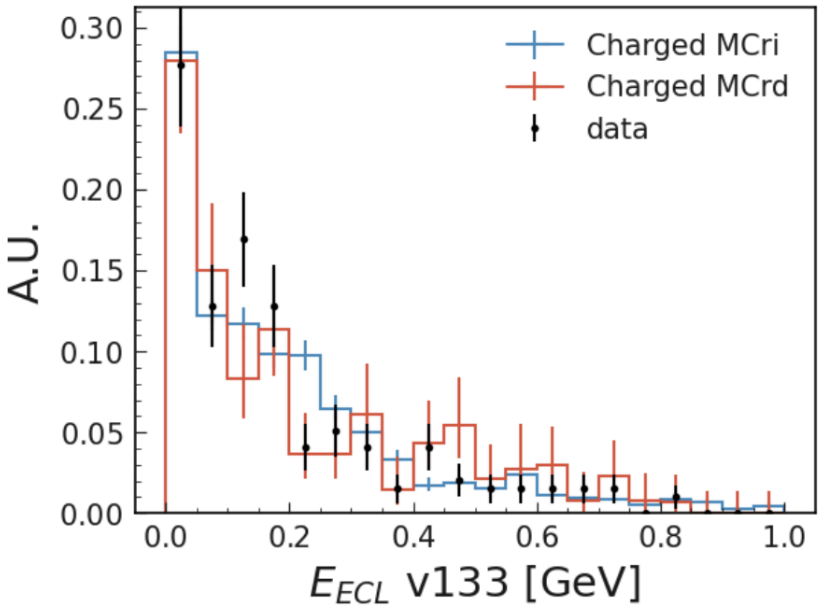


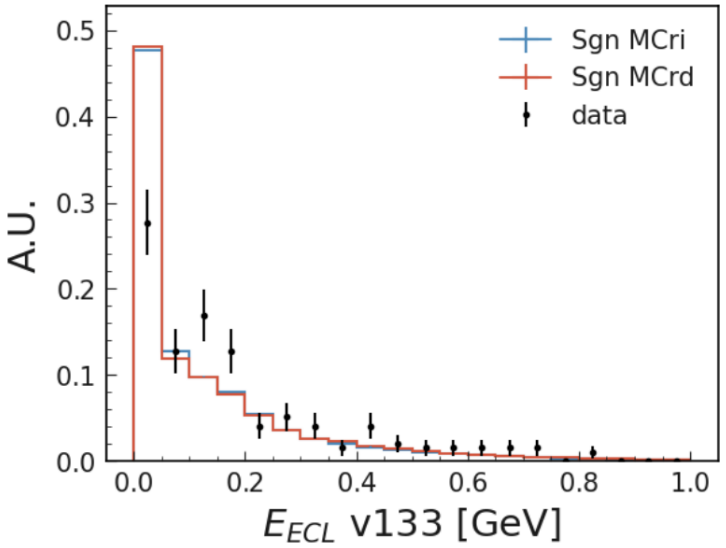
V133 vs V233 in low Eextra region





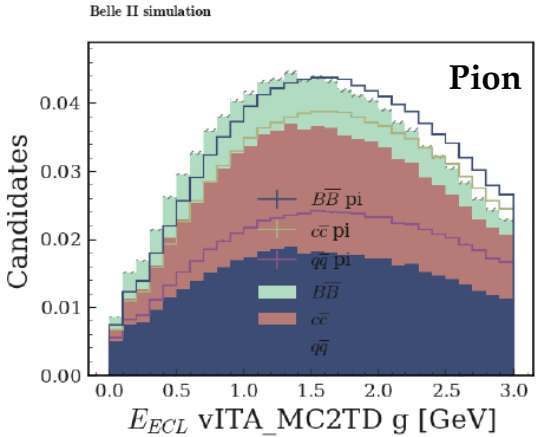
Run dependent MC studies

