Fragmentation fraction ratios with semi-leptonic decays at LHCb with Run3 data

Elisa Minucci

INFN - Laboratori Nazionali di Frascati

Workshop
Semileptonic B decays at the junction of experiment and theory

Torino 12-13 June 2025

Introduction

Knowledge of the production rates of b-hadron in pp collisions at LHC with respect to the decay kinematic (b-hadron η and p_T) is essential in many aspects:

- **To relate theoretical predictions of the bb production cross-section derived from pQCD, to the observed hadrons**
- **To convert the observed $\overline{B}_s^0$ and Λ_b^0 production ratios at LHC into absolute branching fractions:**
  - **Mostly used as a normalization for other measurements when absolute BF is needed, such as for $B_s \rightarrow \mu^+\mu^-$ as a BSM test and for measurements of CKM matrix elements such as $|V_{ub}|$ and $|V_{cb}|$**

Run3 Analysis Goals:

Expand knowledge of the fragmentation functions, extending the analysis to b hadron p_T and η bins with higher statistics and aim for a lower p_T reach.

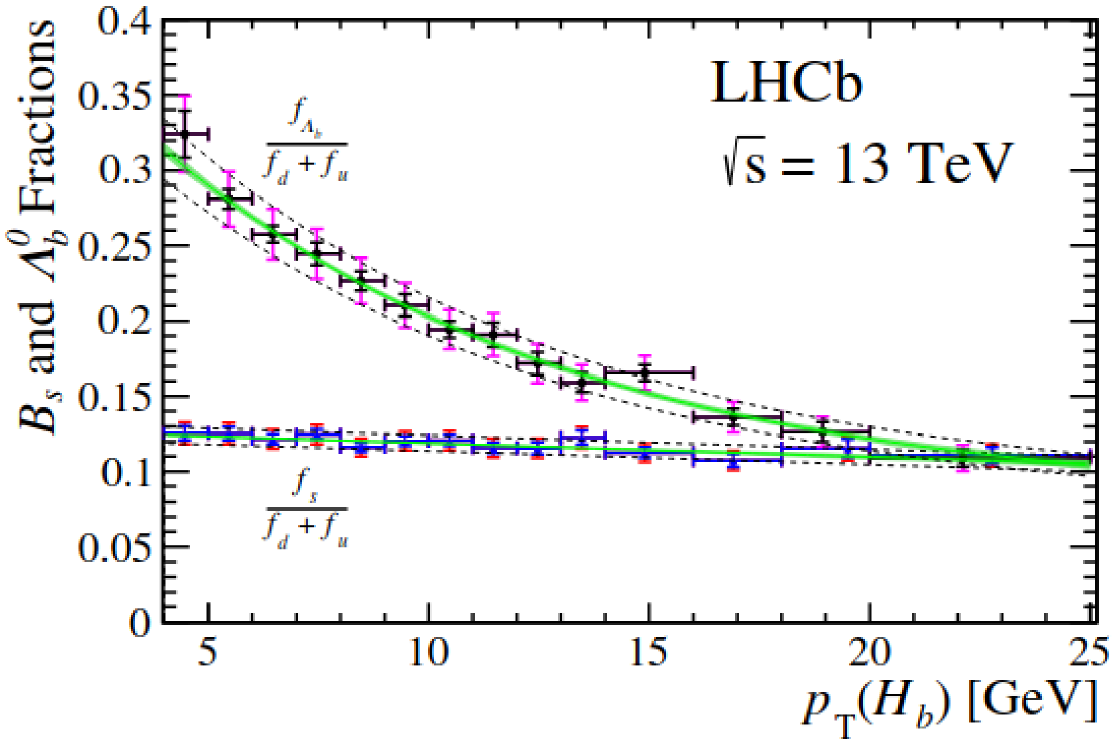
Future extensions:

A. Extend to Bc
B. Fragmentation function shape for other b-baryons

State of the art using inclusive SL decays

Measurement from LHCb with Run2 data @ 13 TeV based on an integrated luminosity of 1.67 fb^{-1}

[Phys. Rev. D 100, 031102(R)]



Average values in the range $4 < p_T(H_b) < 25$ GeV and $2 < \eta < 5$:

$$\frac{f_s}{f_u+f_d} = (0.122 \pm 0.006)$$

Modest linear dependence upon p_T(Hb)

$$\frac{f_{\Lambda_b}}{f_u+f_d} = (0.259 \pm 0.018)$$

Strong dependence on p_T(Hb)

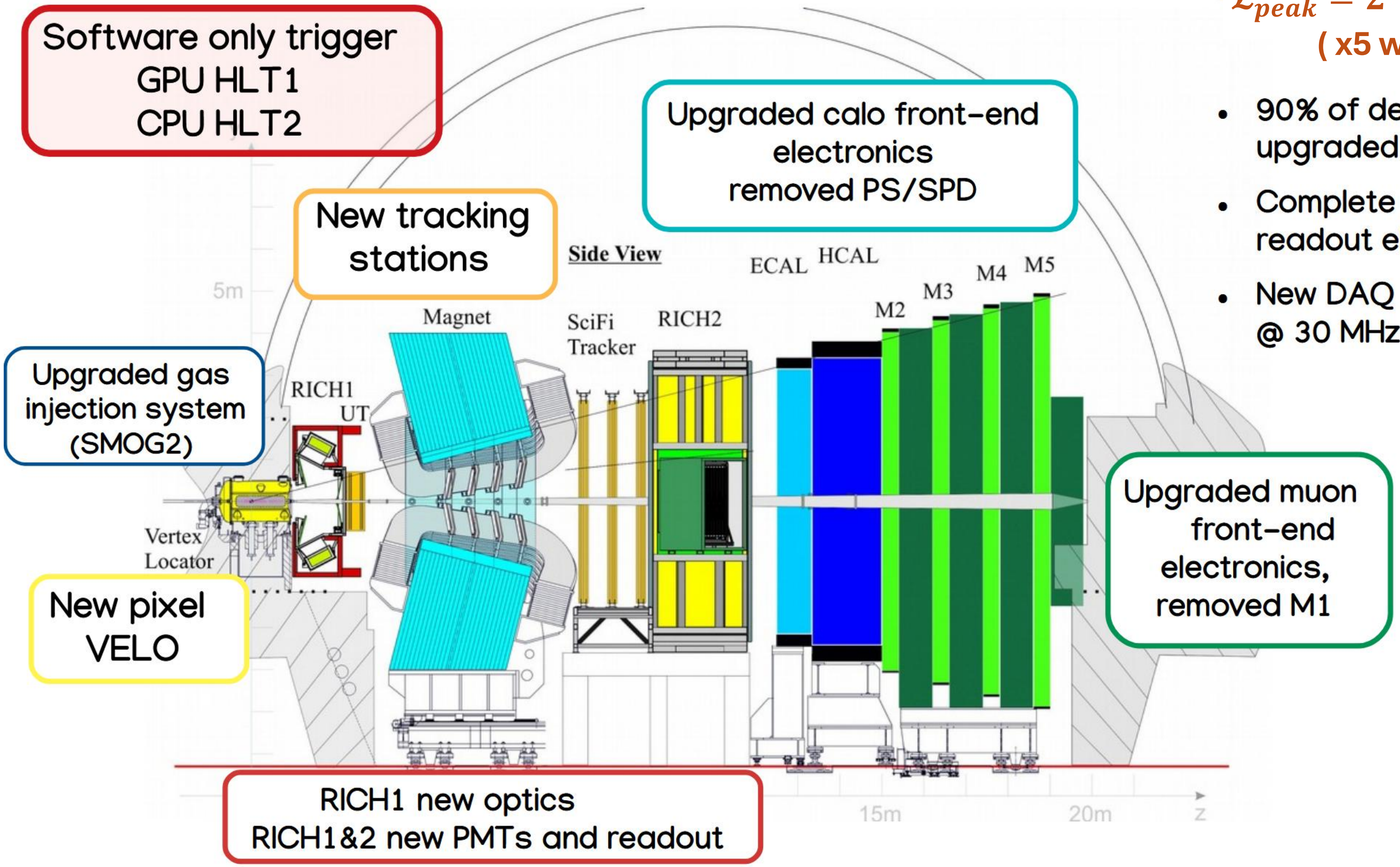
Dependence on pT(Hb) observed also in other analyses: at LHCb using hadronic final states, at CMS and ALICE.

