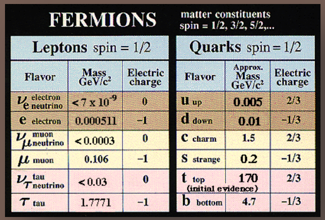
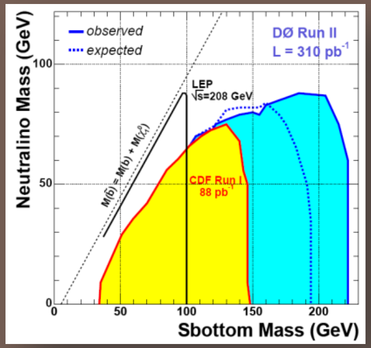
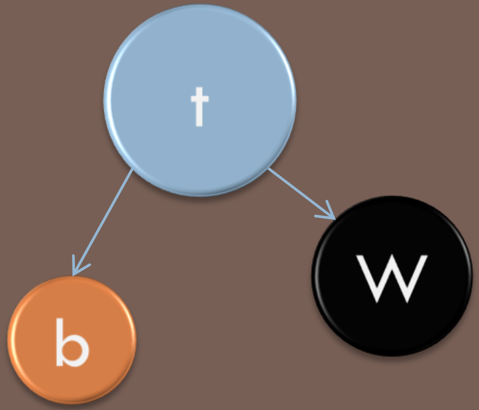


FERMIONS matter constituents spin = 1/2, 3/2, 5/2,...

Leptons spin = 1/2			Quarks spin = 1/2		
Flavor	Mass GeV/c²	Electric charge	Flavor	Approx. Mass GeV/c²	Electric charge
ν_e electron neutrino	$<7 \times 10^{-9}$	0	u up	0.005	2/3
e electron	0.000511	−1	d down	0.01	−1/3
ν_μ muon neutrino	<0.0003	0	c charm	1.5	2/3
μ muon	0.106	−1	s strange	0.2	−1/3
ν_τ tau neutrino	<0.03	0	t top (initial evidence)	170	2/3
τ tau	1.7771	−1	b bottom	4.7	−1/3



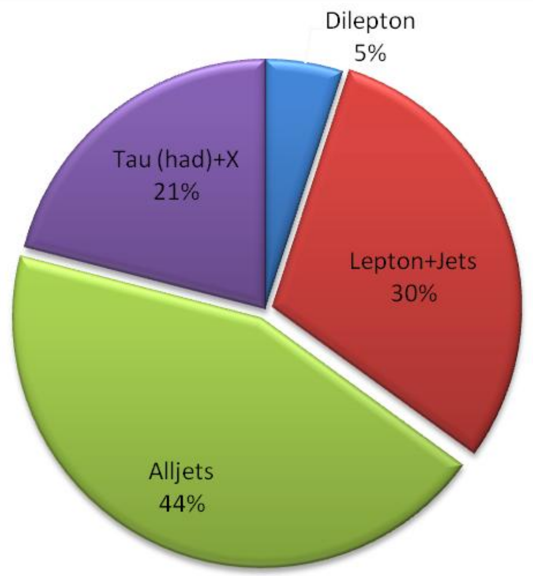


BRING UP B-TAGGING AT DØ

Gordon Watts (Seattle/Marseille)

Introduction

Top Physics: *Almost* Lost Without B-Tagging

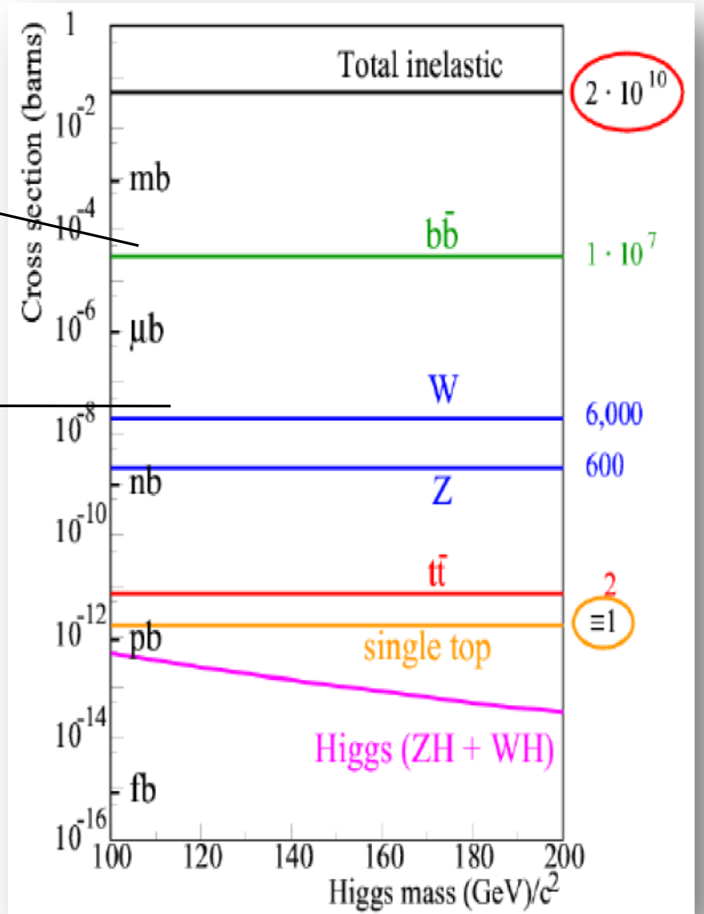


All Jets

Required to reduce QCD background.

Lepton + Jets

Reduce W+Jets background.



