E_ECL studies for $B^+ \rightarrow K^+ \nu \bar{\nu}$ 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

Highlights of event reconstruction:

- Reconstruct $B_{\text{tag}} + K^+$ candidate (KID>0.9)
- Request zero extra (cleaned tracks, $\pi^0$, $K_s$, $\Lambda$) 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

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

Total extra energy in ECL after preselection

Most discriminant variable in final BDT

$B \rightarrow K \nu \bar{\nu}$ HTA, Joint (S)L / EWP mini-workshop, 05/30/2023
Analysis overview

Definition of sidebands

Signal selection

\( k_{\text{aonID}} > 0.9 \)

Pion sideband

\( p_{\text{ionID}} > 0.9 \)

(wrong) Kid sideband

\( p_{\text{ionID}} > 0.5 \) and \( k_{\text{aonID}} > 0.1 \)

Wrong charge sideband

\( B \rightarrow K \nu \bar{\nu} \) HTA, Joint (S)L / EWP mini-workshop, 05/30/2023
Extra energy components

- Charged component: ECL energy associated to extra - remaining tracks, not removed
  
  \[ \text{N\_cleaned\_tracks} = 0 \] (track-cleaning: \( \text{dr} < 2 \text{ cm}, \ |dz| < 4 \text{ cm}, \) in CDC accept., \( n\text{CDCHits}>20 \))

- Neutral component: ECL energy associated to extra - photons

- Nominal extra-photon mask:
  - \( \text{clusterE} > (80, 30, 60) \text{ MeV in (FWD, BRL, BWD) ECL} \)
  - \( \text{beamBackgroundSuppression}>0.5 \)
  - \( \text{hadronicSplitOffSuppression}>0.3 \)

Definition of \( v_{133} \) (nominal mask)
Extra photon mask choice

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

after pre-selection

wrong Btag-K charge

inverted signal K ID (similar behavior in pi sideband)

off-resonance vs qqbar

embedded sample
$E_{\text{ECL}}$ distribution in signal region after unblinding

$E_{\text{ECL}}$ 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 $E_{\text{ECL}}$
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

→ Investigate alternative gammaROE masks with higher energy threshold and removal of beamBackgroundSuppression and hadronicSplitOffSuppression BDTs
### Example of tested masks

<table>
<thead>
<tr>
<th>Mask</th>
<th>Selection</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>v133</td>
<td>E&gt;(80, 30, 60) MeV, BDTs</td>
<td>Nominal mask</td>
</tr>
<tr>
<td>vENE_MC2TD</td>
<td>E&gt;60 MeV, $\theta_{clu}$ in CDC accept., minC2Tdist&gt;50 cm</td>
<td>Removing BDTs and increasing energy threshold</td>
</tr>
</tbody>
</table>

**Pion sideband after preselection and:**
- E>60 MeV
- $\theta$ in CDC accept.

**Other masks**

![Belle II simulation](image)

**Mc-truth matched photons**

![Belle II simulation](image)

**Data**

$B \to K \nu \bar{\nu}$ HTA, Joint (S)L / EWP mini-workshop, 05/30/2023
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

- 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 choose the one which maximize the agreement

\[
E_{\text{ROE}}^n(f_h) = \sum_i E_i^r + f_h \sum_j E_j^n
\]

- \(i \in \text{ECL clusters matched to photons.}\)
- \(j \in \text{ECL clusters not matched to photons.}\)
- \(f_h \equiv \text{scale factor quantifying accuracy of energy calibration.}\)

Same exercise performed for several sidebands and several masks
Photon energy scale correction(II)