LHCb Upgrade I

$\mathcal{L}_{peak} = 2 \times 10^{33}\ cm^{-2}s^{-1}$

(x5 w.r.t. Run1)

- 90% of detector channels upgraded
- Complete replacement of readout electronics
- New DAQ & online system @ 30 MHz



Run3 analysis

In collaboration with:
M.Artuso, A.Paul, M.Rudolph,
W.Vetens, H.Wu
from Syracuse University

The goal is to measure two fragmentation fraction ratios:

$\overline{B}_s^0 (f_s)$ **relative to the sum of** $B^- (f_u)$ **and** $\overline{B}^0 (f_d)$

$\Lambda_b^0 (f_{\Lambda_b^0})$ **relative to the sum of** B^- **and** $\overline{B}^0$

Studying the following final states:

$$\overline{B}_s^0 \rightarrow D_s^+ (\rightarrow K^+ K^- \pi^+) \mu^- \nu X$$
$$\Lambda_b^0 \rightarrow \Lambda_c^+ (\rightarrow p K^- \pi^+) \mu^- \nu X$$
$$B^- \rightarrow D^0 (\rightarrow K^- \pi^+) \mu^- \nu X$$
$$B^- \rightarrow D^+ (\rightarrow K^- \pi^+ \pi^+) \mu^- X$$

$$\overline{B}^0 \rightarrow D^+ (\rightarrow K^- \pi^+ \pi^+) \mu^- \nu X$$
$$\overline{B}^0 \rightarrow D^0 (\rightarrow K^- \pi^+) \mu^- \nu X$$

The available large data sample will allow to perform a measurement across the b-hadron η and p_T exploring new p_T regions. In addition, the study can be performed as a function of the event multiplicity (number of tracks, primary vertices…) with small uncertainty.

The well understanding of semi-leptonic decays from the theoretical point of view and the large branching fractions allows to reach low systematic uncertainties, compared to the measurement using hadronic final states.

Formalism

$$\frac{f_s}{f_u + f_d} = \frac{n_{corr}(\bar{B}_s^0 \to D\mu)}{n_{corr}(B \to D^0\mu) + n_{corr}(B \to D^+\mu)} \frac{\tau_{\bar{B}^0} + \tau_{B^-}}{2\tau_{\bar{B}^0{}_s}} (1 - \xi_s)$$

$$\frac{f_{\Lambda_b}}{f_u + f_d} = \frac{n_{corr}(\Lambda_b^0 \to D\mu)}{n_{corr}(B \to D^0\mu) + n_{corr}(B \to D^+\mu)} \frac{\tau_{\bar{B}^0} + \tau_{B^-}}{2\tau_{\Lambda^0{}_b}} (1 - \xi_{\Lambda_b})$$

Assuming SL widths of b-hadrons are the same (True to $O(1/m^2_b)$ in HQET) allows us to use the measured lifetimes to relate yields to final fragmentation fractions.

The corrections for this assumption are[1]:

$$\xi_s = (1.8 \pm 0.8)\%$$

$$\xi_{\Lambda_b} = -(4.1 \pm 1.6)\%$$

Corrections to Bare Yield

- cross-feed subtraction,
- efficiencies,
- BFs of charm daughters

[1]M. Bordone and P. Gambino, The Semileptonic Bs and Λb Widths (https://arxiv.org/pdf/2203.13107)

Formalism: Exact Expression

Cross feed processes

While most of the D^0 and D^+ come from B^+/B^0, and D^+_s from B^0_s, a smaller contribution from cross-feeds must be subtracted off:

- $\Lambda_b^0 \to D^{0(+)} p(n)\mu^-$ (data-driven).
- $B_s^0 \to D^{0(+)} K^{+(0)}\ \mu^-$ (data-driven).
- $B^+ \to D_s K\mu$

additional subtraction from fs
Computed analytically using PDG BF in Run2 analysis (In principle could also be measured from data):

$$-\frac{\mathcal{B}(B \to D_s K\mu)}{\langle \mathcal{B}_{SL}\rangle}\frac{\epsilon(B \to D_s^+)}{\epsilon(\bar{B}_s^0 \to D_s^+)}$$

$$n_{corr}(B \to D^0\mu) = \frac{1}{\mathcal{B}(D^0\to K^-\pi^+)\epsilon(B\to D^0)} \times$$
$$[n(D^0\mu) - n(D^0K^+\mu^-)\frac{\epsilon(\bar{B}_s^0\to D^0)}{\epsilon(\bar{B}_s^0\to D^0\mathrm{K}^+)} - n(D^0p\mu^-)\frac{\epsilon(\Lambda_b^0\to D^0)}{\epsilon(\Lambda_b^0\to D^0p)}]$$

$$n_{corr}(B \to D^+\mu) = \frac{1}{\epsilon(B\to D^+)} \times [\frac{n(D^0\mu)}{\mathcal{B}(D^+\to K^-\pi^+\pi^+)} -$$
$$\frac{n(D^0K^+\mu^-)}{\mathcal{B}(D^0\to K^+\pi^-)}\frac{\epsilon(\bar{B}_s^0\to D^+)}{\epsilon(\bar{B}_s^0\to D^0\mathrm{K}^+)} - \frac{n(D^0p\mu^-)}{\mathcal{B}(D^0\to K^+\pi^-)}\frac{\epsilon(\Lambda_b^0\to D^0)}{\epsilon(\Lambda_b^0\to D^0p)}]$$

$$n_{corr}(B_s^0 \to D_s^+\mu) = \frac{n(D_s^+\mu)}{\mathcal{B}(D_s^+\to KK\pi)\epsilon(B_s^0\to D_s^+\mu)} -$$
$$\mathrm{N}(\bar{B}^0 + B^-)\mathcal{B}(B \to D_s^+\mu)\frac{\epsilon(\bar{B}\to D_s^+K\mu)}{\epsilon(\bar{B}_s^0\to D_s^+\mu)} + 2\frac{n(D^0K\mu)}{\mathcal{B}(D^0\to K\pi)\epsilon(B_s^0\to D^0K\mu)}$$

$$n_{corr}(\Lambda_b^0 \to D\mu) = \frac{n(\Lambda_c^+\mu)}{\mathcal{B}(\Lambda_c^+\to pK\pi)\epsilon(\Lambda_b^0\to\Lambda_c^+\mu)} + 2\frac{n(D^0p\mu)}{\mathcal{B}(D^0\to K\pi)\epsilon(\Lambda_b^0\to D^0p\mu)}$$

