Update on exclusive *b*-tagging and an introduction to hemisphere efficiency correlation

FCC Performance Meeting

Lars Röhrig[1,2], Kevin Kröninger[1], Romain Madar[2], Stéphane Monteil[2]
In collaboration with Fabrizio Palla

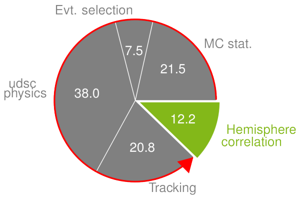
07/25/2023

[1]Department of Physics – TU Dortmund University
[2]Laboratoire de Physique de Clermont – Université Clermont-Auvergne

Introduction & Motivation

- Heavy-quark electroweak observables under study: R_b and A^b_{FB}
- Measurements will be dominated by the systematic uncertainty at FCC-ee (Z-pole)
- Dominant contribution to uncertainty from light- and c-quark physics
- Motivates novel ideas to make use of the Z-pole statistics to bring down syst. uncertainty

Next b-tagger for the FCC-ee

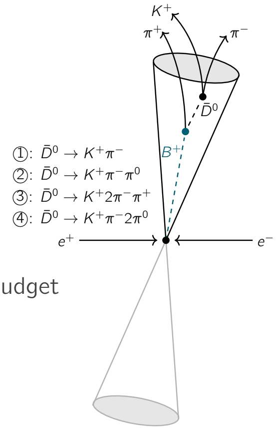


- Goal: elimination of light- and c-physics contamination
- When does it enter? At the b-tagging!
  ↪ Purer b-tagging required

Proposal: b-hemisphere tagger

Identify (the charge of) the hemispheres by exclusively reconstruct b-hadrons. Targets:

- Potential purity: 100 % thanks to the boost, $\overline{\beta\gamma} \approx 6.5$
- Efficiency: 1 %



↪ Removing background introduces an updated systematic uncertainty budget

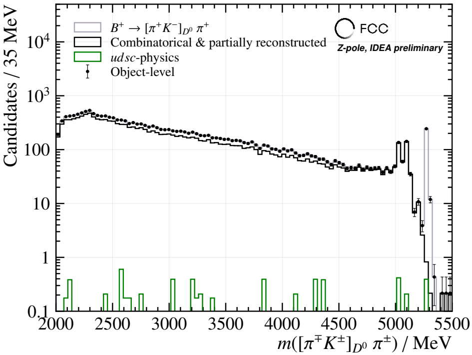
R_b: hemisphere efficiency correlation, A_{FB}^b: QCD corrections

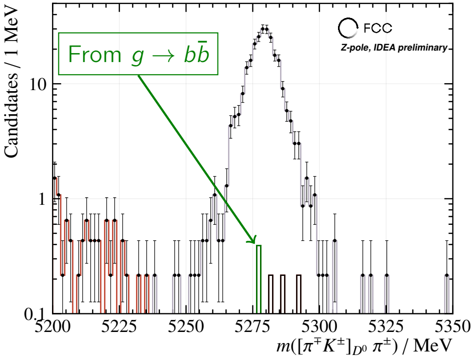
<u>What is needed?</u>

- Estimate purity from `delphes` Fast Simulation (IDEA)
- Estimate hemisphere correlation from Full Simulation (CLD)

What has already been achieved for R_b?

- 200+ b-hadron decay modes sum up to 1.1 % tagging efficiency ✓
- From `winter2023` sample campaign: Purity above 99.6 % for six representative decay modes ✓



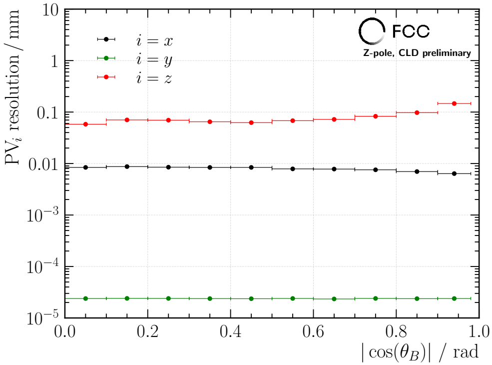


- Simulation of $\sim$ 83 000 CLD-FullSim events with forced decays on both legs with `EvtGen`: $\bar{b} \to B^+ \to [K^+\pi^-]_{\bar{D}^0}\,\pi^+$ and $b \to B^- \to [K^-\pi^+]_{D^0}\,\pi^-$ ✓
- Reconstruction using `VertexFitterSimple` functionalities ✓

