B-Fragmentation Analysis Study

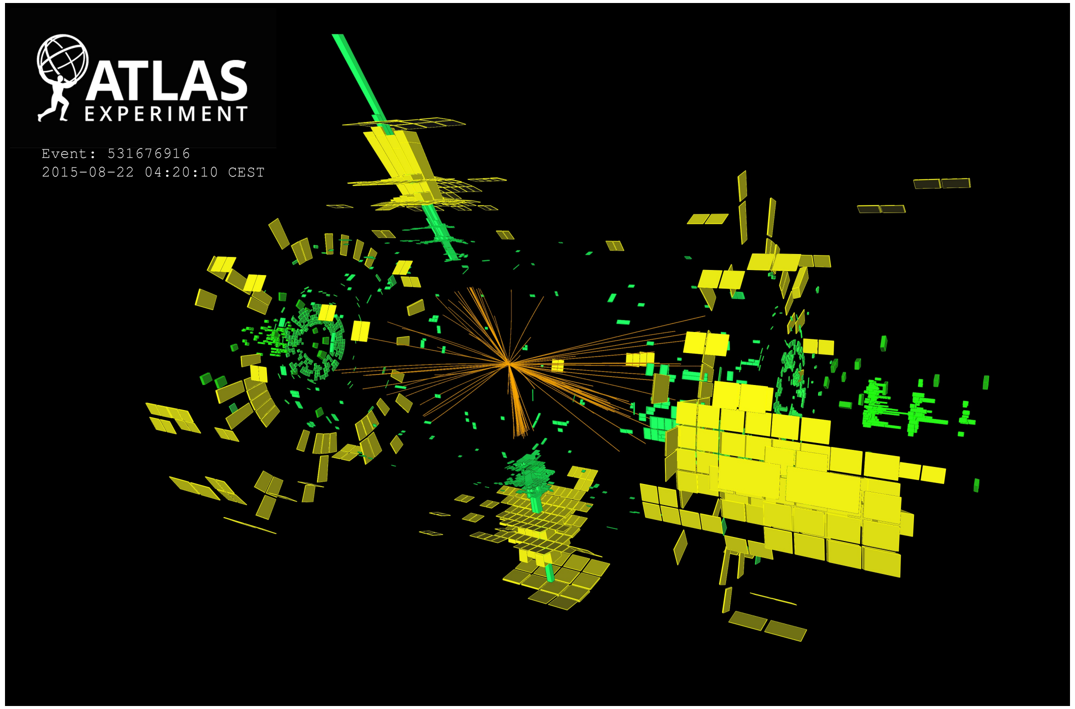
Julian Sennette, Dev Panchal

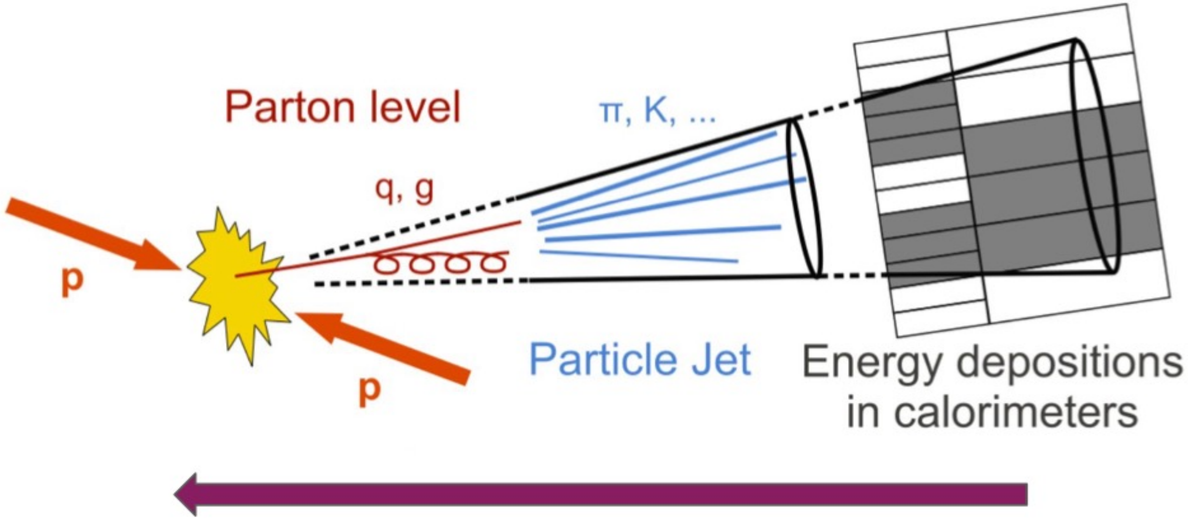
August, 2021

University of Texas At Austin

B Quark in the LHC

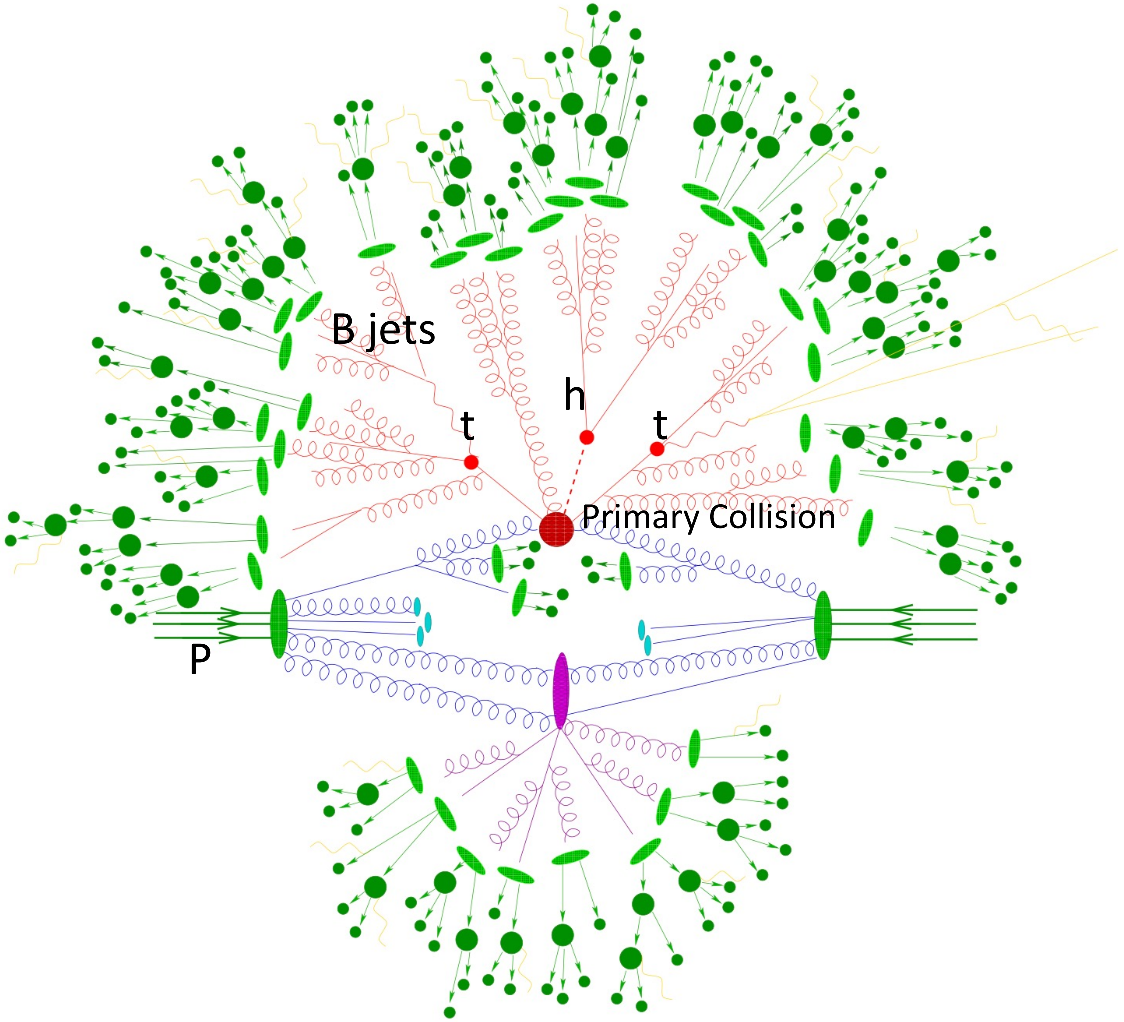
- Bottom (b) Quark: Heavy Quark; 3rd Generation
- Study of the B Quark is important for search of particles beyond the standard model and QCD predictions
- LHC experimental apparatus is equipped with subdetectors for identification and reconstruction of particle and particle paths
- Hadronic Calorimeter and tracking detector provides signal for b-jet information





Hadronization of B-quark

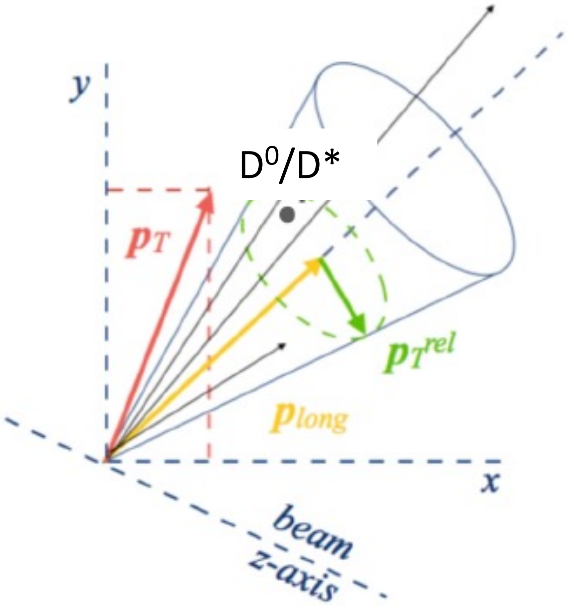
- Top quark pairs (ttbar) are produced in abundance in LHC
- This study focuses on ttbar final decay into b-hadron and hadronization into b-jets
- Hadronization: formation of hadrons from free quarks/gluons
- Fragmentation: Empirical model describing the decay of mesons to observable final state particles



b-fragmentation studies with B-hadron decays

- Studies of $t\bar{t}$ pairs with the final state B-hadron decaying in J/ψ ($b \rightarrow J/\psi \rightarrow \mu^+\mu^-$) or in ($\mu$)$D^0 \rightarrow$ (μ)$K\pi$ or in $D^*(2010)^+ \rightarrow D^0\rho$ mesons offer alternative methods to study m(top) as well as b-fragmentation.
- Motivation:
  - Check/quantify our knowledge of hadronization of b-quarks in hadron collider:
    - Partonic shower generators tuned to LEP results.
  - Current Monte Carlo event generator models and tunes are based on b-fragmentation observables measured at LEP:
    - yields a clean, back-to-back signature of b-quarks.
    - LHC environment is busier.
  - Provide an exclusive measurement of b-fragmentation in ATLAS.