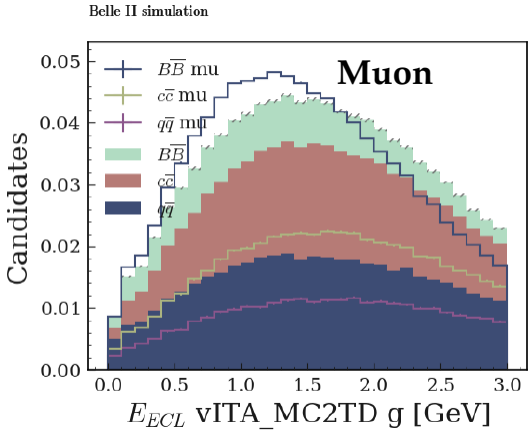


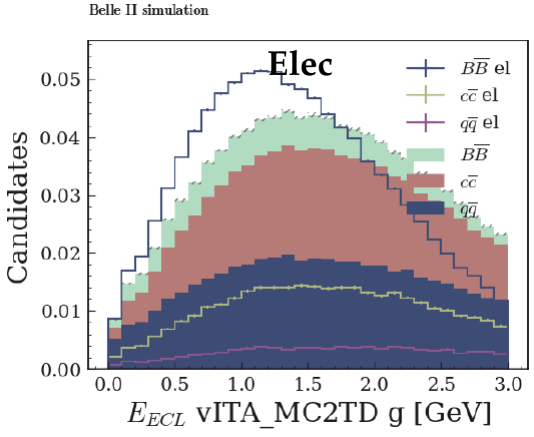


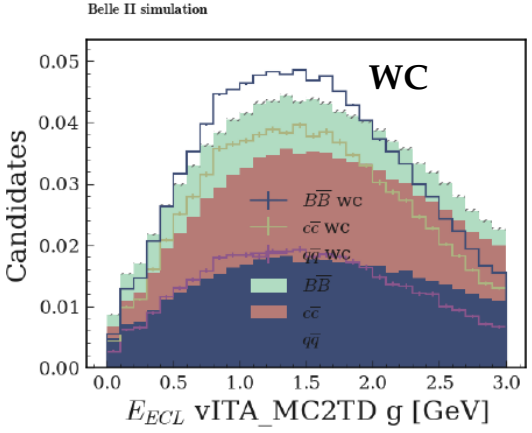
MC15rd does not explain trivially the difference (mainly the peak in the 3rd bin)