FullSim B^+ reconstruction

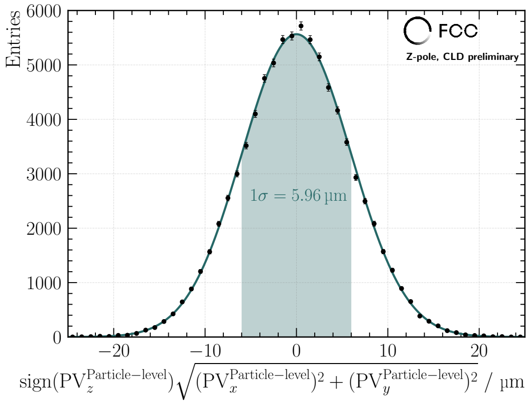
Goal: Bring the correlation of the hemisphere rec. efficiency down: mainly driven by the PV reconstruction

- Handful of custom functions to use `FCCAnalyses` with the FullSim samples
- Tested three approaches to select secondary tracks not coming from PV:
  1. **Shared primary vertex** with the beam spot constraint applied (`VertexFitter_Tk`)
  2. **Inconsistency** with the luminous region by geometrical cuts
  3. **LEP** reconstructed 1 PV/hemisphere (tested, but too less tracks for reco)
- Used a custom `VertexMore`-class to hack the neutral D^0 pseudotrack
- Applied the mass constraints for the D^0 and B^+ vertex fit (no "particle-ID" used)



Alternative secondary track selection

- Track geometries to select secondary tracks by their inconsistency with the luminous region
- Luminous region defined on MC truth by $\bar{d}_0 = \text{sign}(\text{PV}_z^{\text{MC}})\sqrt{(\text{PV}_x^{\text{MC}})^2 + (\text{PV}_y^{\text{MC}})^2}$
- Extract the mean $\mu_{\bar{d}_0} = -0.001\,\mu\text{m}$ and $\sigma_{\bar{d}_0} = 5.961\,\mu\text{m}$ of $\bar{d}_0$
- Evaluate the inconsistency of a track from its transverse and longitudinal impact parameter

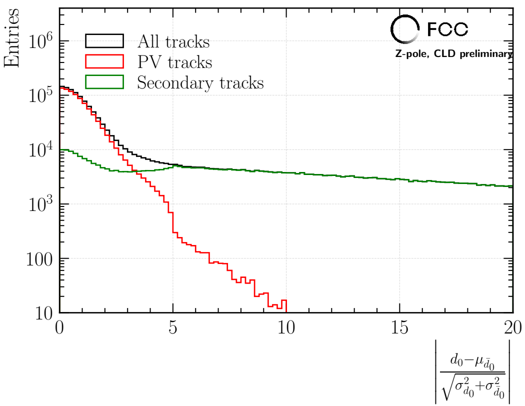


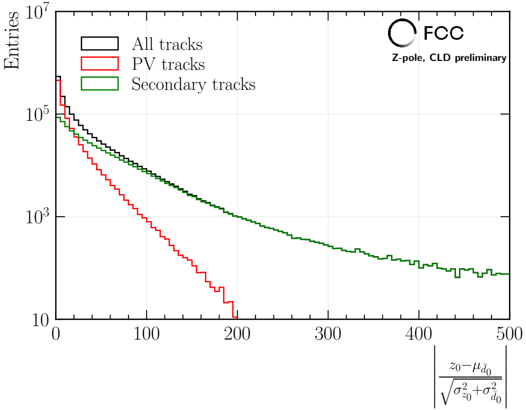
- Define discriminating variables v_1 and v_2:

$$v_1 = \left| \frac{d_0 - \mu_{\bar{d}_0}}{\sqrt{\sigma_{d_0}^2 + \sigma_{\bar{d}_0}^2}} \right| \quad \text{and} \quad v_2 = \left| \frac{z_0 - \mu_{\bar{d}_0}}{\sqrt{\sigma_{z_0}^2 + \sigma_{\bar{d}_0}^2}} \right|$$

Threshold search

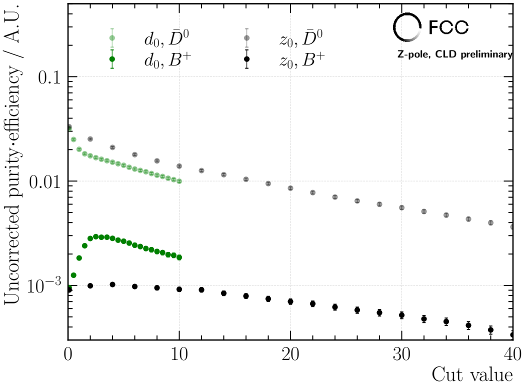
- Find threshold values for v_1 an v_2 to distinguish between primary and secondary tracks
- Evaluation from tracks, that are fitted with `VertexFitter_Tk` to a PV
- Left: v_1 distribution, right: v_2 distribution, separately for all, the PV and the secondary tracks