Analysis Strategy



- Obtain a clean sample of b-jets from leptonic top mass decay with b-meson hadronization to (μ)D^0→ (μ)Kπ or $D^*(2010)^+$→D^0ρ.
- Identify D^0 or $D^*(2010)^+$ candidates using techniques used for top mass analysis.
- Construct moments of the D^0 or $D^*(2010)^+$ with respect to the b-jet.
    - using calorimetric information from jets:

$$\mathrm{z}_T^{\mathrm{calo}} = \frac{p_T(\mathrm{meson})}{p_T(\mathrm{calo\ jet})}$$

$$\mathrm{z}_L^{\mathrm{calo}} = \frac{\vec{p}(\mathrm{meson}) \cdot \vec{p}(\mathrm{calo\ jet})}{|p(\mathrm{calo\ jet})|^2}$$

$$\mathrm{z}_{rel}^{\mathrm{calo}} = \frac{|\vec{p}(\mathrm{meson}) \times \vec{p}(\mathrm{calo\ jet})|}{|p(\mathrm{calo\ jet})|^2}$$

    - using ghost-associated charged tracks:

$$\mathrm{z}_T^{\mathrm{ch}} = \frac{p_T(\mathrm{meson})}{p_T(\mathrm{ch,\ jet})}$$

$$\mathrm{z}_L^{\mathrm{ch}} = \frac{\vec{p}(\mathrm{meson}) \cdot \vec{p}(\mathrm{ch,\ jet})}{|p(\mathrm{ch,\ jet})|^2}$$

$$\mathrm{z}_{rel}^{\mathrm{ch}} = \frac{|\vec{p}(\mathrm{meson}) \times \vec{p}(\mathrm{ch,\ jet})|}{|p(\mathrm{ch,\ jet})|^2}$$

Analysis Selections

Detector-level selections: AnalysisTop 21.2.150, TOPQ1 derivation

$D^*(2010)^+ \rightarrow D^0\rho$:

- Lepton Selection:
  - PID: LHTight for elections, Medium for Muons
  - e-isolation: PLVTight
  - μ-isolation: PflowTight_FixedRad
  - exactly one lepton with $p_T > 25$ GeV for 2015
    $p_T > 27$ GeV for 2016–2018
- Jet Selection:
  - ≥ 4 jets with $p_T > 25$ GeV and $|\eta| < 2.5$, EMPFlow jet container
  - ≥ 1 b-tagged jet with DL1r
- MET > 20 GeV and $m_T(W) > 40$ GeV

- $D^*(2010)^+$ candidate:
  - $p_T > 8$ GeV
  - $|y(D^*)| < 2.1$
  - 0.14 GeV < $(m(D^*) - m(D^0))$ < 0.15 GeV
- Kaon and pions:
  - Three tracks with $p_T > 3$ GeV
  - Require one track to be π^+
  - ΔR (track, track) < 0.4, ΔR (tracks, jet) < 0.5

Particle-level selections:

Exactly one lepton with $|\eta| < 2.5$ and $p_T > 25$ GeV
One $D^*(2010)^+$ candidate
At least four jets with $|\eta| < 2.5$ and $p_T > 25$ GeV

Analysis Selections

Detector-level selections: AnalysisTop 21.2.150, TOPQ1 derivation

$(\mu)D^0 \rightarrow (\mu)K\pi$:

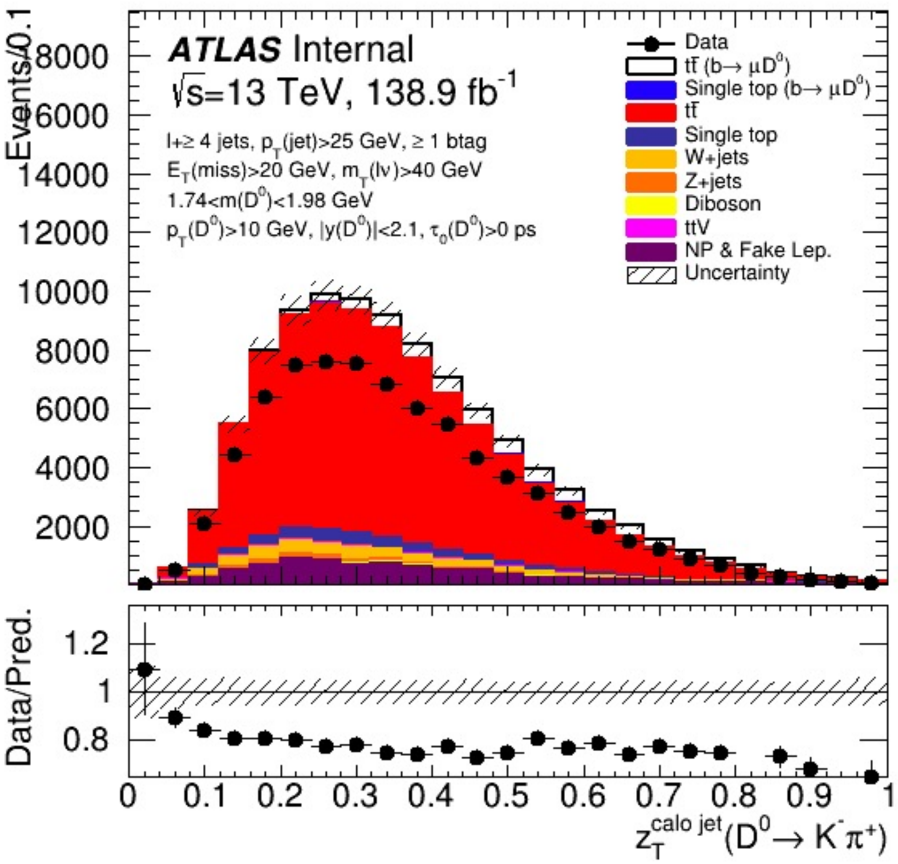
- Lepton Selection:
  PID: LHTight for elections, Medium for Muons
  e-isolation: PLVTight
  μ-isolation: PflowTight_FixedRad
  exactly one lepton with $p_T > 25$ GeV for 2015
  $p_T > 27$ GeV for 2016–2018
- Jet Selection:
  ≥ 4 jets with $p_T > 25$ GeV and $|\eta| < 2.5$, EMPFlow jet container
  ≥ 1 b-tagged jet with DL1r
- MET > 20 GeV and $m_T(W) > 40$ GeV

