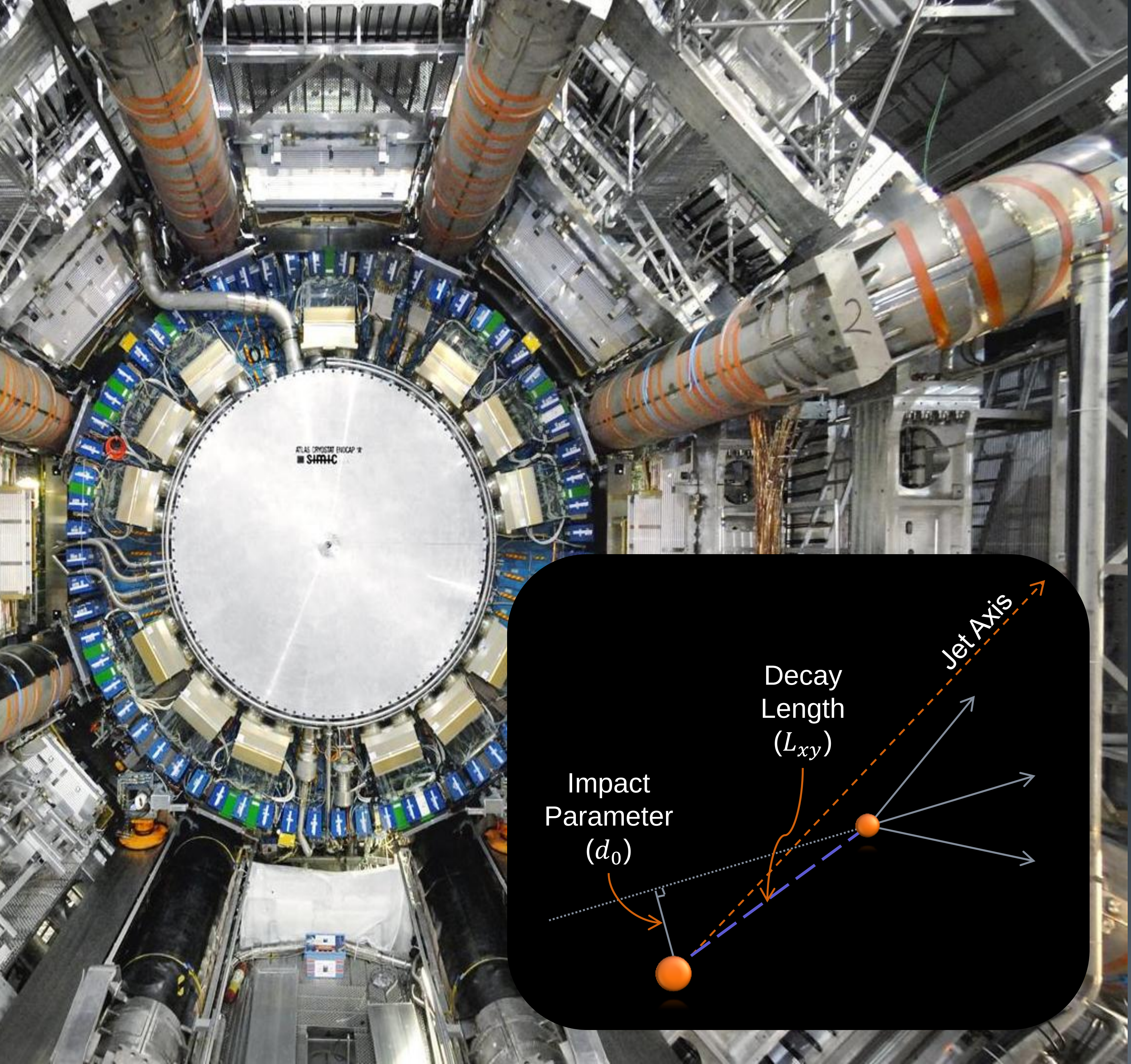
b-tagging in ATLAS

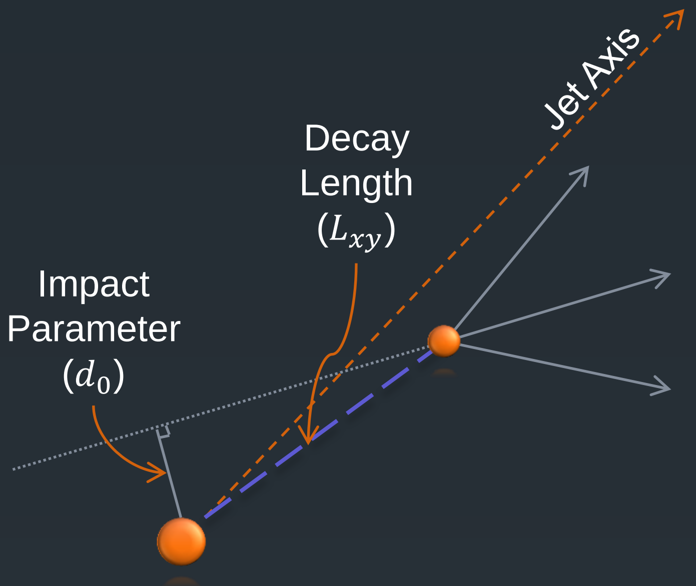
G. Watts (UW/Seattle for the ATLAS Collaboration)



DPF 2011

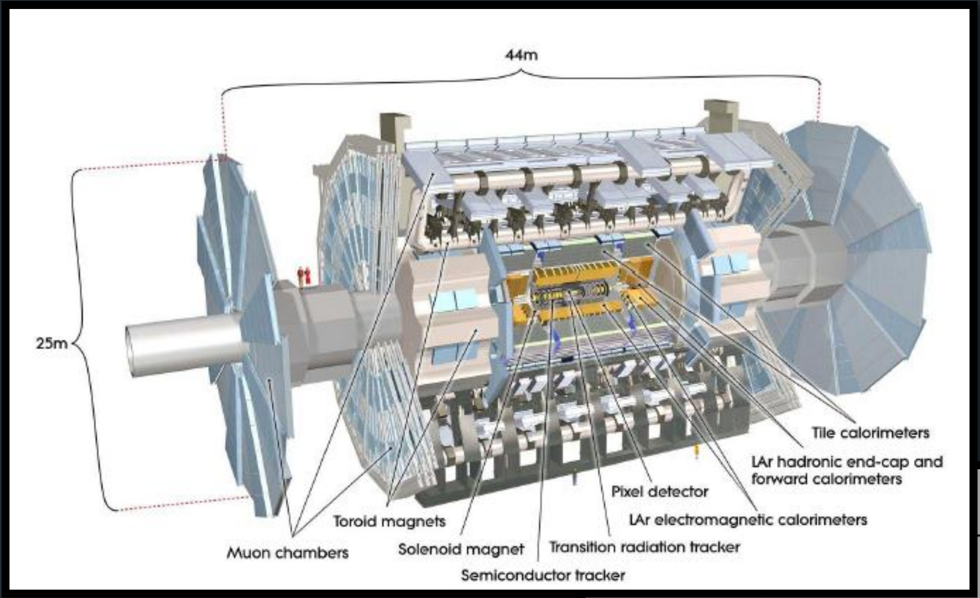
B-tagging at ATLAS

- The Detector
- The Algorithms
- The Calibration



High Mass: $m_B \sim 5$ GeV
Long Lifetime:
$c\tau \sim 470 \mu m \ (B^+, B^0, B_s)$
$\sim 390 \mu m \ (\Lambda_b)$
For 50 GeV Bottom
$L_{xy} \sim 5$mm, $d_0 \sim 500 \mu m$

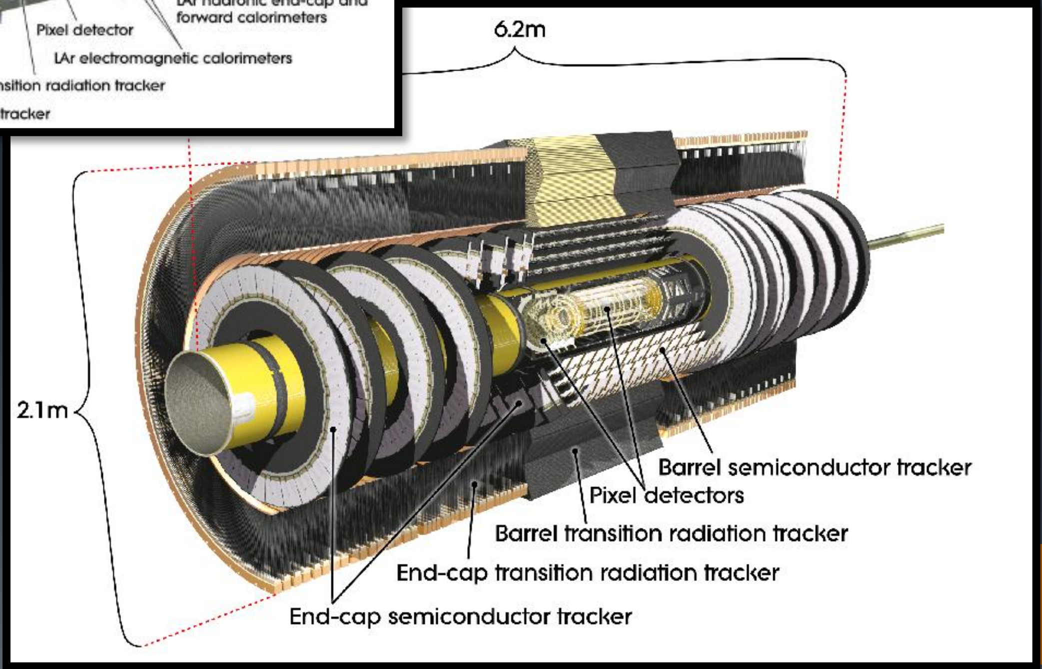
The Raw Materials



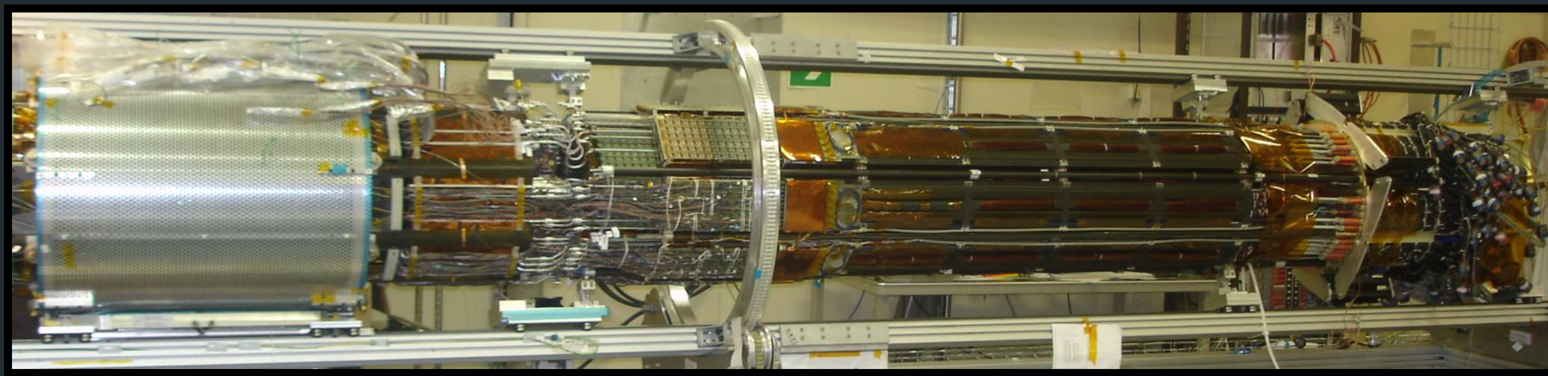
2 T Magnetic Field

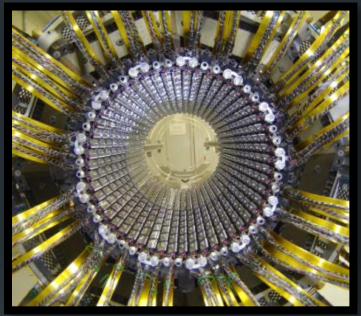
$$\frac{\sigma_{p_T}}{p_T} = 0.05\% \, p_T \oplus 1\%$$

