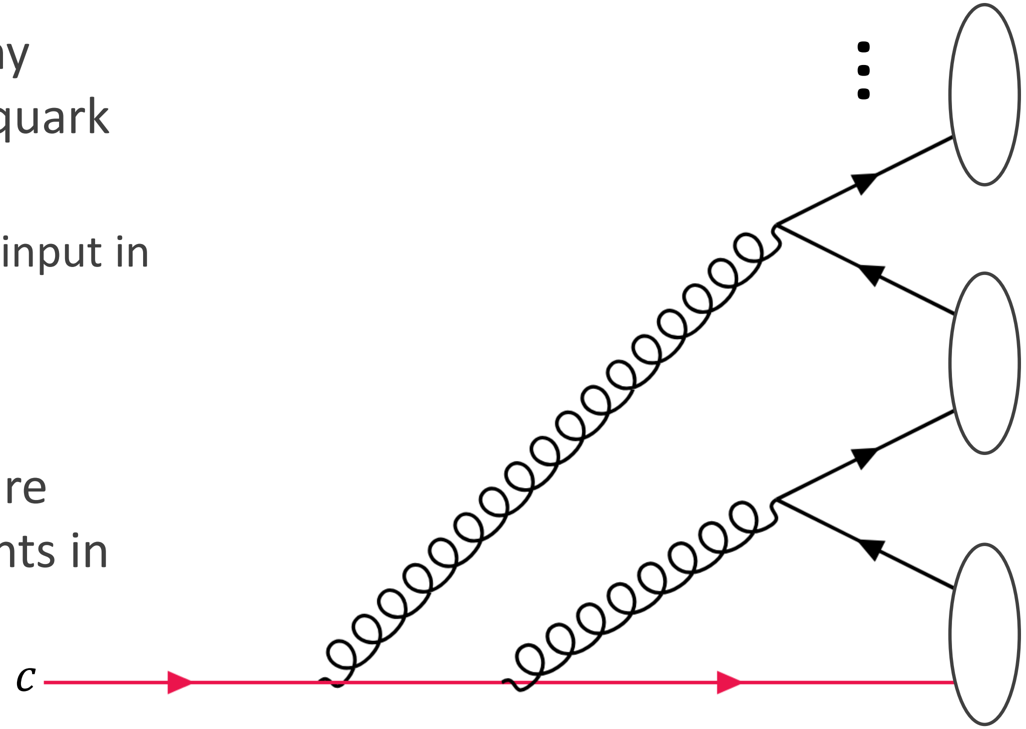
# Measuring $D^{*\pm}$ production in jets at the ATLAS Experiment

Eric Liu

# Motivation

- Understanding charm quark fragmentation can aid in many studies that involve charm hadrons, such as $H \rightarrow c\bar{c}$, top quark decays, BSM and SUSY searches.
  - Specifically, charm fragmentation functions are an important input in MC event generators for charm processes.



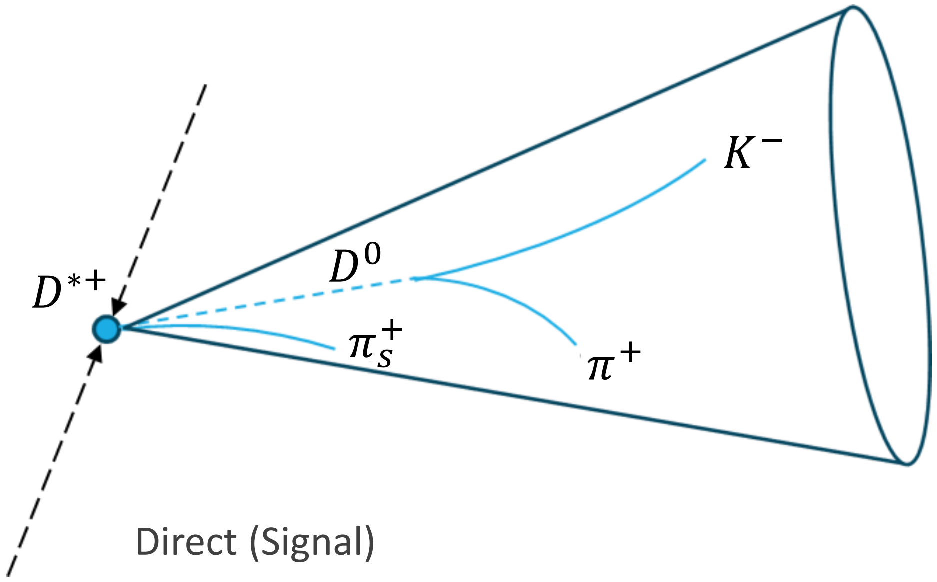

- Most measurements of charm fragmentation properties are from $ee/ep$ data, beneficial to compare with measurements in $pp$ environments.

- One important property is the fragmentation function which quantifies the energy carried by the charm hadron relative to the original charm quark.

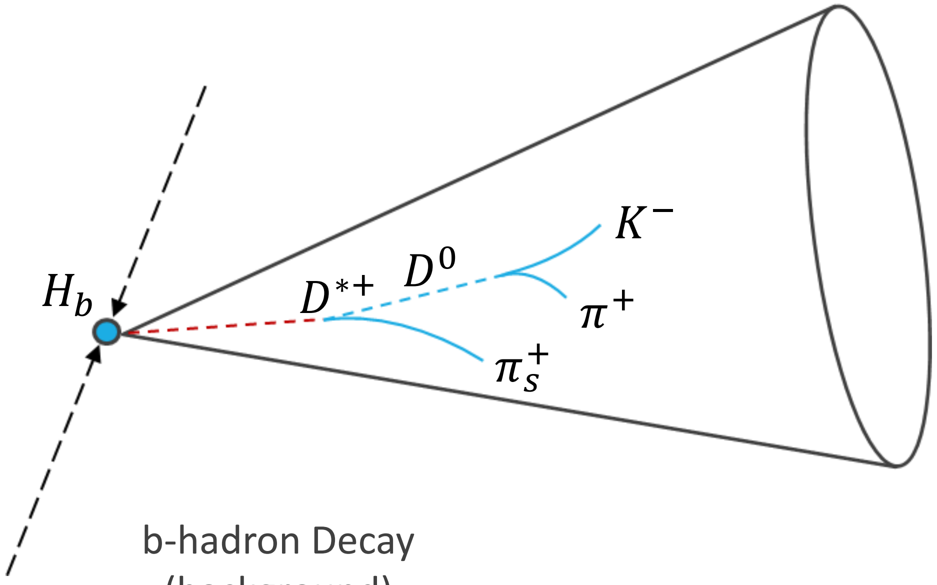
- The measurement of the production $D^{*\pm}$ mesons $(c\bar{d}/d\bar{c})$ within charm jets can be used to provide experimental information on the fragmentation function. (Henceforth denoted as $D^*$).