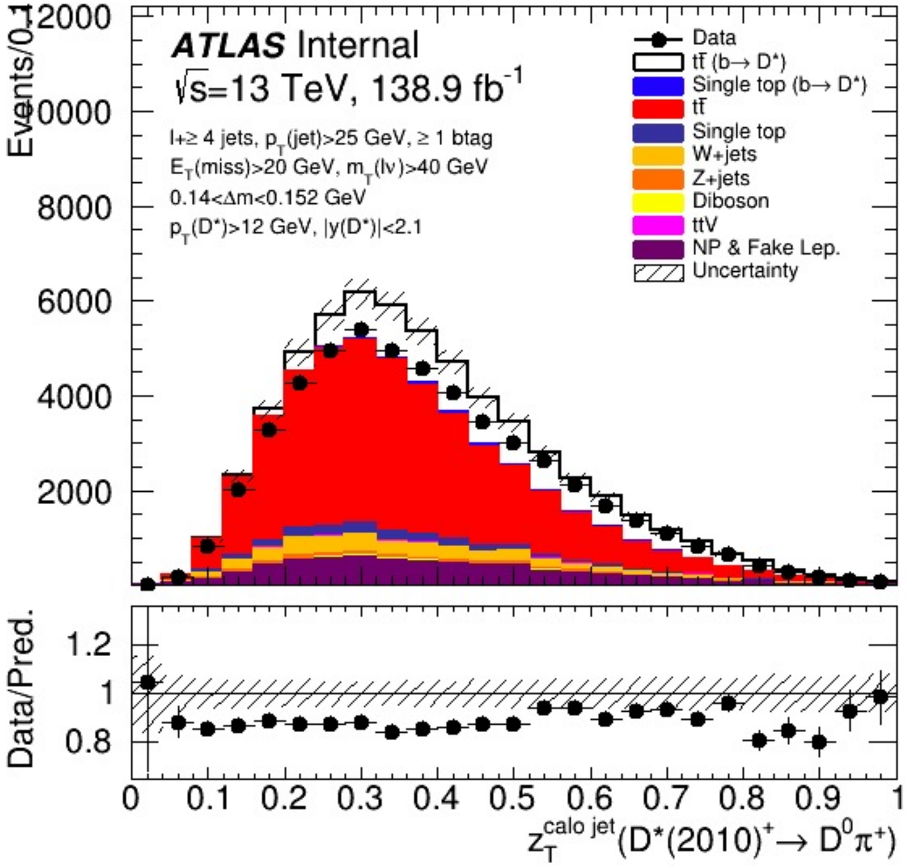
- D^0 candidate:
  $p_T > 8$ GeV
  $|y(D^0)| < 2.1$
  $1.74 \text{ GeV} < m(D^0) < 1.98$ GeV
  $\tau(D^0) > 0$ ps
- Kaon and pion:
  Two opposite sign tracks with $p_T > 3$ GeV
  ΔR (track, track) < 0.4, ΔR (tracks, jet) < 0.5
- Soft-μ:
  one soft-μ with LowPt PID
  same charge of the soft-μ and kaon
  ΔR (tracks, soft-μ) < 0.3

Particle-level selections:

Exactly one lepton with $|\eta| < 2.5$ and $p_T > 25$ GeV
One D^0 candidate
At least four jets with $|\eta| < 2.5$ and $p_T > 25$ GeV