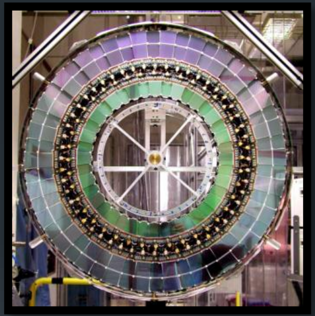
Resolution of Pixel Hit: $10\mu m$

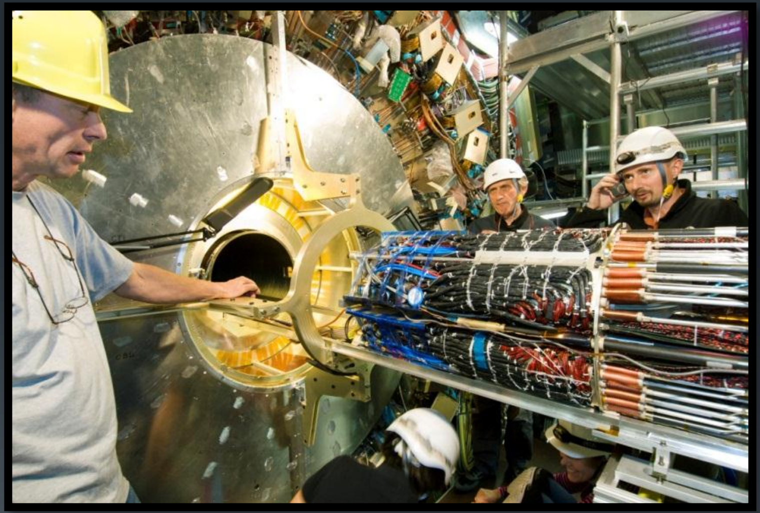


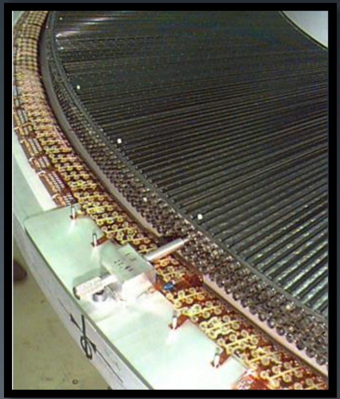
Inner Detector

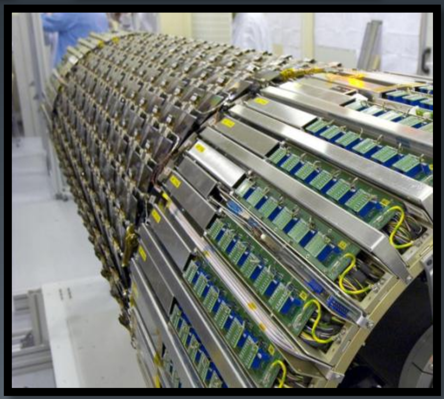




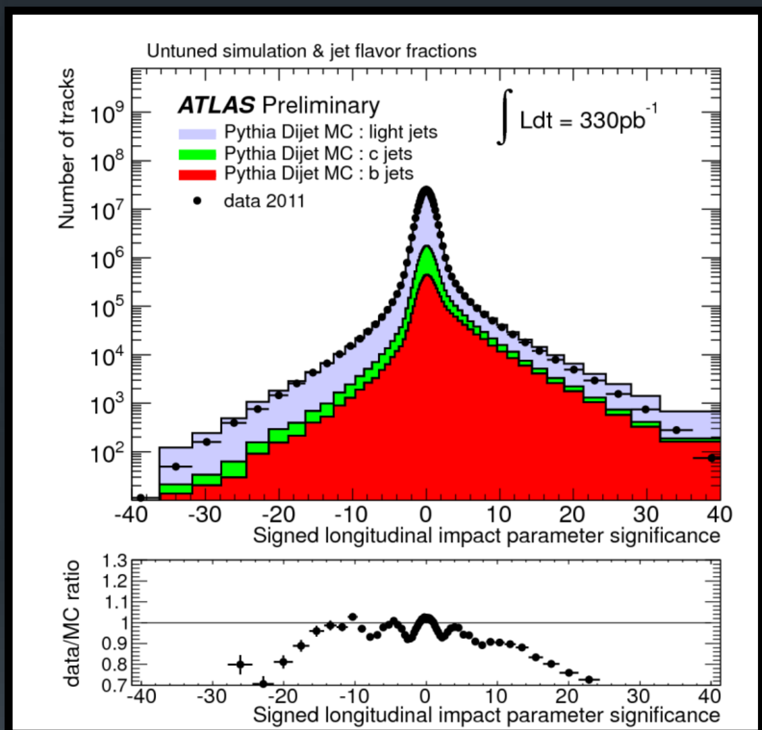


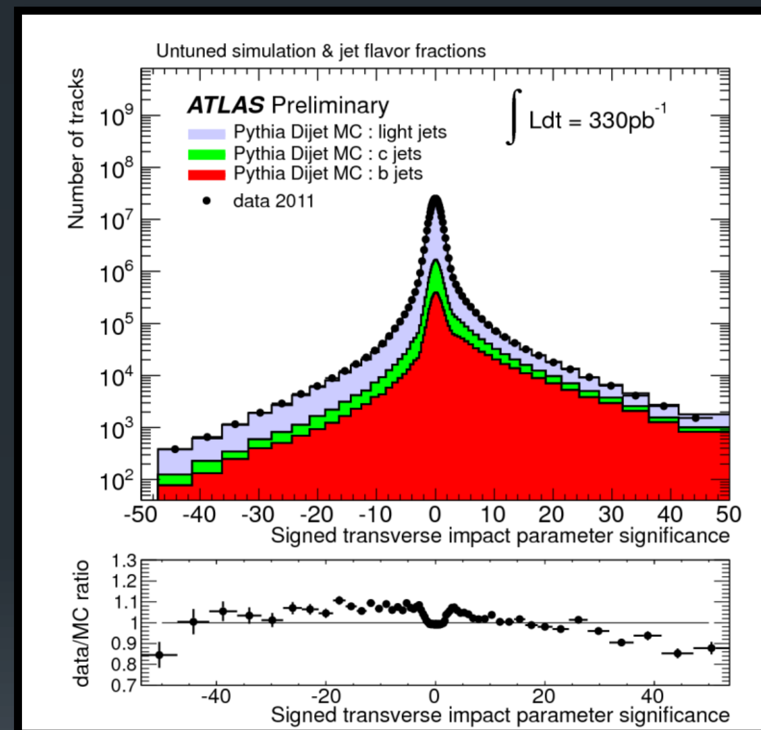






Signed Impact Parameter



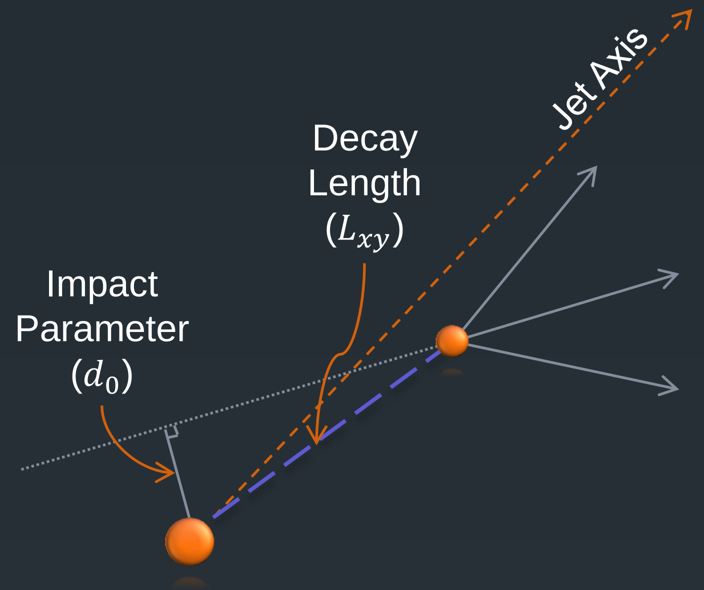


$$\frac{d_0}{\sigma_{d_0}}$$

The Algorithms

Algorithm Types at ATLAS

- Impact Parameter Type
- Secondary Vertex Reconstruction Type
- Combined



High Mass: $m_B \sim 5$ GeV
Long Lifetime:
$c\tau \sim 470 \mu m \ (B^+, B^0, B_s)$
$\sim 390 \mu m \ (\Lambda_b)$
For 50 GeV Bottom
$L_{xy} \sim$5mm, $d_0 \sim 500 \mu m$

Commissioning

QCD Jet Events (all flavors)

Single Jet Triggers
Jets: $p_T > 20$ GeV and $|\eta| < 2.5$

} About 10 million jets

Data/Monte Carlo Comparisons

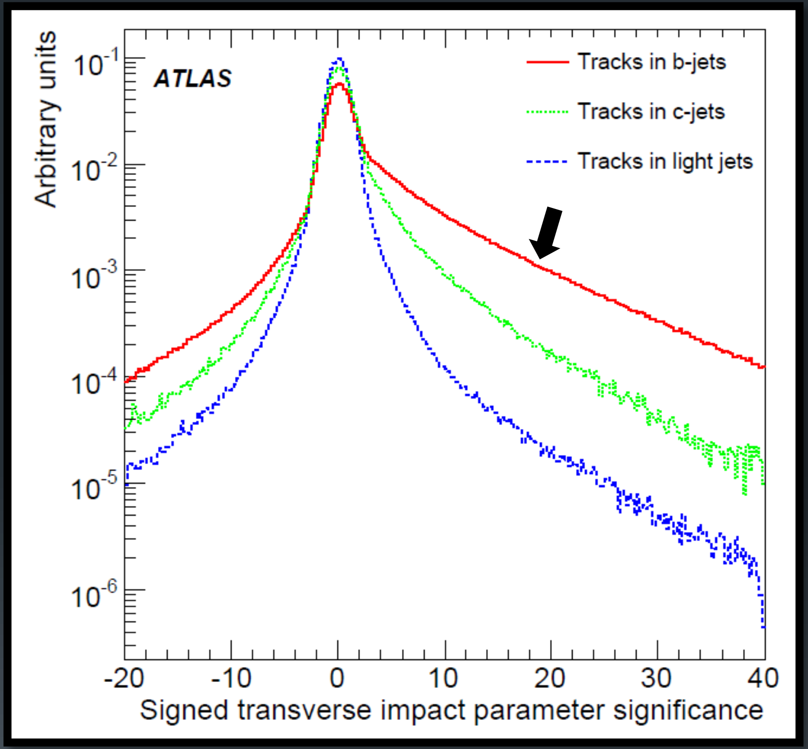
Tagger Input Variable Comparisons
Tagger Result Comparisons
Use standard Pythia QCD Monte Carlo