# Process under investigation

- Decay chain: $D^{*+} \rightarrow D^0 \pi_s^+$, $D^0 \rightarrow K^- \pi^+$ (and the charge conjugate)
- Expect $D^*$ mesons from two main production modes:
  - "Direct": $D^*$ produced at the $pp$ vertex (signal)
  - "b-Hadron $(H_b)$ decays": $H_b$ produced at the $pp$ vertex, subsequently decaying into a $D^*$ (background)
- The main background is the combinatoric background from the track selection and reconstruction.
- Analysis will look at the inclusive/multijet events using a high statistics dataset.




Direct (Signal)




b-hadron Decay (background)

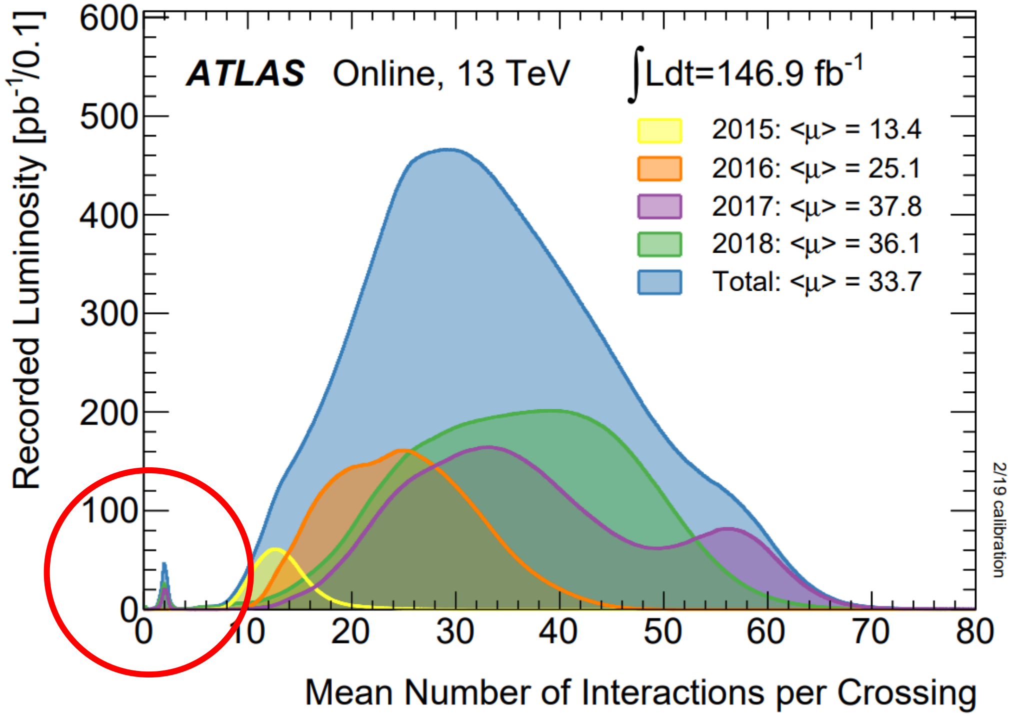
# Dataset + Trigger

- The last ATLAS publication on this study of $D^*$ meson production was published back in 2012, using Run 1 data ($\sqrt{s} = 7$ TeV, $\sim 0.30$ pb$^{-1}$) [1].

- This investigation will be using the Run 2 low pileup $pp$ collision data from 2017 + 2018 ($\sqrt{s} = 13$ TeV, $\mu \sim 2$).
  - Employ a single jet trigger with a 20 GeV threshold, collecting a dataset of $\sim 177$ pb$^{-1}$.

- Advantages of the low pileup dataset:
  - Availability of low threshold triggers
  - Lower combinatoric background
  - Less ambiguity on reconstructing primary vertex (important for $D^0$ decay length reconstruction)

**N.B.** the results shown in this talk represents only a small subset of the data, around 1%.

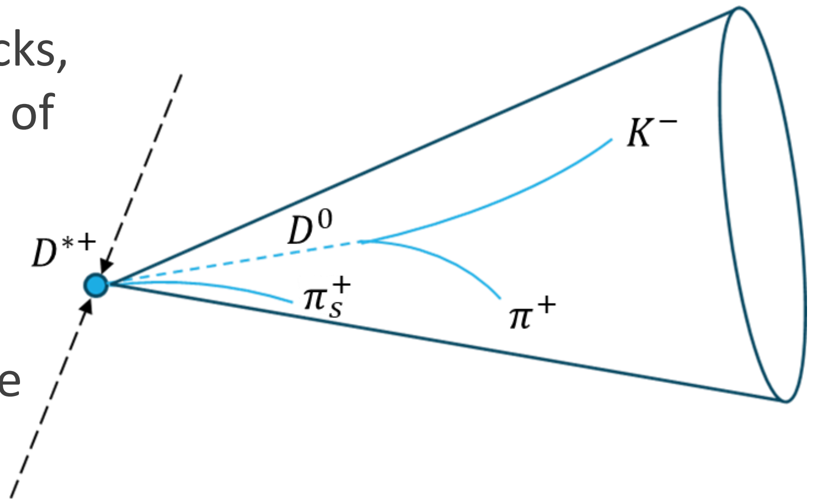



[1]: Phys.Rev.D 85 (2012) 052005 (arXiv:1112.4432)

# Event Reconstruction + Cut Selection

- The $D^0$ candidate is formed from the selection of two oppositely charged tracks, which is then combined with a third track with a charge opposite to the track of the kaon hypothesis to form the $D^*$ candidate.

- A $D^*$ candidate passing the kinematic selection is then only considered for the final selection if coincides with a jet of radius 0.4.




| | |
|---|---|
| Jet Selection | $p_T^{jet} > 20$ GeV, $\lvert\eta_{jet}\rvert < 2.5$ |
| $D^0$ Vertex Fit Quality<br>$D^0$ Mass | $\chi^2 < 10$<br>$\lvert m(K^{\mp}\pi^{\pm}) - m(D^0)\rvert < 72$ MeV |
| $\pi_s$ | $p_T > 500$ MeV |
| $D^*$<br>$D^*$ + Jet Separation<br>$D^*$ Mass Difference ($\Delta$m) | $p_T > 7.5$ GeV<br>$\Delta R(D^*, jet) < 0.4$<br>$139 \text{ MeV} < \lvert m(K^{\mp}\pi^{\pm}\pi_s^{\pm}) - m(K^{\mp}\pi^{\pm})\rvert < 180 \text{ MeV}$ |