Detector-level plots





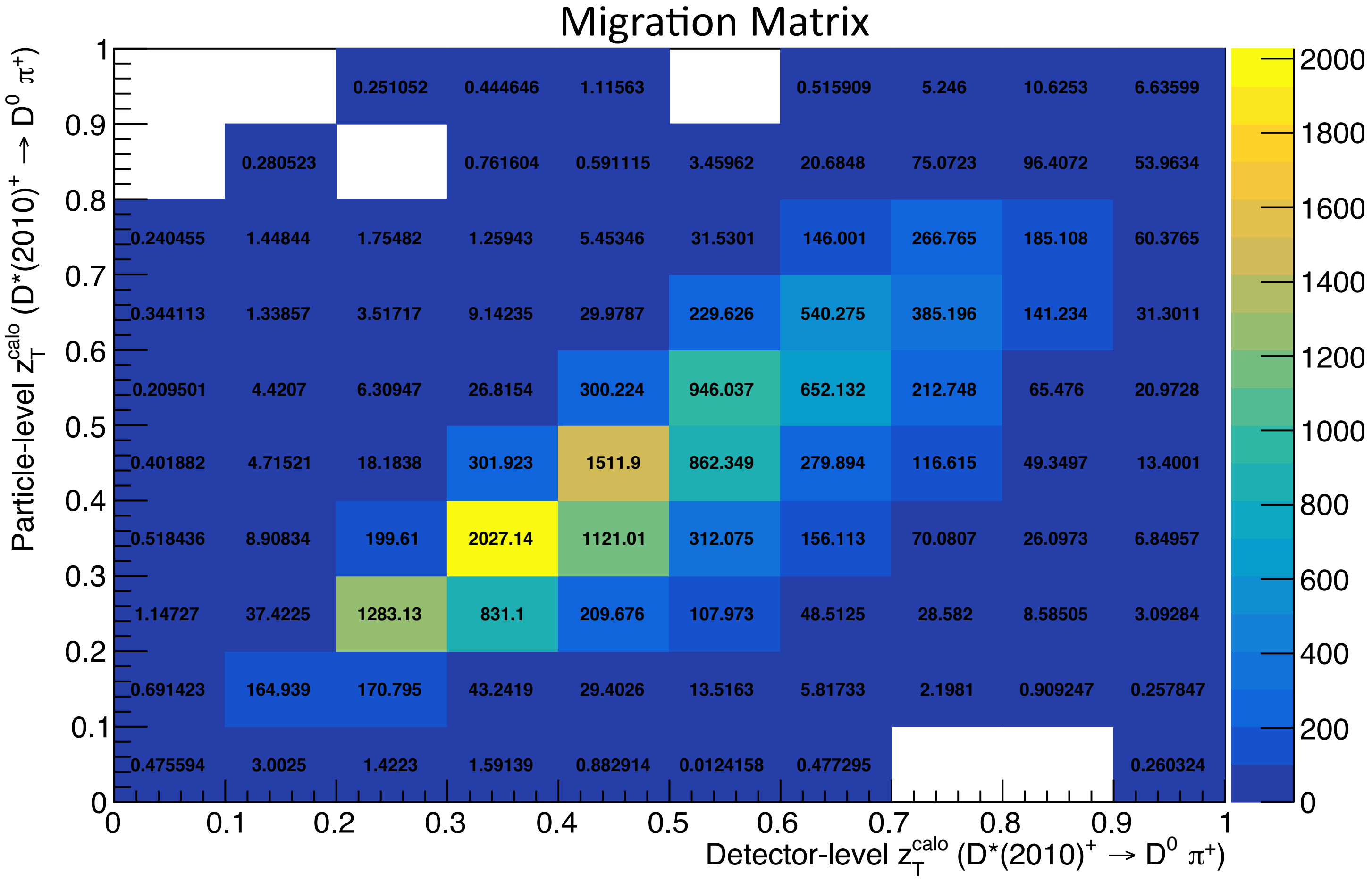
Observe bad Data/MC agreement, with ~2x more MC events than Data. Investigation of this discrepancy is on-going…

Uncertainty = statistical & cross-section uncertainties

Unfolding Procedure

- Concept: Estimate the “truth-level” spectrum of an observable if measured with an ideal detector and infinite statistics
- Distribution of any observable is distorted due to experimental limitations—detector resolution, efficiencies, statistics, …
- Goal of unfolding is to un-smear the detector effects to obtain a “truth-level” spectrum.
- The detector smearing is described by a migration matrix.
- The *(i,j)th* element of the migration matrix corresponds to the number of events of an observable at particle level in bin *i* to detector level in bin *j.*