Note: important to persist extra tracks other than the signal → a MVA isolation tool for charge tracks has been implemented within the SL WG.

2024 data set

Block	Polarity	Fills	Integrated Luminosity (pb^{-1})	μ_{avg}
2	up	9945-9978	604	4.4
1	up	9982-10056	1178	4.4
5	up	10059-10102	1162	4.4
6	down	10104-10138	941	4.4
7	down	10197-10213	737	5.3
8	up	10214-10232	435	5.3

AP Tuples complete for Data

Online selection: different data streams for the different final states.
Offline selection: additional selections made to harmonise the selections among the different final states and to reduce background

Background sources

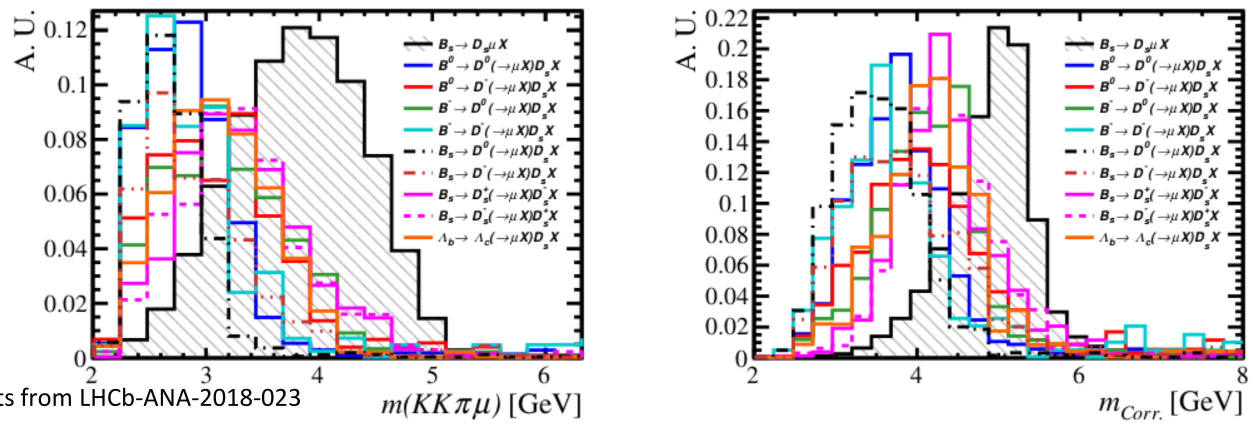
- **b-hadron double charm decays $B \rightarrow DDX$, with 1 charm hadron decaying semi-leptonically**

  suppressed due to D lifetimes

  <u>Strategy for background estimate:</u>

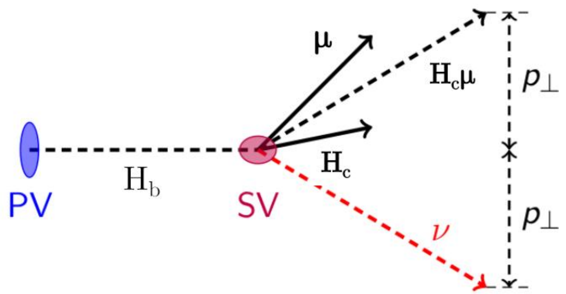
  Study M(charm+μ) and M_{corr}(charm+μ) distribution from dedicated MC samples

all the channels normalized to unit area



plots from LHCb-ANA-2018-023

$$m_{corr} = \sqrt{m^2_{H_c\mu} + p_T^2} + p_T$$



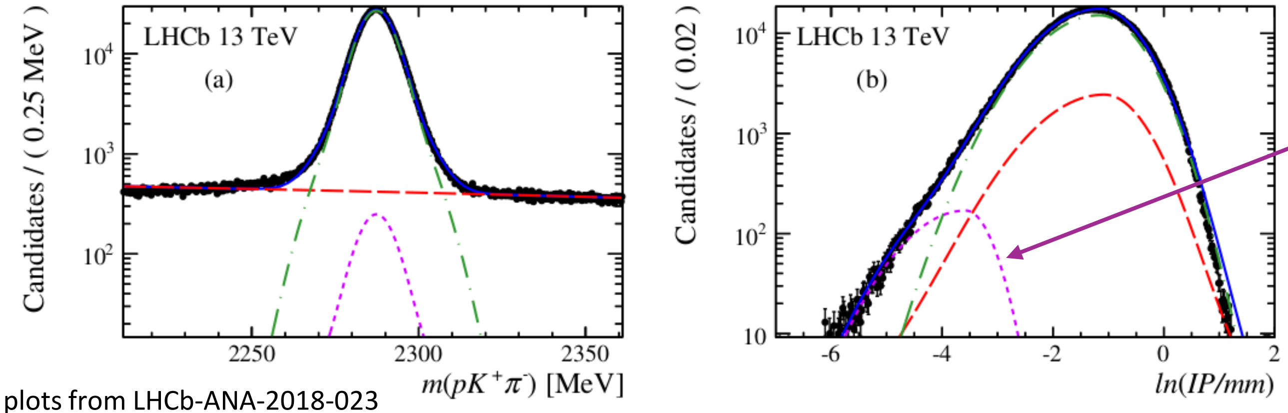
Run2 analysis: double charm background was ~6% for the B_s channel, ~1% for the Λ_b channel and below 1% for the B^- and B^0 channels

Background sources

Prompt-charm background

Strategy for background estimate:

ln(IP/mm) of the charm hadron is a powerful discriminant as demonstrated by the Run 2 analysis:



plots from LHCb-ANA-2018-023

prompt charm background in case of Λ_c
ln(IP/mm) shape is determined by examining a prompt enriched sample in MC

- Run2 analysis: contribution was ~ 1% for all channels
- A cut on the ln(IP/mm) was applied in the selection and a 1d fit was used to determine the signal yield
- Run3 analysis: exploring the possibility to perform a 2d fit on the charm hadron mass and ln(IP/mm)