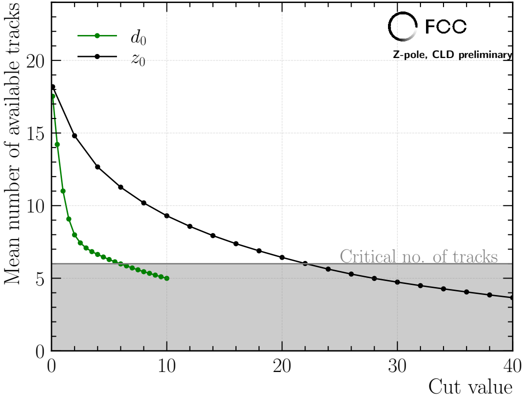




Threshold search

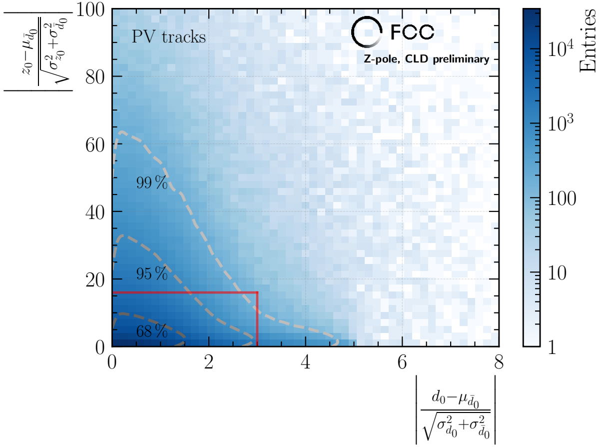
- Scanning $v_{1/2}$ and accepting tracks as secondaries, if $v_{1/2}$ is above threshold
- Evaluating the product of efficiency and purity + the mean number of secondary tracks
- Take maximum of d_0 ($\bar{D}^0$), since B^+ constant over wide range: $v_1 \overset{!}{>} 3$
- v_2 threshold taken to match mean number of secondary tracks at v_1 cut: $v_2 \overset{!}{>} 16$





Threshold search

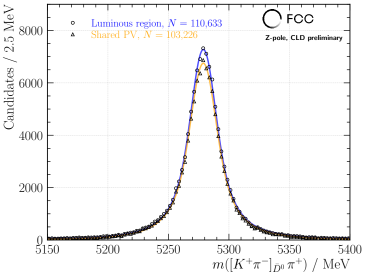
- Scanning $v_{1/2}$ and accepting tracks as secondaries, if $v_{1/2}$ is above threshold
- Evaluating the product of efficiency and purity + the mean number of secondary tracks
- Take maximum of d_0 ($\bar{D}^0$), since B^+ constant over wide range: $v_1 \overset{!}{>} 3$
- v_2 threshold taken to match mean number of secondary tracks at v_1 cut: $v_2 \overset{!}{>} 16$

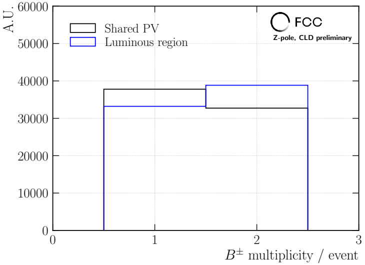


Reconstruction efficiencies

		Shared PV / %	Inconsistency with LR / %
Before reconstr.	D^0	87.00	90.55
	$\bar{D}^0$	87.15	90.40
	D^0 and $\bar{D}^0$	63.28	69.42
	B^+	69.20	78.05
	B^-	69.35	78.02
	B^+ and B^-	43.04	58.17
After reconstr.	D^0	74.42	75.47
	$\bar{D}^0$	74.77	75.48
	B^+	59.72	62.96
	B^-	59.46	62.76
	B^+ and B^-	38.50	44.85

- $D^0/\bar{D}^0$ efficiencies similar, $B^\pm$ differs since π-track close to the PV and could be catched by PV vertex fit



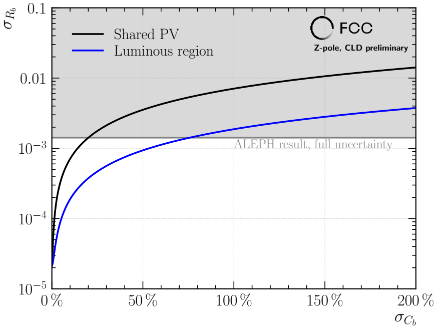


Hemisphere correlation

- Hemisphere (efficiency) correlation C_b: do we catch the other hemisphere, if the first one is tagged?