Tevatron Cross Sections

Typical efficiency currently

Fake Rate	0.5%	1%
b-Efficiency	48%	55%

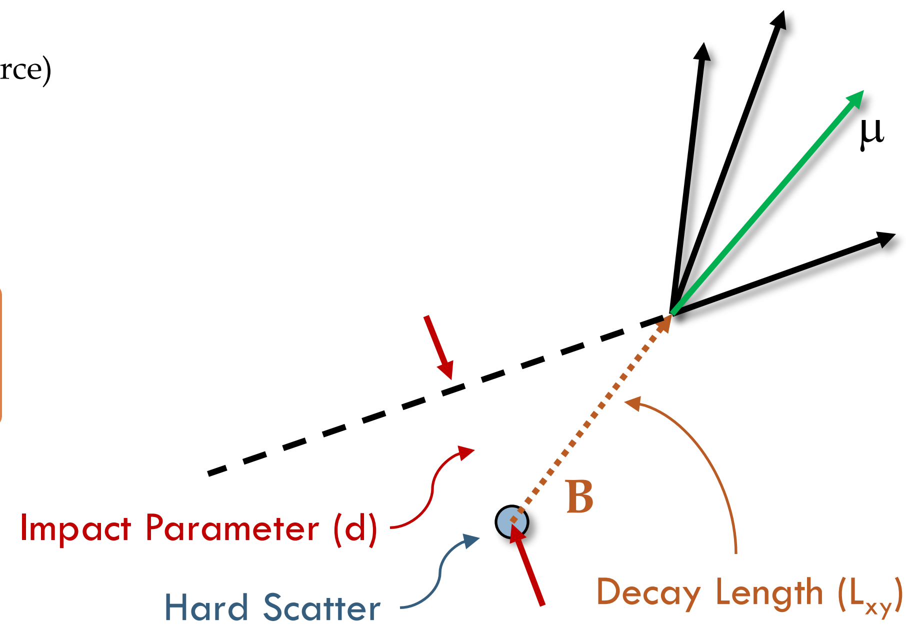
Much looser cuts possible for double tag channels…

Relative cross sections change at the LHC…

B Decay Basics

- B hadrons have a long lifetime (decays via weak force)
  - $c\tau \sim 450\ \mu m$
  - Can decay as far as 3mm away
  - Average of 4.5 tracks in decay

Tracks with large impact parameters (d)
Vertex Reconstruction (L_{xy})



- B hadrons decay to a muon about 15% of the time
  - Soft Lepton Searches
  - μ is the only viable one
    - e – high fake rate
    - τ – just hard

Efficiency is low, however…

Soft muons in jets

Can also use Jet Mass, N_{trk}, etc.

Using it in an analysis brings up other issues…

Monte Carlo

Running on MC

Tracking performance the same?

Directly tag on MC and scale to data performance?

Roll the dice using data-derived probabilities?

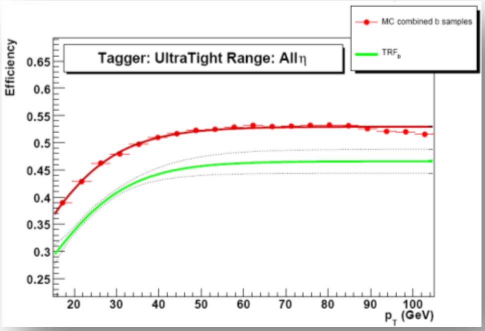
Even if MC is good enough, detector conditions change!

Calibration

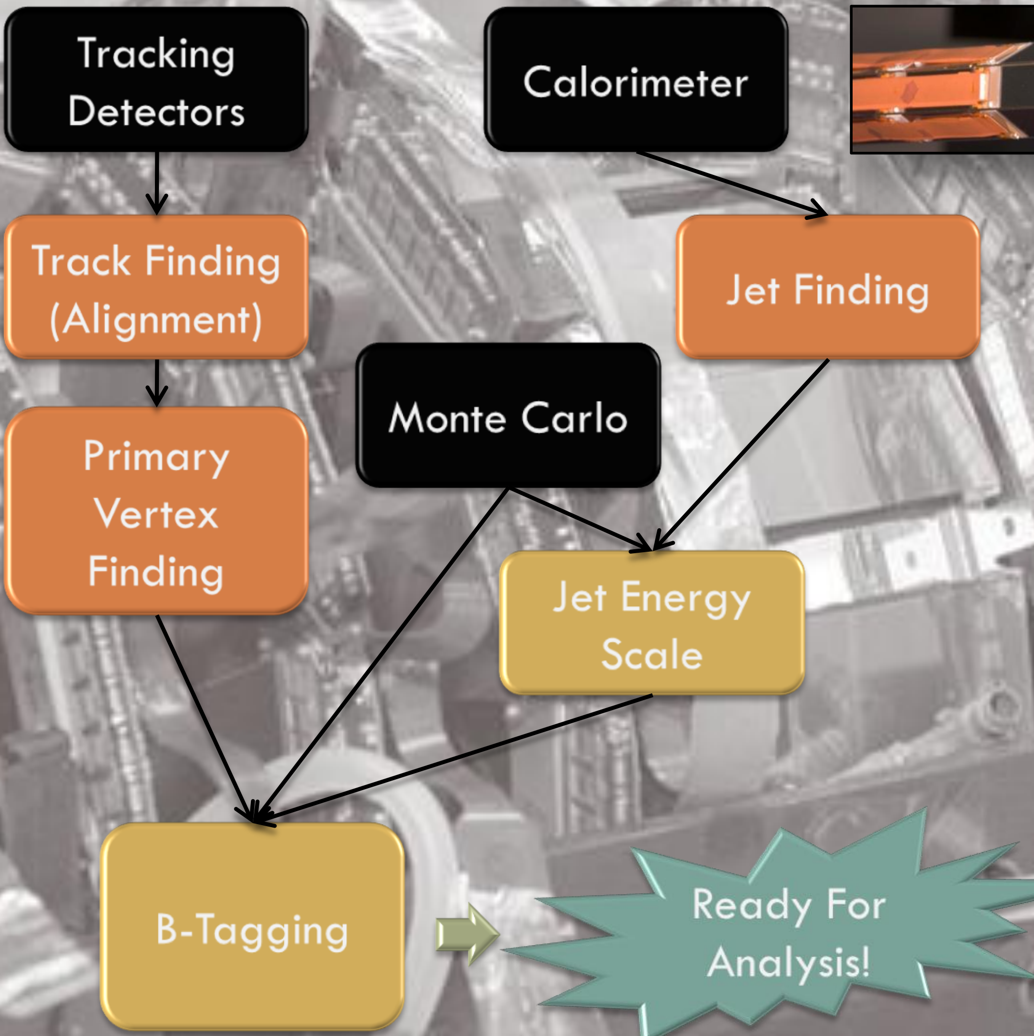
Where do you find a sample of known b-content??