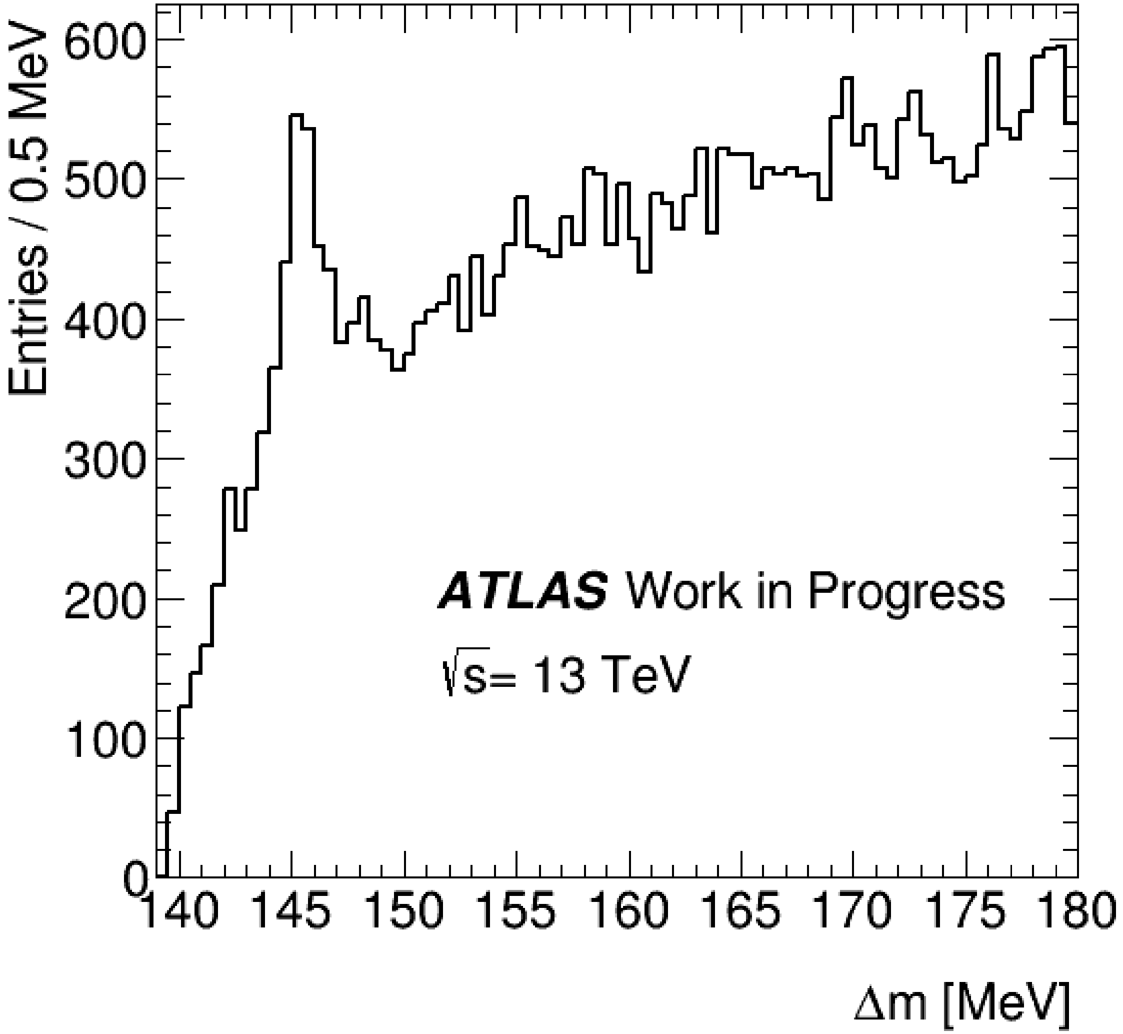
# First Look - Δm Distribution

- The mass difference $\Delta m = \left| m(K^{\mp}\pi^{\pm}\pi_s^{\pm}) - m(K^{\mp}\pi^{\pm}) \right|$ was chosen as the observable as it cancels some of the mass resolution effects from the $D^0$ reconstruction.

- Signal expected at $m(D^*) - m(D^0) \sim 145.4\ \mathrm{MeV}$

- Initially, enforced a cut on the $D^0$ transverse decay length significance, $\frac{L_{xy}}{\sigma_{L_{xy}}} > 3.$
  - $D^0$ lifetime $(\sim 0.4\ \mathrm{ps})$ sufficiently long to be measurable given the detector resolution.
  - Three sigma significance cut for the reconstruction of a displaced vertex.

- This cut helps to remove a large proportion of the combinatoric background by selecting the real $D^0$ candidates.

