B-Tagging Performance

Carl Gwilliam
(on behalf of the b-tagging group)

UNIVERSITY OF
LIVERPOOL

ATLAS Experiment

9th January 2008
ATLAS UK Physics Meeting, Durham

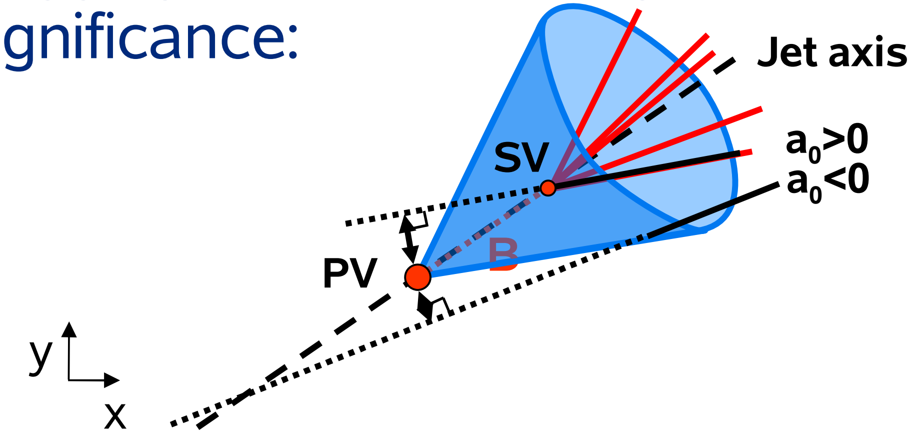
Introduction

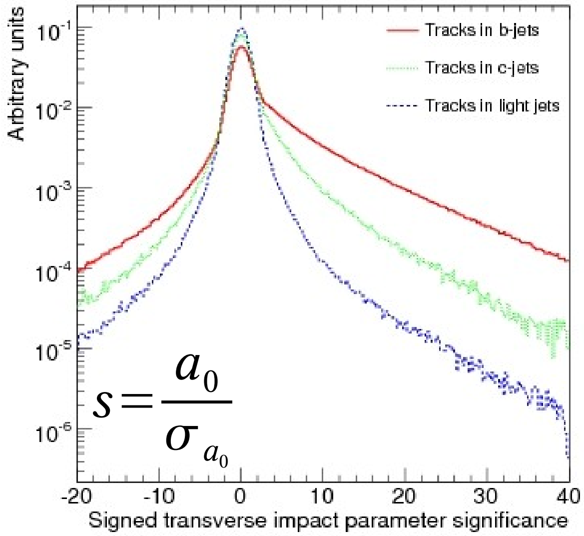
- Ability to tag jets from heavy flavours important for many physics processes
  - From precision measurements of top to Higgs & new physics searches
- Lifetime of b-hadrons (cτ ~470μm) => displaced vertex (~mm)
- Many taggers developed, based on:
  - Secondary vertices and impact parameters (2D & 3D)
  - Soft leptons (electron or muon)
- New taggers, such as JetFitter & Topological Vertex Finder arriving
- Key ingredients for b-tagging:
  - Tracking (crucially b-layer and pixels) : primary vertex and IP resolution
  - Jets finding: axis
  - Lepton ID for soft lepton taggers

} Need to understand these components in first data!

2D Impact Parameter Tagger

- Tagging based on the signed transverse IP significance:

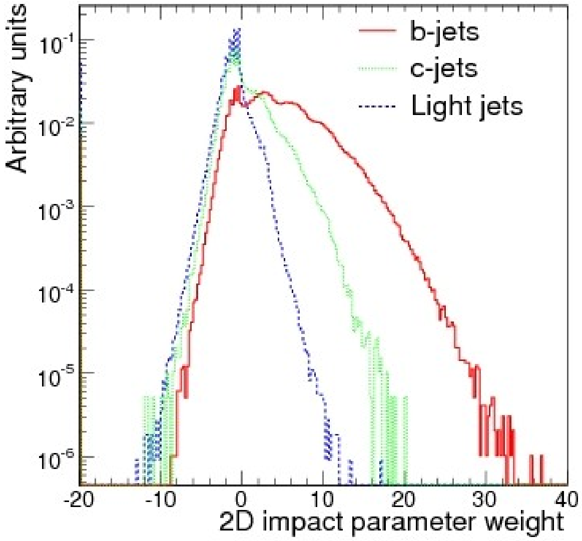




- Log-likelihood ratio used to construct a single weight from predefined reference histograms for b and light hypotheses.

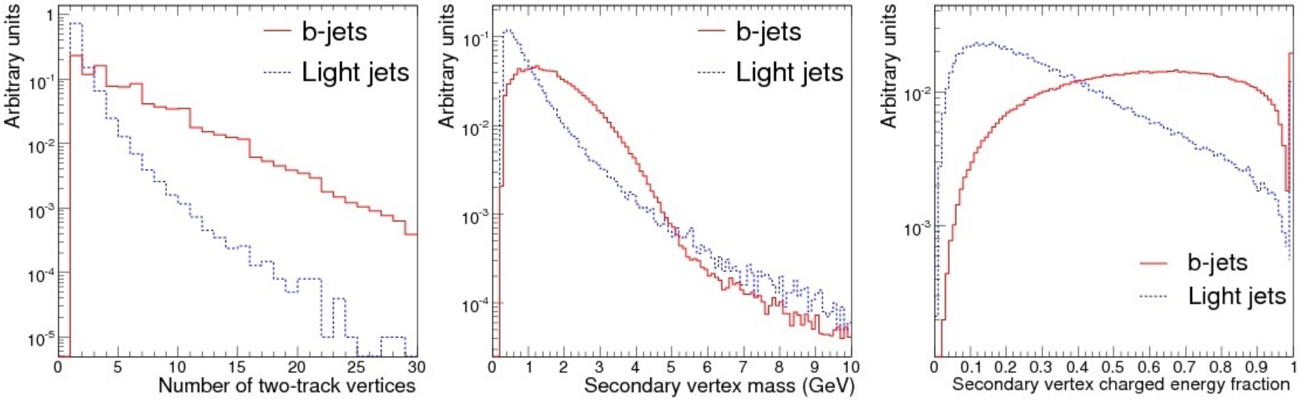
$$W_{jet}=\sum_{i=1}^{N_{tr}}\ln\frac{b(S_i)}{u(S_i)}$$

  - Powerful & easy to combine non-corr. variables
  - PDFs may prove difficult to determine in data

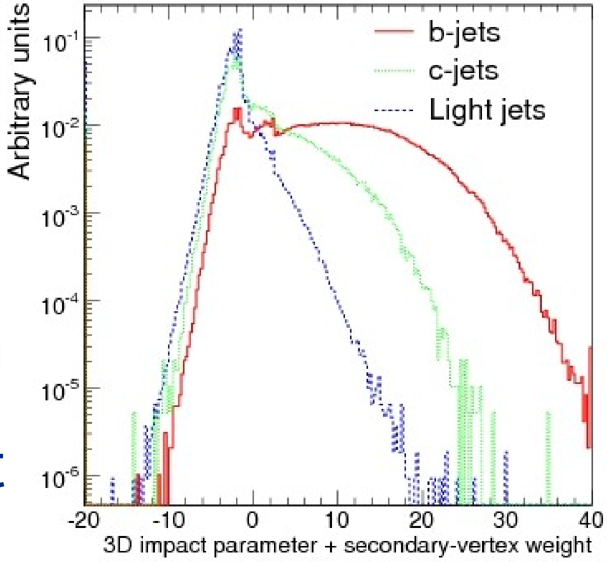


Secondary Vertex (SV) + IP3D Tagger

- The discrimination between b & light jets can be improved by reconstructing an inclusive secondary vertex.
- Discriminating variables which have low correlation with IP method:



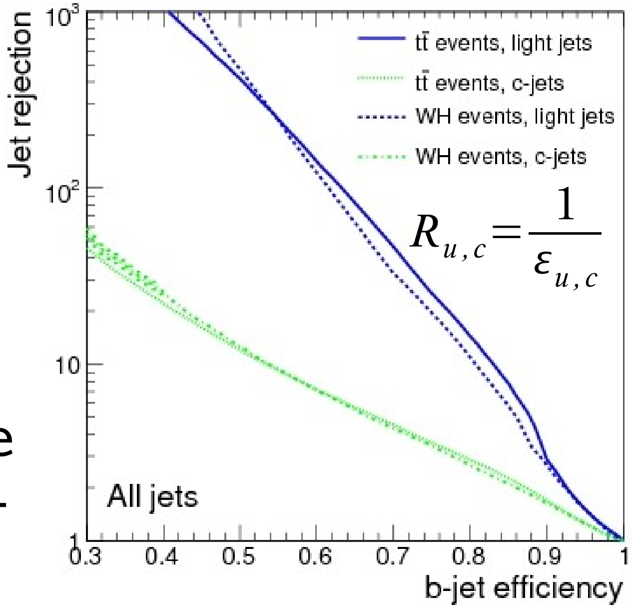
- Combined into single weight with 3D impact parameter tagger for max performance:

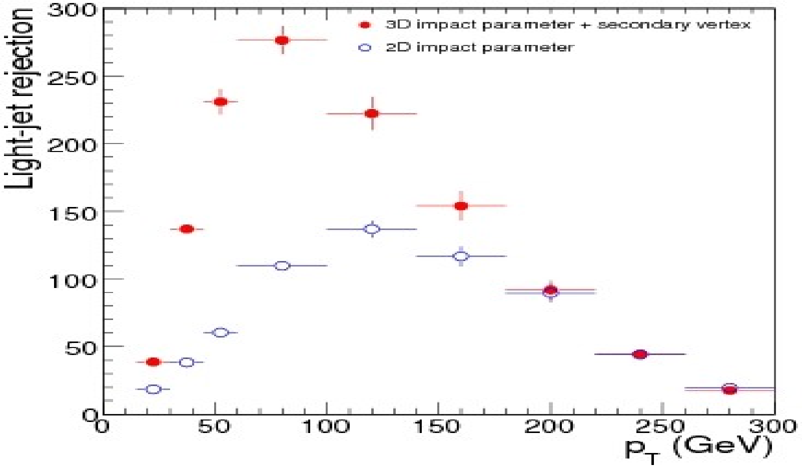


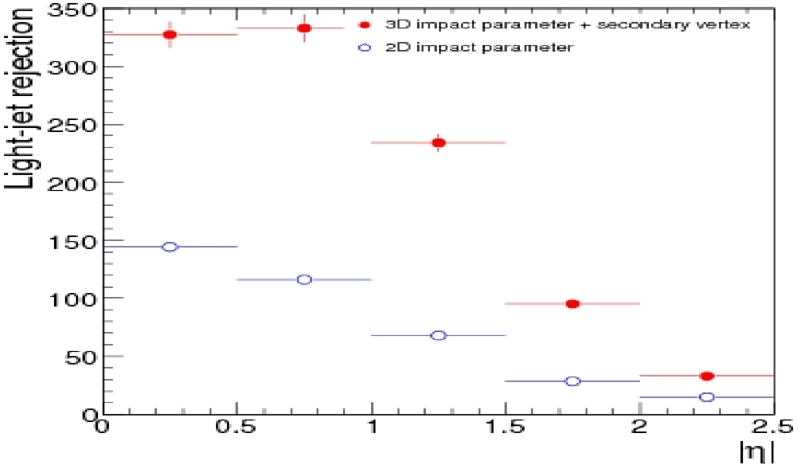
$$W_{jet} = W_{tracks} + W_{vertex} = \sum_{i=1}^{N_{track}} \log \frac{b(a_{0_i}, Z_i)}{u(a_{0_i}, Z_i)} + \log \frac{b(M,F,N)}{u(M,F,N)}$$

SV + IP3D Tagger Performance

- Default tagger => rejection ~100 for ε=60%
  - But requires knowledge of SV efficiency and several (multi-dim) PDFs. Early data?
- Performance strongly depends on P_t & η
  - High P_t: B/D can decay beyond the b-layer + pattern recognition difficult in v. dense jet core
  - Low P_t/ High η: increased multiple scattering + material interactions (GEANT description)





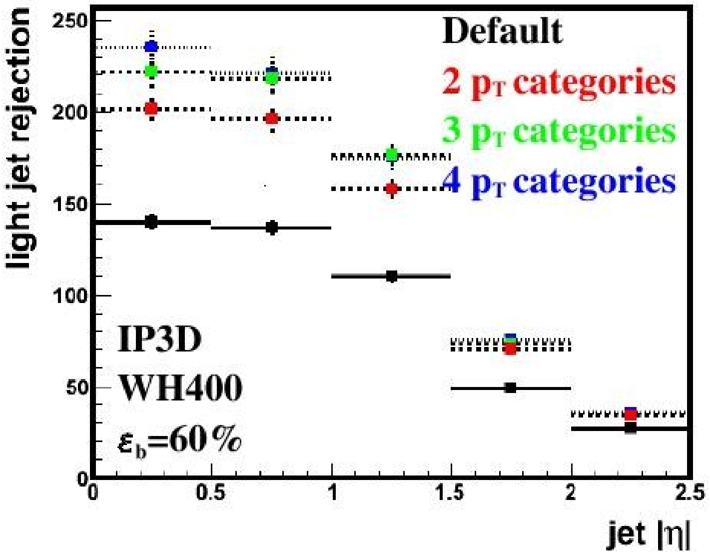


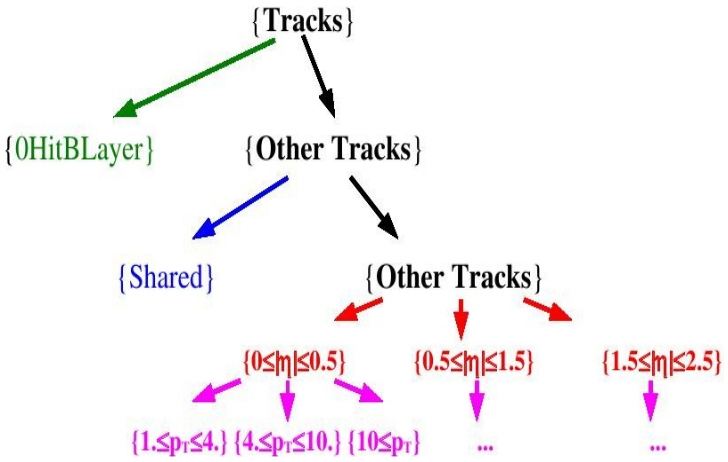
Performance: Track Categories

- Introduce different track categories with dedicated PDFs:

$$w_{\text{jet}} = \sum_{j=1}^{N_{\text{cat}}} \sum_{i=1}^{N_{trk}} \ln \frac{b_j(S_i)}{u_j(s_i)}$$

- Possible categories are:
  - Good + shared hit tracks (already used)
  - No b-layer hit (rejected by default)
  - P_t and η
- Introducing P_t categories can give ≈60% increase in rejection
  - 3 bins seems optimal, 4th has little effect
- Can define 11 track categories, while keeping adequate stats in each
  - Up to 70% increase in rejection
  - Effect mostly due to P_t categories though

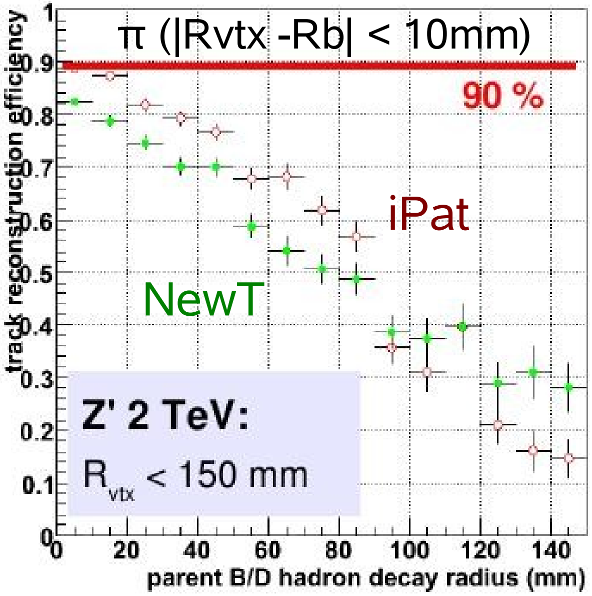




Marcel Vos, Remi Zaidin et al.