$$C_b = \frac{\varepsilon_{\text{double}}}{\varepsilon^2_{\text{single}}} \quad \text{with} \quad \varepsilon_{\text{double}} = \frac{N_{B^+\&B^-}}{N^{\text{generated}}_{B^+\&B^-}} \quad \text{and} \quad \varepsilon_{\text{single}} = \frac{N_{B^+} + N_{B^-}}{\left(N^{\text{generated}}_{B^+\&B^-}\right)^2}$$

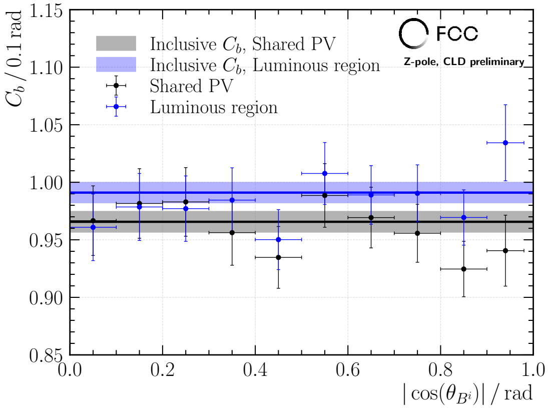
- Result: $C_b^{\text{Shared PV}} = 0.966 \pm 0.009$ and $C_b^{\text{Luminous}} = 0.991 \pm 0.009$ (ALEPH: $C_b = 0.962 \pm 0.003$), uncertainties are the statistical ones
- Estimate the impact of the relative hemisphere uncertainty on the full R_b uncertainty



- Evaluate the sources and dependences of the hemisphere correlation

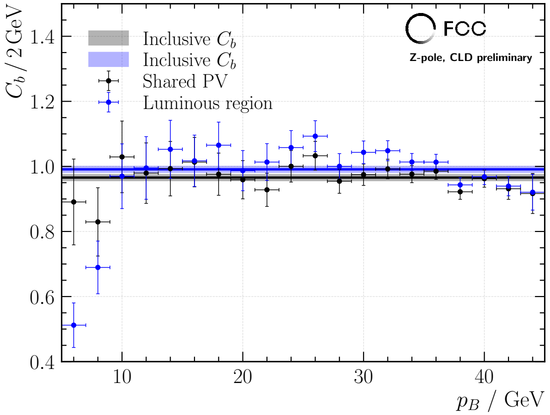
Hemisphere correlation computation – differential (preliminary)

- Differentially evaluated in bins of $\cos(\theta_B) = \frac{p_z}{|\boldsymbol{p}|}$ (angle wrt. the beam)
- No clear dependence within the statistical uncertainties, since reconstruction close to the beam affects both, $\varepsilon_{\text{double}}$ and $\varepsilon_{\text{single}}$



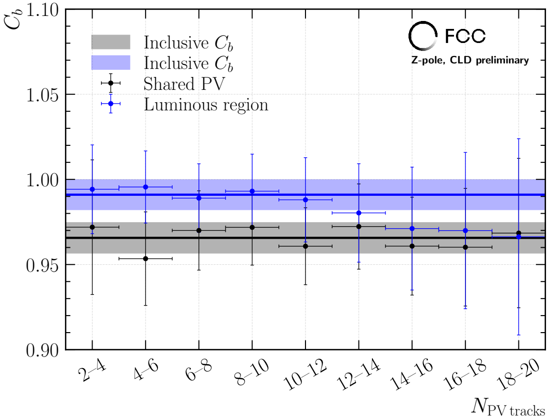
Hemisphere correlation computation – differential (preliminary)

- Differentially evaluated in bins of p_B
- Low momenta tail: $\rightarrow$ close to the PV: affects the reconstruction since tracks are taken as PV tracks
- From FastSim studies: $p_B > 20\,\text{GeV}$ to reduce gluon splitting contamination in the signal region



Hemisphere correlation computation – differential (preliminary)

- Differentially evaluated in bins of $N_{\text{PV tracks}}$ (number of primary tracks or non-secondary)
- Decrease towards higher $N_{\text{PV tracks}}$ for the LR approach (though consistent within uncertainties)



Conclusions and Outlook

- *b*-tagger based on exclusive *b*-hadron reconstruction foreseen to improve on syst. uncertainty
- Target purity and efficiency in reach and tested
- Hemisphere correlations for R_b: get indepedent from the PV
- Select tracks outside the luminous region seems promising to reduce the correlation
  ↪ Preliminary results, still WIP