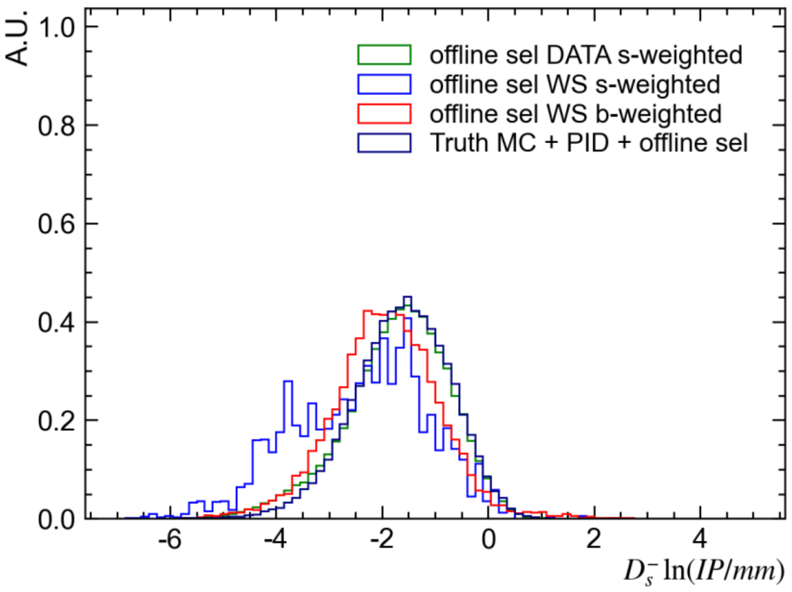
Background sources

- **Combinatorial background**
  - combinations of tracks from pp → bbX events, where one b hadron decays into a charm hadron and the other b hadron decays semi-muonically
  - fake muon

Strategy for background estimate:

Data driven estimate, using the sample collected with the same signal selection, but with the muon track having the wrong sign. Expected contribution O(1%)

Prompt Charm background

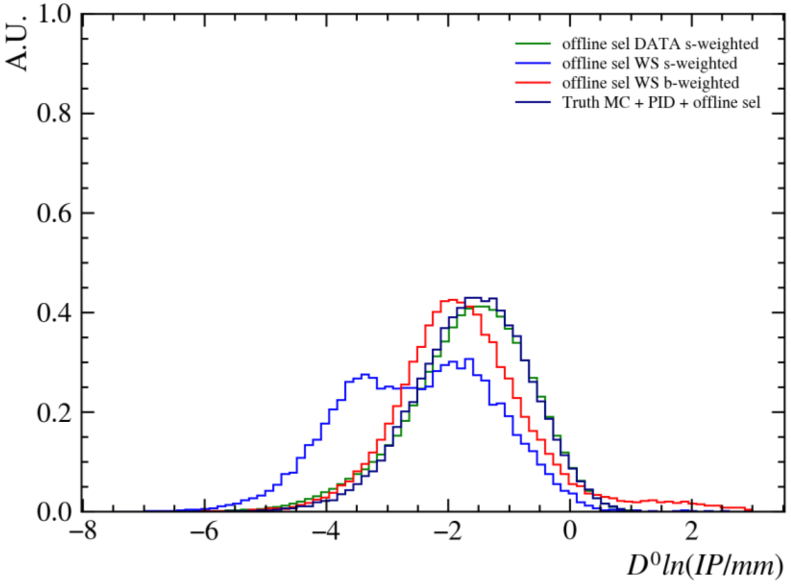


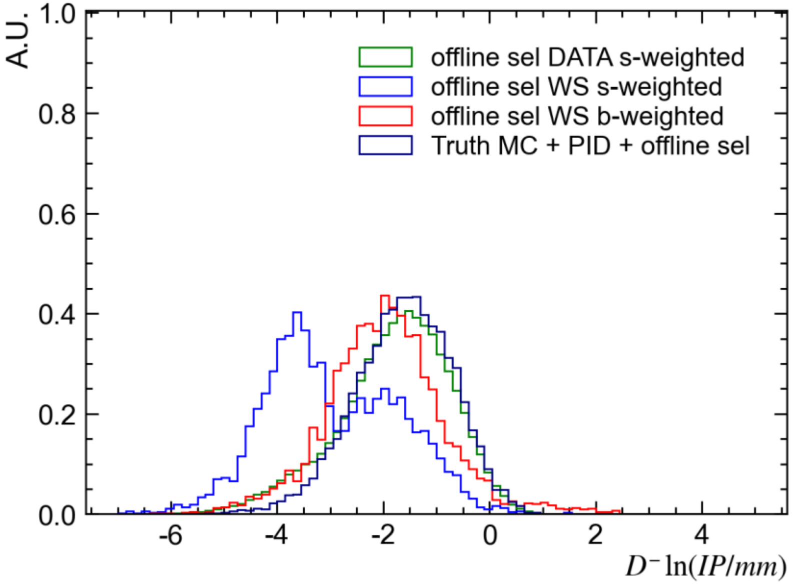
- highly suppressed by B FD z > 2mm cut, but still remainder most prominent in $D^+\mu^-$ channel. Slight presence also in $D^0\mu^-$ before this cut.

- Highly Suppressed in runs I or II

- Currently being further investigated:
  - the multiplicity of tracks or vertices has little correlation with the prompt-charm background

- Find the right template, including as a background in yield extraction and adding a log(IP) dimension is one potential avenue.





Signal Yields

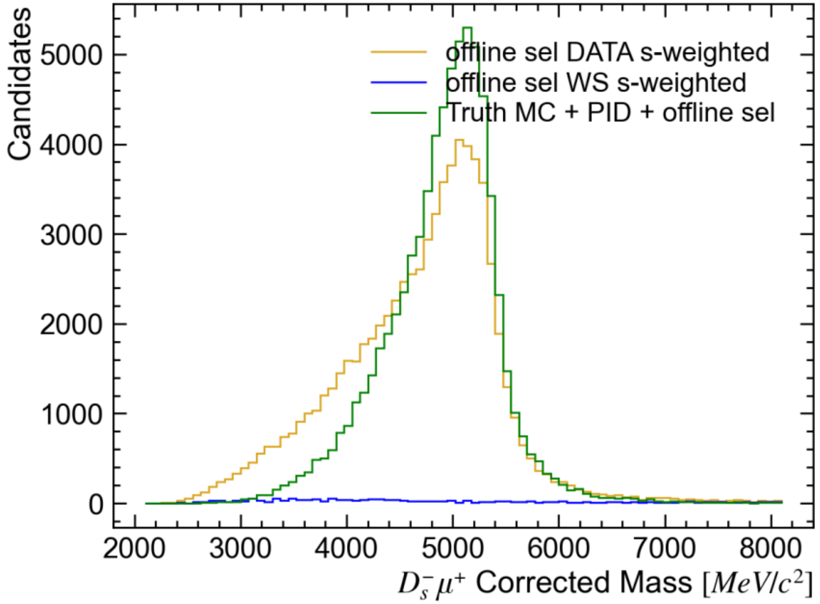
Procedure: Still developing, considering either: 2D fit with log(IP) and M_{corr} on the s-weighted data, or a 3D fit including the charm mass for unweighted data.