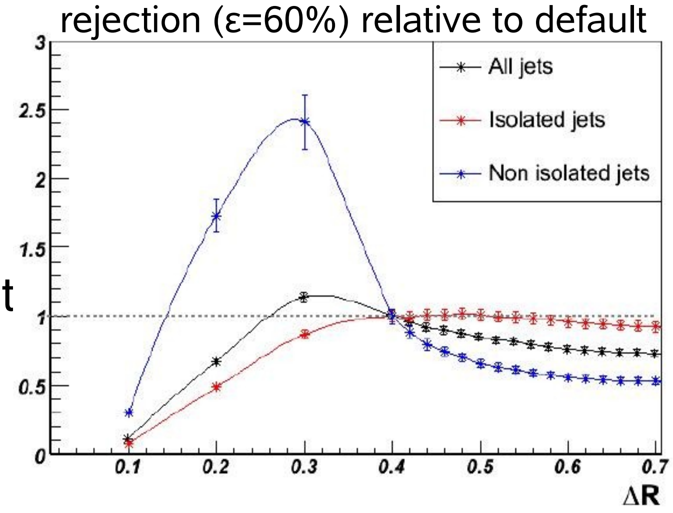
Performance: High P_t B-tagging

- Important for channels like SUSY,little H,Z'...
- Tracking Challenges:
  - Large average decay length:
    - Miss B-layer (L=γcτ > 50mm for 600 GeV B-had)
    - Decay just in front of b-layer => no time to separate
  - High mult + more collimated jet => dense core
    - Difficulties in pattern recognition, especially challenging for tracks from displaced vertices



- Since jets are more collimated can try to optimise track-jet association of ΔR = 0.4
  - ΔR = 0.3 optimal for **non**-isolated jets
    - Degradation due to taking tracks from nearby jet
    - Optimal results depend on event topology
  - Similar performance with variable cut:

$$\Delta R = a_0 + \exp(a_1 + a_2 p_t)$$



Misalignment

- B-tagging performance is v. sensitive to the ID performance and hence misalignments.
- Samples:
  - Initially, use 2 sets of random pixel-only residual misalignments
    - Initial: misalignment after 100pb^{-1}
    - Final: misalignment after several years

    } Syst shifts for layer/disk/overall and random shifts for modules
  - Eventually, use constants defined from actual alignment algorithms
    - Will include the impact of systematics in the alignment procedure
- Error scaling:
  - Iteratively correct hit measurement errors to allow proper track finding:

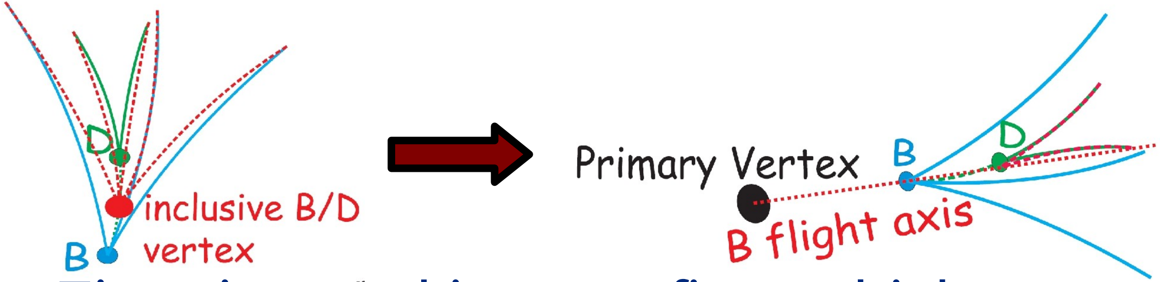
    $$\sigma'^2 = a^2 \sigma^2 + c^2$$

    } a = effects correlated to track reco performance
    c = effects correlated to misalignments
  - Have had several technical difficulties but hopefully now on the way.
- Initial results suggest the effect on b-tagging could be large and only partially recovered by error scaling. Work still in progress ...

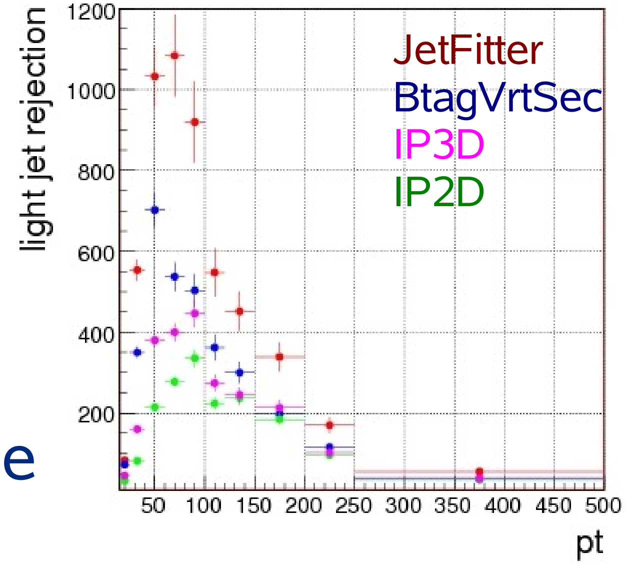
Jet Fitter (New)

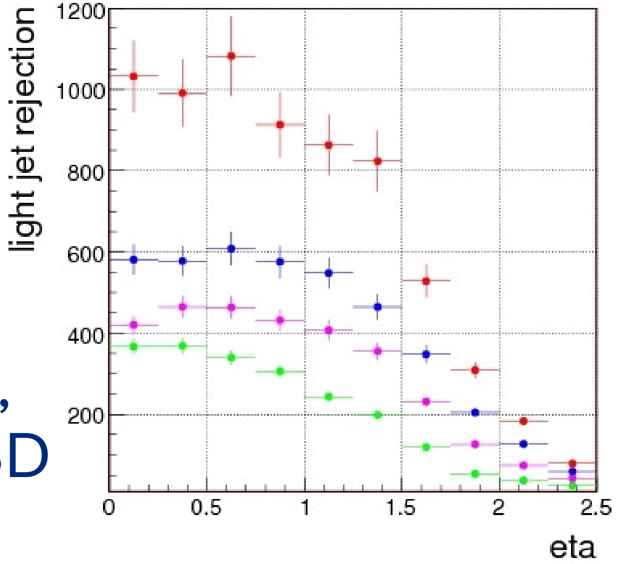
Christian Weiser
Giacinto Piacquadio

- Don't really have single SV but a B/D cascade.



- JetFitter is a multi-vertex fitter which uses this topological info to constrain tracks to the B-Hadron flight axis.
  - Can reconstruct incomplete topologies
    - In principle even 1-prong B & 1-prong D (≈10%)
  - Cascade topology increases discrimination against light jets
- Split into several categories (#vertices, #trk on flight axis) and create PDFs based on M, E/Ejet & Ntrk. Additionally combine with IP3D
  - ≈50% improvement in rejection for ε=60%





Towards early data …

Jet Probability Tagger

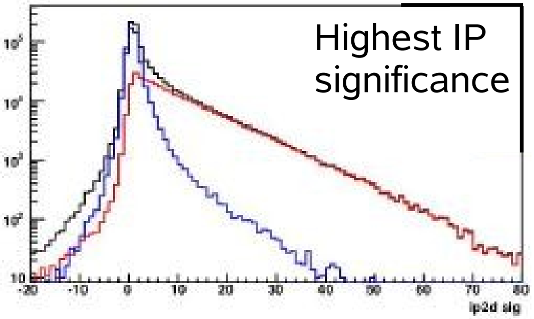
- Jet probability tagger à la Aleph/LEP:
  - Compatibility of the tracks with the primary vertex
  - Fit negative side of signed IP with a resolution function (gaus+2exp)
  - Apply to positive signed impact parameters => track probability
  - Combine into a single jet probability:

$$P_{jet} = \Pi \cdot \sum_{j=0}^{N_{tr-1}} \frac{(-\ln \Pi)^j}{j!}, \quad \Pi = \prod_{i=1}^{N_{tr}} P_{tr_i}$$