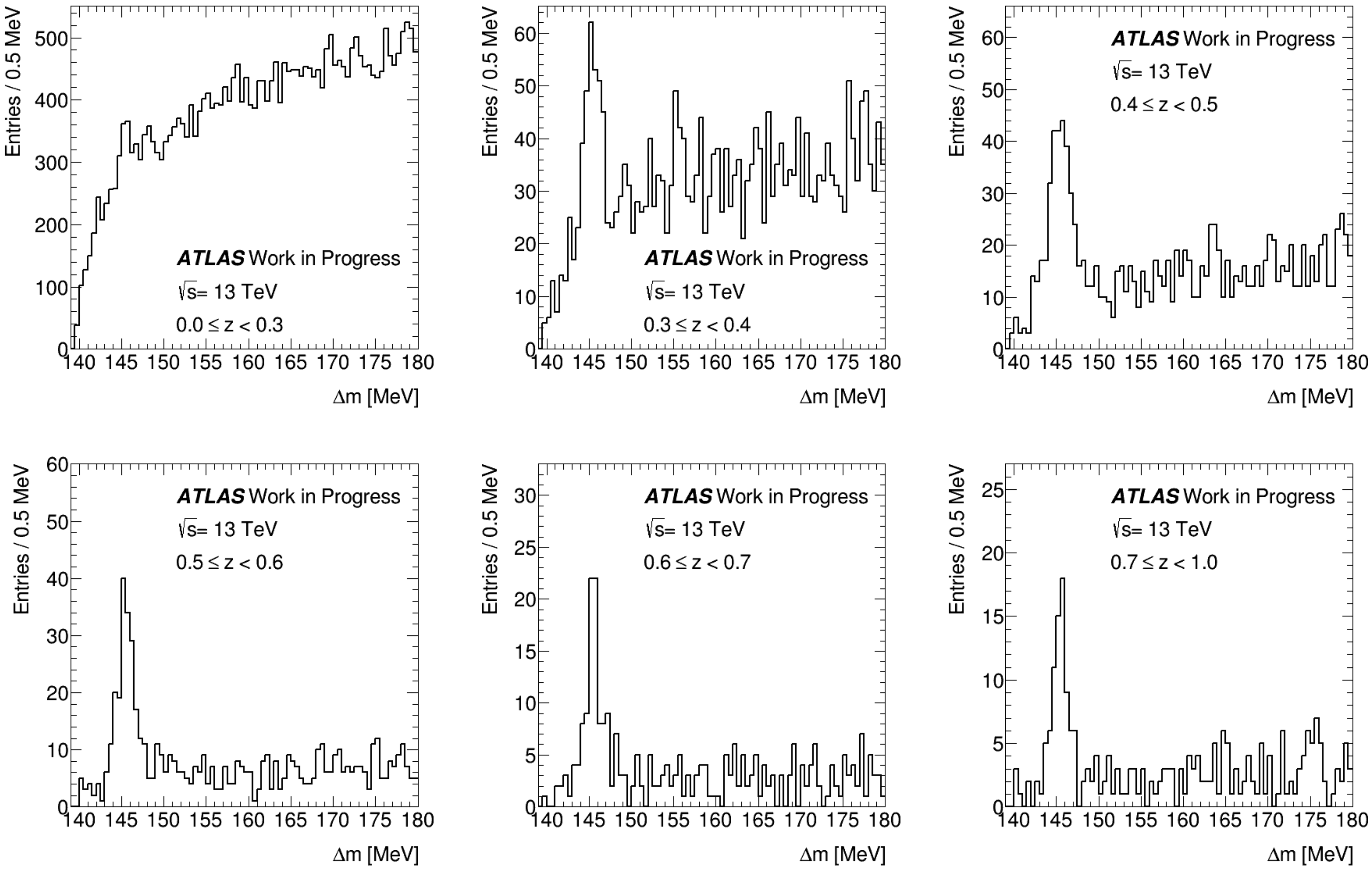


# First Look – Fragmentation Variable

- To study the fragmentation function quantitatively, introduce the fragmentation variable, $z$.
- $z = \dfrac{p_{\parallel}(D^*)}{E_{jet}}$

$p_{\parallel}(D^*)$: $D^*$ momentum parallel to jet axis

$E_{jet}$: Jet energy



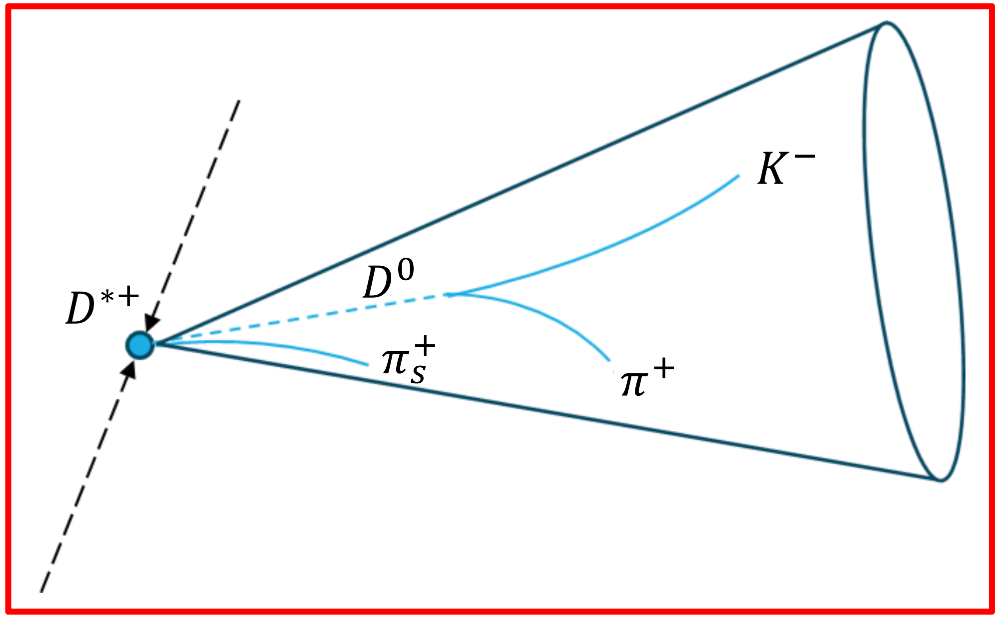
# First Look – Issue with the $L_{xy}$ Cut

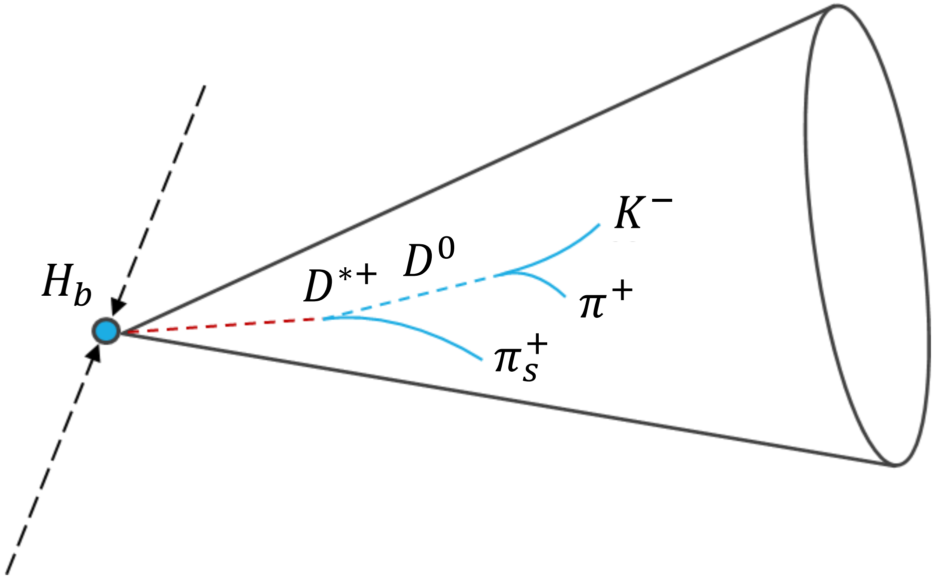
- For the fragmentation function, only want the $D^*$ that come from the direct production.

- The previous Δm distributions show an admixture of $D^*$ from both direct/$H_b$ production modes.
  - Cannot use the $D^*$ counts from these distributions.

- The $L_{xy}$ cut only reduces the combinatoric background component – no effect on the separation of the production modes.