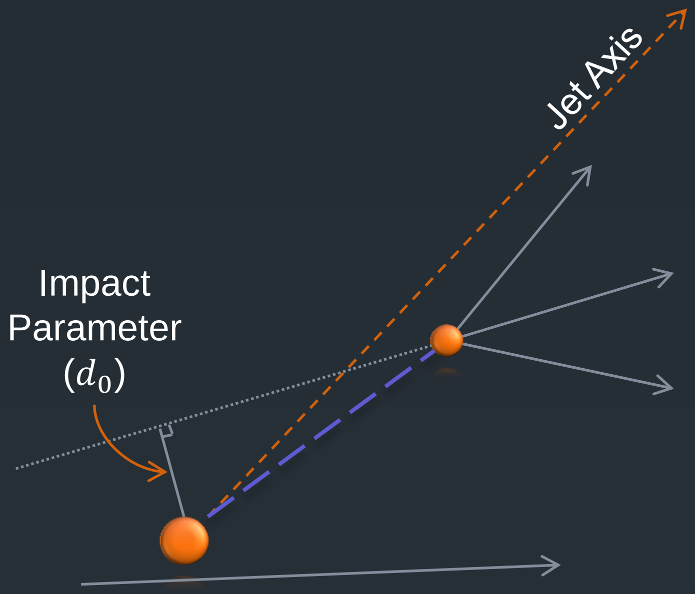
- Trigger Prescales (p_T and η)
- Pile-up (in-time and out of time)
- Pre-tag distributions normalized to data

Calibrations

Impact Parameter Algorithms

$\frac{d_0}{\sigma_{d_0}}$ Heavy Flavor Tracks Have an exceptional distribution





$$P_b\left(\frac{d_0}{\sigma_{d_0}}\right), P_c\left(\frac{d_0}{\sigma_{d_0}}\right), P_{light}\left(\frac{d_0}{\sigma_{d_0}}\right)$$

Combine P_{light} for each track near a jet

Impact Parameter Details

JetProb

Uses just the probability distribution of light quark jets transverse impact parameter (derived in MC)

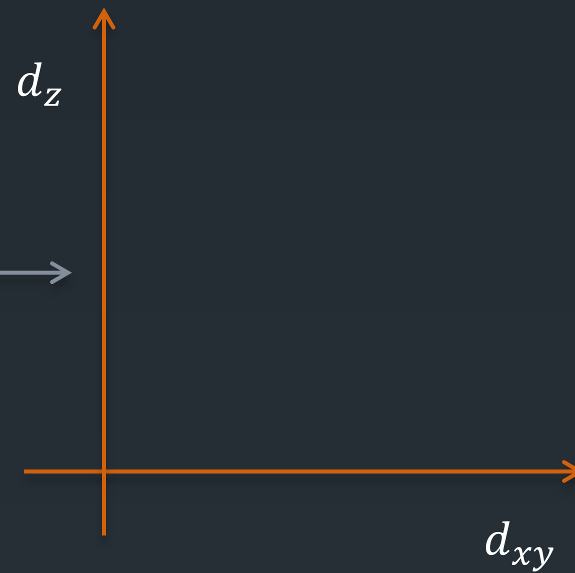
Uses the transverse impact parameter, the longitudinal impact parameter, and a likelihood ratio comparing both bottom quark and light quark jets

IP3D

IP3D

Likelihood Based Algorithm

Based on 2D distribution

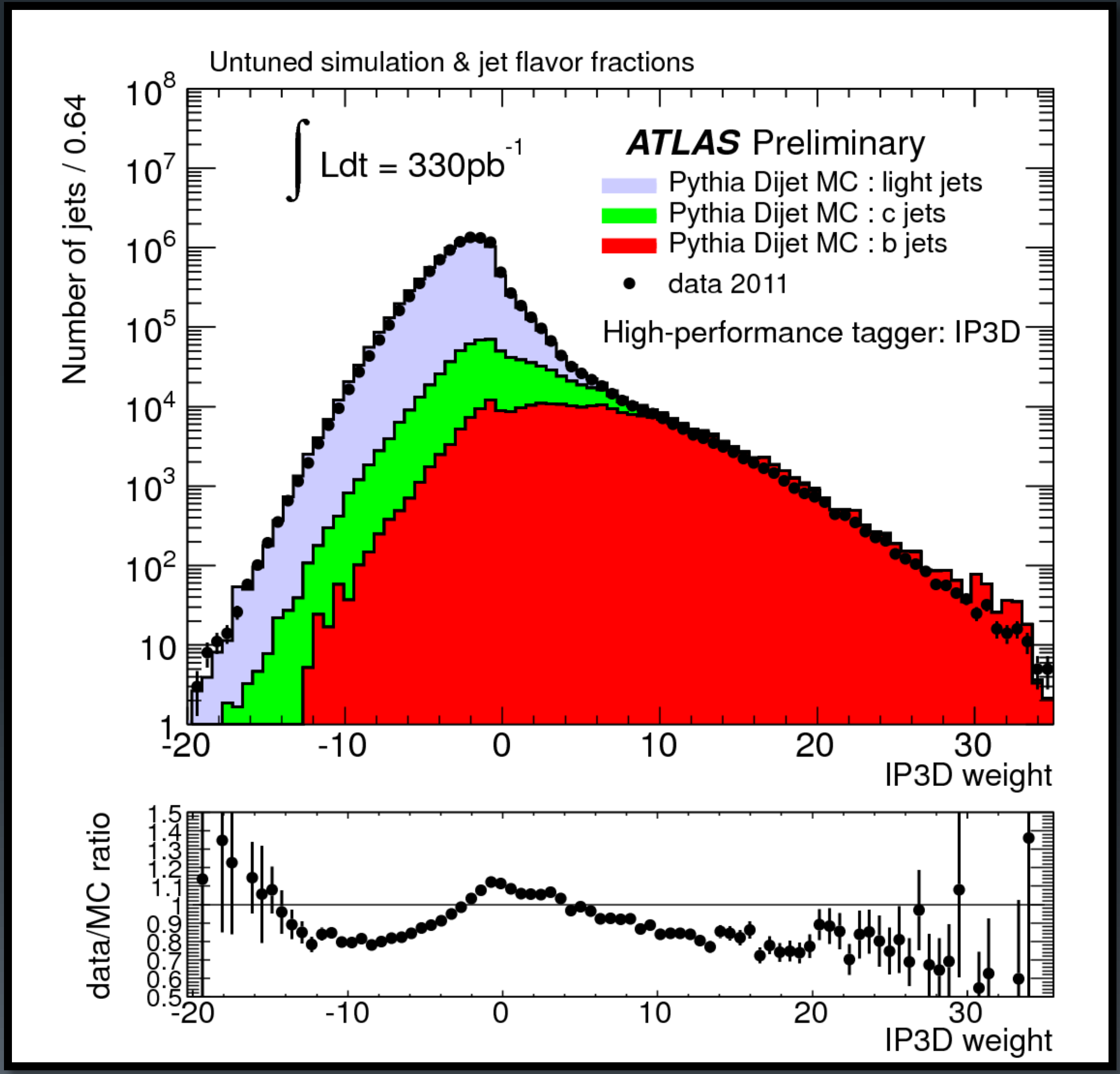


$$W_J = \sum_{i}^{N_{Tracks}} \ln W_i = \sum_{i}^{N_{Tracks}} \ln \frac{b(S_i)}{u(S_i)}$$

The ratio of probabilities for bottom (b) or light (u) for a particular (d_{xy}, d_z) pair

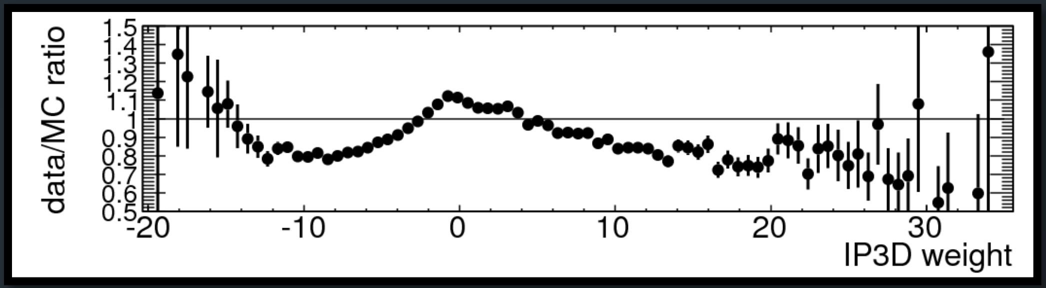
Good Monte Carlo Model Modeling

Tag Weight



G. Watts (UW/Seattle)

Tag Weight

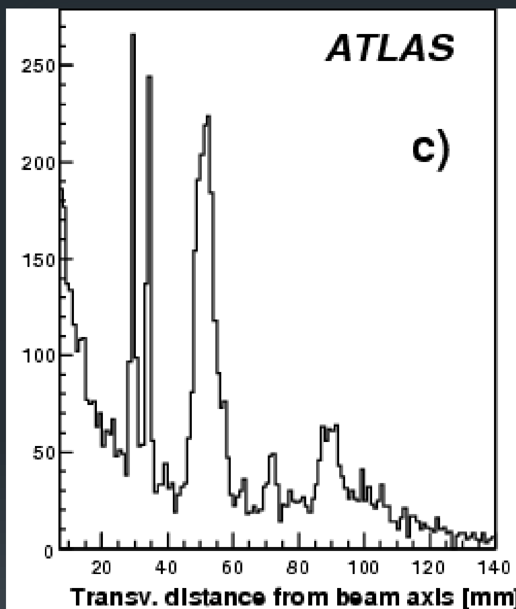


Impact Parameter Differences
Bottom, Charm, Light Fraction

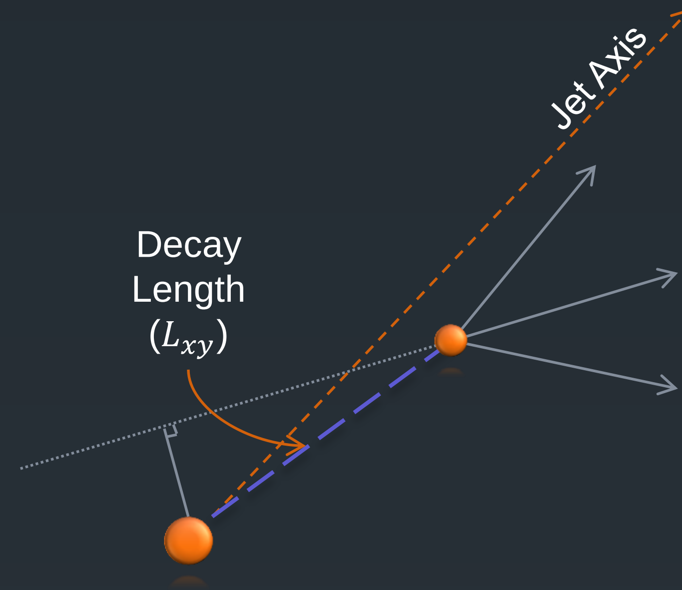
Calibration

Vertex Reconstruction Algorithms

$\frac{L_{xy}}{\sigma_{L_{xy}}}$ Heavy Flavor decays some distance from the primary vertex