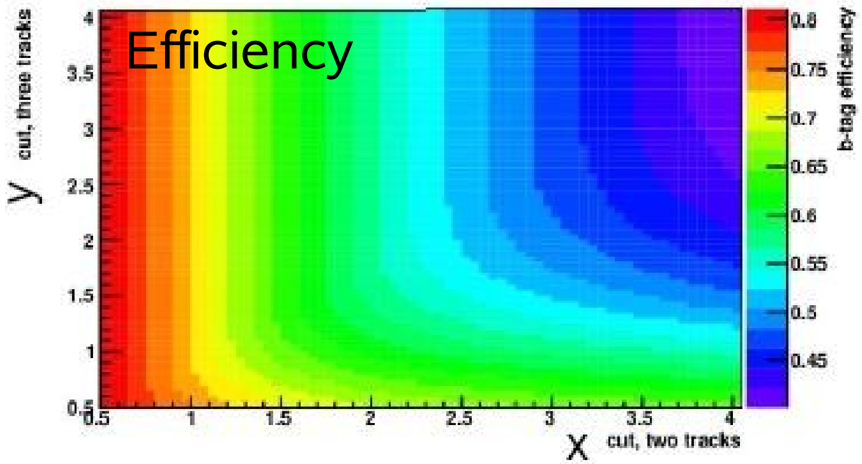
- Expect poorer performance than IP2D BUT will be easier to understand and calibrate in early data (no PDFs to understand).
- Available now but performance still needs to be evaluated

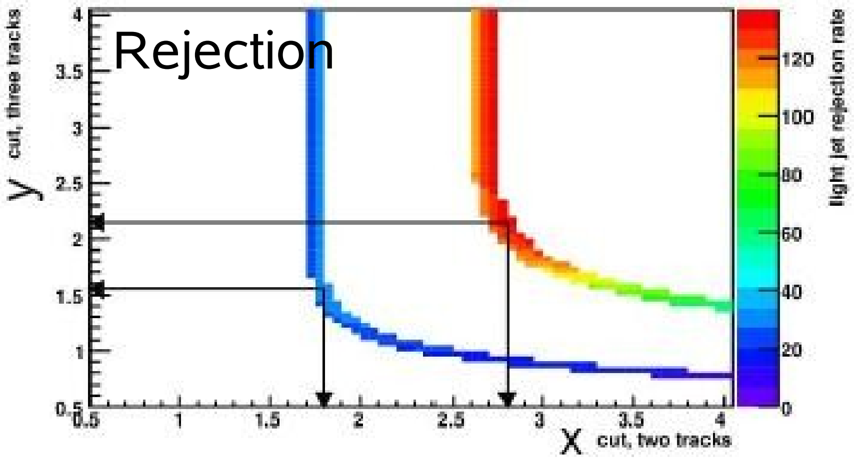
CSIP Tagger

- Simply count IP significances (used in D0 at start of data taking)
  - jet tagged if 2 tracks with IP significance above X or three above Y
- Tune cuts for ε = 50/60% and look at rejection:
  - x = 1.8, y = 1.55 for ε = 60% (c.f. x=2, y=3 at D0)
  - Corresponding light-jet rejection is ~60


Highest IP significance

- Performance is about ½ that of IP2D BUT easier to understand in early data (no PDFs).


Efficiency


Rejection

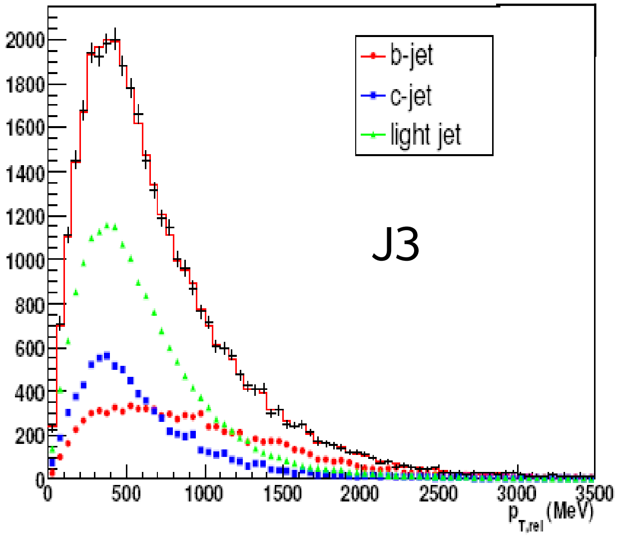
F. Rizatdinova, G. Watts et al.

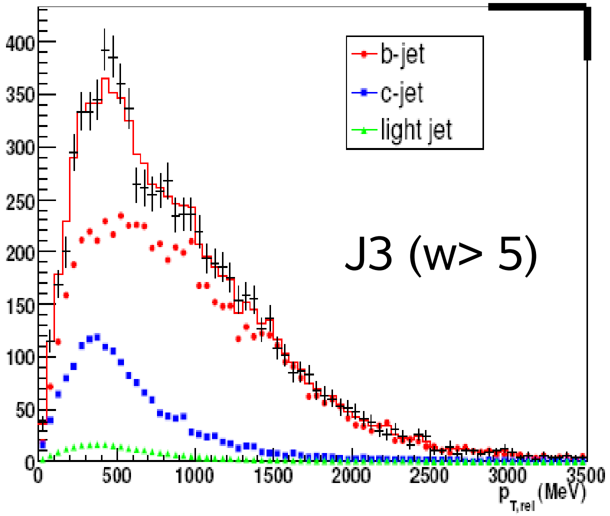
Efficiency from Data: Dijets

- Need to measure b-tag efficiency from data
- One method is to use dijets à la Tevatron:
  - Abundant statistics (particularly at low P_t)
  - Errors/biases well understood from Tevatron
  - Calibration sample dissimilar to phys sample
- Select enriched b-jet sample by requiring soft-μ (maybe e) within jet (ΔR<0.4).
- 1. Relative P_t
  - μ's from light,c & b have diff P_t wrt jet axis
  - Use MC templates to fit fraction of b-jets in sample before and after tagging:

$$\varepsilon_{b\to\mu} = \frac{N_{\mu}^{tag} F_{b\to\mu}^{tag}}{N_{\mu} F_{b\to\mu}}$$

  - ε(SV+IP3D) = 0.68 ± 0.04 c.f. 0.70 from truth
  - Rely on MC for templates & assume $\varepsilon_{b\to\mu} = \varepsilon_b$





F. Rizatdinova, G. Watts et al.

Efficiency from Data: Dijets

- 2. System 8:
  - Data driven approach developed at DØ
  - Select two samples with different b content
    - μ-in-jet sample (n)
    - Same sample with tagged away-side jet (p)
  - Tag μ-jet with 2 **independent** taggers
    - Lifetime-based tagger (LT)
    - Soft muon tagger (SMT)
    - Both (DT)
  - Solve system of 8 equations for ε(LT)
    - ε(SV+IP3D) = 0.67 ± 0.06 c.f. 0.68 from truth
    - Good stability wrt cut on soft-μ tagger
  - Some issues at high P_t due to poorer b/c separation of SMT => need higher SMT cut to ensure samples distinct => less stats
- Need to study triggers (MU4/MU6+JT10)

$$n = n_b + n_{l,c}$$
$$p = p_b + p_{l,c}$$

untagged sample

$$n^{LT} = \varepsilon_b^{LT} n_b + \varepsilon_{l,c}^{LT} n_{l,c}$$
$$p^{LT} = \varepsilon_b^{LT} p_b + \varepsilon_{l,c}^{LT} p_{l,c}$$