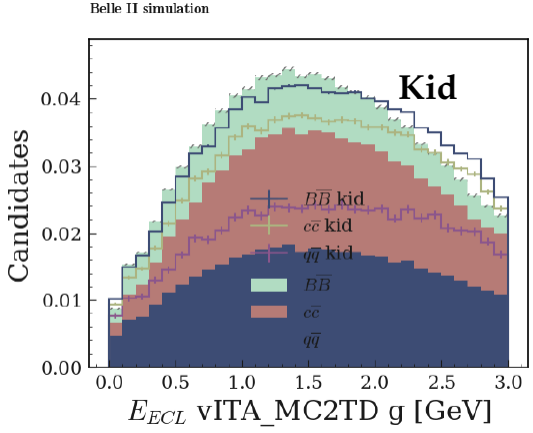
Sideband comparison vITA_MC2TD



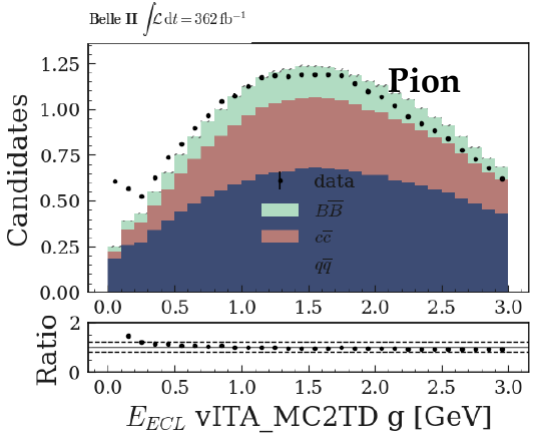


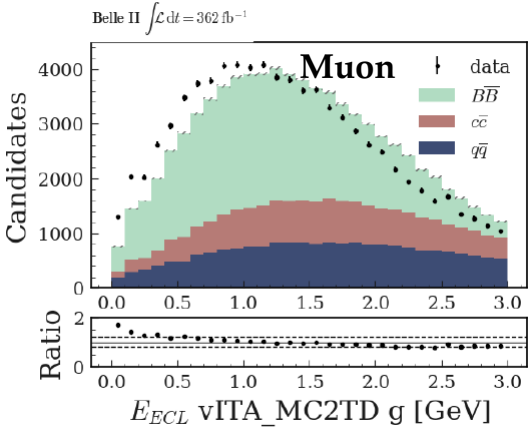


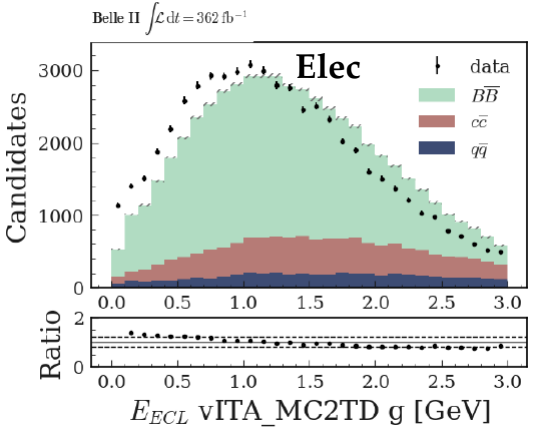


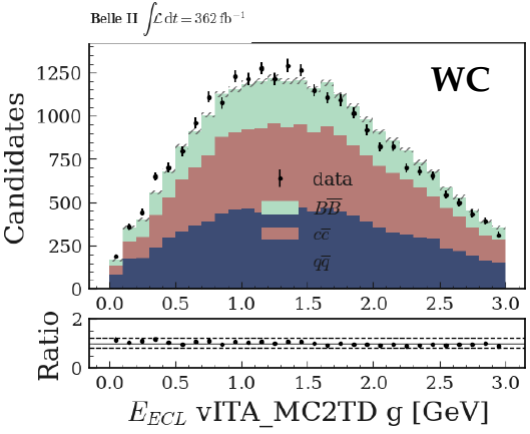


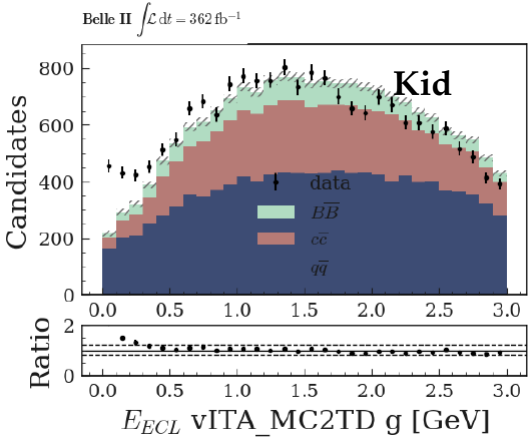
Sideband comparison vITA_MC2TD