Remove all 2-track vertices consistent with Λ, K_s, conversions, material interactions



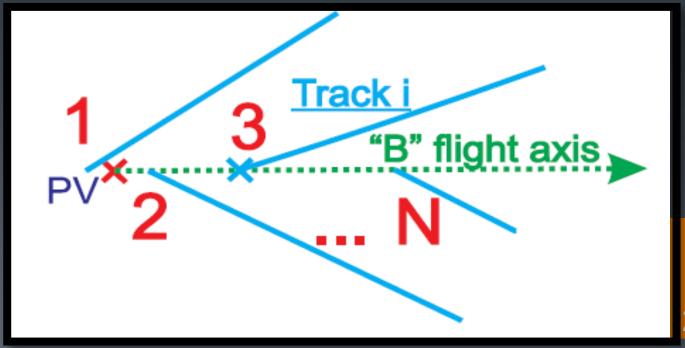
Remaining displaced tracks fit to common vertex
Remove the worst tracks

Vertex Reconstruction Details

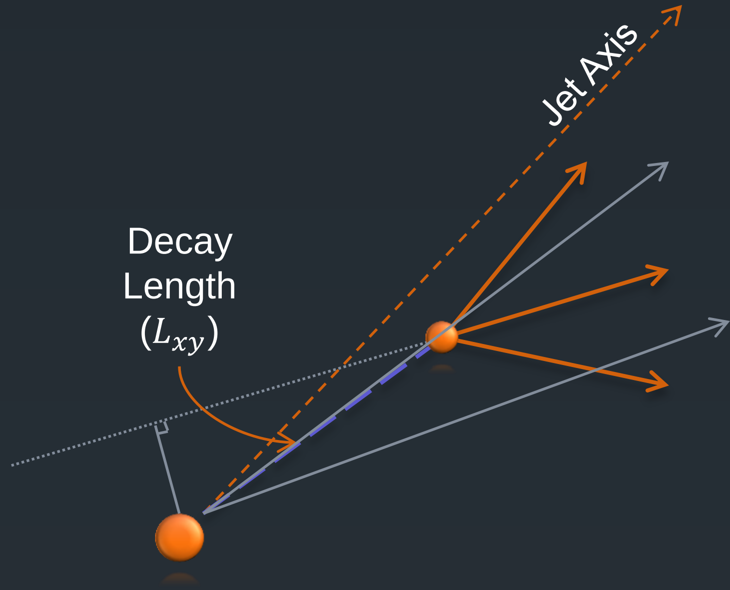
SV0 The decay Length Significance

SV1 Likelihood Ratio with PDF's from Monte Carlo

JetFitter
Topological approach
Constrain B/D decays to single axis, not vertex
b/c separation



SV1



Reconstructs a secondary vertex

Likelihood based

Ratio #1 (2D):

- Mass of all tracks attached to the secondary vertex
- Ratio of the energy of tracks associated with the vertex to those near the jet

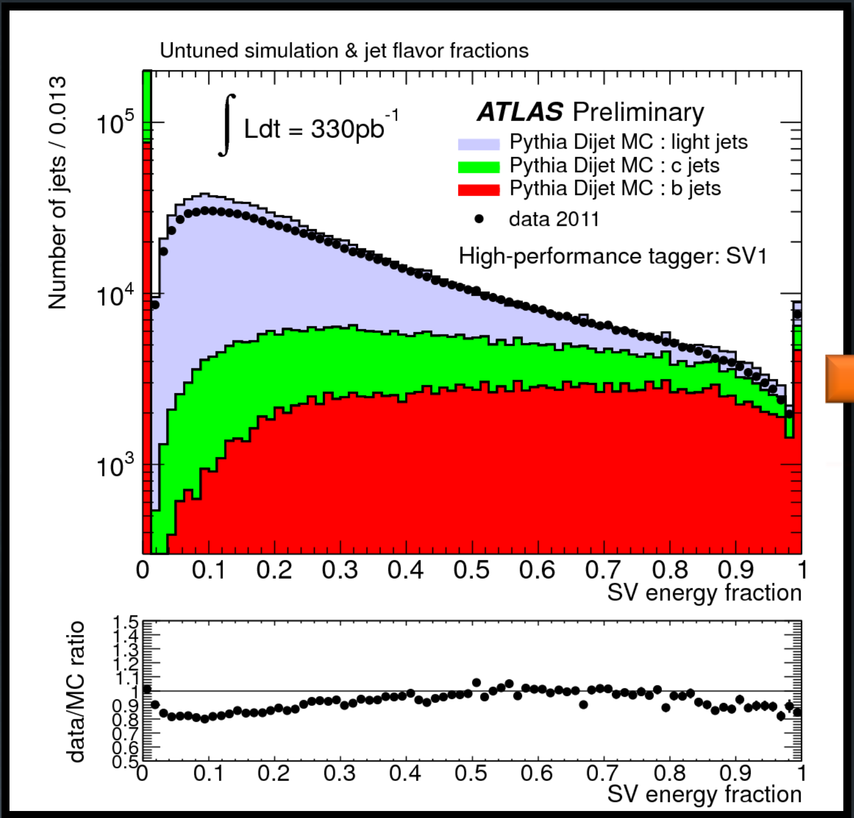
Ratio #2 (1D):

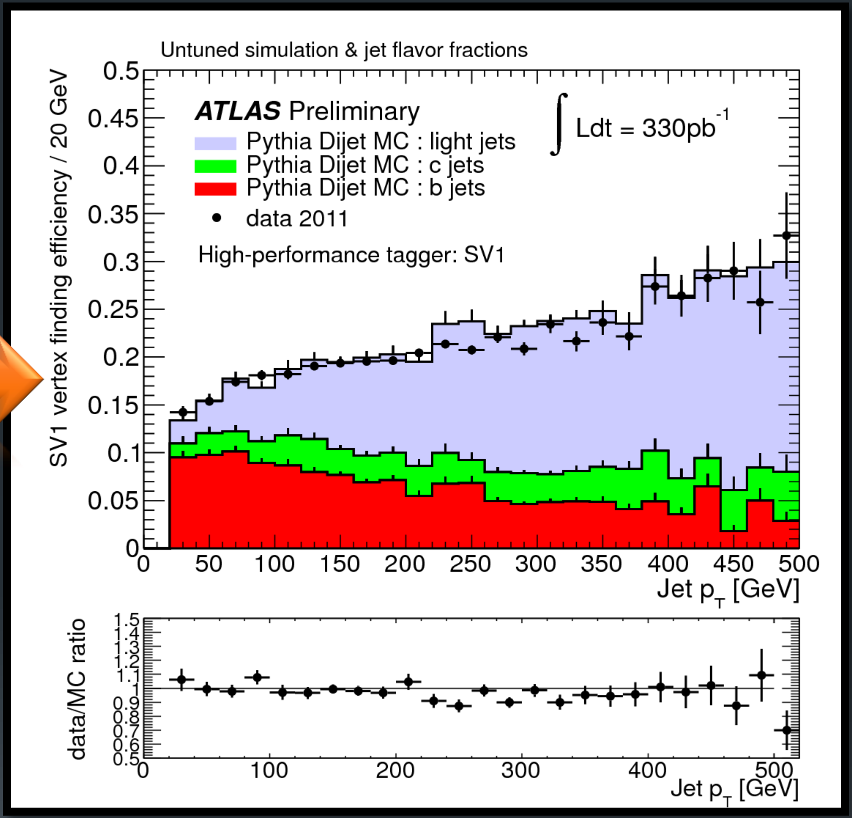
- # of 2-track vertices found near jet

SV2 algorithm: a single 3D likelihood. Not commissioned due to statistics!

Note that $L_{xy}/\sigma_{L_{xy}}$ is not used directly in this algorithm

SV1

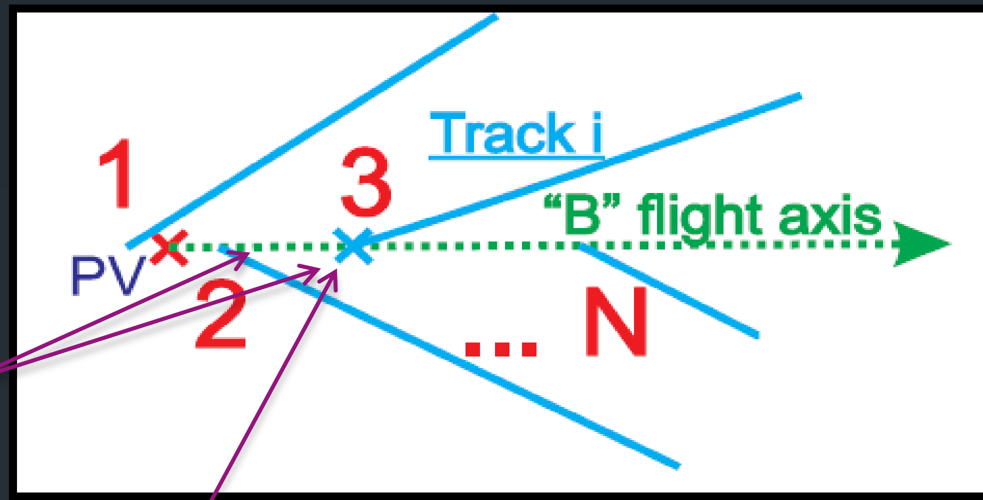




For light dominated samples we see discrepancies as large as 10%

Jet Fitter

2 Vertex Hypothesis!



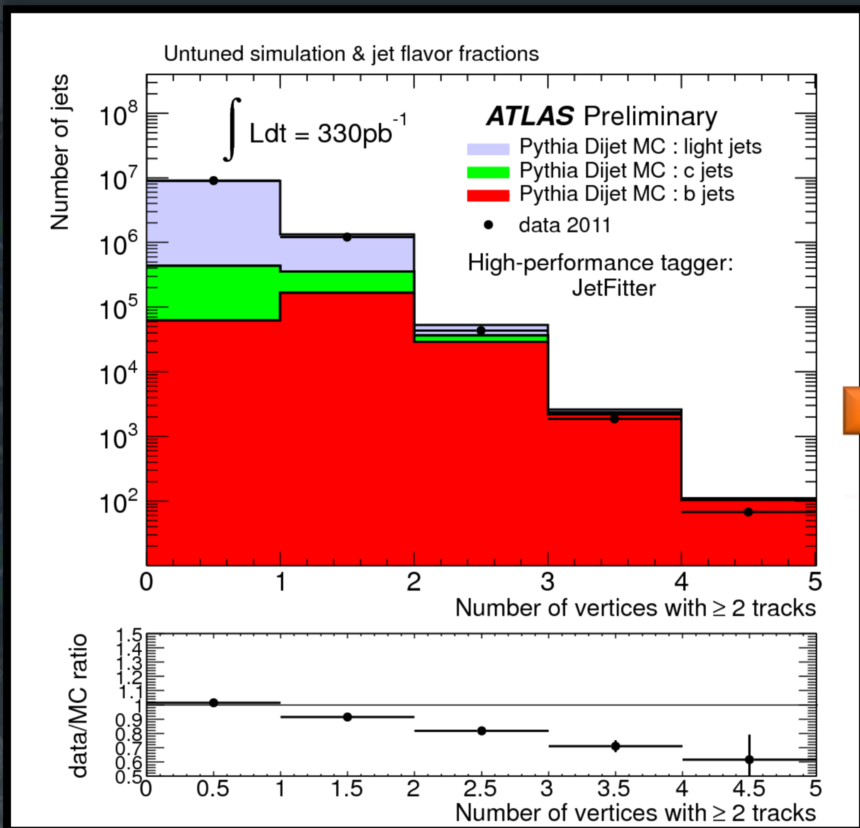
Attempt to reconstruct multiple vertices

