# Jet with Heavy Flavor

M. Shapiro

Sept 6, 2023

# Overview

- Processes with heavy flavor (HF) playing an increasingly important role in particle physics measurements
- Measurements involving HF provide important input to:
  - Insightful studies of QCD
  - Measurements of heavy particles that decay to HF
  - Constraints on SM couplings
  - Searches for BSM physics
- These studies require the ability to identify and measure HF jets

# Outline

- Jet Fragmentation
- HF Jet Tagging
- Examples of Using HF Jets for Physics

# Hadronization and Fragmentation Functions

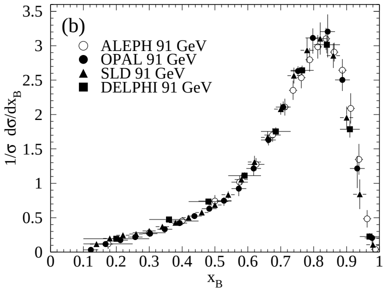
- Define distribution of hadrons using a "fragmentation function":
  - Define $D_q^h(x)$ as probability that a quark $q$ will fragment to form a hadron that carries fraction $x = E_h/E_q$ of the initial quark energy
  - We cannot predict $D_q^h(x)$
    - Measure them in one process and then ask are they universal
  - Like PDF's the $D_q^h(x)$ exhibit scaling violations as a function of $q^2$
- Parameterization of $D_q^h(z)$ essential for Monte Carlo programs used to predict the hadron level output
- Also important for modern NLO and NNLO calculations, some of which incorporate fragmentation into calculated observable
- In both cases, parameterization of fragmentation depends critically on theoretical approach

# Heavy Quark Fragmentation: $B$ hadrons

- Heavy flavored mesons retain a large fraction of momentum of intial quark
- For (N)NLO calculations:
  - In limit of very large quark mass, fragmentation peaked near $z = 1$:

$$D_Q(x, \mu^2) = \delta(1 - x)\Big|_{\mu^2 = m_Q^2}$$

  - Large perturbative corrections can be resummed over powers of $\alpha_S \log(m_Q/p_T)$ and to NNLO accuracy
  - Inclusion of non-perturbative effects by convoluting perturbative result with a phenomenological non-perturbative form.
- For Monte Carlos, introduce phenomological form for $D_Q(x, \mu^2)$ and fit to experimental measurements



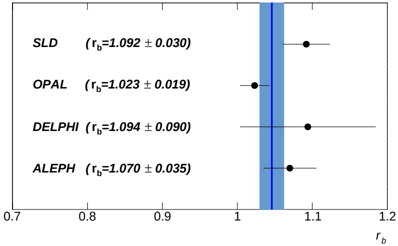

- For $b$-quarks good data exists from LEP/SLC
- Small differences between experiments can be treated as a systematic uncertainty
- Decay mode most commonly used: $B \to D^{(*))}\ell\nu$
  - Small correction to reco-level value needed to account for unmeasured $\nu$

# Fragmentation in Pythia

- In PYTHIA8 HF fragmentation fn given by the Lund-Bower function

$$f(x) = \frac{1}{x^{1+r_q b m_q^2}} (1-x)^a \, e^{-b m_T^2/x}$$

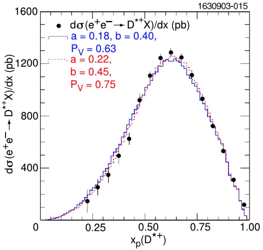
- The $r_q$ parameter can be tuned for each HF species to improve agreement with measured $D_q^h(x)$ distribution
  - Fragmentation fn applied in MC AFTER gluon radiation
  - Fit to same data needs different parameters for different $\alpha_S$ values
  - ATLAS uses the A14 tune of PYTHIA which has $\alpha_S = 0.127$ while Monash uses $\alpha_S = 0.1365$
  - This has a big effect on input parameter for fragmentation




- Tuning A14 to LEP data moved $r_b$ from Monash value of 0.855 to a new value of 1.05
- This changed measured top mass by a few 100 MeV!