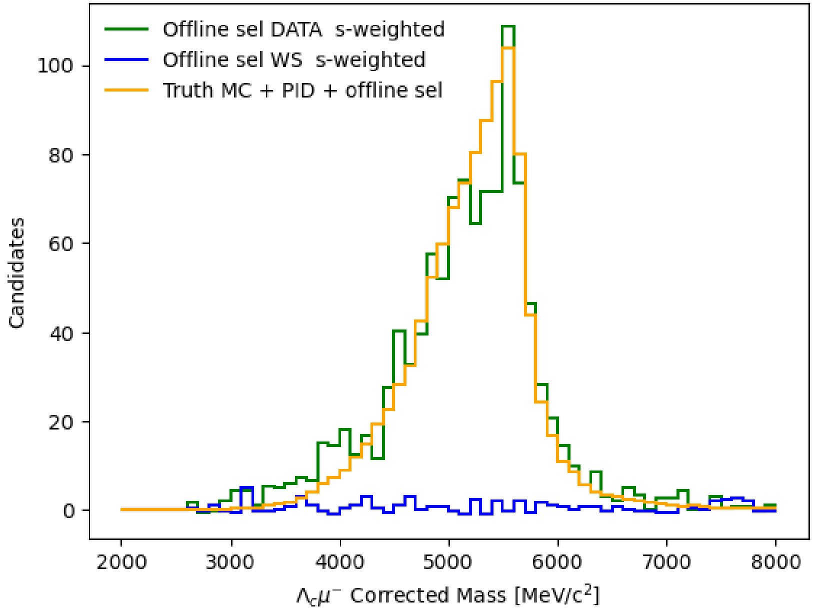
Fit:

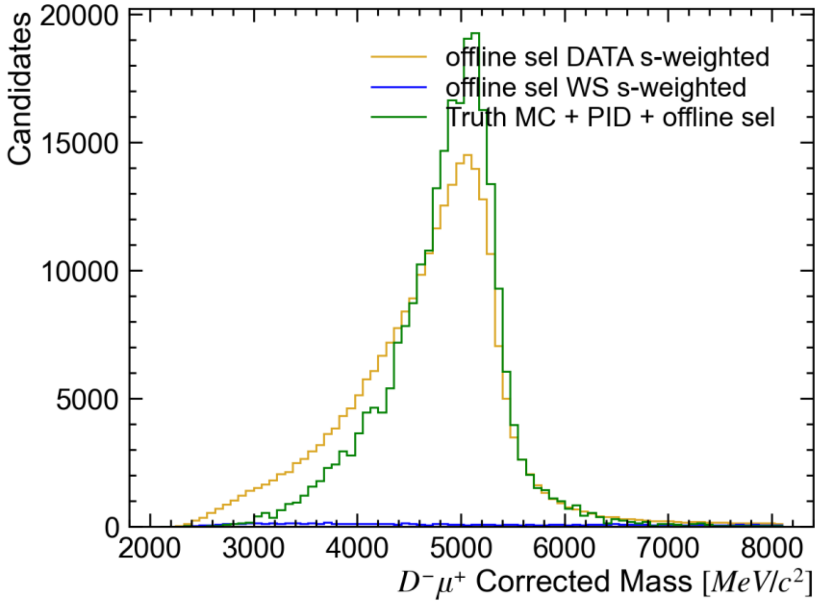
- **Signal: Signal MC Template,**

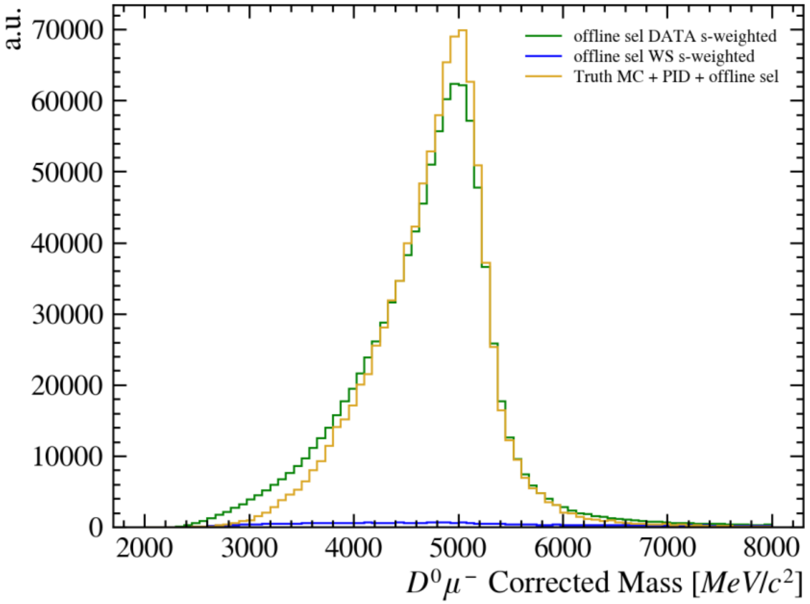
- **backgrounds:**
  - **DD: MC Templates,**
  - **Combinatoric: Removed with s-weight from D mass fit,**
  - **Prompt Charm: TBD,**
  - **Other charm+mu: WS Template**

Fit done in 10 bins of the b-hadron p_T, 3 bins of eta.