Single Track Vertex is possible now

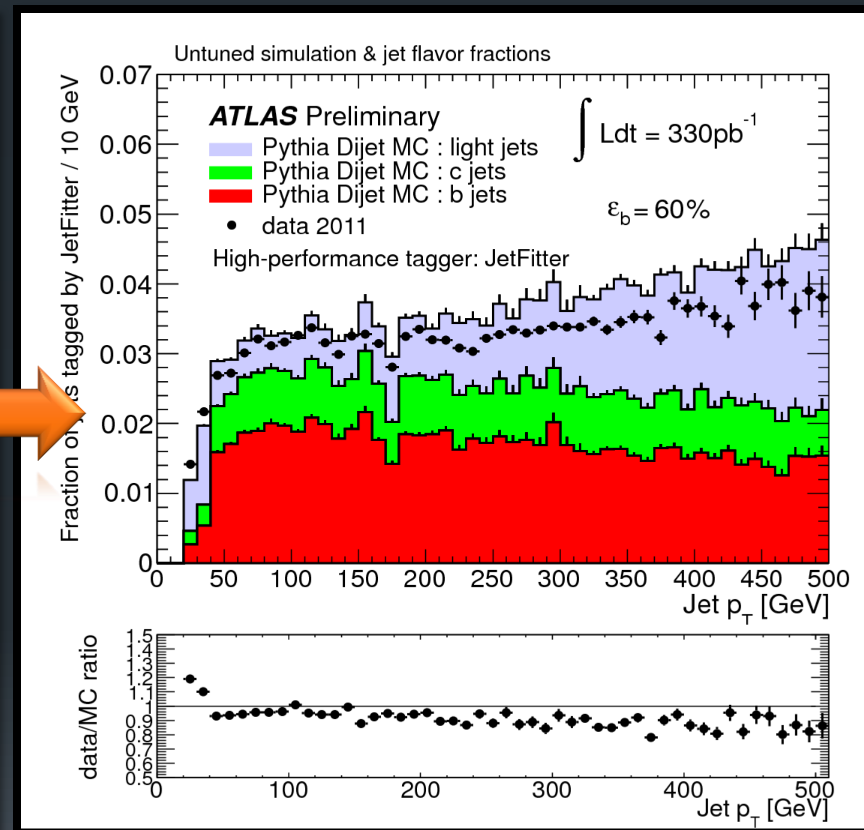
Likelihood ratio based on:

- Topology: # of single track vertices, # of two track vertices, and total number of tracks
- Energy Fraction of tracks (SV1) and mass of all tracks in the verticies
- Average decay length significance

Jet Fitter

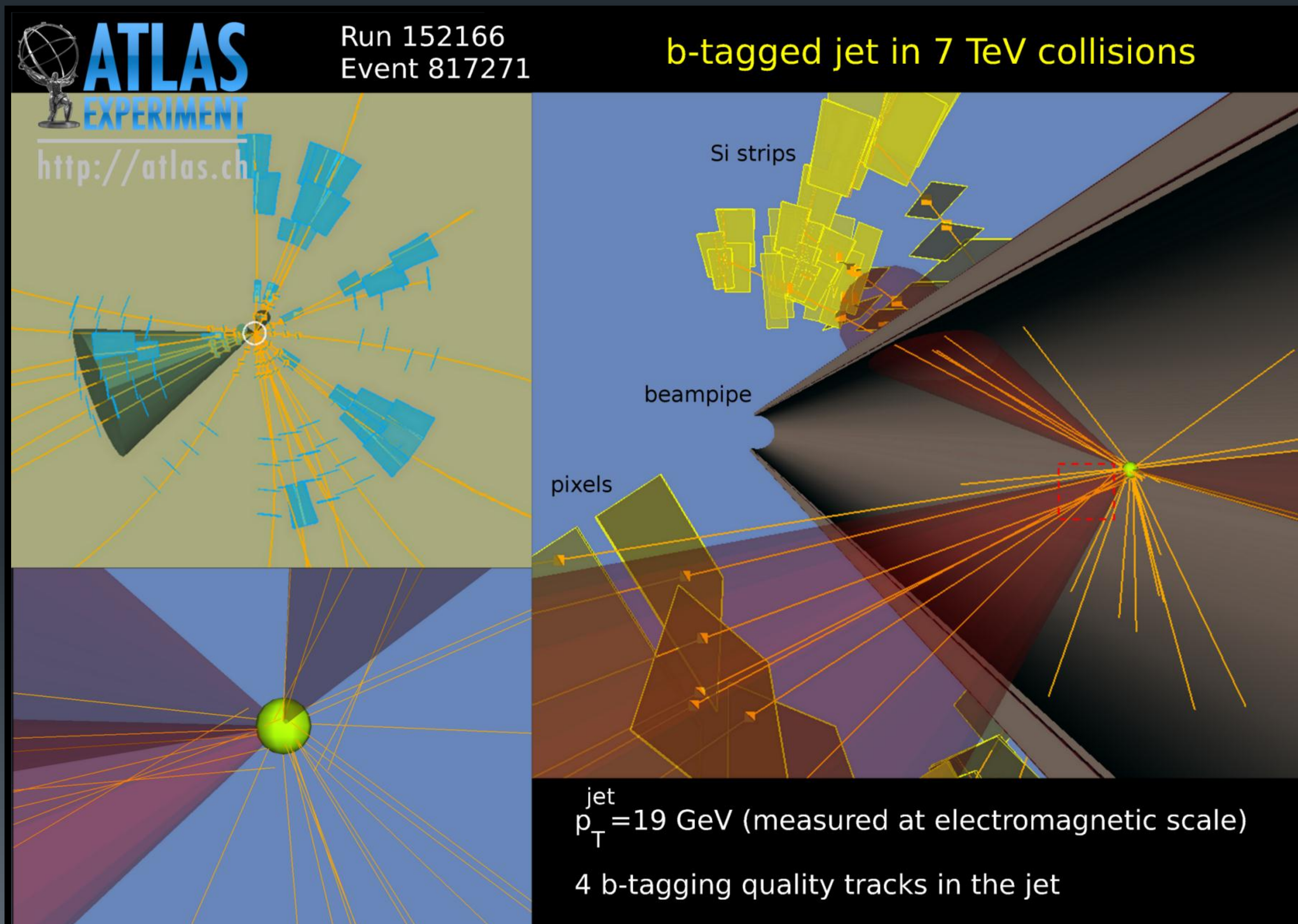


Untuned simulation & jet flavor fractions
Number of jets
∫ Ldt = 330pb⁻¹
ATLAS Preliminary
Pythia Dijet MC : light jets
Pythia Dijet MC : c jets
Pythia Dijet MC : b jets
data 2011
High-performance tagger: JetFitter
Number of vertices with ≥ 2 tracks
data/MC ratio
Number of vertices with ≥ 2 tracks



Untuned simulation & jet flavor fractions
Fraction of jets tagged by JetFitter / 10 GeV
ATLAS Preliminary
Pythia Dijet MC : light jets
Pythia Dijet MC : c jets
Pythia Dijet MC : b jets
data 2011
High-performance tagger: JetFitter
∫ Ldt = 330pb⁻¹
ε_b = 60%
Jet p_T [GeV]
data/MC ratio
Jet p_T [GeV]

b-tagging at 7 TeV



The Algorithms - Performance

Monte Carlo Based Performance

Efficiency

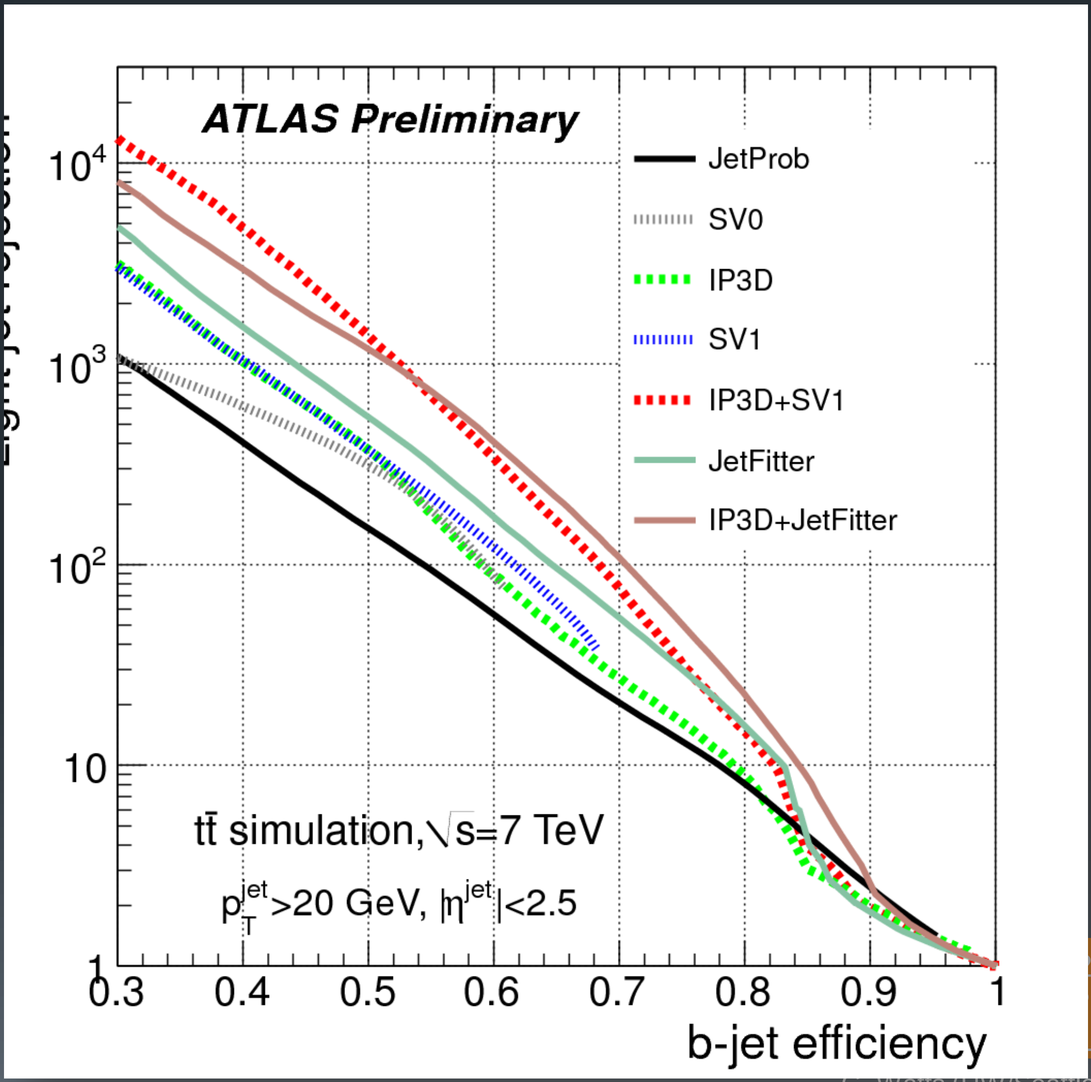
$$\epsilon_b = \frac{Tagged}{Good\ Jets}$$