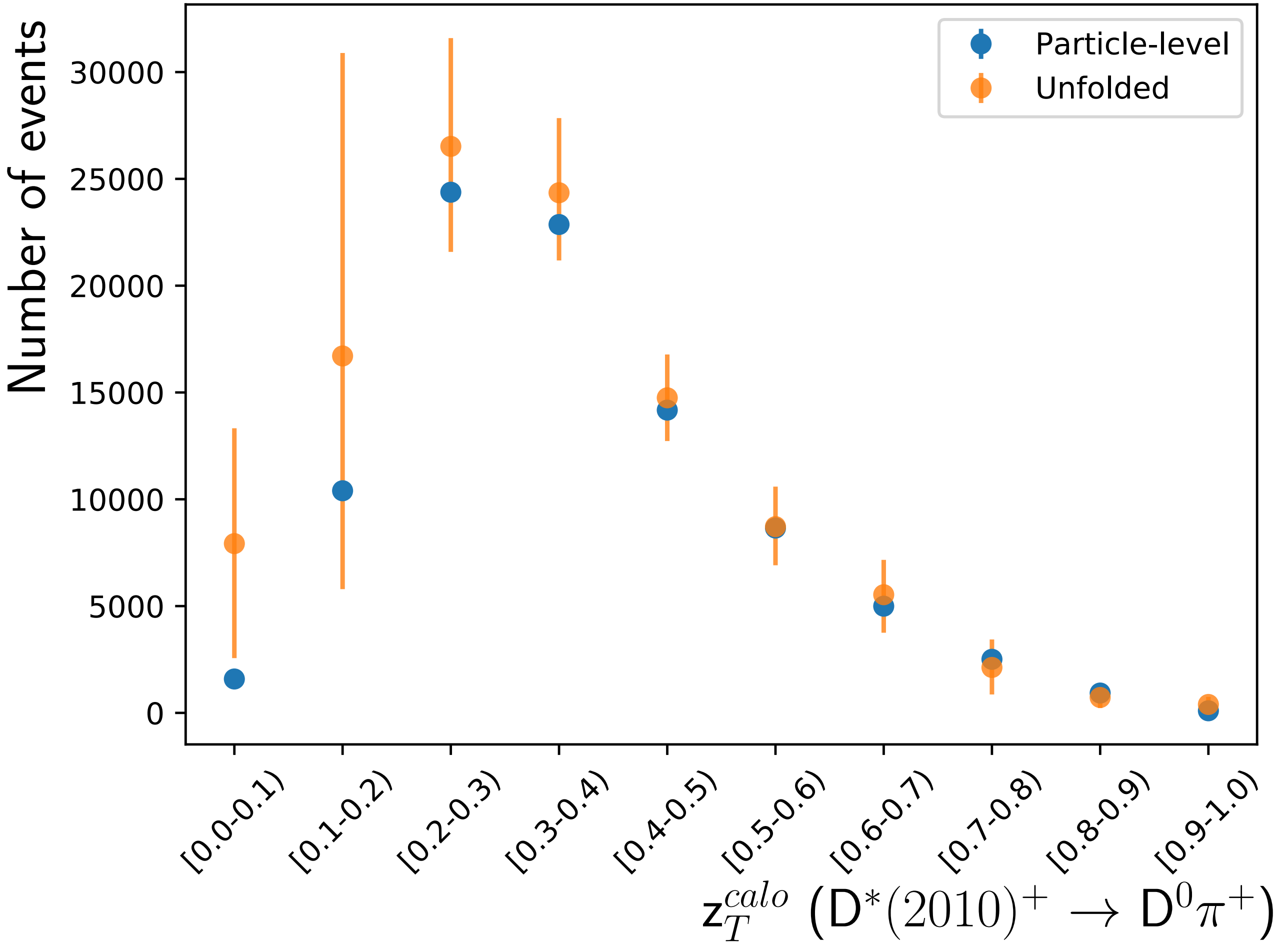
Migration Matrix



		0.251052	0.444646	1.11563		0.515909	5.246	10.6253	6.63599
	0.280523		0.761604	0.591115	3.45962	20.6848	75.0723	96.4072	53.9634
0.240455	1.44844	1.75482	1.25943	5.45346	31.5301	146.001	266.765	185.108	60.3765
0.344113	1.33857	3.51717	9.14235	29.9787	229.626	540.275	385.196	141.234	31.3011
0.209501	4.4207	6.30947	26.8154	300.224	946.037	652.132	212.748	65.476	20.9728
0.401882	4.71521	18.1838	301.923	1511.9	862.349	279.894	116.615	49.3497	13.4001
0.518436	8.90834	199.61	2027.14	1121.01	312.075	156.113	70.0807	26.0973	6.84957
1.14727	37.4225	1283.13	831.1	209.676	107.973	48.5125	28.582	8.58505	3.09284
0.691423	164.939	170.795	43.2419	29.4026	13.5163	5.81733	2.1981	0.909247	0.257847
0.475594	3.0025	1.4223	1.59139	0.882914	0.0124158	0.477295			0.260324