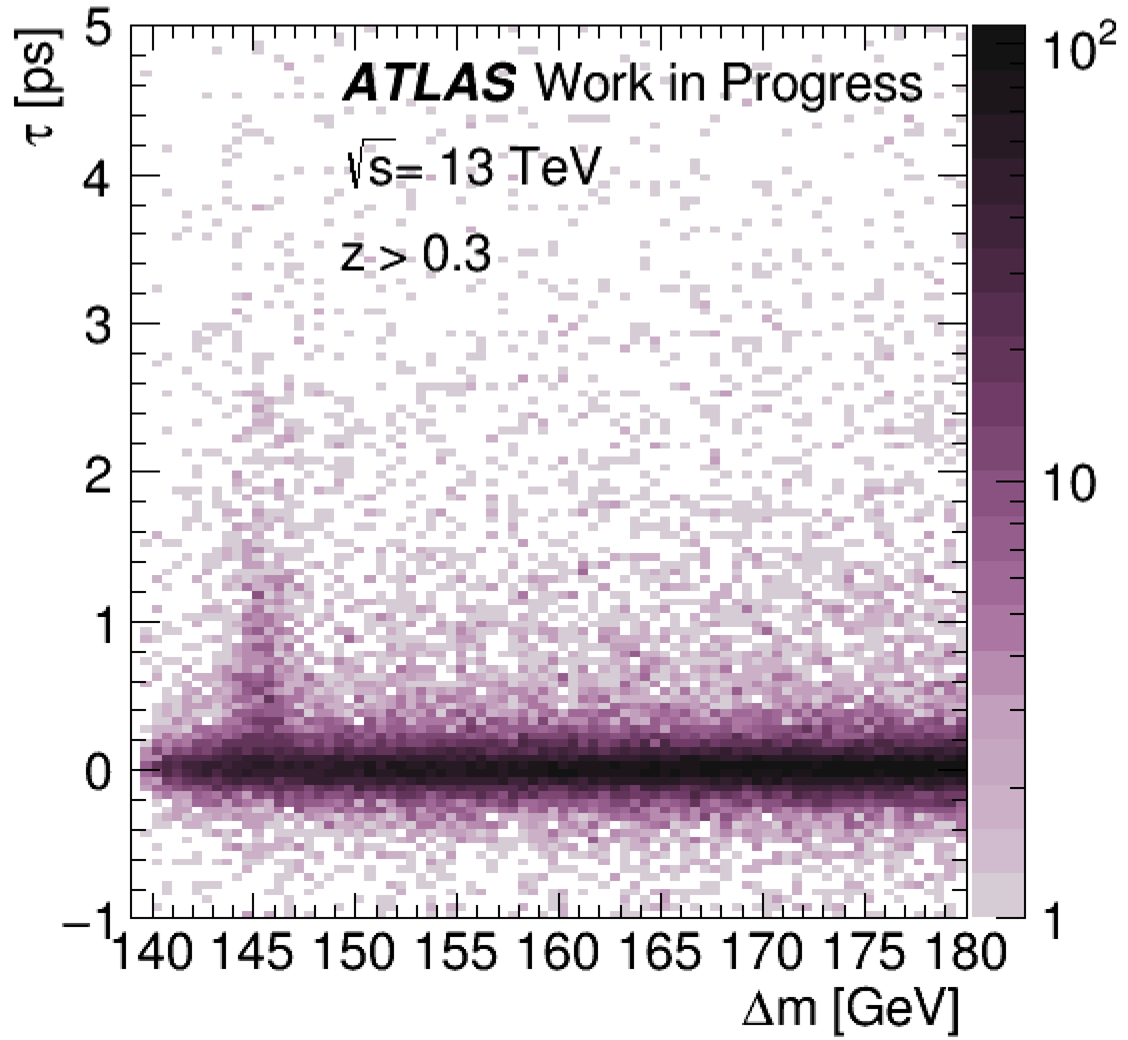
➢ There is no single cut that can be applied at the selection stage to effectively separate the direct/$H_b$ contributions.



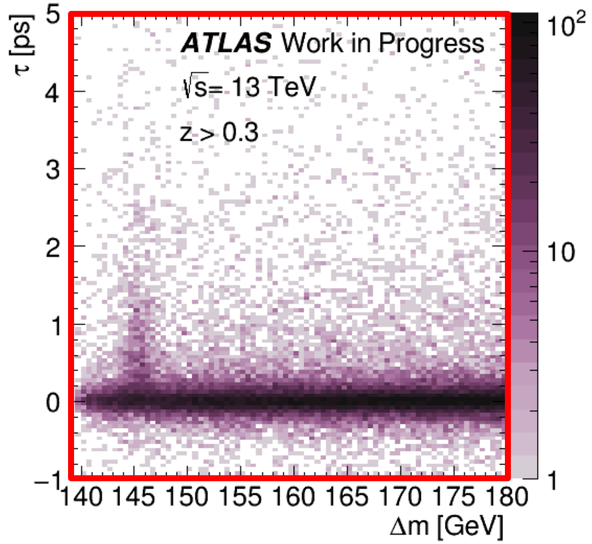
# Fitting – 2D Fit

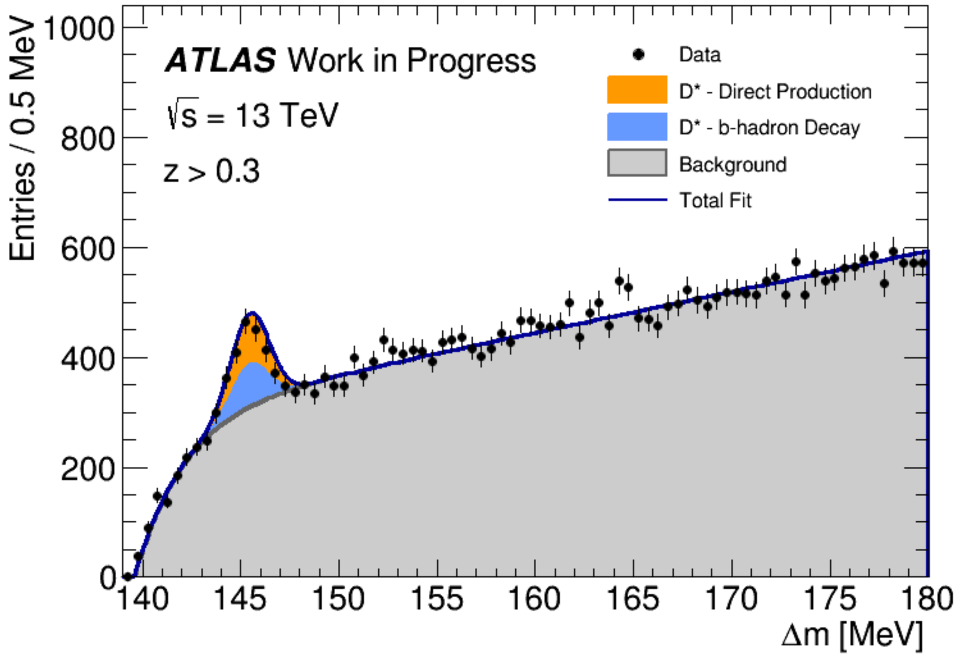
- Solution: Employ a 2D fit to independently estimate the number of $D^*$ from the direct and $H_b$ contributions.
- Introduce a new variable present in the analysis in which the direct/$H_b$ component PDFs are not-degenerate → $D^0$ pseudo-proper decay time ($\tau$)
  - Computed from the $D^0$ $L_{xy}$: $\tau(D^0) = m(D^0) \cdot \frac{L_{xy}(D^0)}{p_T(D^0)}$
  - For direct: this is $D^0$ lifetime, $\sim 0.4$ ps
  - For $H_b$: this is a convolution of the $H_b$ and $D^0$ lifetimes
- Construct 2D PDFs in $\Delta$m and $\tau$ for the direct/$H_b$/background components and fit to the 2D dataset via an extended maximum likelihood fit. (PDF details in backup)
- Advantage: all the data are used, and no assumptions are required for the $H_b$/background components of the fit.