Fake Rate

$$f_l = \frac{Tagged}{Good\ Jets}$$

Rejection

$$r_l = \frac{1}{f_l}$$



Calibrations

Despite (or in fact of) the MC/Data Agreement
Tagging Input Variables
Even Tagger Weights

Calibration takes care of any final discrepancies by providing a
Scale Factor

Bottom, charm, light

The Efficiency Techniques

All public results are from Moriond 2011

DiJet Events

p_T^{rel} — Use the p_T^{rel} distribution of the muon to determine the relative bottom fraction of a dijet sample

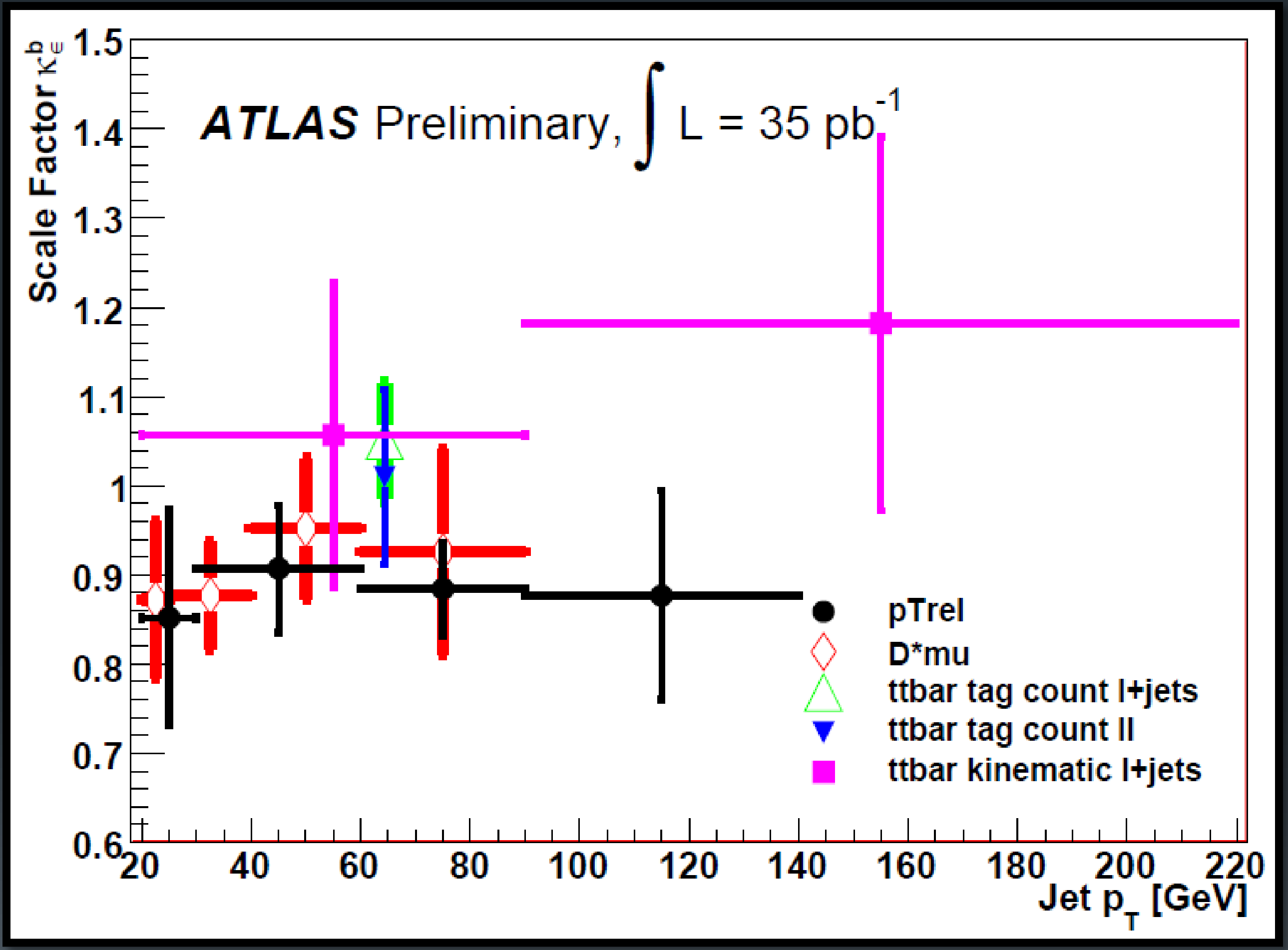
$D^*\mu$ — Exclusive reconstruction of $D^*\mu$ state makes for a very pure bottom quark sample of jets

$t\bar{t}$ Events

Tag Counting — Measure ϵ_b and $\sigma_{t\bar{t}}$ in a simultaneous fit

Kinematic Method — Count the number of tags in a very pure selection of $t\bar{t}$ events

The ϵ_b Results



The Fake Techniques

All public results are from Moriond 2011

DiJet Events – Only available source

SV0 Mass	Use the SV0 Vertex Mass distribution to predict the light/b/charm composition of the sample before and after tagging.
Negative Tag	Light quark tags are due to tracking resolution effects and are just as likely to appear in front and behind the primary vertex.