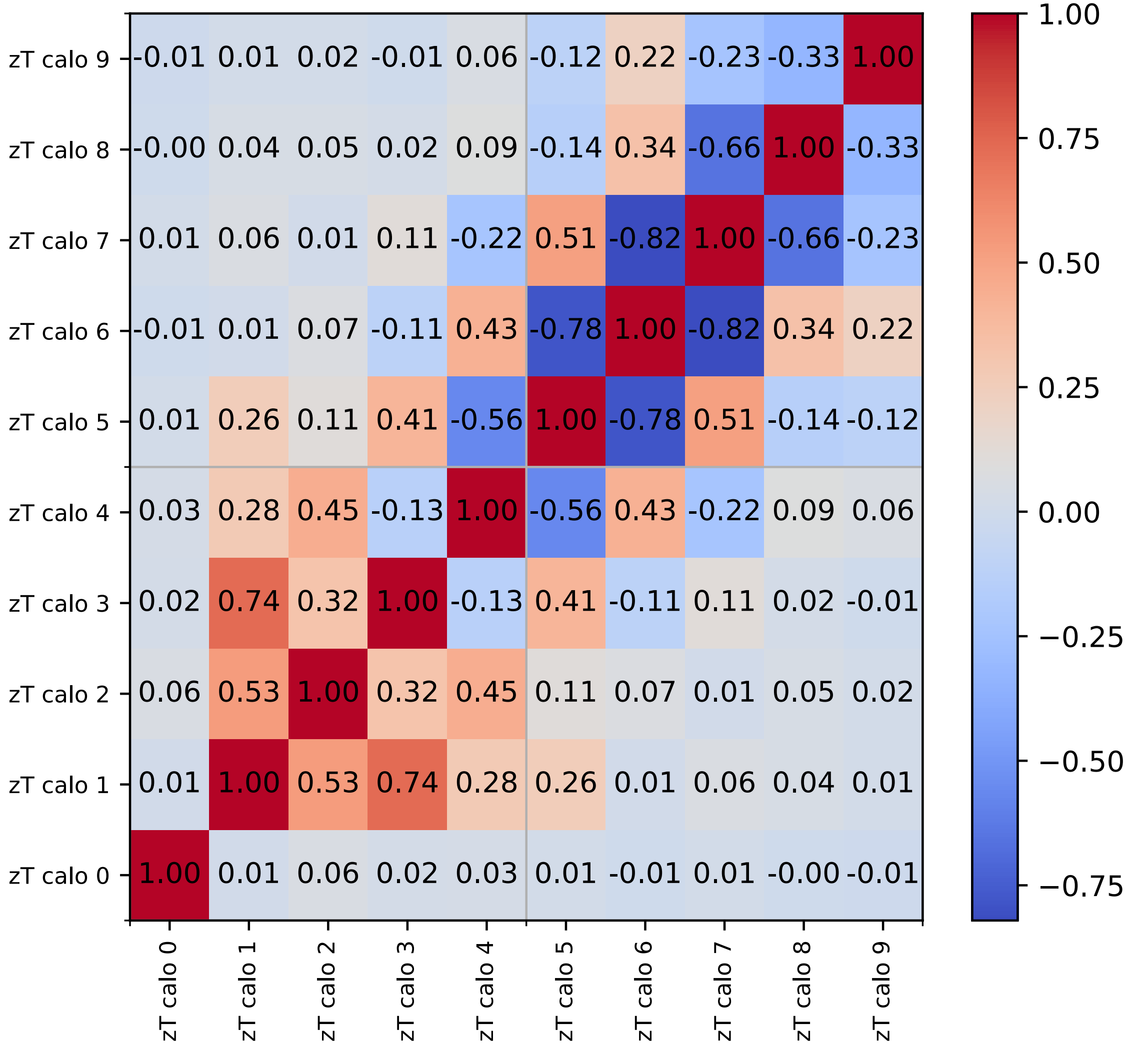
Closure Test

- The first step of unfolding is to perform a closure test.
- The aim of the closure test is to retrieve the distribution of observables at particle level from the ones at detector level.
- Use nominal ttbar sample Powheg+Pythia 8 as the Asimov data.
- Performed a closure test with uniform binning for zT moment
    - [0.1, 0.2, 0.3, 0.4 ,0.5, 0.6, 0.7, 0.8, 0.9, 1.0]



Closure Test

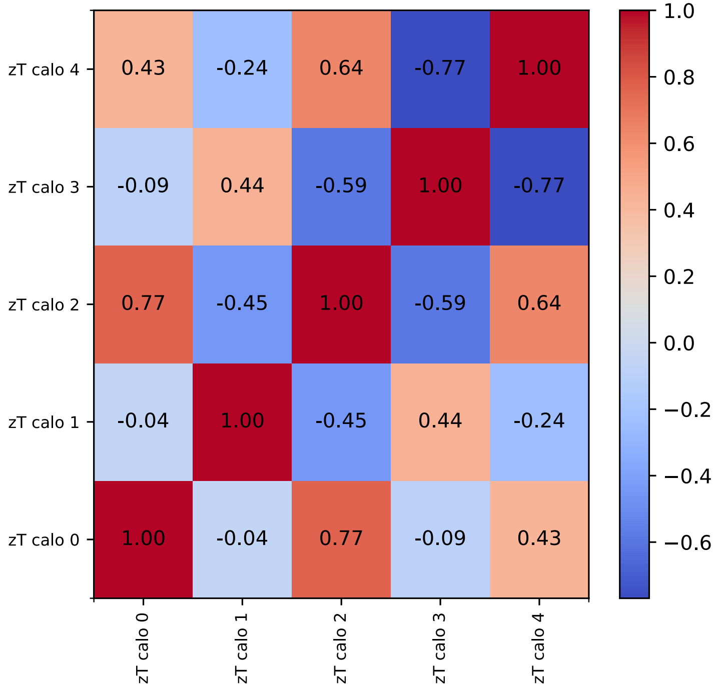
- The unfolded distribution does not match the particle level distribution:
  - We observe large bin by bin correlations in the unfolded distribution as evident from this correlation matrix
- Our choice of uniform binning does not give a good unfolding distribution
  - We observe migration from particle level to detector level distribution
  - This migration causes large bin by bin correlations in the unfolded distribution