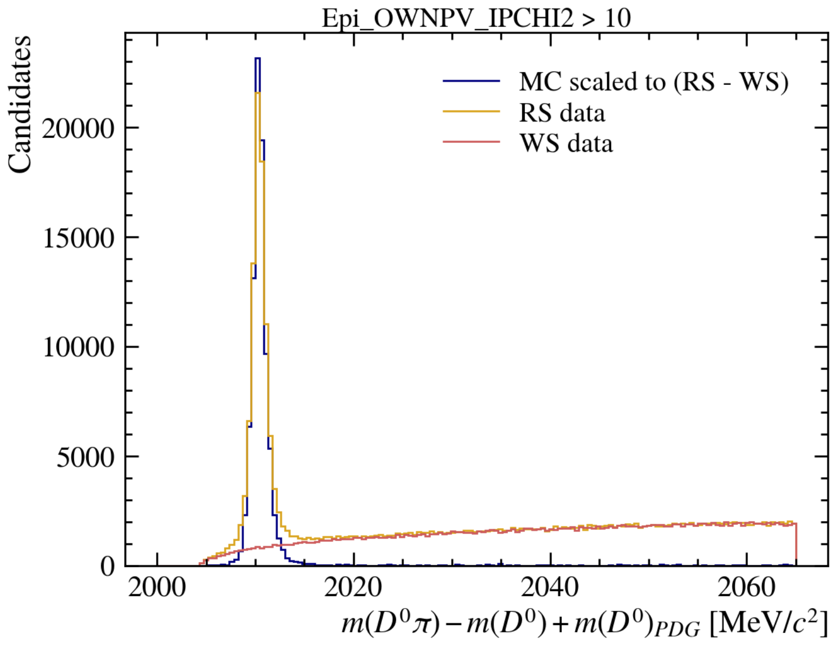


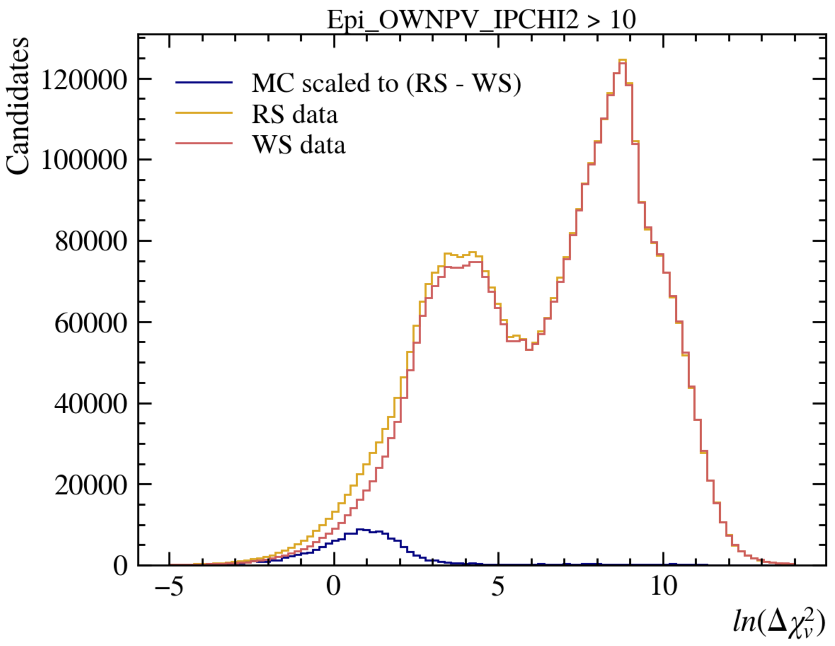






Cross-feed Yields





Note: RS and WS here refer to the sign difference between the D daughter hadrons, i.e.: $K^+\pi^-$ versus $K^+\pi^+$.

- **These plots have an additional cut of the extra particle impact parameter CHI2 with respect to the best primary vertex**
- **More Background characterization to be done.**
- **Goal: 2D fit to M(D0h) – M(D0) & log(ΔIPχ2) of h to Dμ vertex,**
- **s-weight removes combinatorial background.**

Procedure in development for $D^0\pi^\pm$, D^0p, & $D^0K^\pm$

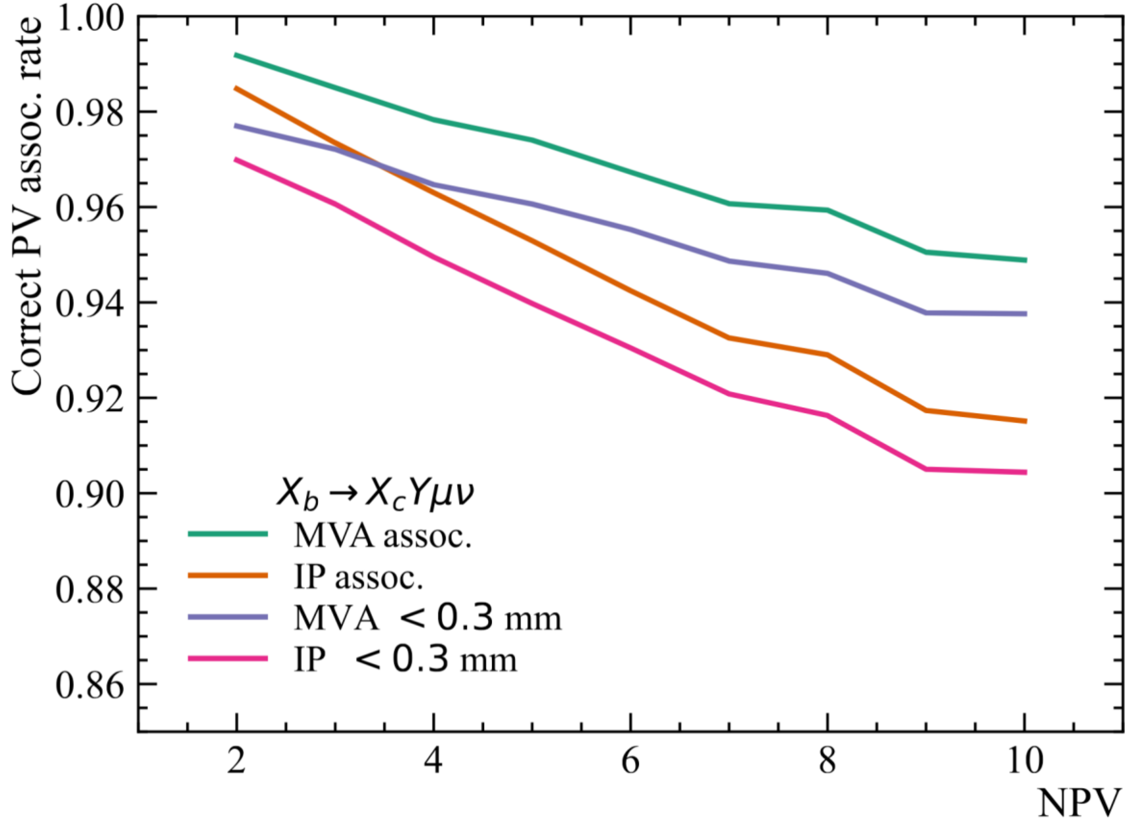
Primary vertex association

Wrong PV association gets worse with the number of Primary vertices (nPV)

- **BDT Can greatly improve this WRT IP association**

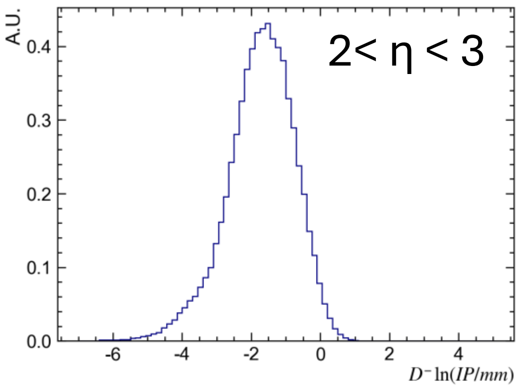
- **To properly reconstruct our binning variables, need to know what the true PV is**