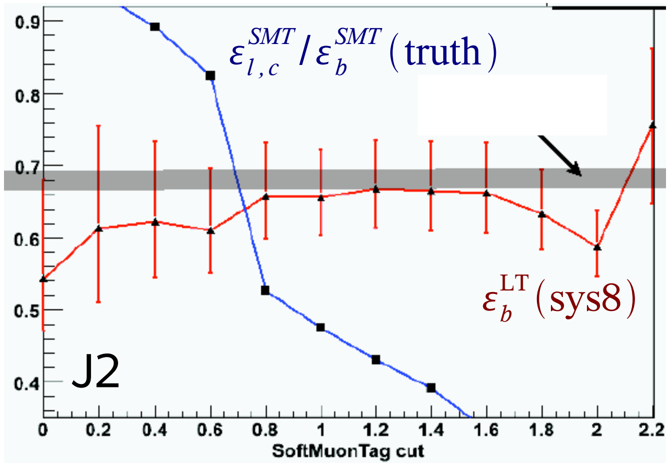
LT sample

$$n^{SMT} = \varepsilon_b^{SMT} n_b + \varepsilon_{l,c}^{SMT} n_{l,c}$$
$$p^{SMT} = \varepsilon_b^{SMT} p_b + \varepsilon_{l,c}^{SMT} p_{l,c}$$

SMT sample

$$n^{DT} = \varepsilon_b^{LT} \varepsilon_b^{SMT} n_b + \varepsilon_{l,c}^{LT} \varepsilon_{l,c}^{SMT} n_{l,c}$$
$$p^{DT} = \varepsilon_b^{LT} \varepsilon_b^{SMT} p_b + \varepsilon_{l,c}^{LT} \varepsilon_{l,c}^{SMT} p_{l,c}$$

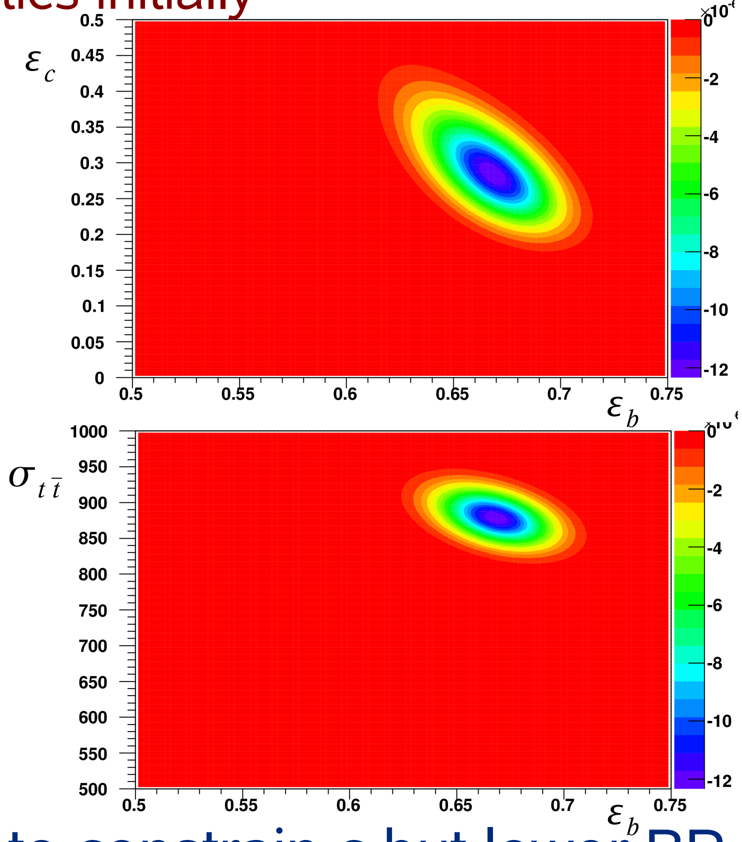
DT



Efficiency From Data: Top Events

- Top cross section 100x larger @ LHC => Can use to measure ε(b)
  - Well known HF content + calib sample closer to phys (inclusive b-decays)
  - More complex event reco + limited statistics initially
- 1. Counting Method (Henri Bachacou):
  - Select “lepton + jet” channel with:
    - $N_{jets} \geq 4$ with $E_t > 30\,GeV$
  - Count events with 1, 2, 3 tagged jets
    - To 1st order: $N_{1-tag} = 2N_{event}\,\varepsilon_b(1-\varepsilon_b)$

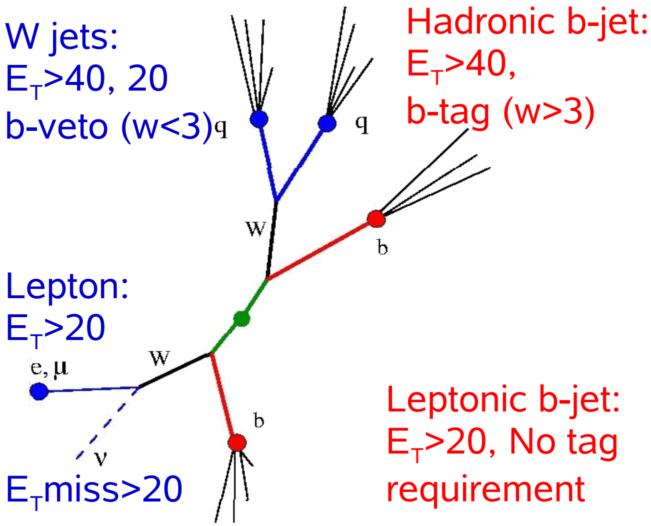
      $N_{2-tag} = N_{event}\,\varepsilon_b^2$
    - In fact need to take into account light (elsewhere) & c jets (from 3 tag events)
  - Perform likelihood fit to ε(b), ε(c) & σ(tt) => ≈ 3-4%(stat) + 2-3%(syst) in 100 pb^{-1}



- Dilepton channel => less bkg & no need to constrain c but lower BR

Efficiency From Data: Top Events

- 2. Kinematics (Richard Hawkings):
  - Try to select approx pure b-jet sample
  - "Tag" hadronic side using b-tagging:
    - Select 1 b-jet (t) and 2 non b-jets (W)
    - Require $M_{jj} \sim M_W$ and $M_{bjj} \sim M_t$
  - "Probe" leptonic side b-jet (unbiased):
    - Study b-tag weight for events with $M_{l\nu j} \sim M_t$
    - Estimate bkg by fitting $M_{l\nu j}$ in M_{bjj} sideband with b-jet veto on leptonic side
  - Measured $\varepsilon(b)$ compares well with truth & get ≈ 4% (abs) statistical error for 100 pb^{-1}.
  - Data driven => should be insensitive to MC
  - But fit struggles to converge for 100 pb^{-1} => will probably need to wait for more data.



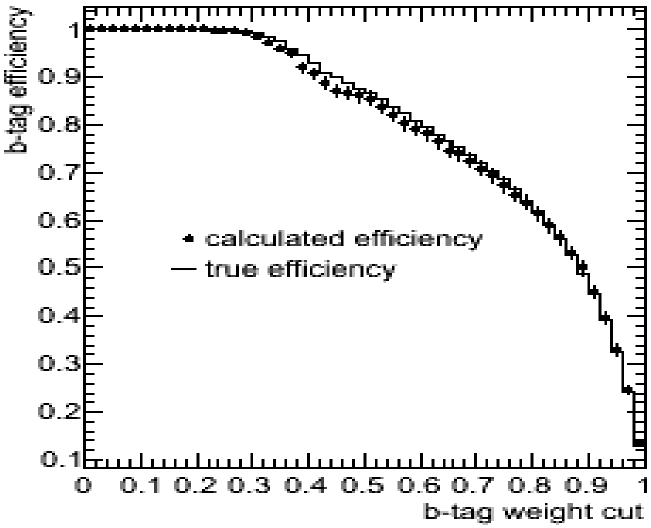
W jets:
$E_T>40$, 20
b-veto (w<3)

Hadronic b-jet:
$E_T>40$,
b-tag (w>3)