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

Results for Other masks

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

<table>
<thead>
<tr>
<th>Selection</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ( E \geq (80, 30, 60) ) MeV, BDTs</td>
<td>Nominal mask</td>
</tr>
<tr>
<td>2. ( E &gt; 100 ) MeV, BDTs</td>
<td>Increasing energy threshold</td>
</tr>
<tr>
<td>3. ( E &gt; 60 ) MeV, gamma(_\theta) in CDC accept., ( \text{minC2Tdist} &gt; 50 ) cm</td>
<td>Removing BDTs and increasing energy threshold</td>
</tr>
<tr>
<td>4. ( E &gt; 100 ) MeV, gamma(_\theta) in CDC accept., ( \text{minC2Tdist} &gt; 50 ) cm</td>
<td></td>
</tr>
</tbody>
</table>

- 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

\[ B \rightarrow K \nu \bar{\nu} \] HTA, Joint (S)L / EWP mini-workshop, 05/30/2023
“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

\[ E_{\text{Extra corr}} = E_{\text{Extra}} + \text{NEExtra}\_corr - \text{NEExtra} \]

After preselection + BCS

The effect is very small
## Summary of corrections derived from piID sidebands

<table>
<thead>
<tr>
<th>selection</th>
<th>fraction of unmatched photons in piID sideband</th>
<th>correction factor from piID sideband</th>
</tr>
</thead>
<tbody>
<tr>
<td>v133 E&gt;(80, 30, 60) MeV, BDTs</td>
<td>15.9%</td>
<td>+17%</td>
</tr>
<tr>
<td>vITA E&gt;100 MeV, CDC acc</td>
<td>37.4%</td>
<td>+2%</td>
</tr>
<tr>
<td>v233 E&gt;100 MeV, BDTs</td>
<td>14.7%</td>
<td>-14%</td>
</tr>
</tbody>
</table>

➢ 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, $\pi^0$, Ks, $\Lambda$ 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)

\[
\begin{align*}
  \times 10^5 & \\
  \text{+25\% unmatched photons} & & \text{nominal} \\
\end{align*}
\]\n
\[
\begin{align*}
  \text{clusterE} & > (80, 30, 60) \text{ MeV in (FWD, BRL, BWD)} \\
  \text{beamBackgroundSuppression} & > 0.5 \\
  \text{hadronicSplitOffSuppression} & > 0.3
\end{align*}
\]
Eextra V111

\[
\text{"[clusterReg == 1 and clusterE>0.080] or [clusterReg == 2 and clusterE>0.030] or [clusterReg == 3 and clusterE>0.060]\"''}
\]

\[
\text{"abs(formula(clusterTiming / clusterErrorTiming))<2.0 and abs(clusterTiming) < 200\"''}
\]

\[
\text{"minC2TDist>20\"''}
\]
Eextra V233

"clusterE>0.1"

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

"minC2TDist>20"
V133 vs V233 in low Eextra region

Belle II $\int \mathcal{L} dt = 362 \text{ fb}^{-1}$

nominal

$100\text{MeV}$ energy threshold

$B \rightarrow K\nu\bar{\nu}$ HTA, Joint (S)L / EWP mini-workshop, 05/30/2023
Run dependent MC studies

MC15rd does not explain trivially the difference (mainly the peak in the 3rd bin)
Sideband comparison vITA_MC2TD

- **Pion**
  - Belle II simulation
  - Candidates vs. $E_{ECL}$ vITA_MC2TD g [GeV]

- **Muon**
  - Belle II simulation
  - Candidates vs. $E_{ECL}$ vITA_MC2TD g [GeV]

- **Elec**
  - Belle II simulation
  - Candidates vs. $E_{ECL}$ vITA_MC2TD g [GeV]

- **WC**
  - Belle II simulation
  - Candidates vs. $E_{ECL}$ vITA_MC2TD g [GeV]

- **Kid**
  - Belle II simulation
  - Candidates vs. $E_{ECL}$ vITA_MC2TD g [GeV]
Sideband comparison vITA_MC2TD

Pion

Muon

Elec

WC

Kid
Pion sideband

E<250 MeV

E>250 MeV
MC signal region, no mask
MC signal region, no mask