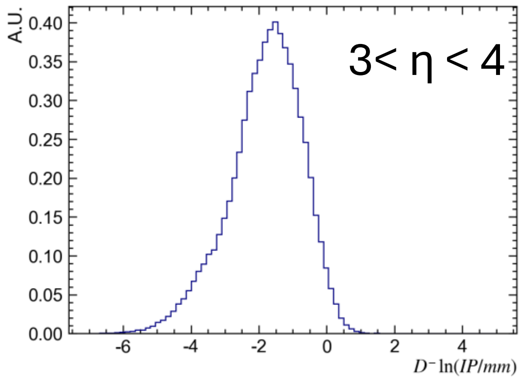
- **PV mis-association can also strongly affect the corrected mass.**

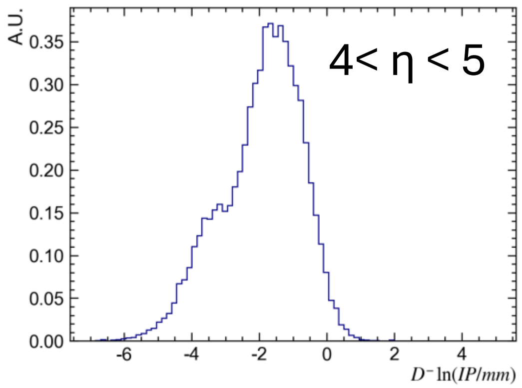


Performance of the BDT PV association algorithm compared to the default minimum IP association, as a function of the number of reconstructed PVs. Two definitions of correct rate are used: in one the requirement is that the reconstructed PV closest to the true position is chosen in the event; in the other the association is considered correct if the chosen PV position is within 300 µm of the true position in the z direction without regard for any other PV positions.

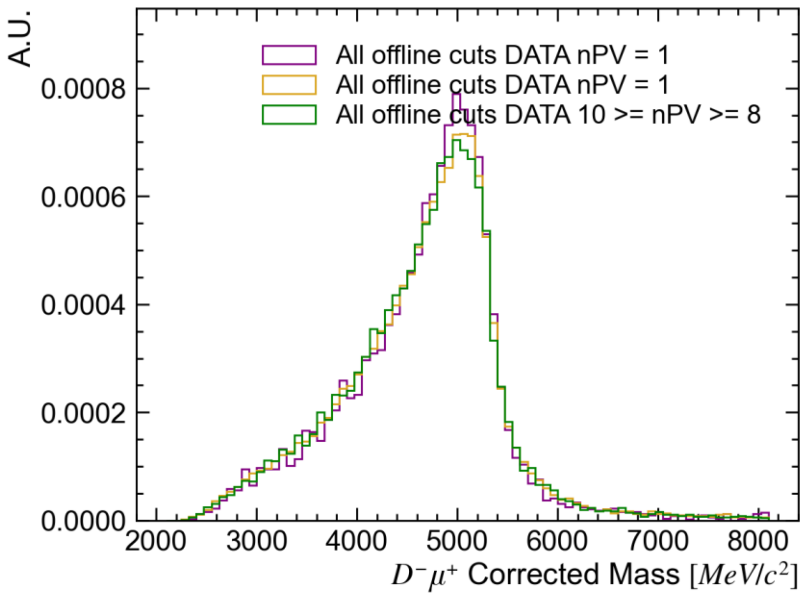
Binning

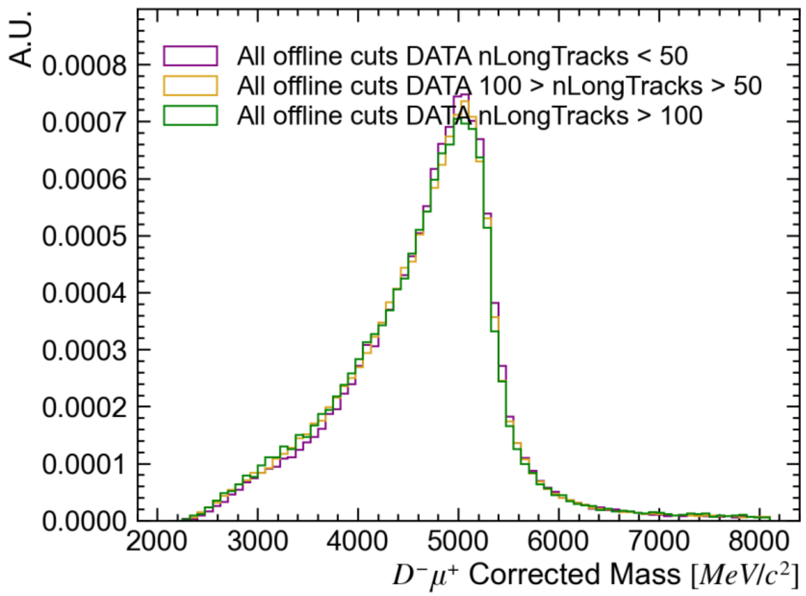






D^+ log(IP) in eta bins
Below: Corrected mass for different nPV regions (left) and nLongTracks regions (right)





Binning procedure in place for η to carry out yield extraction. p_T and nPV/nLongTracks forthcoming.

- **Prompt charm background highly correlated with eta**
- **Minimal nPV correlation with corrected mass.**

fs/fd with exclusive decays

In collaboration with:
M.Rotondo, P. De Simone,
K.Vos, J.Ruiz Vidal

$$\frac{N(B_s \to D_s^+ \mu\nu)}{N(B_d \to D^+ \mu\nu)} = \frac{f_s}{f_d} \cdot \frac{\tau_s}{\tau_d} \cdot \frac{\Gamma(B_s \to D_s^+ \mu\nu)}{\Gamma(B_d \to D^+ \mu\nu)} \cdot \frac{\epsilon_s}{\epsilon_d}$$

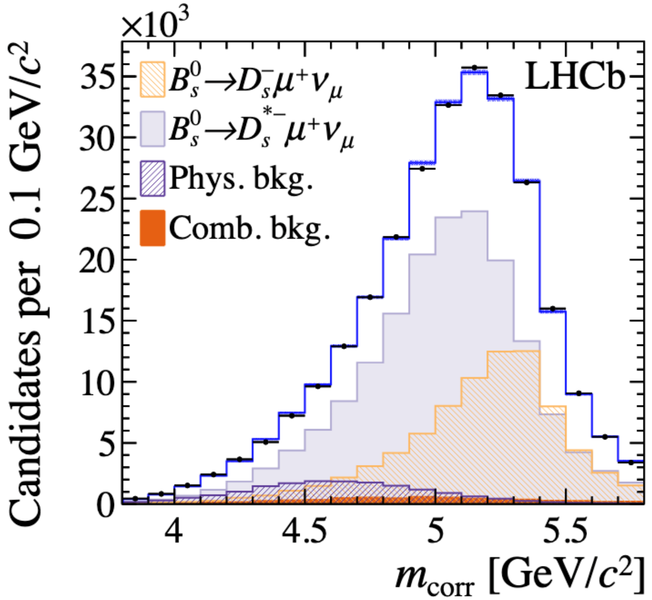
Fitting M_{corr}

PDG

LQCD

Bs: McLean et al. PRD 101 (2020) 7, 074513
Bd: MILC PRD 92 (2015) 3, 034506
B: H. Na et al Phys.Rev.D 92 (2015) 5, 054510

Simulation
+
Corrections from control sample



LHCb, Vcb measurement
PRD 101 (2020) 7, 072004

PROS

- Direct measurement of fs/fd
- Simultaneous extraction of Bs→Ds* and B→D*: additional constrain on fs/fd
- Measurement can be optimized at high q2, where LQCD has better uncertainties

CONS

- Theory error is not yet competitive with inclusive SL

Conclusion

The b-fragmentation fraction analysis using inclusive semileptonic decays is well-established and positioned for a publication in a short period

- Next steps:
    - Finalise offline selection,
    - Work on the background estimation
    - Include corrections (PID,tracking…)
    - Systematic checks with f+/f0 and D0 $\rightarrow$ K3π.