Lepton:
$E_T>20$

Leptonic b-jet:
$E_T>20$, No tag
requirement

$E_T$miss>20



- Alternatively, perform kinematic fit on lepton, EtMiss & 4 leading jets

M. Saleem, A. Khanov et al.

Mistags from Data

- Mis-reconstruction/resolution effects can cause light-jets to be mistagged as b's.
- The mistag rate can be extracted from dijet data using the negative tagging efficiency.
  - The tagger weights are recalculated for only:
    - Tracks with -ve impact parameters
    - Vertices with -ve decay length significance

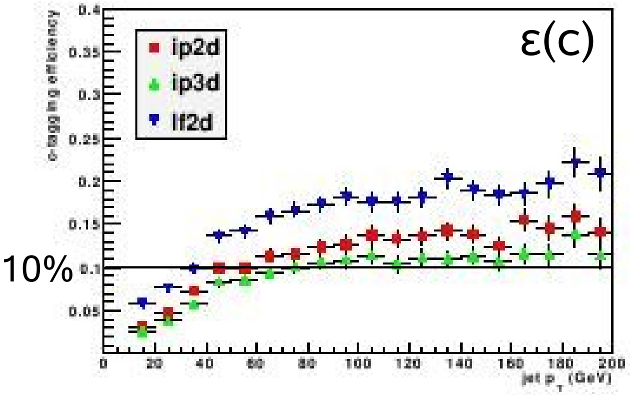
    and the usual weight cut applied to extract $\varepsilon_{-}^{\text{data}}$
- This is then corrected to obtain mistag eff

$$\varepsilon_l(P_t,\eta)=\varepsilon_{-}^{\text{data}} K_{\text{HF}} K_{\text{LL}}$$

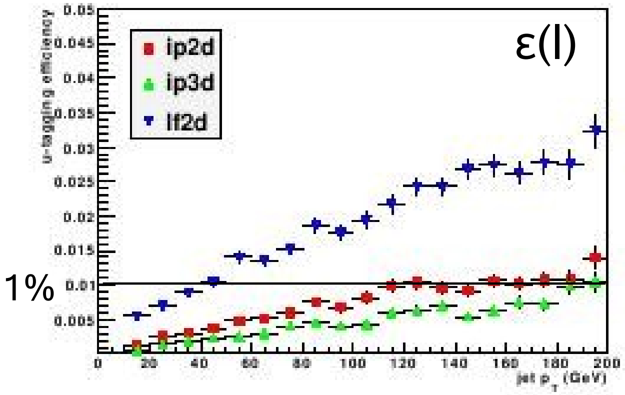
  - K_{HF} = HF contamination to -ve tags (eg $g\rightarrow b\bar{b}$)
  - K_{LL} = Long-live particles (Λ, Ks) in light jets

} Determined from MC: $K_{\text{HF}} K_{\text{LL}} \sim 1$

- Need to understand how to combine IP & SV mistag rates together.



ε(b) = 60% at Pt = ∞



Summary

- Many high performance tagging algorithms have been developed
  - SV1+IP3D, JetFitter, ...
- Performance in low/high P_t and high η regions being studied
  - Track categories, pattern recognition within dense jets, ...
- Understanding tracking and misalignment effects will be crucial
  - misalignment in progress
  - many other tracking effects to think about (ganged pixels, outliers, ...)
- Need to evaluate simple taggers for use in first data NOW
  - JetProb, CSIP, ...
- Will need to measure the REAL performance (eff/mistag) in data
  - Dijets initially, followed by top events

Backup Slides …

UK Involvement

- Liverpool (CG, Barry King, Andrew Mehta)
  - Effects of uncertainties in MC modelling on b-tagging efficiency
    - Fragmentation functions (Peterson, Lund, Lund-Bowler with various params)
    - Decay multiplicity, Lifetimes, Branching fractions

ATL-COM-PHYS-2007-062

=> Main effect due to multiplicity and, to lesser extent, fragmentation:

$$\varepsilon_b = 0.72 \pm 0.002(\mathrm{stat}) \pm 0.021(\mathrm{syst}) \text{ for } w > 3 \ (t\bar{t} \text{ sample})$$

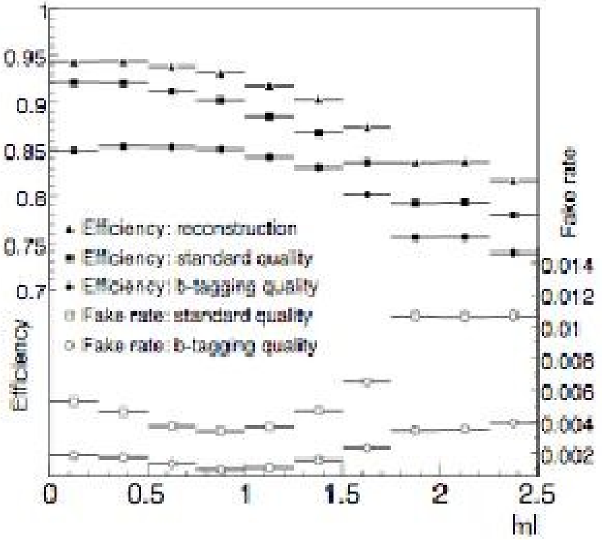
  - Code up Delphi-based IP & SV taggers as independent cross check.
    - Comparing performance with standard taggers => Aim for ATLAS note
- Glasgow (Stan Thompson)
  - B-tagging performance using different jet algorithms
    - Cone (0.4,0.7), Kt (0.3, 0.4, 0.6), MidPoint (0.4, 0.7)

=> Not much variation observed over the various algortihms

- Sheffield (Kirill Prokofiev)
  - Primary vertex finders
  - Vertexing Interface

Other Tracking Issues

- Other issues that need investigation are:
  - Resolution of ganged pixel ambiguities
    - Exploiting difference in expected cluster size.
  - Refitting of outliers in b-layer with scaled errors
    - E.g. try using cluster size as error
  - Tuning of shared hits
    - Track is currently accepted if doesn't share >2 clusters with a previous track
    - Instead of removing shared hits, can try error scaling or reclustering
  - Splitting of large clusters
    - Large clusters can be caused by merging clusters from 2 tracks
    - Try splitting them or scaling the errors
  - Special tracking for dense jets
    - Normal tracking fails to resolve double-track resolution effects
    - Try modern pattern techniques like MultiTrackFitter