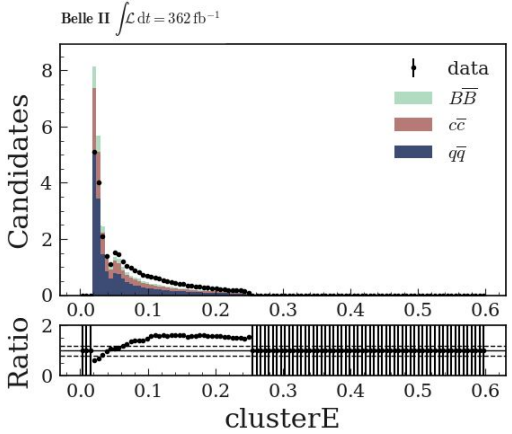


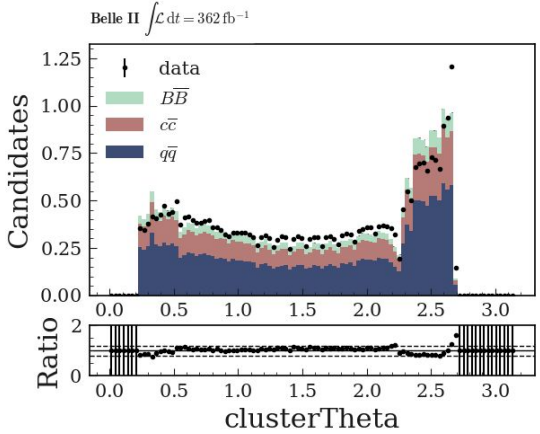


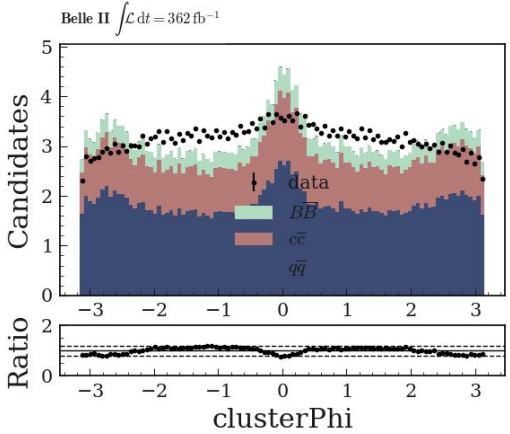


Pion sideband

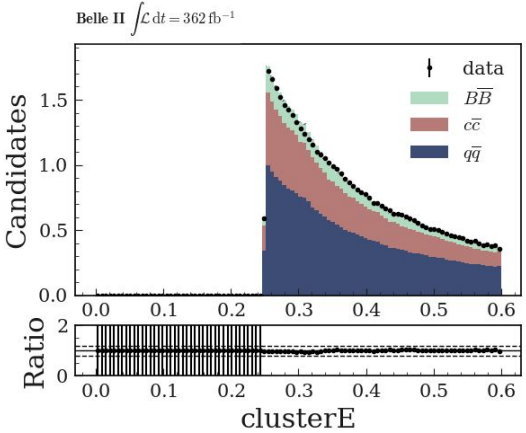
E<250 MeV

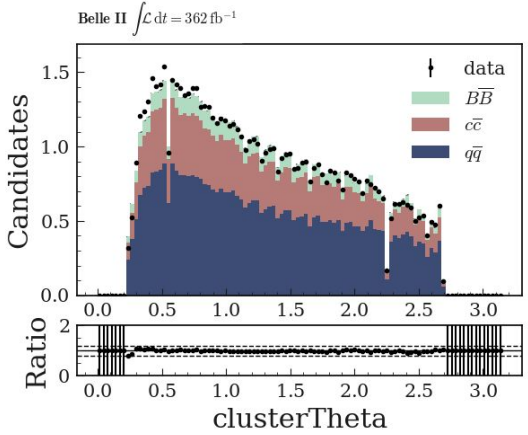


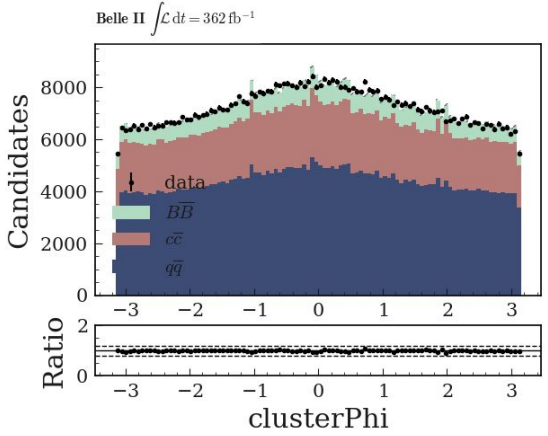




E>250 MeV

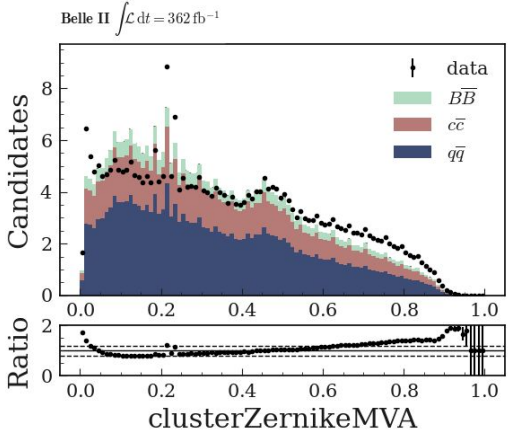