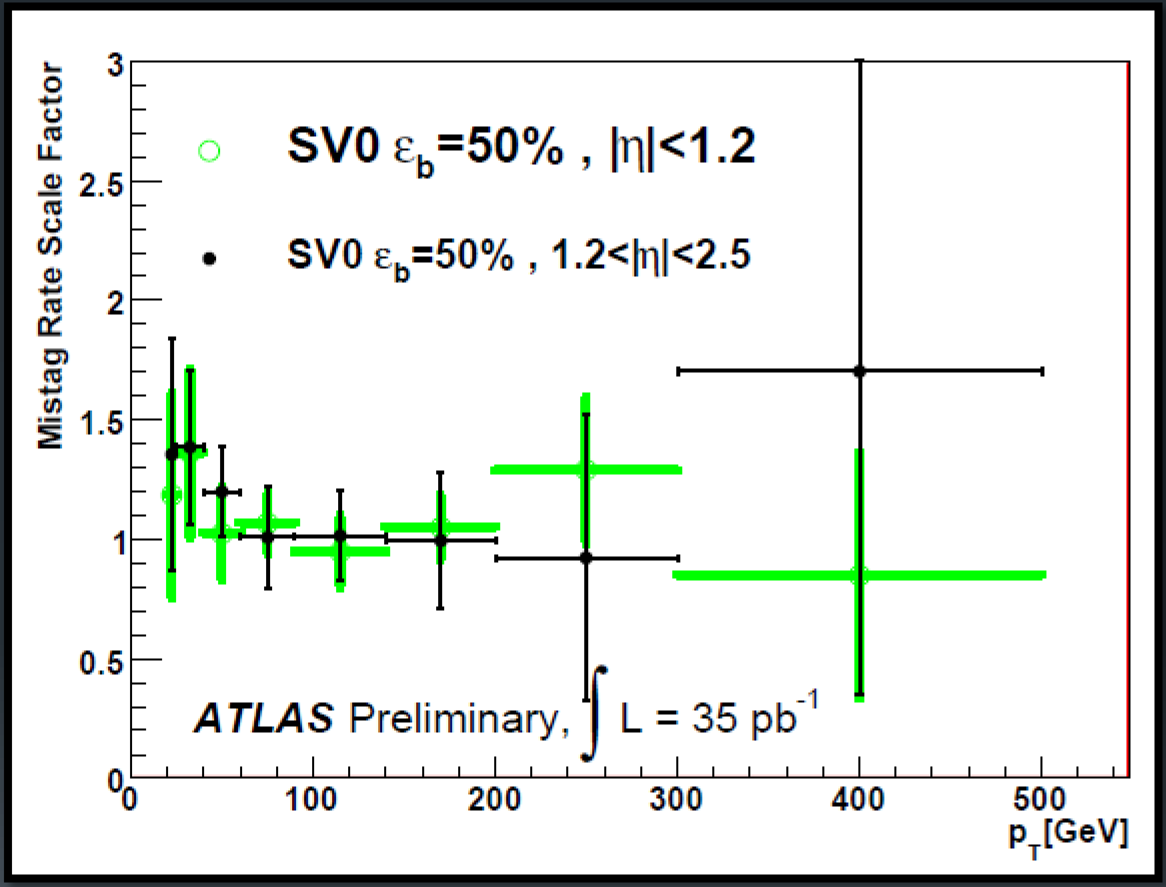
Fake Rate Combination

Results are compatible

Assume uncorrelated errors

Weighted combination

Split in η regions

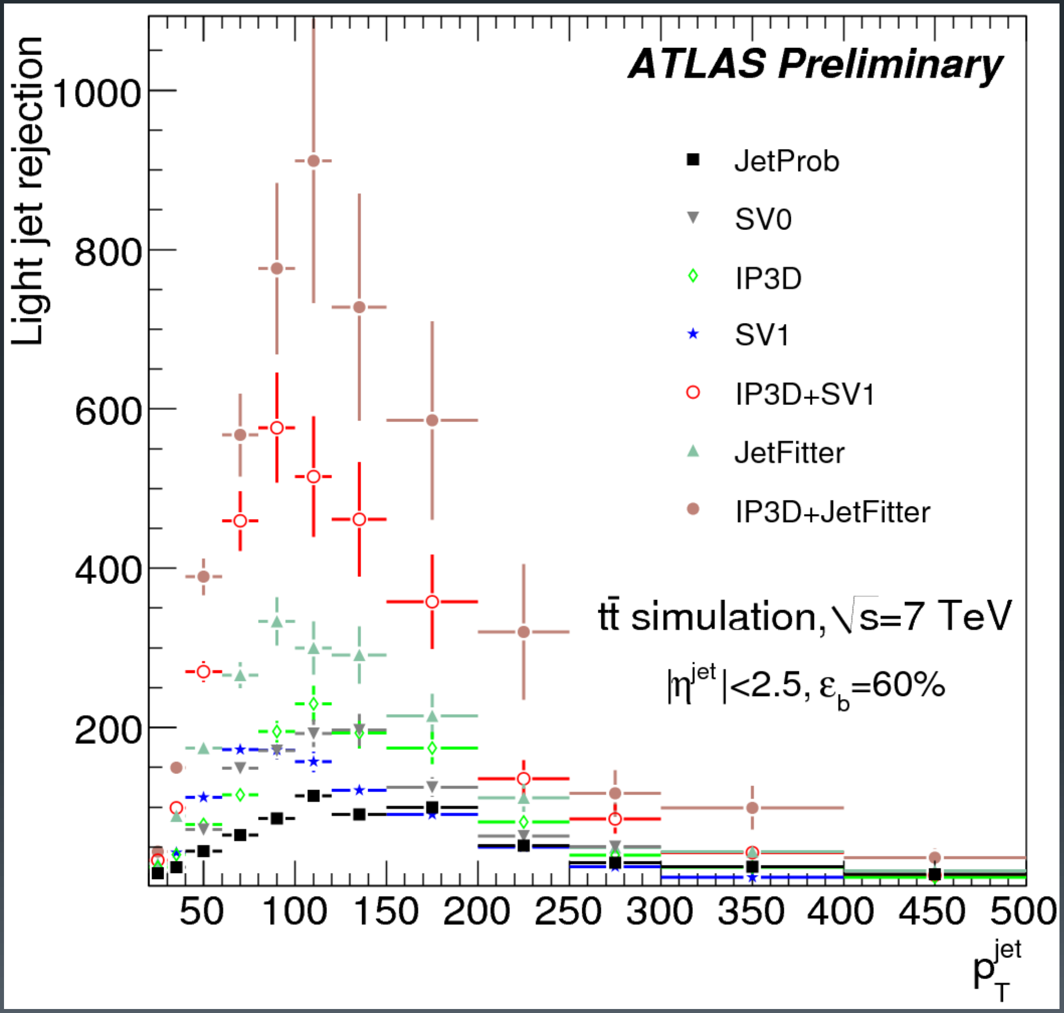


G. Watts (UW/Seattle)

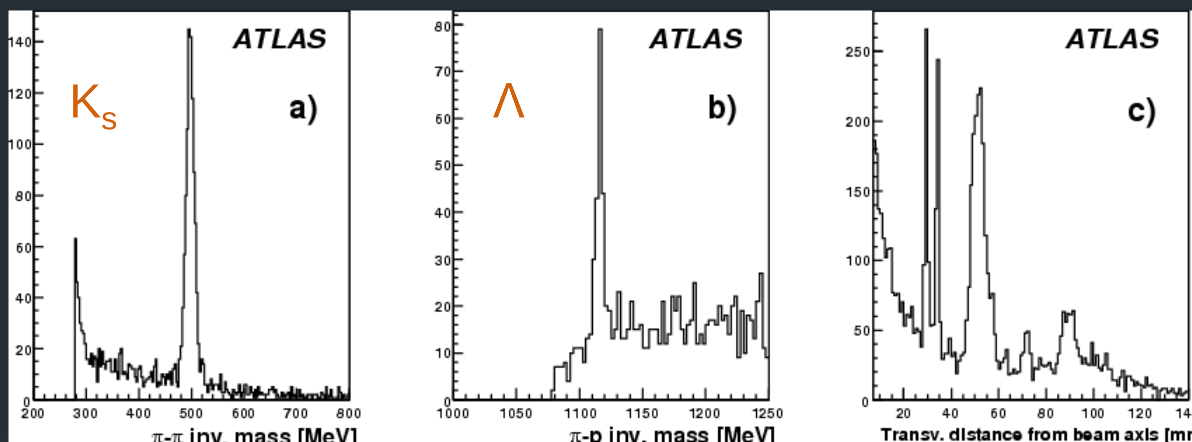
Conclusions

- Many results from ATLAS for Winter 2010 and Sumer 2011 are now using b-tagging and b-tagging calibrations
  - $t\bar{t}$ cross section measurements
  - SUSY heavy flavor searches (for example).
- Everything shown here can be found in public notes on the ATLAS b-tagging results pages.
- This is just the beginning
  - Full $t\bar{t}$ methods
  - System 8
  - Tagging very high p_T jets
  - Full weight calibration

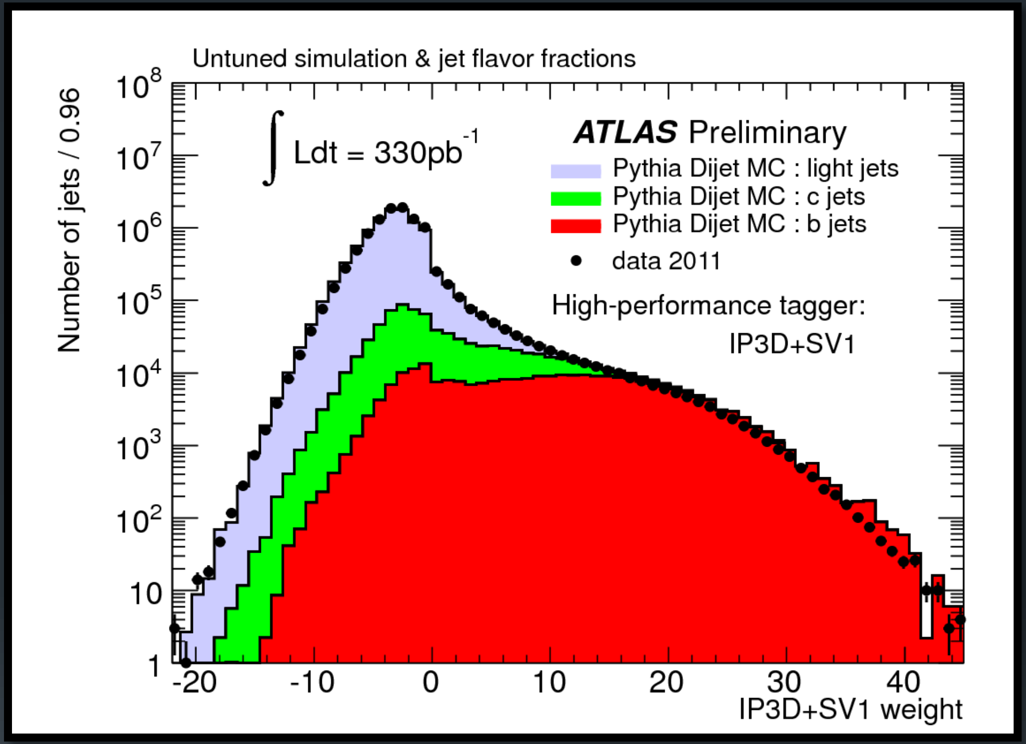
Performance as a function of p_T



Removing Material Interactions, etc.

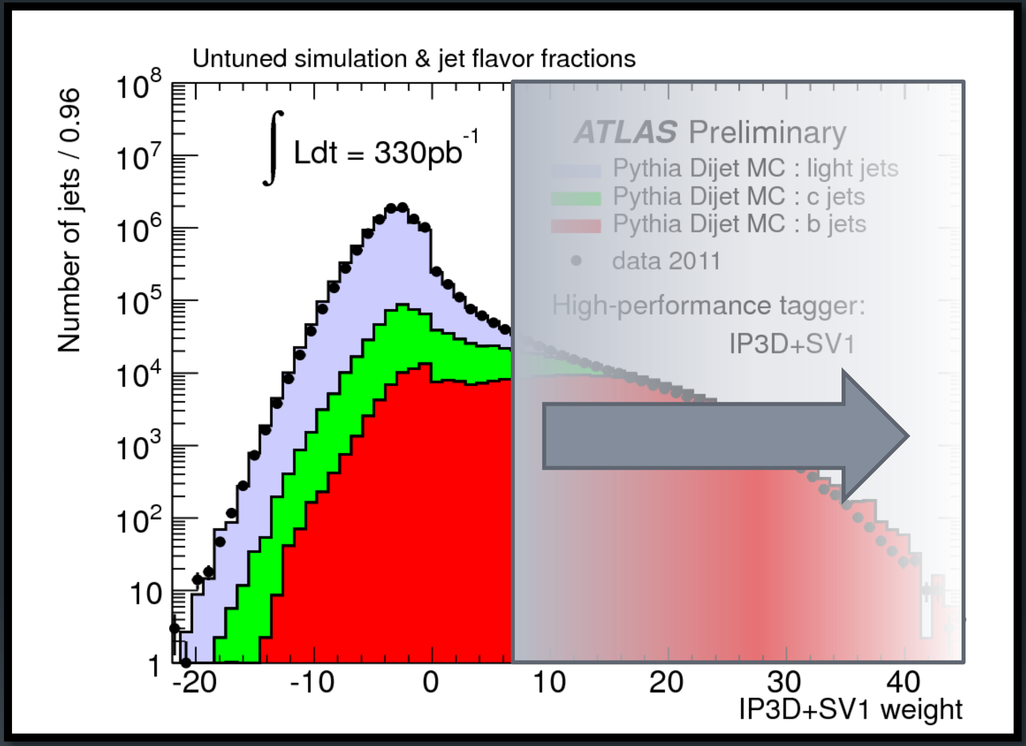


Calibrate The Weight



Input to a Decision Tree?

Calibrate Operating Points

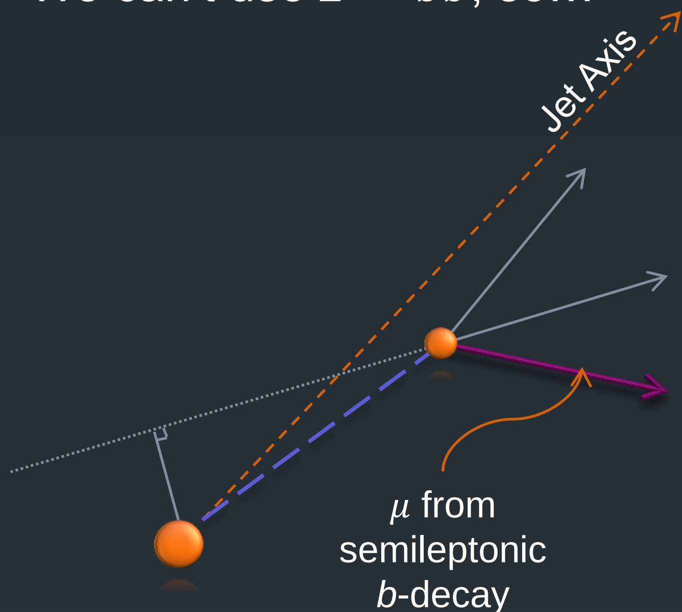


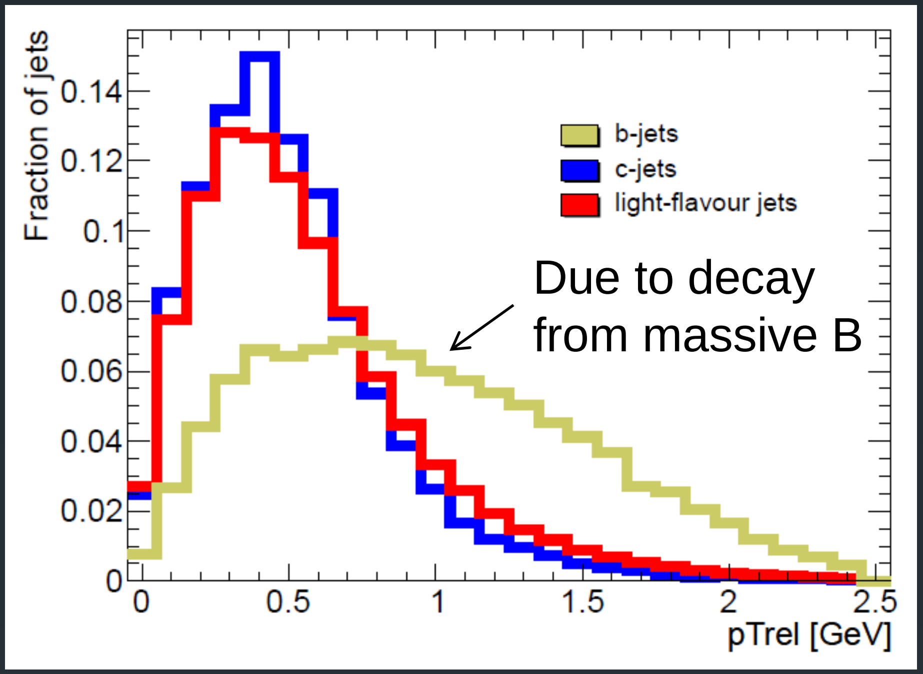
$$W_{jet} > 6.5$$

Two approaches…

Dijet Events

We can't use $Z \rightarrow b\bar{b}$, so…





p_T^{rel} of muon with respect to the jet axis

Fit the p_T^{rel} distribution in data with Monte Carlo templates before and after tagging