The fs/fd analysis using exclusive decays is at an early stage but will profit of several studies performed for the inclusive analysis.

Thanks for your attention

Additional Material

Offline selection

Track Selections

Muon:

- $p > 6\text{GeV/c}$, $p_T > 1\text{GeV/c}$, $IPCHI2 > 9$,
- $PID\_MU > 0$, $IsMUON == 1$,

Kaon:

- $p > 5\text{GeV/c}$, $p_T > 300\text{MeV/c}$, $IPCHI2 > 10$,
- $PID\_K > 4$,

Pion:

- $p > 5\text{GeV/c}$, $p_T > 300\text{MeV/c}$, $IPCHI2 > 10$,
- $PID\_K < 2$,

Proton:

- $p > 8\text{GeV/c}$, $p_T > 250\text{MeV/c}$, $IPCHI2 > 4$,
- $PID\_P > 0$,
- $PID\_P - PID\_K > 0$,

All Tracks

- $TRACKCHI2DOF < 2.5$,
- $TRACKGHOSTPROB < 0.3$,
- $ProbNN\_\{track\} > 0.3$

Other Offline Selections:

Composite Hadrons:

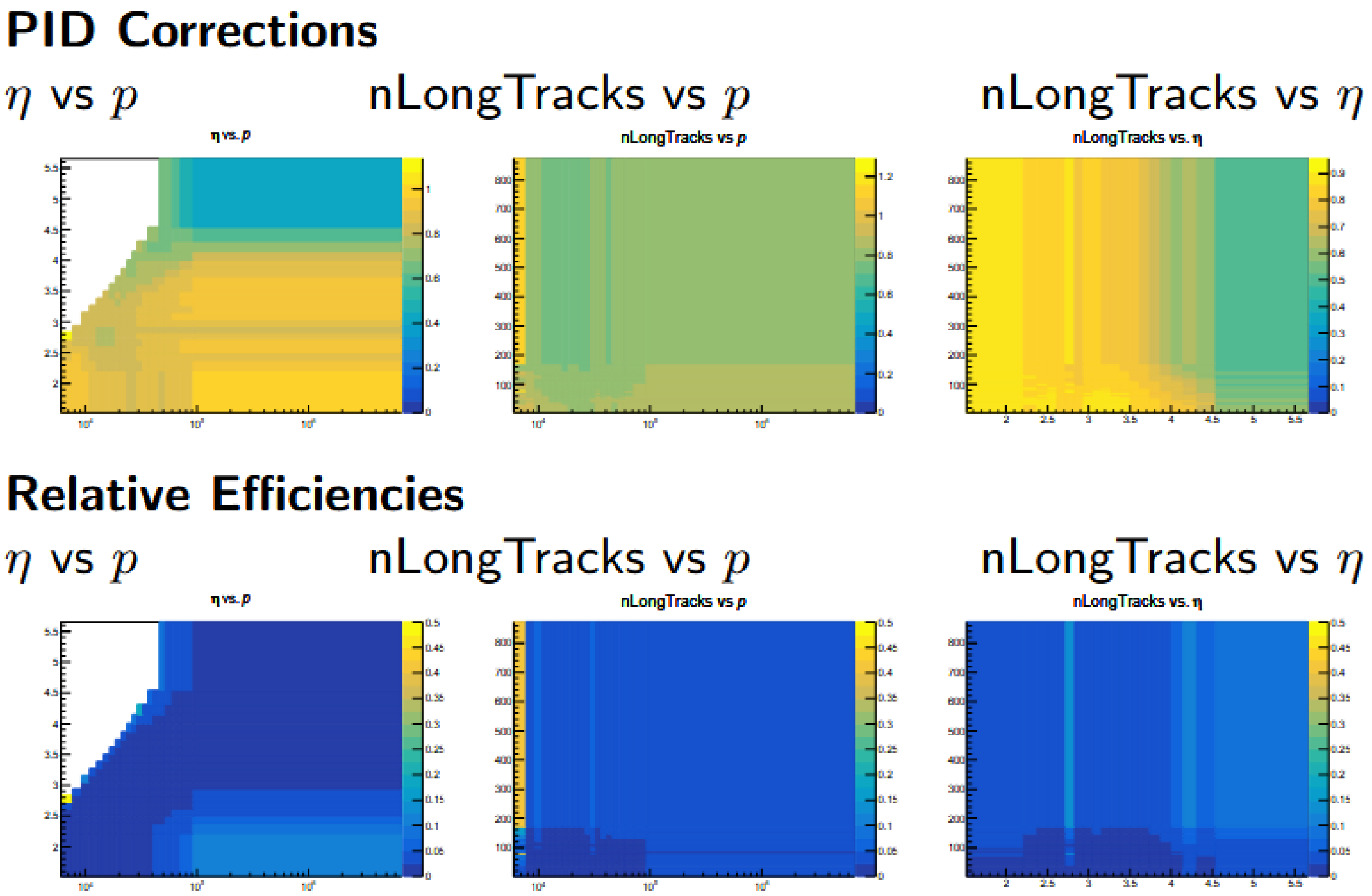
- $\{B\}\_ENDVERTEX\_CHI2DOF < 4$,
- $\{D\}\_ENDVERTEX\_CHI2DOF < 4$,
- $\{B\}\_ENDVERTEX\_VZ - \{D\}\_ENDVERTEX\_VZ < 0$,
- $Delta\_openingsAngle(\mathrm{K}^{\pm}\,\mu^{\pm}) > 0.005$,
- $\{B\}\_FD\_Z > 2\,\text{mm}$,

Event-Level:

- $nPVs < 11$,

Corrections

- **Data-driven PID, tracking and trigger HLT1/HLT2 corrections obtained using dedicated tools available at LHCb (PIDCalib, TrackCalib, TISTOS method)**
- **Final total efficiencies determined after including the data-driven corrections based on final state kinematics.**



Calibration performed using the recently released PIDCalib2

• p, η, nLongTracks bins determined such that each bin has comparable and reasonable statistics,

• Pre-selections and PID requirements in backup slides.