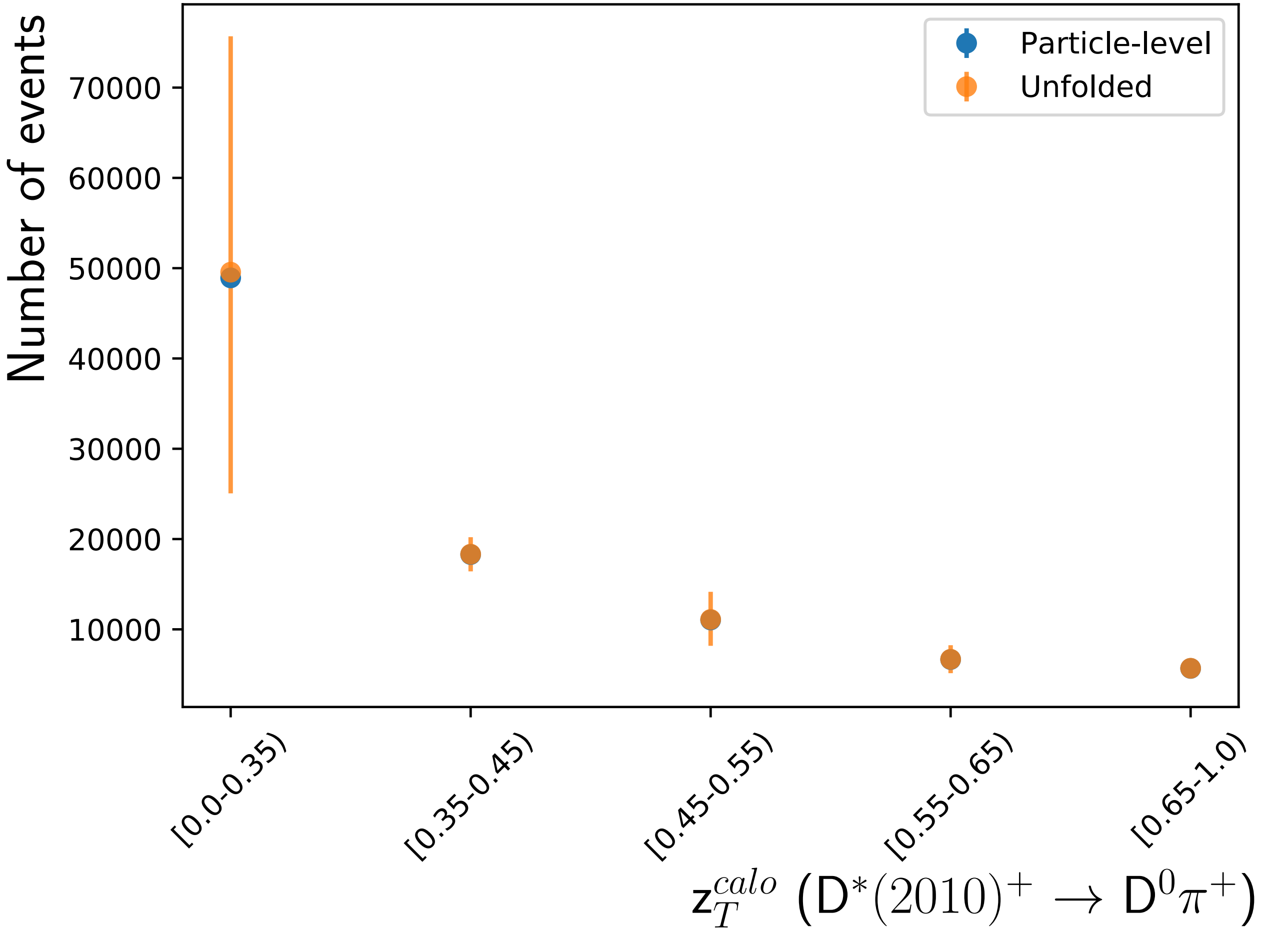
	zT calo 0	zT calo 1	zT calo 2	zT calo 3	zT calo 4	zT calo 5	zT calo 6	zT calo 7	zT calo 8	zT calo 9
zT calo 9	-0.01	0.01	0.02	-0.01	0.06	-0.12	0.22	-0.23	-0.33	1.00
zT calo 8	-0.00	0.04	0.05	0.02	0.09	-0.14	0.34	-0.66	1.00	-0.33
zT calo 7	0.01	0.06	0.01	0.11	-0.22	0.51	-0.82	1.00	-0.66	-0.23
zT calo 6	-0.01	0.01	0.07	-0.11	0.43	-0.78	1.00	-0.82	0.34	0.22
zT calo 5	0.01	0.26	0.11	0.41	-0.56	1.00	-0.78	0.51	-0.14	-0.12
zT calo 4	0.03	0.28	0.45	-0.13	1.00	-0.56	0.43	-0.22	0.09	0.06
zT calo 3	0.02	0.74	0.32	1.00	-0.13	0.41	-0.11	0.11	0.02	-0.01
zT calo 2	0.06	0.53	1.00	0.32	0.45	0.11	0.07	0.01	0.05	0.02
zT calo 1	0.01	1.00	0.53	0.74	0.28	0.26	0.01	0.06	0.04	0.01
zT calo 0	1.00	0.01	0.06	0.02	0.03	0.01	-0.01	0.01	-0.00	-0.01

Closure Test with Updated Binning

- We would like to reduce this bin migration by choosing a different set of binning
  - Our first choice of rebinning is the following [0, 0.35, 0.45, 0.55, 0.65,1]
  - This choice of binning was an attempt to reduce bin migrations, however, this is not the optimal binning



	zT calo 0	zT calo 1	zT calo 2	zT calo 3	zT calo 4
zT calo 4	0.43	-0.24	0.64	-0.77	1.00
zT calo 3	-0.09	0.44	-0.59	1.00	-0.77
zT calo 2	0.77	-0.45	1.00	-0.59	0.64
zT calo 1	-0.04	1.00	-0.45	0.44	-0.24
zT calo 0	1.00	-0.04	0.77	-0.09	0.43



Conclusions and Next Steps

- ❖We have observed a better closure test with updated binning
- ❖However, the first unfolded bin has large uncertainty and strong bin to bin correlations
- ❖We are currently investigating this bin migration and will try to reduce it by choosing a different set of particle level selections
- ❖We will also perform unfolding for the D^0 decay channel

BACKUP

Migration Matrix with Updated Binning