# Heavy Quark Fragmentation: Charm hadrons

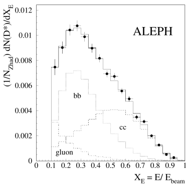
CLEO

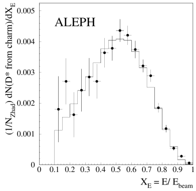



- Measurements published for $D^+$ and $D^{*+}$
- Rivet routine CLEO_2004_S5809304
- Mean values of $x_p = p/p_{max}$:

$$D^+: \quad 0.582 \pm 0.008 \pm 0.004$$
$$D^{*+} \quad 0.611 \pm 0.007 \pm 0.004$$

LEP



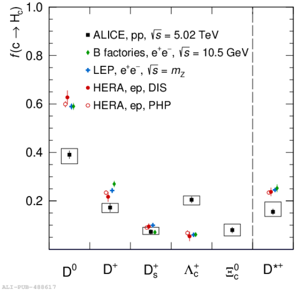

- Most precise measurements from ALEPH arXiv:hep-ex/9909032v2
- Problem: Contributions from $B$ decay and gluon splitting
- Attempt to isolate $c$-fragmentation, but resulting histogram NOT provided
- Quoted mean value of $x_E \equiv E_D/E_{beam}$:

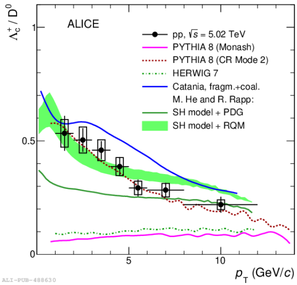
$$< x_E >= 0.4778 \pm 0.0046 \pm 0.0061$$

Not clear if quoted mean agrees with the left fig above?

# Is HF Hadronization Universal?

- Has always been assumed that fragmentation is universal
- But there are reasons this might not be true:
  - Color flow in hadron collisions much more complicated
  - Final state partons can interact with remnants from initial hadrons
  - Possible presence of coherent effects
- Recent results from ALICE show a higher-than-expected baryon production rate at low $p_T$
- Interesting result, not-yet fully understood theoretically, that demonstrates need to test fragmentation models in the same phase space as the physics measurements being performed



# Identifying (aka Tagging) HF Jets

- HF jet defined as a jet containing one or more HF hadrons
  - Typically don't include quarkonia (which has charm and bottom number zero)
  - Higher mass HF states decay strongly to lighest states of same flavofr
  - Lightest states decay weakly
- Tagging strategy depends on properties of these weakly decaying states

Bottom Hadrons

| Species | Mass GeV | $c\tau$ $\mu$m | semileptonic BR |
|---|---|---|---|
| $B^+$ | 5.279 | 491 | 11% |
| $B^0$ | 5.279 | 455 | 10.3% |
| $B_s$ | 5.366 | 456 | 9% |
| $\Lambda_B$ | 5.619 | 441 | 10.9 |

Charm Hadrons

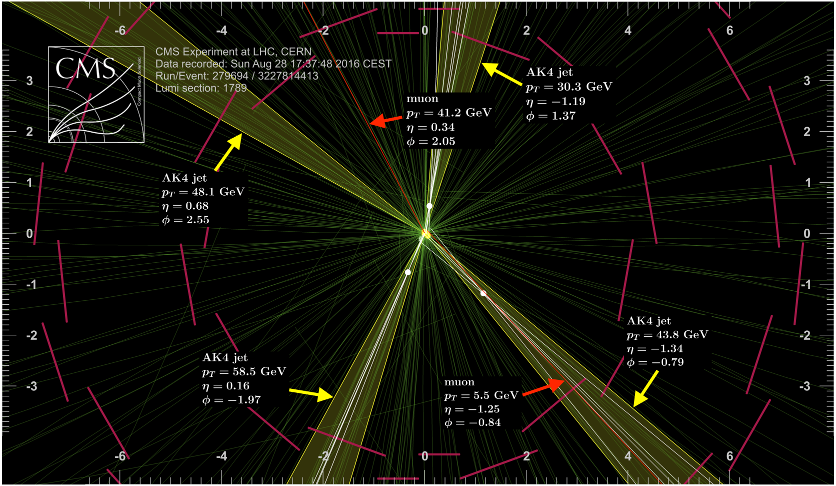
| Species | Mass GeV | $c\tau$ $\mu$m | semileptonic BR |
|---|---|---|---|
| $D^+$ | 1.870 | 309 | 16% |
| $D^0$ | 1.865 | 123 | 6.5% |
| $D_s$ | 1.968 | 151 | 6.3% |
| $\Lambda_C$ | 2.286 | 60.4 | 3.9% |

Important characteristics:

- States with mass $\sim 1.8$ GeV for charm and $\sim 5.2$ GeV for botto
- Long lifetime
- Large semileptonic BR