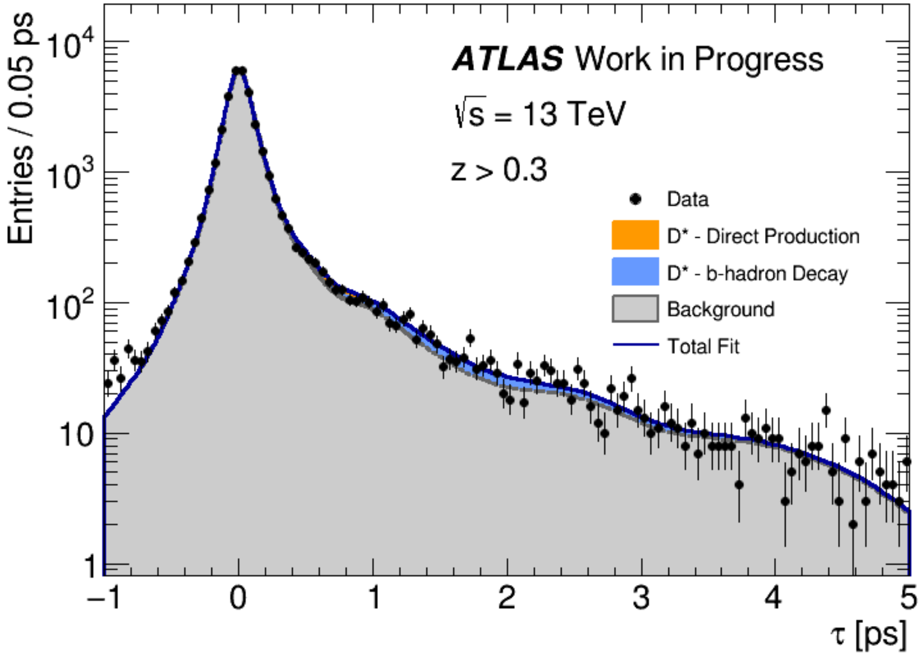



2D histogram of the $[\Delta m, \tau]$ distributions for the $D^*$ candidates in data

# 2D Fit Results

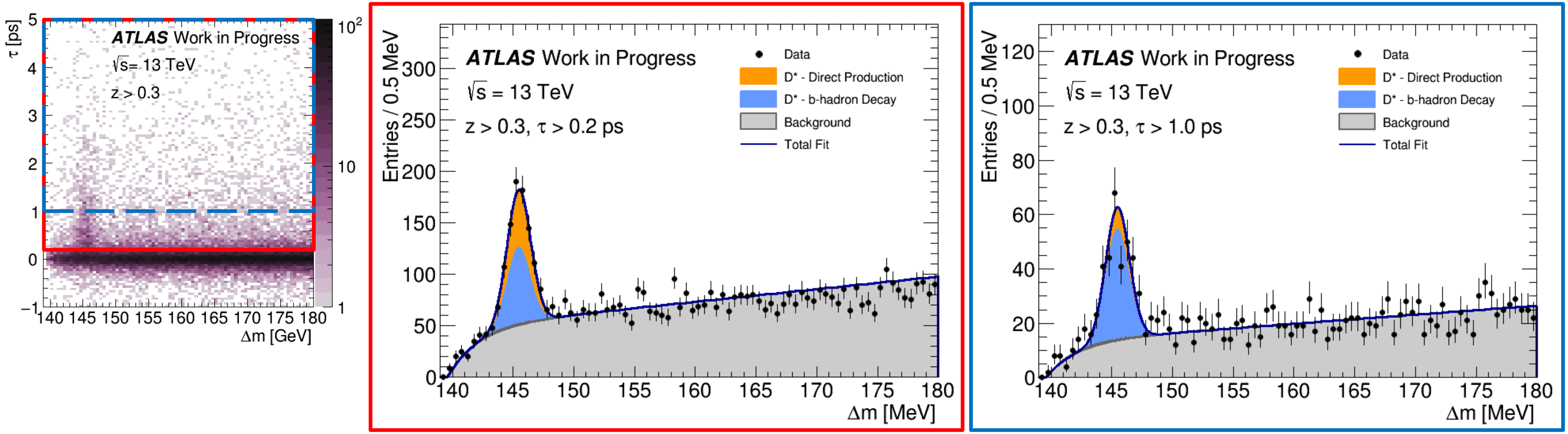



- Full 2D fit shown in the mass difference and lifetime components for the whole range of $\Delta m$ and $\tau$.
- $z > 0.3$ shows the distribution with reduced combinatoric background.
- Pseudo-proper decay time, $\tau$, background is modelled in a data-driven approach using a (Gaussian) Kernal Density Estimate (KDE).
- Can make projections in $[\Delta m, \tau]$ to see how the fit and distributions above evolve.