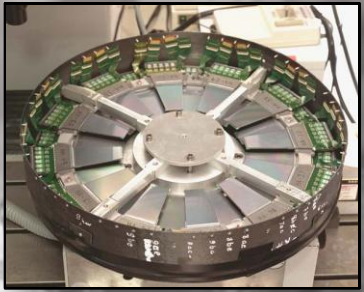
Calibrate as a function of η and jet p_T, etc.

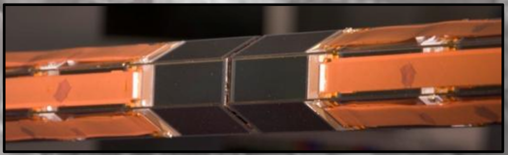
- Jet energy scale!
- Does your calibration sample have statistics in right bins for your signal?

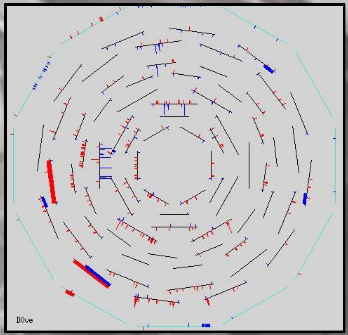


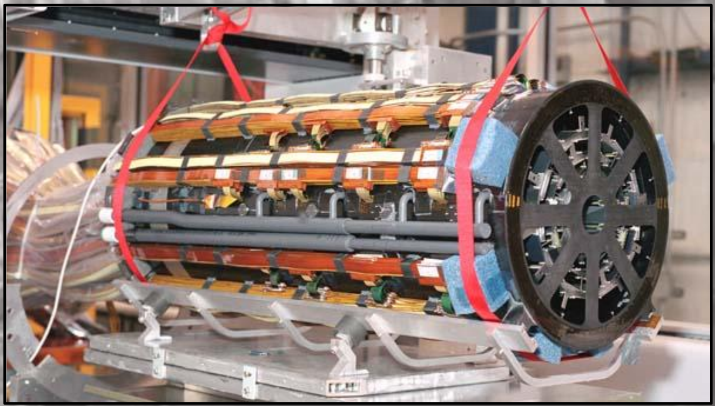
The Food Chain

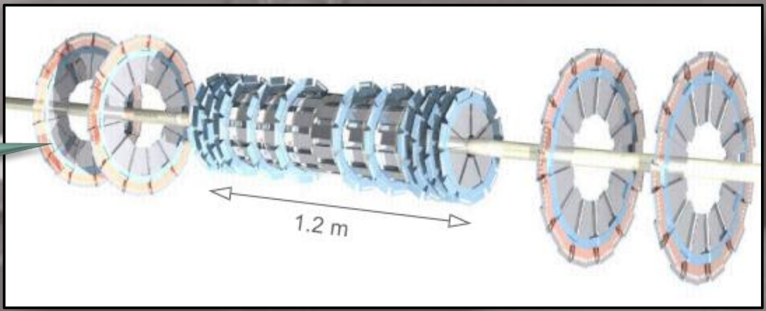










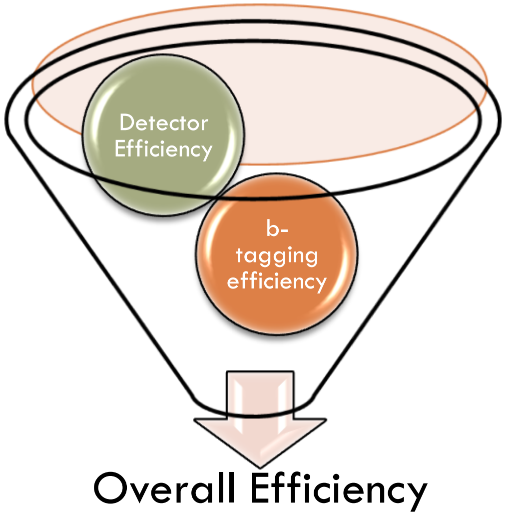


Taggability

We discovered this after about a year of work trying to get calibration results to make sense on data…

Detector conditions change

- Silicon wafers turned off for short periods of time, etc.
- Monte Carlo never reflects these exactly.
- But these have a direct impact on b-tagging efficiency.



Taggability separates these two effects.

1. Require a jet to be tagable
2. Apply b-tagging

→ Removes detector effects, sample, trigger, and selection specific effects

Taggability must be determined for each analysis after trigger selection, and then applied to the Monte Carlo.