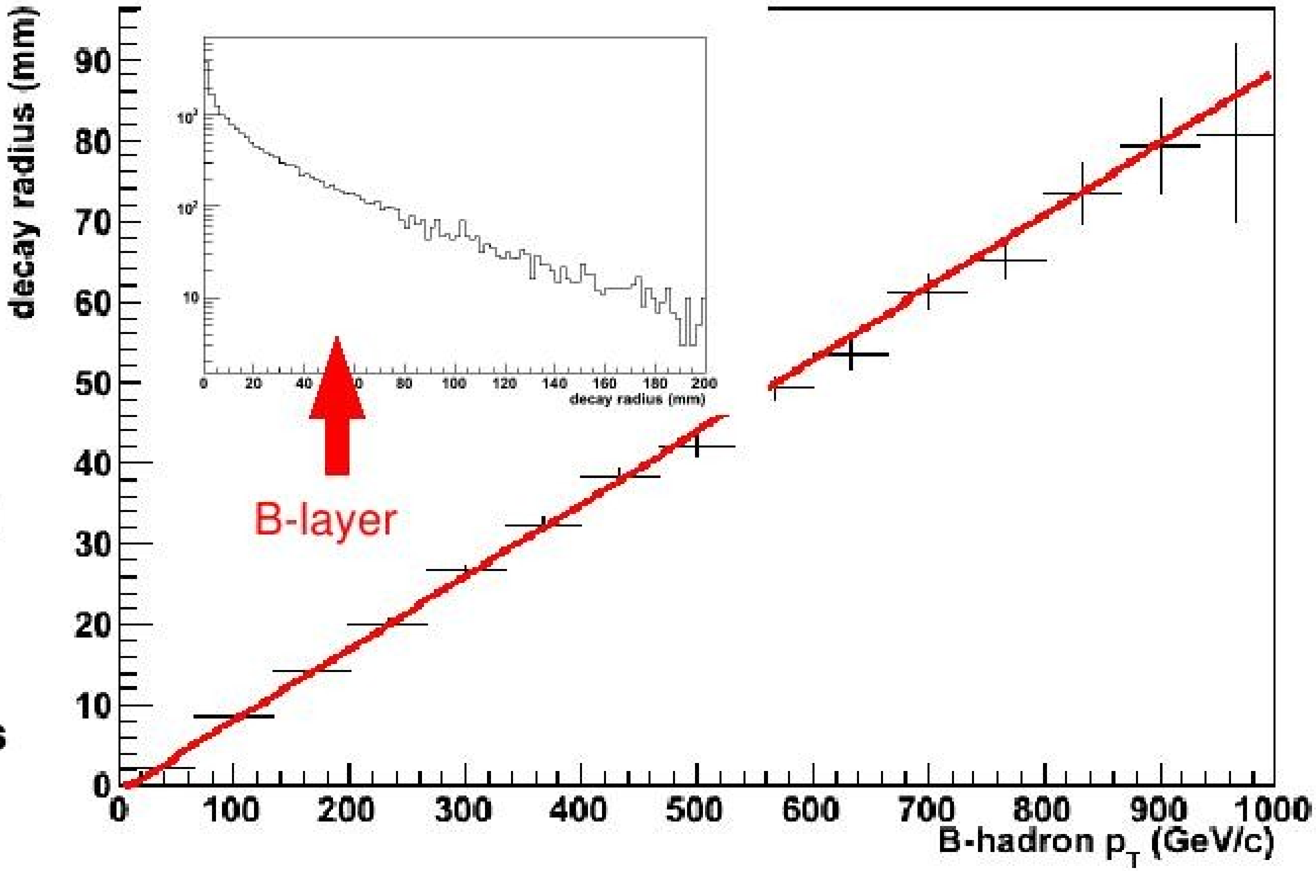
High Pt B-tagging: Decay Radius

$L = c\,\tau\,\gamma$

Average decay radius of B hadrons versus B-hadron transverse momentum

Insert plot: decay radius distribution for B-hadrons in Z'->bb events ($m_{Z'}$ = 2 TeV)

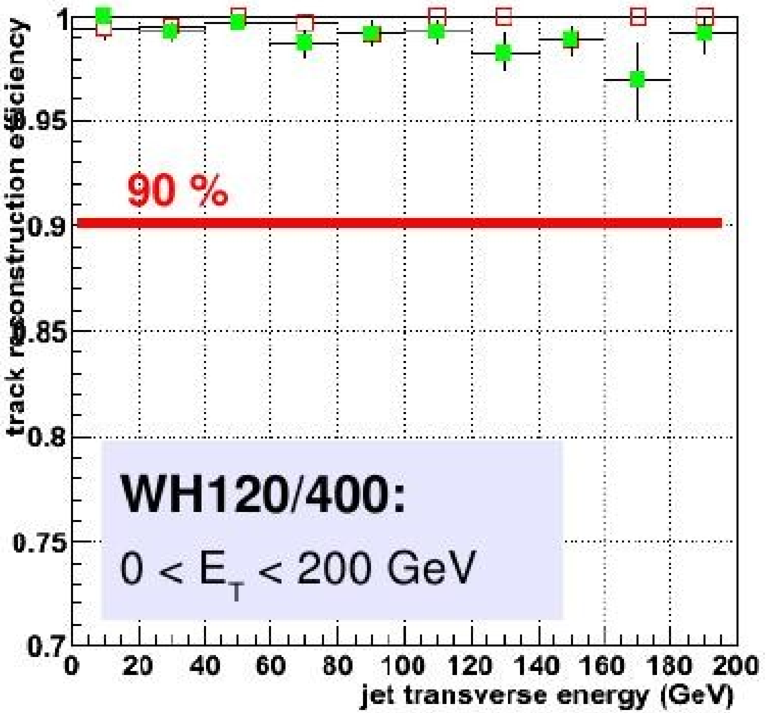
L no longer << B-layer radius

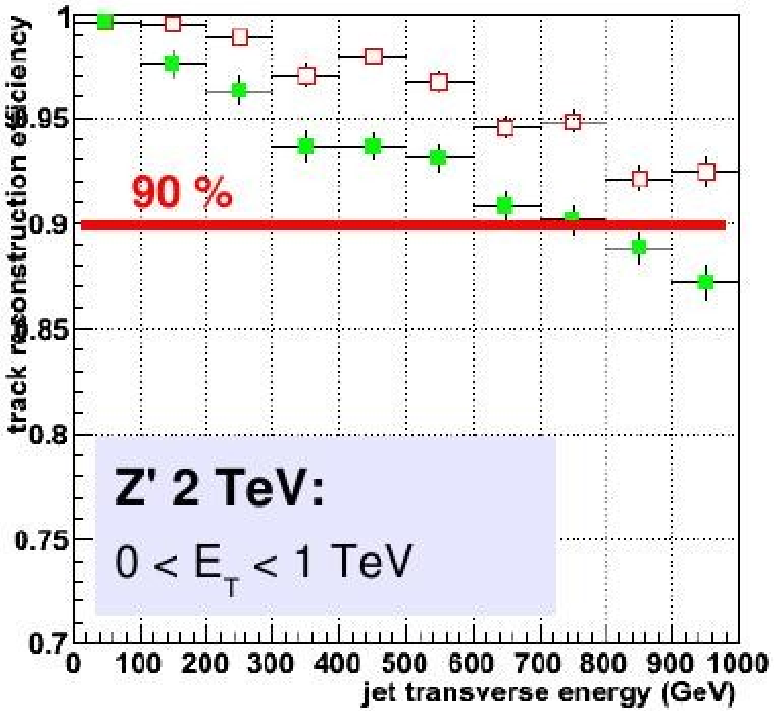


Marcel Vos

High Pt B-tagging: Prompt Tracks

$\pi^{\pm}$ with no decay, and R_{vtx} < 10 mm





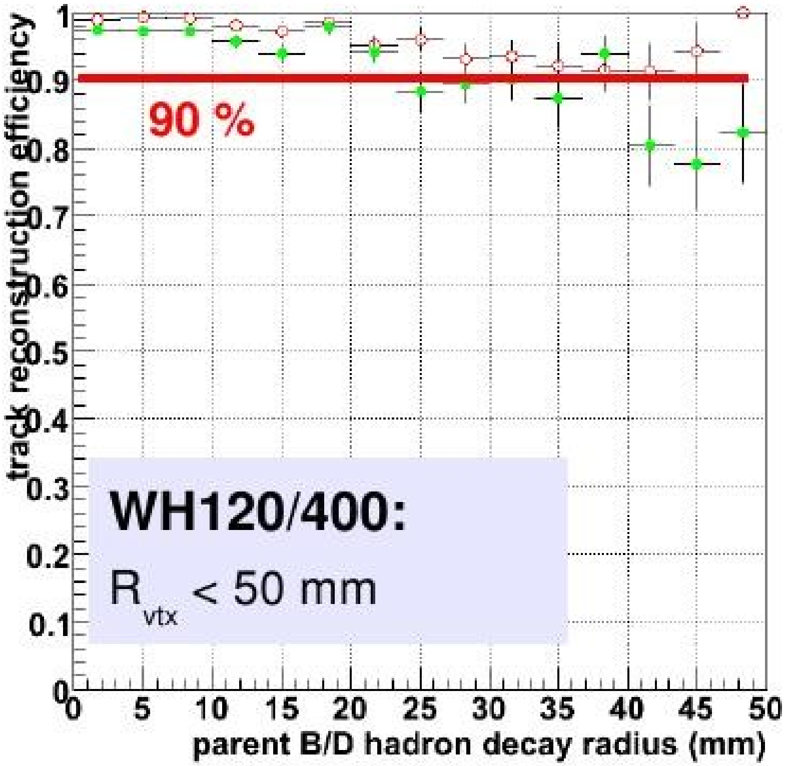
Reconstruction efficiency for “good tracks” from InvertedTruthMap (A. Gaponenko)