These properties define how to tag HF jets

# HF jets at the LHC



# Track Impact Parameters for HF Decay products

- IP defined as distance of closest approach of reconstructed track to primary vertex
  - At LHC, often use transverse impact parameter since beamspot is small is $x$-$y$ direction and long in $z$
- IP given by

$$d_0 = \gamma\beta c\tau \sin\phi$$

  where $\tau$ is HF proper decay time and $\psi$ is angle between secondary vertex and HF-hadron direction of flight
- $\sin\phi$ is $\propto 1/\beta\gamma$, so $< d_0 > \propto c\tau$, independent of HF-hadron momentum
  - One advantange of IP tagging: Does not depend on knowledge of HF-hadron momentum spectrum
- IP can be signed to be positive if track consistent with coming from vertex with positive decay distance and negative otherwise
- Can construct likelihood function for track IP for primary tracks (distribution depends on multiple scattering and uncertainty on primary vertex position)
  - Overall likehood constructed as product of likelihoods of all tracks
- Product likelihood is one option to use for HF-tagging

# Vertex Tagging

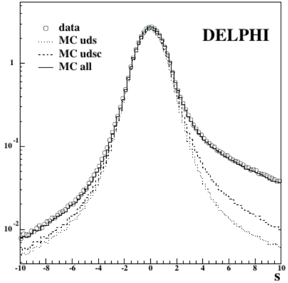
- Rather than treating tracks independently, can start with large IP tracks and ask if they are consistent with coming from a single vertex
- Vertex constrained fit: vary track parameters within uncertainties on fited parameters to find best vertex position and associated track parameters for the tracks
- Position of secondary vertex and its uncertainty returned from the fit
- More sophisticated algorithms can ask if more than one vertex is present
  - Either from multiple HF in jet (gluon splitting) or from $b \to c \to light$
  - ML techniques such as graphical neural nets perfect for this application
- Can calculate mass of the secondary vertex.
  - Helps to separate $B$ and $D$

# HF tagging with leptons

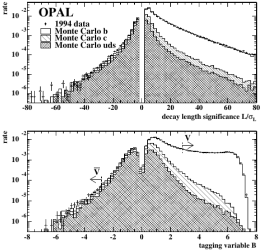
- Leptons from $W$ and $Z$ decays tend to be isolated (not near any jets) and have high $p_T$
- Leptons from HF tend to be inside jets which momentum distribution that depends on the HF hadron momentum
- Background leptons inside jets come from hadron decays (eg $\pi^+ \to \mu^+ \nu_\mu$ and $\pi^0 \to \gamma e^+ e^-$)
- Electrons from photon conversion also a source of background
- Leptons from HF decay will have non-zero IP and transverse momentum relative to the jet axis
  - IP distribution depends on $c\tau$
  - $p_T^{rel}$ depends on mass of HF hadron
  - Can separate signal from background and bottom from charm by fitting shape of $d_0$ and/or $p_T^{rel}$ distributions (or defining signal and background likelihoods)
- Here again, ML techniques can really help with the separation