# 2D Fit Results – Slices in $\tau$

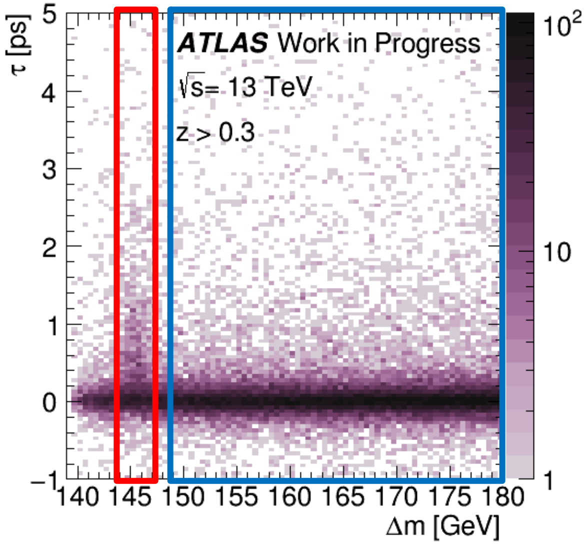


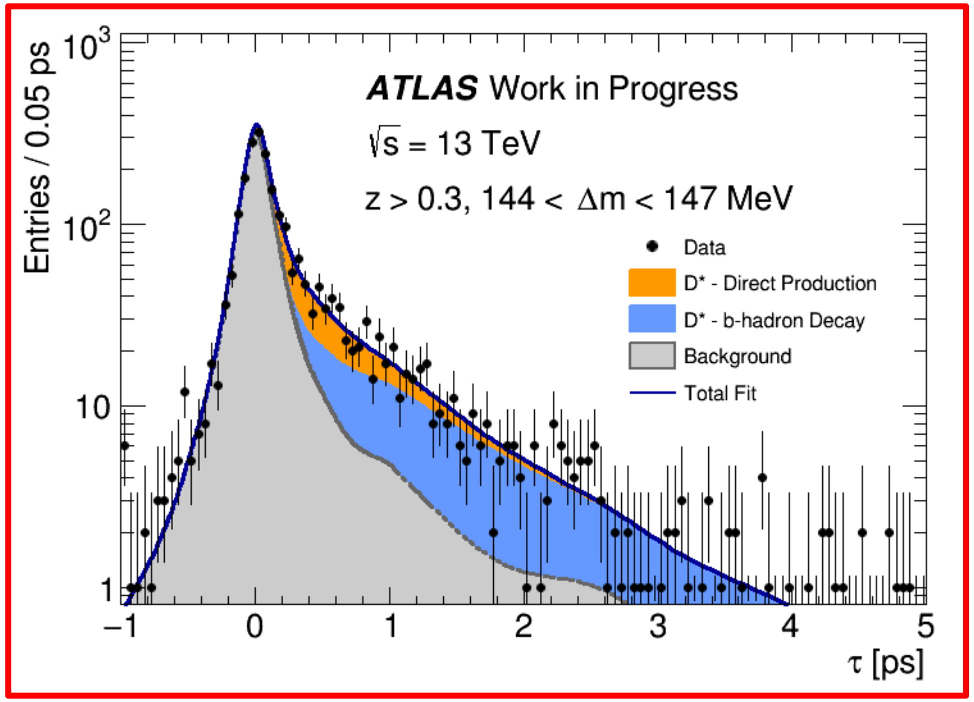

- Mass difference plots with the same single fit as before, but with different projections in $\tau$.

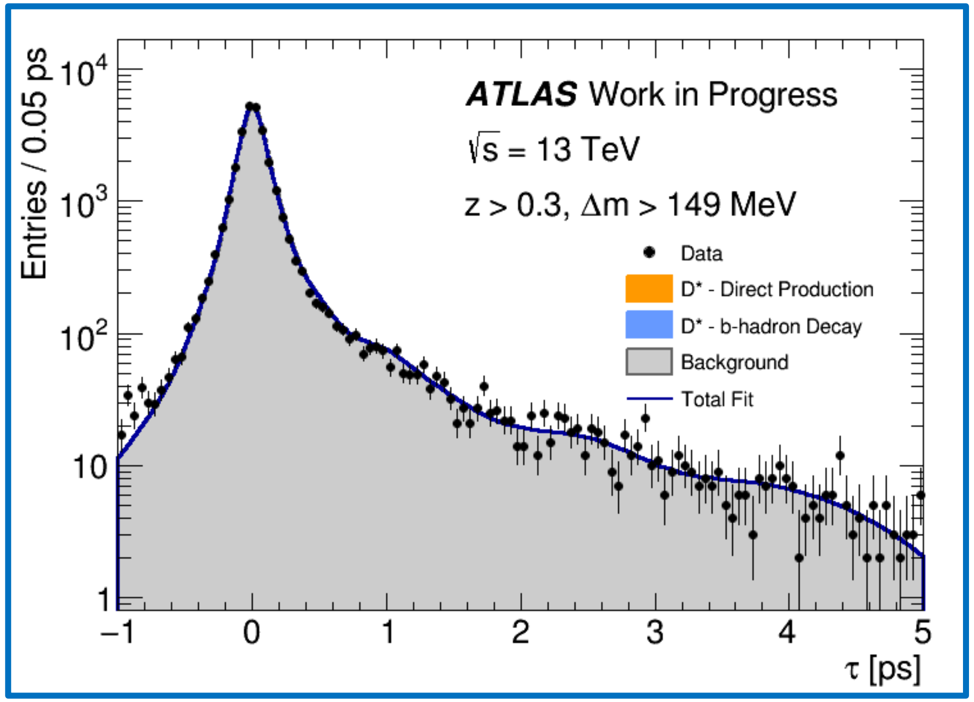
□ $\tau > 0.2$ ps → slice removes much of the combinatoric background.

□ $\tau > 1.0$ ps → region dominated by $D^*$ from $H_b$ decays.

# 2D Fit Results – Slices in Δm



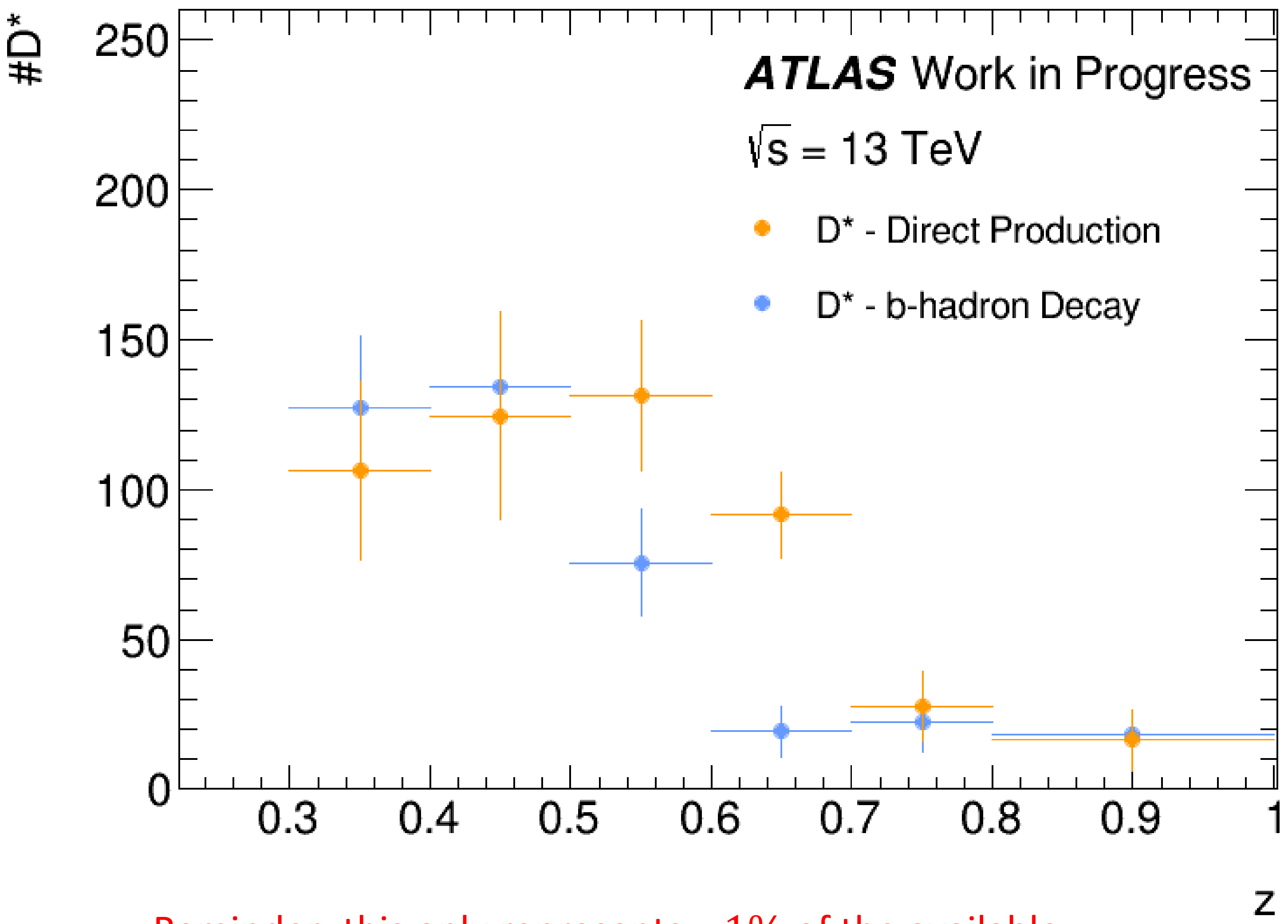

- Lifetime plots with the same single fit as before, but with different projections in Δm.