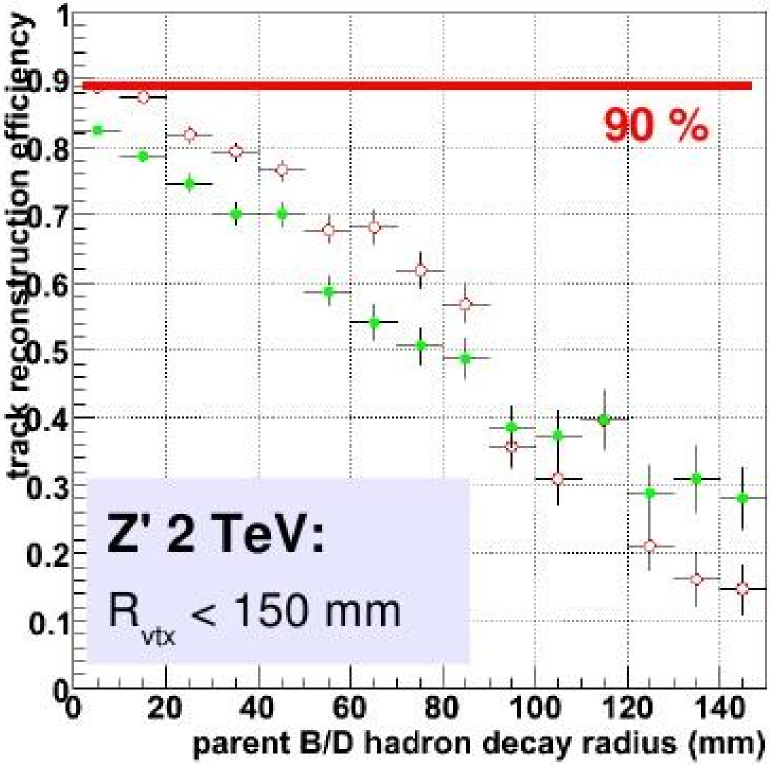
Efficiency >90% inside highest p_T jets!

IpatRec (red) performs slightly better than New Tracking (green)

High Pt B-tagging: Displaced Vertices

$\pi^{\pm}$ with no decay, $| R_{vtx} - R_{B\text{-}decay} | < 10$ mm



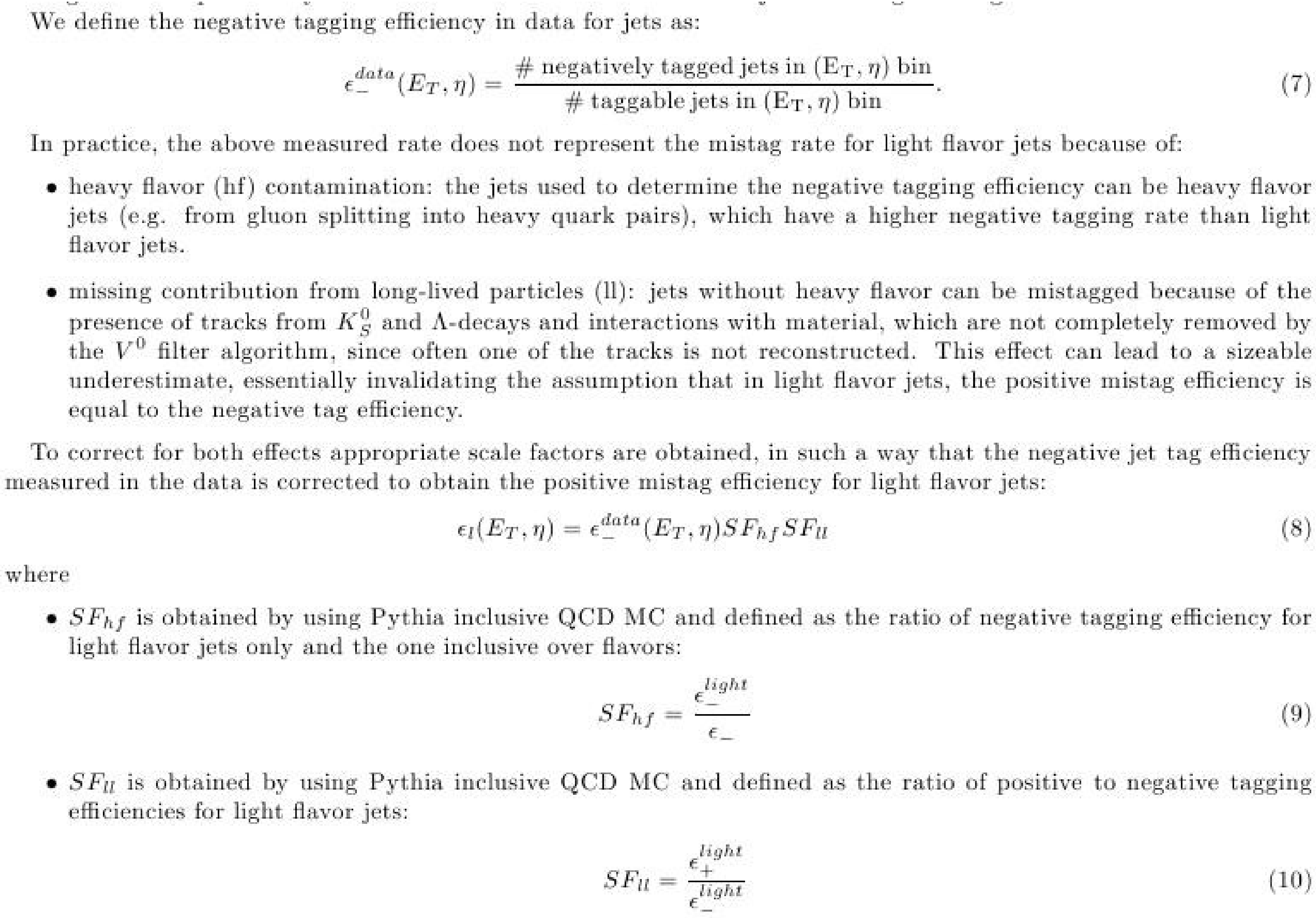


Tracks from displaced vertices present a real challenge!

IpatRec (red) performs slightly better than New Tracking (green)

F. Rizatdinova, G. Watts et. al.

Mistags From Data: Calculating



We define the negative tagging efficiency in data for jets as:

$$\epsilon_{-}^{data}(E_T,\eta) = \frac{\#\ \text{negatively tagged jets in } (\mathrm{E_T},\eta)\ \text{bin}}{\#\ \text{taggable jets in } (\mathrm{E_T},\eta)\ \text{bin}}. \qquad (7)$$

In practice, the above measured rate does not represent the mistag rate for light flavor jets because of:

- heavy flavor (hf) contamination: the jets used to determine the negative tagging efficiency can be heavy flavor jets (e.g. from gluon splitting into heavy quark pairs), which have a higher negative tagging rate than light flavor jets.
- missing contribution from long-lived particles (ll): jets without heavy flavor can be mistagged because of the presence of tracks from K_S^0 and Λ-decays and interactions with material, which are not completely removed by the V^0 filter algorithm, since often one of the tracks is not reconstructed. This effect can lead to a sizeable underestimate, essentially invalidating the assumption that in light flavor jets, the positive mistag efficiency is equal to the negative tag efficiency.

To correct for both effects appropriate scale factors are obtained, in such a way that the negative jet tag efficiency measured in the data is corrected to obtain the positive mistag efficiency for light flavor jets:

$$\epsilon_l(E_T,\eta) = \epsilon_{-}^{data}(E_T,\eta) SF_{hf} SF_{ll} \qquad (8)$$

where

- SF_{hf} is obtained by using Pythia inclusive QCD MC and defined as the ratio of negative tagging efficiency for light flavor jets only and the one inclusive over flavors:

$$SF_{hf} = \frac{\epsilon_{-}^{light}}{\epsilon_{-}} \qquad (9)$$

- SF_{ll} is obtained by using Pythia inclusive QCD MC and defined as the ratio of positive to negative tagging efficiencies for light flavor jets:

$$SF_{ll} = \frac{\epsilon_{+}^{light}}{\epsilon_{-}^{light}} \qquad (10)$$