Jet $E_T > 15$ GeV
$N_{trk} \geq 2$, $p_T > 0.5$ GeV, $\Delta R < 0.5$
1 track $p_T > 1$ GeV, SMT hits

Central jets: 98% tagable

The Algorithms

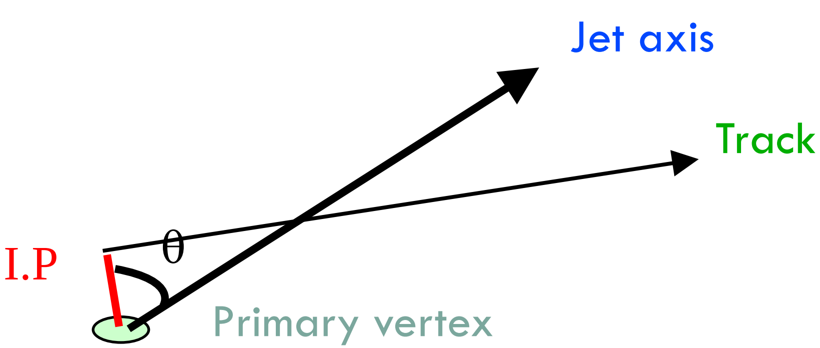
DØ started with three…

CSIP Algorithm

Counting Signed Impact Parameter

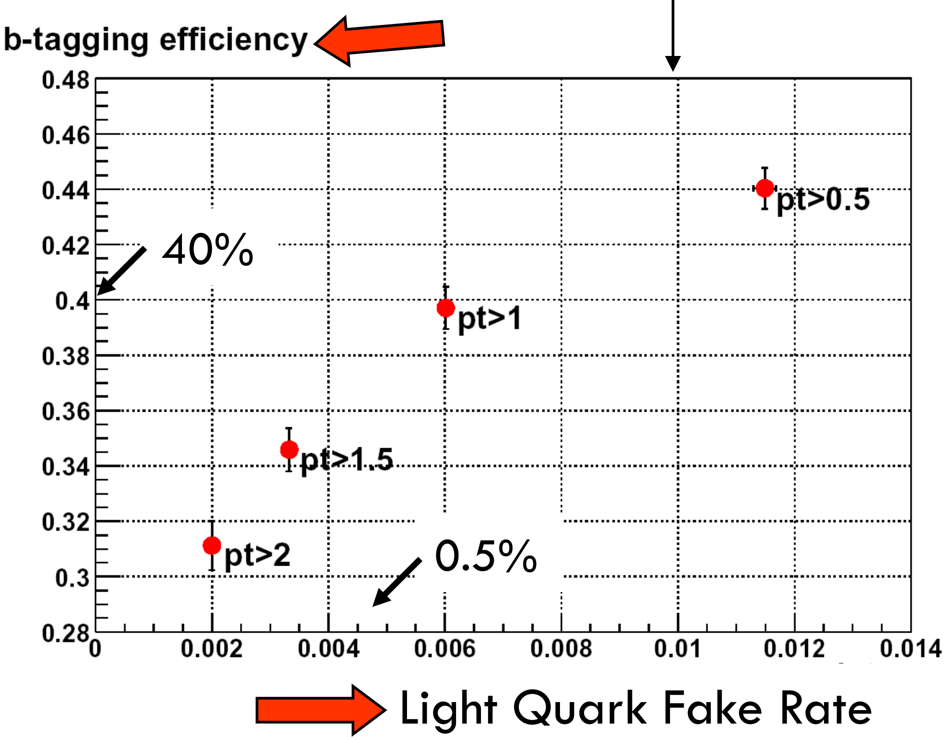
Based on Impact Parameter Significance

$S(IP) = IP/\sigma(IP)$



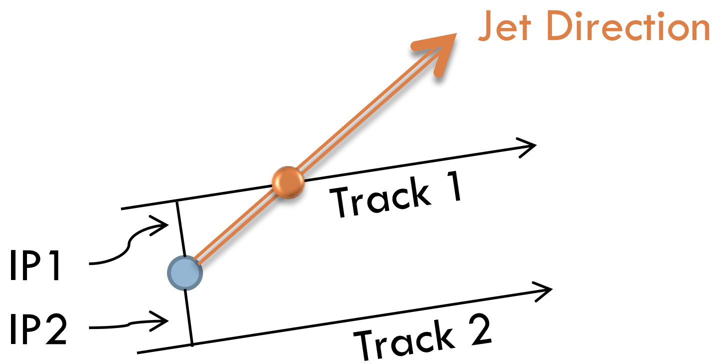
Requirements to tag a jet:
- at least 2 tracks with S(IP) > 3
- or at least 3 tracks with S(IP) > 2

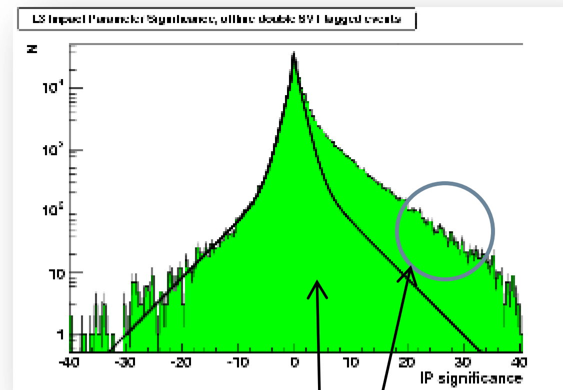
Per Tagable Jet Rates Measured in Data!



Signed Impact Parameter

Negative impact parameter definition.





Track 1 could originate inside the jet core, IP1 > 0
Track 2 could not, IP2 < 0

Negative impact parameter tracks are due to resolution effects.

- There is no physics that would generate a negative impact parameter.
- Resolution effects should contribute equally to positive and negative impact parameter tracks.

Predict the positive IP tracks due to resolution effects – the rest will be due to long lived particles!

JLIP Algorithm

Jet Lifetime Impact Parameter

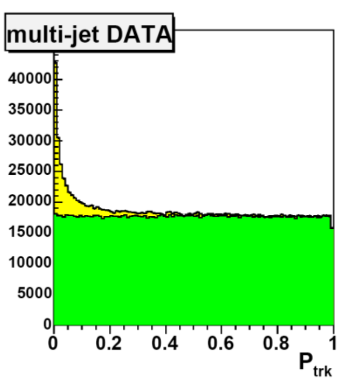
- Based on Impact Parameter Significance
- Use IP<0 tracks to construct flat probability distribution in IP.