# Heavy Flavor Tagging Methods at LEP (I)

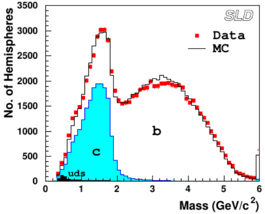
IP Significance for single tracks



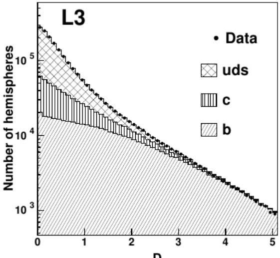

Decay Length Significance




Vertex Mass from secondary Vertex

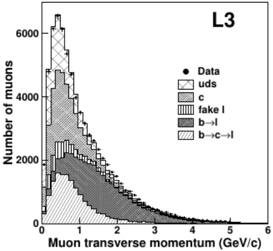



IP Significance from track product likelihood

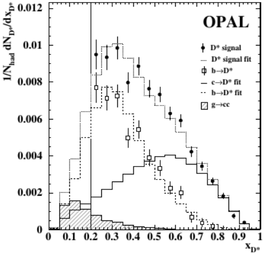


# Heavy Flavor Tagging Methods at LEP (II)

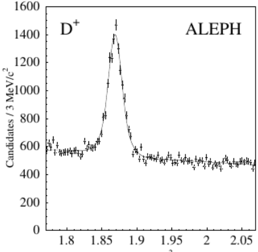
Muon $p_T^{rel}$




Fraction of jet momentum carried by HF hadron




Fully reconstructed hadrons



# Example of using HF tagging: $R_b$ and $R_c$ Measurements at LEP

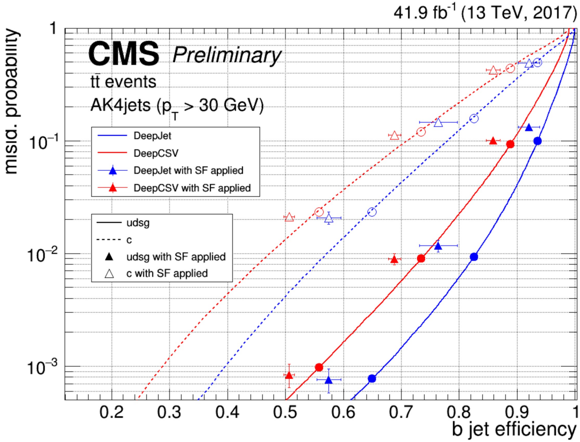
- Double Tag method (two hemispheres)

$$
\begin{aligned}
f_s &= \epsilon_b R_b + \epsilon_c R_c + \epsilon_{uds}(1 - R_b - R_c) \\
f_d &= \epsilon_b^{(d)} R_b + \epsilon_c^{(d)} R_c + \epsilon_{uds}^{(d)}(1 - R_b - R_c) \\
\epsilon_f^{(d)} &= (1 + C)\epsilon_f^2
\end{aligned}
$$

  where $f_s$ and $f_d$ are fraction of single and double tagged events and $C$ is a small correction due to correlation between hemispheres

- Note: Requires simulation for the $\epsilon$'s and independent measurement of $R_c$
- Multitag method
  - Employ several tags and independent categories to refine the measurement

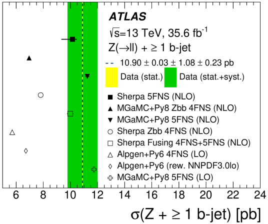
# New Developments: ML based tagging



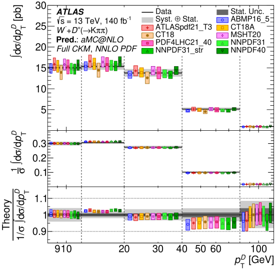

- Example of latest-and-greatest from CMS
- Similar plots available from ATLAS
- Detailed discussion of algorithms and their calibration would be good topic for a student presentation

# Using HF jets to study QCD (Some examples)

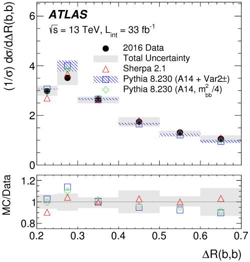
## $Z + b-$jet production



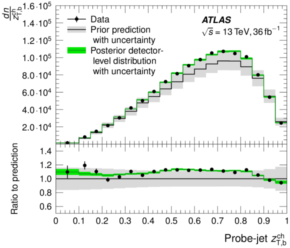

## $W + c$-jet production




## $g \to b\bar{b}$

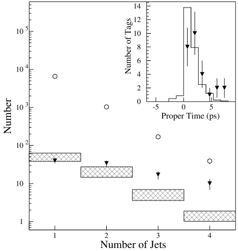



## $b$-fragmentation from $t\bar{t} \to W * +bW^{-}b$

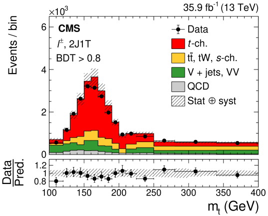


# Particles that Decay to HF jets

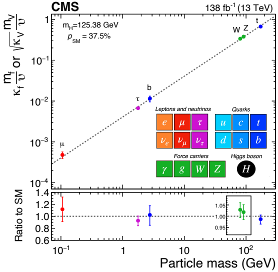
## Discovery of Top (CDF)



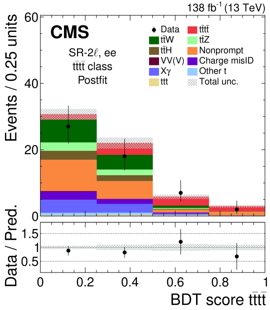

## Top Mass Measurement




## $h \to b\bar{b}$




## 4 top production $t\bar{t}t\bar{t}$



# Conclusions

- HF jets important for many physics measurements
- QCD studies of HF interesting in their own right
- Sophisticated tagging algorithms exist together with techniques to calibrate the efficiency and purith
- Lots of good topics for students talks in this area