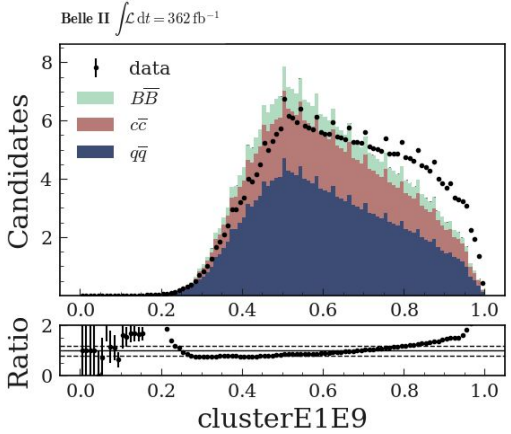


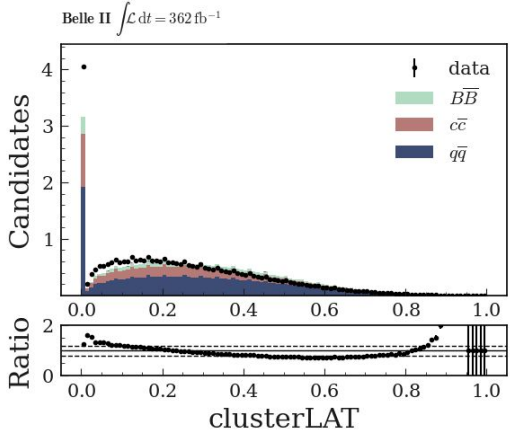


Pion sideband

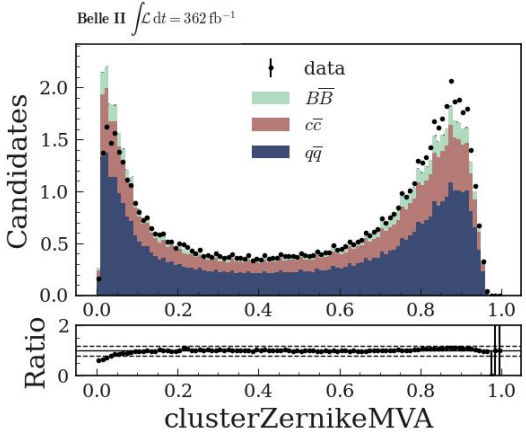
E<250 MeV

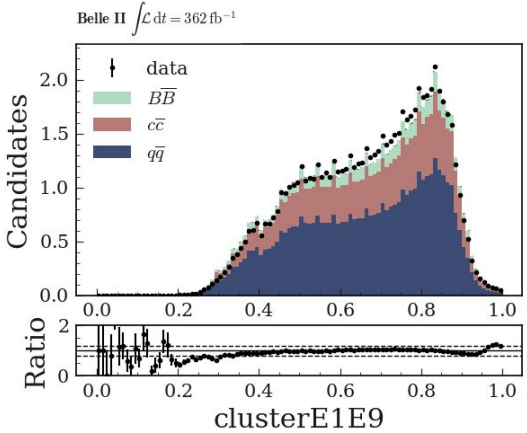


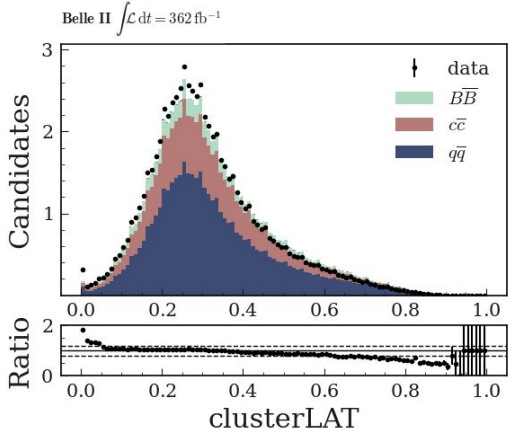




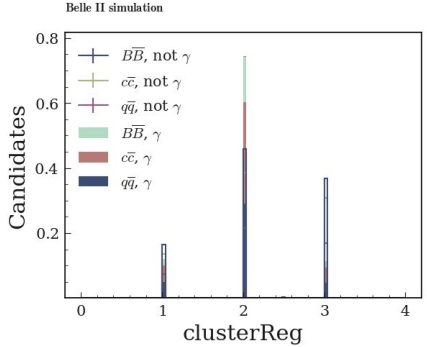
E>250 MeV

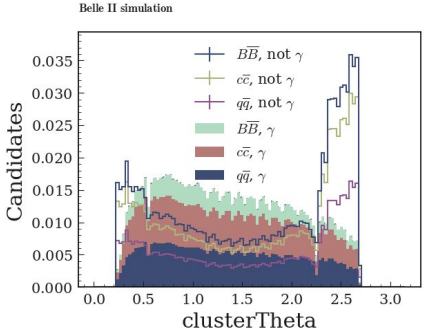