□ $144 < \Delta\text{m} < 147$ MeV → $D^*$ signal region.

□ $\Delta\text{m} > 149$ MeV → background region of the Δm distribution.

# Towards a Fragmentation Function

- Perform the 2D fit to the data in bins of $z$.
  - Extract the yield of $D^*$ from the direct/$H_b$ modes.
- $z < 0.3$ is currently dominated by background with our sample dataset.
- Expect the $D^*$ from the $H_b$ decays to have a softer $z$-spectrum compared to the direct production.
- Need to unfold this distribution to acquire the fragmentation function.
  - Consider resolution in $z$, reconstruction and selection efficiencies, etc.




Reminder: this only represents ~1% of the available dataset and has corrected for experimental effects

# Summary

- Measurement of $D^*$ production within jets, focusing on the direct yield to perform a measurement of the charm quark fragmentation function.

- Implementing a two-dimensional fit to separate the direct and b-hadron $D^*$ contributions.

- Acquired the $D^*$ yield as a function of the fragmentation variable as the first step towards the fragmentation function.

- Starting to look at MC samples and soon the unfolding strategy to acquire the fragmentation function.

# Backup

# Fitting – 2D Fit Component PDFs

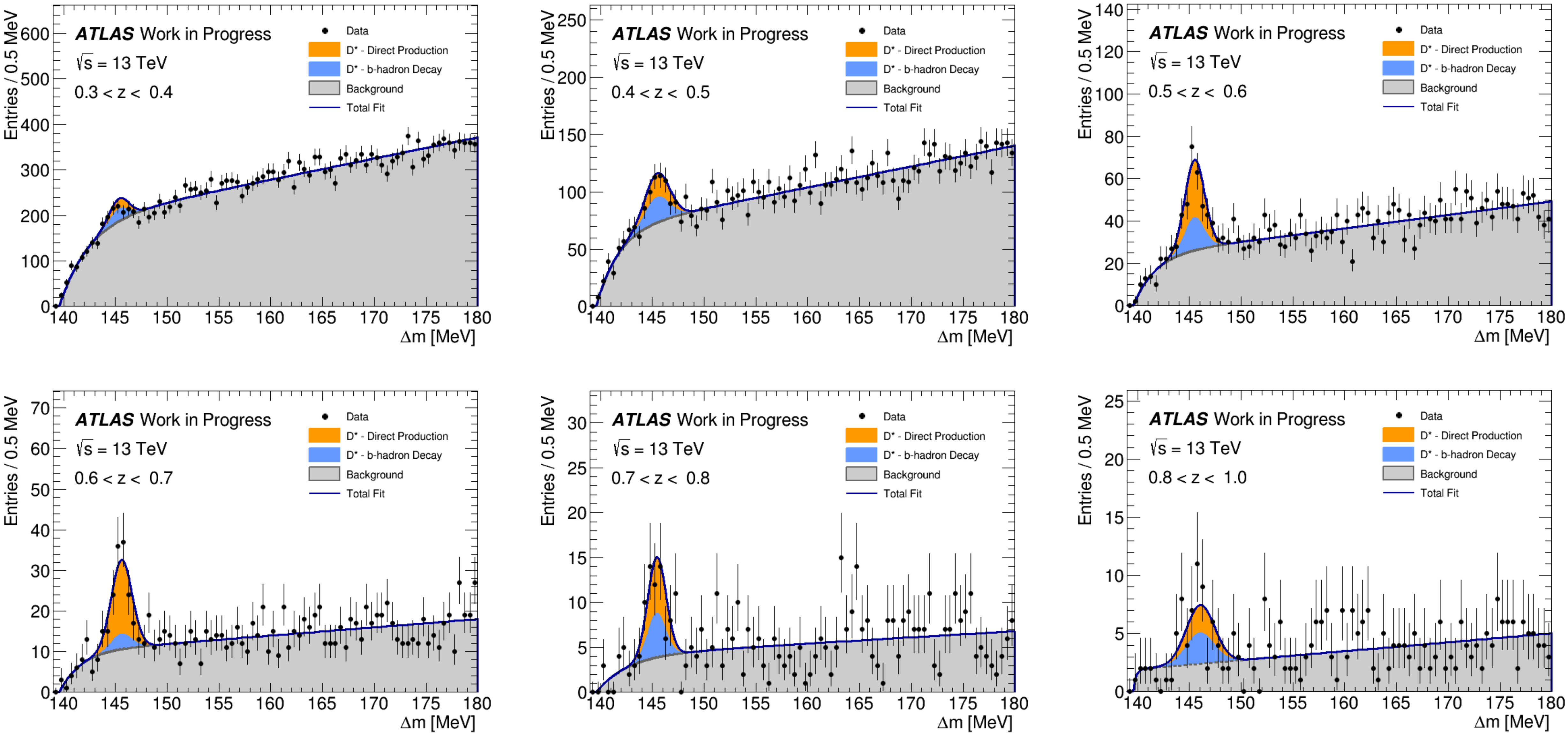
Δm Fit Model

- **Direct/$H_b$ $D^*$ signal PDF**: Gaussian
- **Background PDF**: implemented ROOT class "RooDstD0BG"
  - $\mathrm{RooDstD0BG}(\Delta m \mid m_0, A, B, C) = \left(1 - e^{-\frac{\Delta m - m_0}{C}}\right) \cdot \left(\frac{\Delta m}{m_0}\right)^A + B\left(\frac{\Delta m}{m_0} - 1\right)$

$\tau(D^0)$ Fit Model

- **Direct Component PDF**: Exponential decay (lifetime fixed at the $D^0$ PDG value) convoluted with a Gaussian resolution function (mean fixed to zero)
- **$H_b$ Component PDF**: Describe the $H_b \to D^*$ with a Gaussian resolution convoluted with an exponential decay (lifetime free). This function is then convoluted with another exponential decay (lifetime fixed at the $D^0$ PDG value) for the subsequent $D^0$.
- **Background PDF**: (Gaussian) Kernel Density Estimate (KDE) using sideband data taken from a background-enriched region of the Δm distribution
- → An analytical model for the $\tau$-background is being developed but is currently unstable due to the low statistics.

# 2D Fit Results – Bins of *z*, Mass Difference Plots



# 2D Fit Results – Bins of $z$, Pseudo-Proper Decay Time Plots