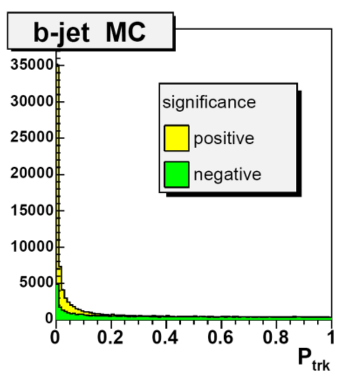
Use probability distributions

P(Track from PV)

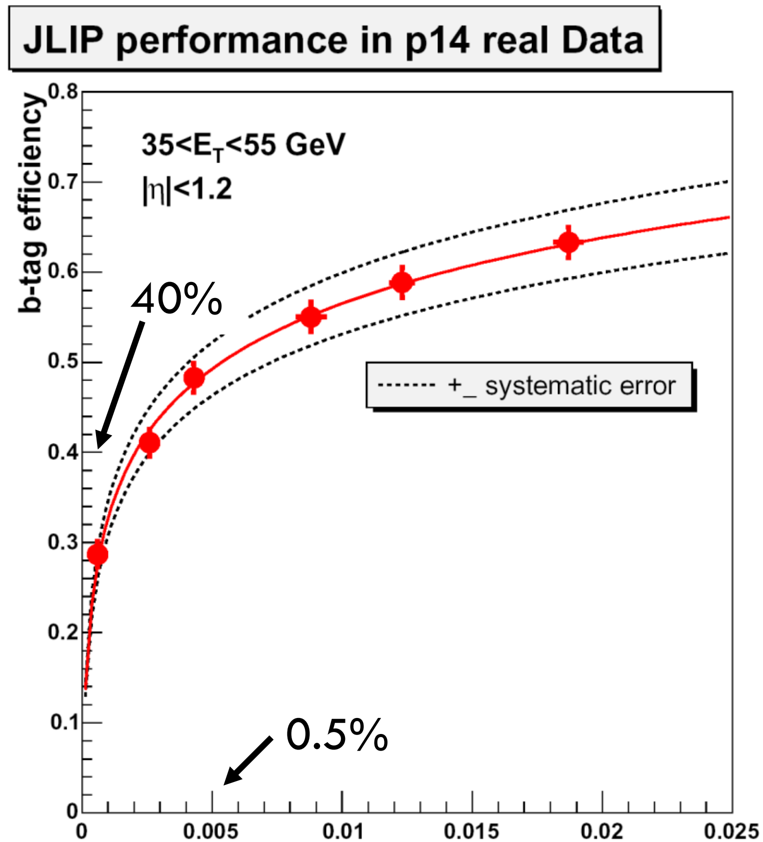
Defined for each class of tracks

# of SMT Hits, p_T, etc.





Each jet assigned P(light quark)