Relative change in bottom template fraction gives you tagging efficiency

p_T^{rel} Systematic Errors

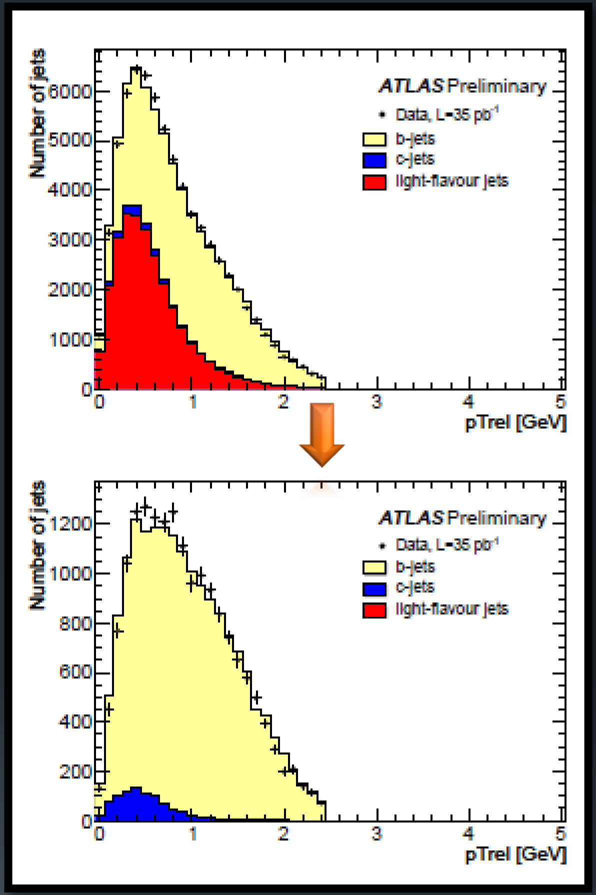
b-decay, fake muons, fragmentation, b-production, jet energy scale, Monte Carlo statistics, Scale Factor for inclusive jets, c-production, gluon splitting

+

Jet Axis Resolution

High energy jets have small angular separation between muon and jet axis

Method fails for high energy jets without modification



$t\bar{t}$ Methods

Kinematic Method

Flavor composition from Monte Carlo
High purity Lepton + Jets sample
Requires knowing the fake rate

Tag Counting Method

Measure ϵ_b and $\sigma_{t\bar{t}}$ in simultaneous fit
Requires knowing the fake rate and flavor composition of all jets in $t\bar{t}$ events (Monte Carlo)
Lepton+Jets and Dilepton Events are analyzed

Largest systematic uncertainties:

- W+Jets Background Normalization
- $t\bar{t}$ normalization (Kinematic)
- QCD Flavor composition

And a large number of other methods waiting in the wings

$D^*\mu$ Calibration

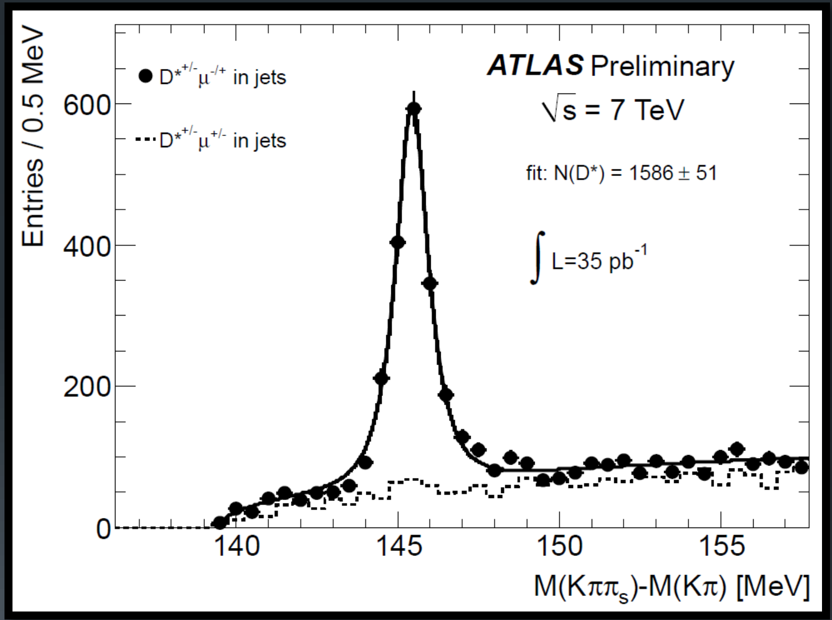
Exclusive Reconstruction of a B final state → Very pure sample of B's, with unique topology

$$b \to X\mu^- D^{*+} \to X\mu^- D^0(\to K^-\pi^+)\pi^+$$

$D^0 \to K^-\pi^+$ — Two opposite signed tracks

$D^{*+} \to D^0\pi^+$ — Additional track is added with proper sign

$\mu^- D^{*+}$ — Look at both sign muons – signal is clear



Account for fakes from $b \to DD$, gluon splitting, charm splitting, fake muons

Use technique similar with p_T^{rel} to extract tagging efficiency

Fake Rate

The better the algorithm → The more pure the sample required

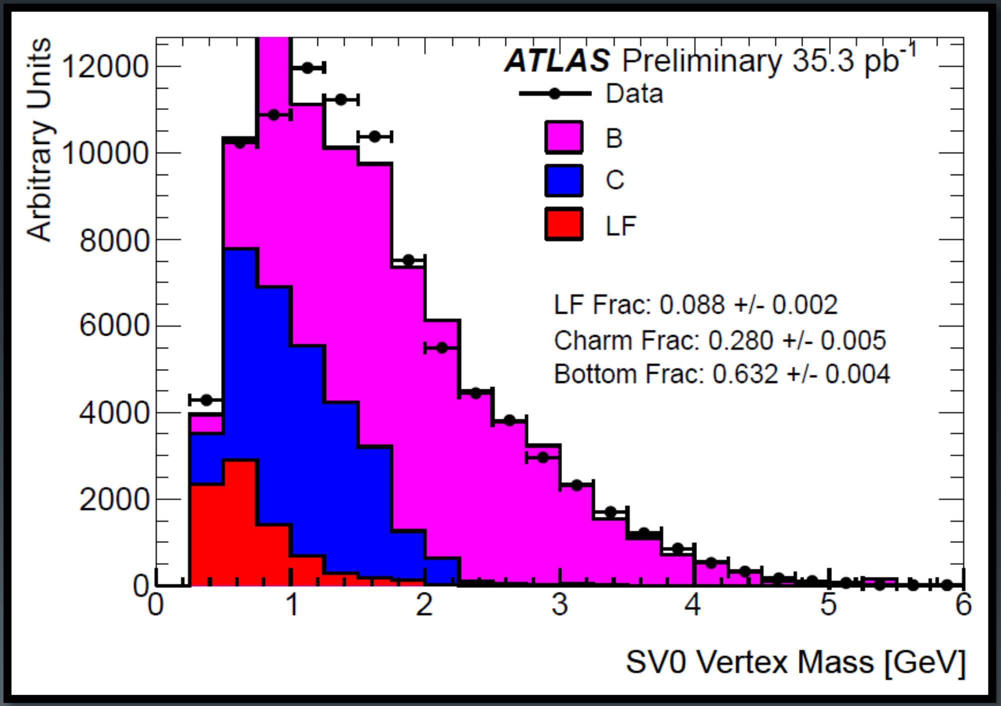
Sources of light jet tags:

- Long lived particles
- Material Interactions
- Detector Resolution

SV0 Mass Method

Requires a “tag”

$$\epsilon_l = \frac{N_l^{tag}}{N_l} = \frac{N_l^{tag}}{N_{data} - \frac{N_b^{tag}}{\epsilon_b} - \frac{N_c^{tag}}{\epsilon_c}}$$



Systematic Errors

ϵ_b, ϵ_c, statistics, shape of the mass templates

Negative Tag Rate

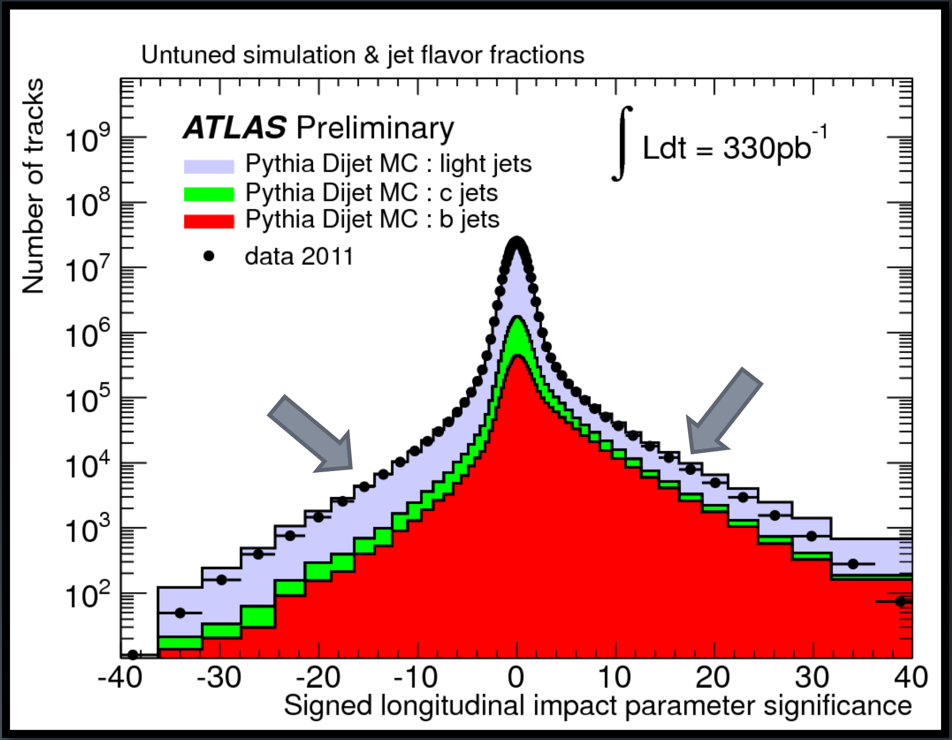
Resolution effects that lead to large d_0 should be symmetric about zero.

"Negative Tags" should look a lot like positive tags.

Negative tag uses jets with negative signed d_0.

Two Corrections:

- Heavy Flavor negative tag rates aren't the same as light negative tag rates
- Secondary interactions, etc., cause light tags but have no negative tag counter part.



G. Watts (UW/Seattle)