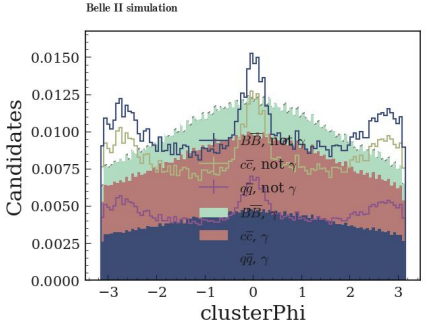


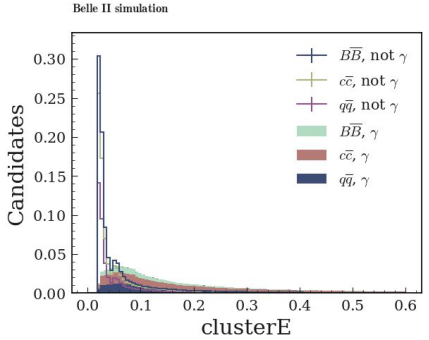


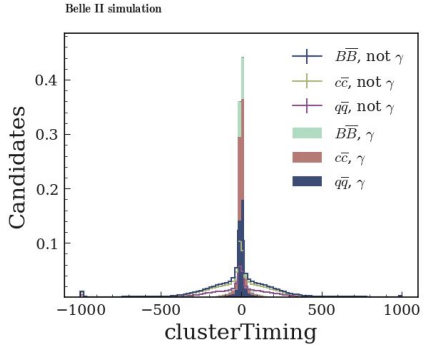
MC signal region, no mask

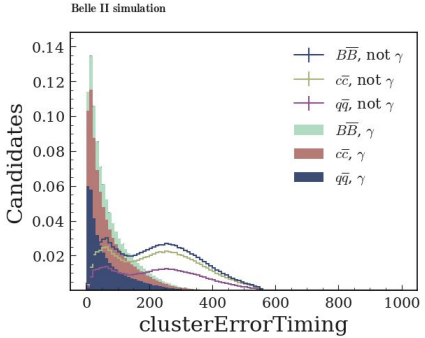


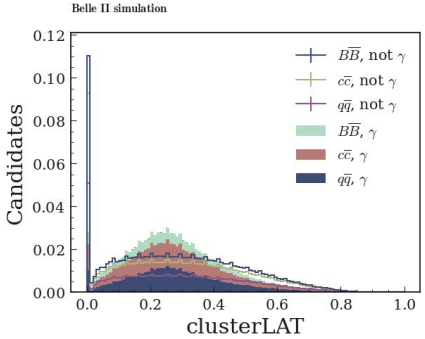


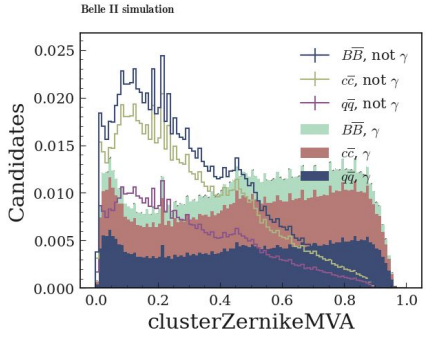












MC signal region, no mask