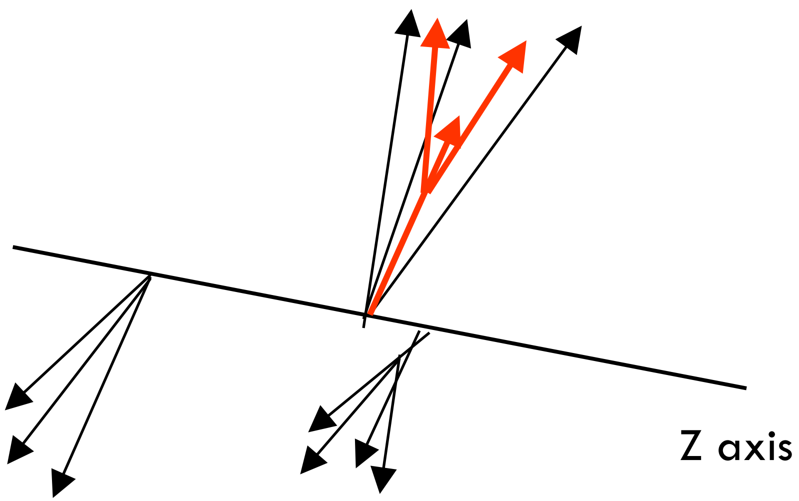
Light Quark Fake Rate

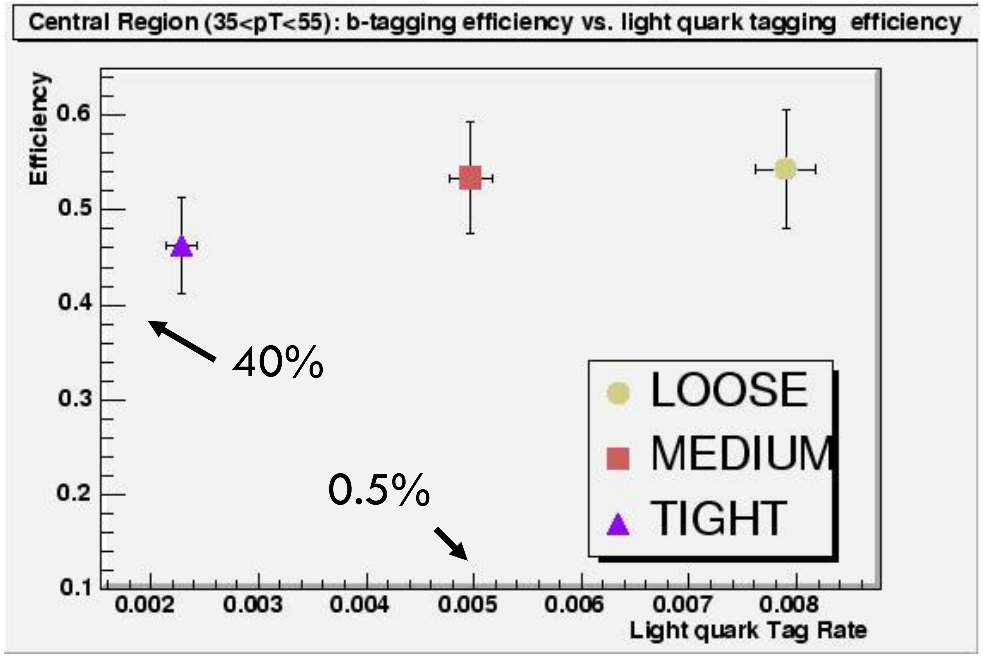
SVT Algorithm

Secondary Vertex Tagger

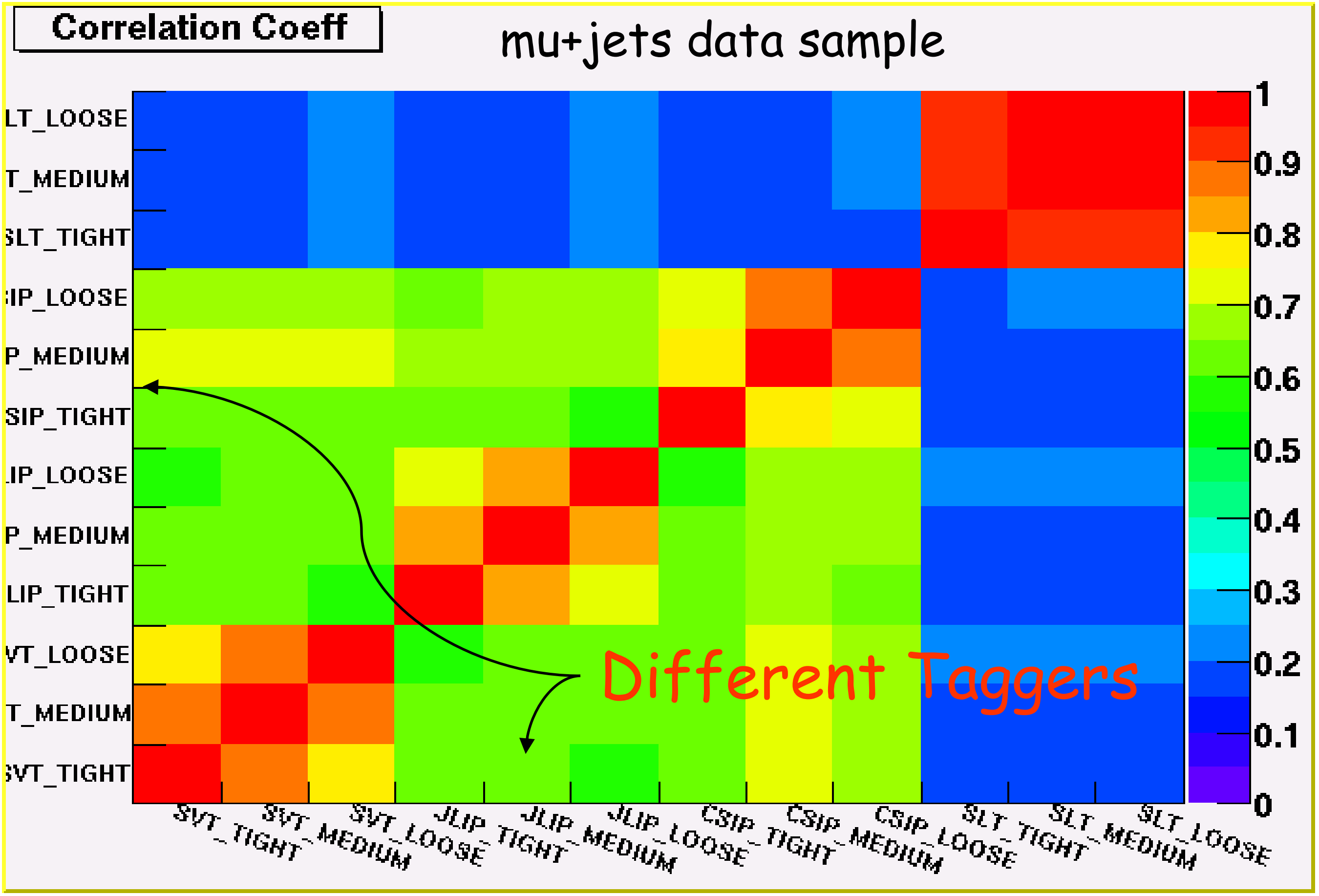
- *Reconstruct* vertices using displaced tracks.
- Cut on decay length significance.

$S(L_{xy}) = L_{xy}/\sigma(L_{xy})$.





That is what we started with…



70% Correlated (signal)

30% Correlated (fake)

Uncorrelated on background, correlated on signal

Combine with a simple NN

Gain of more than 20% in efficiency

General Comments On The Taggers

b-ID group's first efforts on b-tagging:

"What the!?"

Understanding our results meant understanding tracking, vertexing, etc.

→ We ended up working in other groups instead of on b-tagging!

Primary vertexing is a good example.

Part way into the run we discovered that the PV error was distorted along the direction of the b-jet

Tracks with larger than normal impact parameter were being used in the PV fit

→ b-tagging was the only one who cared about the PV at this level!

General Comments On The Taggers

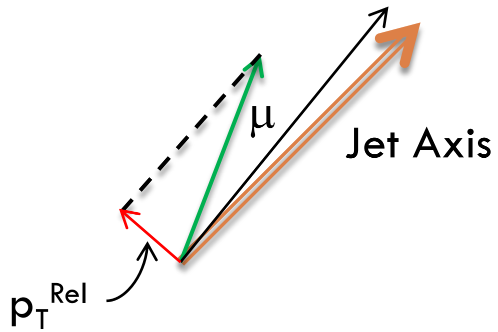
- The physics motivated taggers are best when you start
  - Direct feedback to tracking, PV people.
    - JLIP's resolution curves were excellent cross-checks of detector performance, for example.
  - Too hard to understand what happens in a NN during startup.
- How do you define a jet as a b-jet in MC?
  - To this day we don't totally understand why light quark jets in ttbar have a higher tag rate than Wjj light quark jets.
- When it is time to go to a multi-variate tagger use something more robust than a simple neural network.
  - Especially something resistant to noise!
- Simple tagger (CSIP) perform almost as well as the sophisticated taggers
  - It wasn't until the rest of the detector had really been tuned up!
  - Start simple…
- People have use combination of muon tagger and IP based tagger
  - Split data into orthogonal samples.

Calibration

Calibration: Efficiency

p_T^{Rel} Template Method

B mesons more massive than average light quark mesons…



Use MC to determine p_T^{Rel} distributions of muons for light, charm, and bottom jets.

QCD DiJet Events

Single Tag Method

$n = n_b + n_f$

$n_{tagged} = \varepsilon_b * n_b + \varepsilon_f * n_f$

Use p_T^{Rel} Fits to determine $\frac{n_b}{n}$ and $\frac{\varepsilon_b * n_b}{n_{tagged}}$

Double Tag Method

Similar, but start with a sample of tagged events

DØ dropped this as a tag measurement method. We did not understand how to determine the systematic errors on the MC templates.

Calibration: Efficiency

System 8 Method

8 equations, 8 unknowns (a.k.a. System D)

$$
\begin{aligned}
n &= n_b + n_l \\
p &= p_b + p_l \\
n^{CSIP} &= n_b \varepsilon_{btag}^{CSIP} + n_l \varepsilon_{non-btag}^{CSIP} \\
p^{CSIP} &= p_b \varepsilon_{btag}^{CSIP} + p_l \varepsilon_{non-btag}^{CSIP} \\
n^{SLT} &= n_b \varepsilon_{btag}^{SLT} + n_l \varepsilon_{non-btag}^{SLT} \\
p^{SLT} &= p_b \varepsilon_{btag}^{SLT} + p_l \varepsilon_{non-btag}^{SLT} \\
n^{both} &= n_b \varepsilon_{btag}^{CSIP} \varepsilon_{btag}^{SLT} + n_l \varepsilon_{non-btag}^{CSIP} \varepsilon_{non-btag}^{SLT} \\
p^{both} &= p_b \varepsilon_{btag}^{CSIP} \varepsilon_{btag}^{SLT} + p_l \varepsilon_{non-btag}^{CSIP} \varepsilon_{non-btag}^{SLT}.
\end{aligned}
$$

Two uncorrelated taggers on two samples of different b-content.

- Measure single tag and also double tag rate
- 8 unknowns, 8 equations

Tagger #1: Muon tagger.
Tagger #2: Tagger under study.

DØ uses this technique to this day to calculate the efficiency

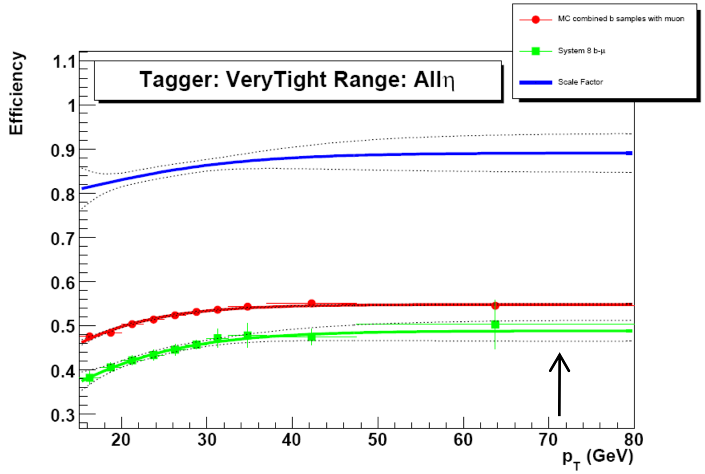
Monte Carlo does enter this calculation

- We have to determine what the tag rate is for a B jet w/out a muon (a ratio).
- Charm is determined by using the charm-to-bottom tagging ratio from MC and the data derived efficiency.

Calibration: Efficiency

Efficiency (and fake rate) is parameterized in p_T and η.

- Partition the data sets into bins of p_T or η.
- Fit to get shape in p_T or η.
- Combine fits to get 2D.



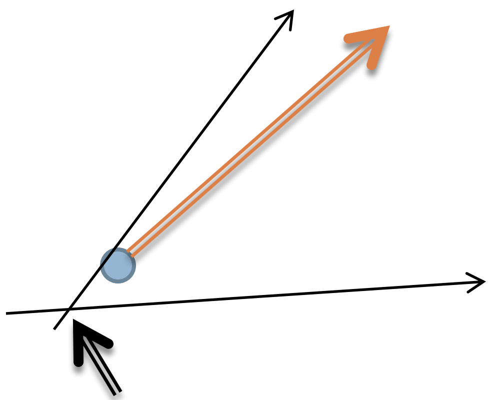
System 8 is statistics hungry.
Dijet sample runs out of statistics at high p_T's!
What do you do with jets there?

Be consistent with calibration and application of tagging!

Calibration: Fake Rate

Negative Tag Rate

- Tags behind the PV are due to mis-measured tracks
- Tags in front of PV due to mis-measured tracks will occur at the same rate



Use a large sample of QCD events

There are some tricks…

- A b-jet is slightly more likely to have a negative tag than a light quark jet
- Asymmetry caused by tag definition: what do you do if a jet has both a positive and negative tag?
- How do you define a negative tag for a tagger like JLIP?
  - Make sure definition isn't too asymmetric.

Use MC Scale Factors

Comments On Calibration

- Don't measure things twice: we use two different data samples to measure the fake rate
  - Give different fake rates.
  - Take difference as a systematic error.
- Took DØ much more time to calibrate the algorithms than it did to write them and test them on Monte Carlo.
- Triggers
  - At DØ we have soft lepton triggers that gather enough data
- Sample sizes: millions of events.
- Food Chain Consequences
  - Often b-tagging is the last thing to be certified!
  - b-tagging had to run on both our raw data format and root-tuple format.
  - Non trivial amount of infrastructure code to support this!
- Instantaneous luminosity balance your calibration samples!
  - Or understand and parameterize the trend!
- What Jet Energy Scale should be used?
  - Closure tests didn't work with JES w/muon.

Conclusions

- DØ's tagging experience
  - First 4 years of running with three competing algorithms.
  - All had similar performance – analysis's preference often made the decision.
  - Finally combined…
  - Calibration is a huge effort – shear # of events mean it can take more than a month start-to-end for someone who knows what they are doing.
- We learned a lot along the way
  - Many of our lessons are already in the proto-type LHC taggers.
  - And many of our people are active in LHC – which will hopefully make the time between collisions and a well understood tagger short!
- Do a better job at external documentation!

Backup Slides

(some stolen out-right…)

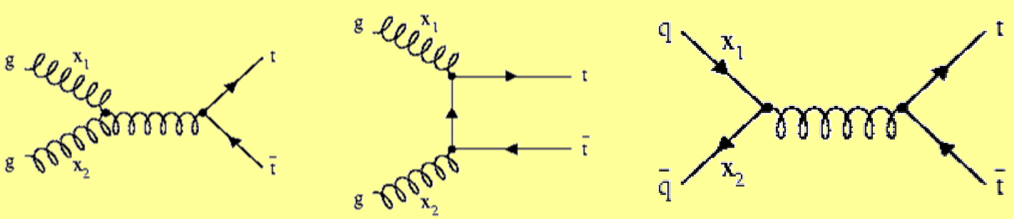
Top pair production at LHC

Low lumi = 10 fb^{-1}/y

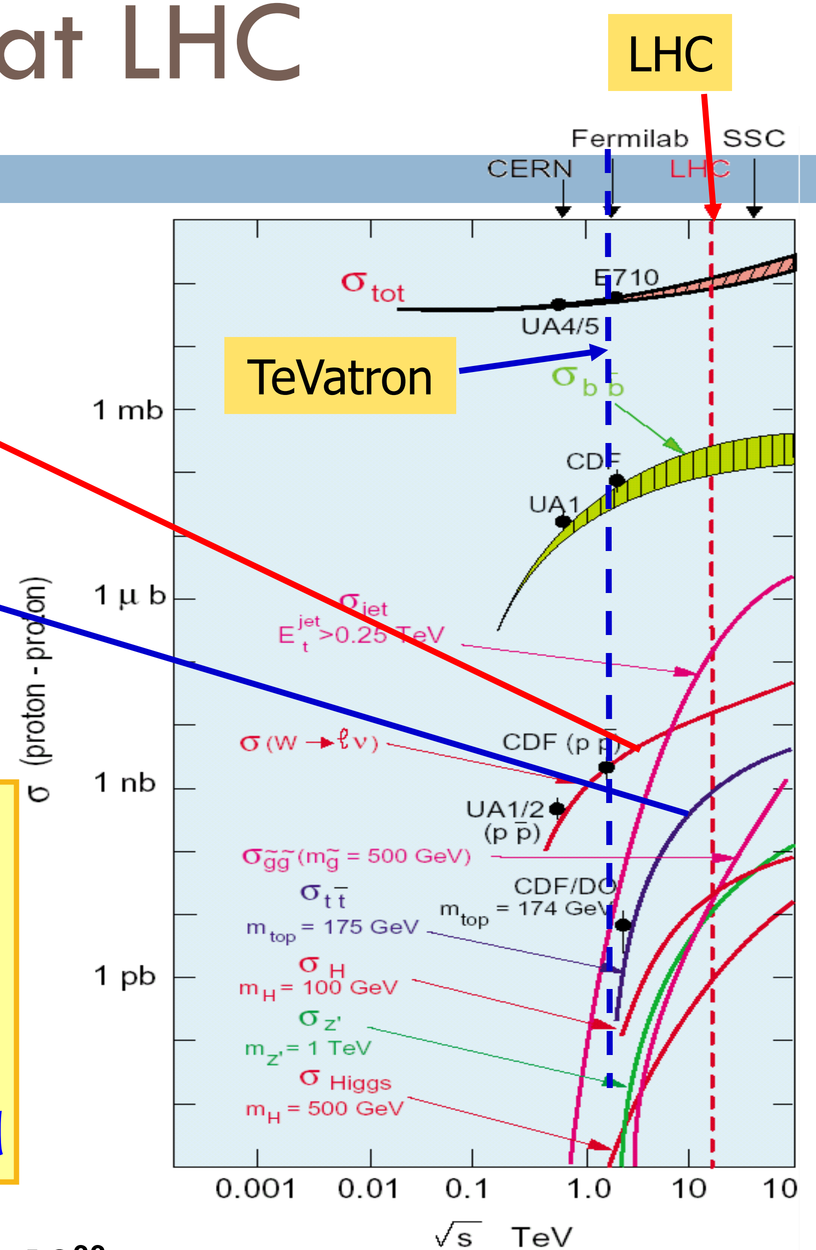
Process	σ(pb)	N/s	N/year	Total collected before start of LHC
$W \to l\nu$	3×10^4	30	10^8	10^4 LEP / 10^7 FNAL
$Z \to ee$	1.5×10^3	1.5	10^7	10^7 LEP
$t\bar{t}$	830	1	10^7	10^4 Tevatron
$b\bar{b}$	5×10^8	10^6	10^{13}	10^9 Belle/BaBar ?

→ LHC top factory !

$$\hat{s} = s x_1 x_2 \quad ; \quad x_1 x_2 \sim 10^{-3}$$



~90% gg (Opposite @ FNAL) ~10% qq



LHC start up in April 2007 @ L=10^{33}