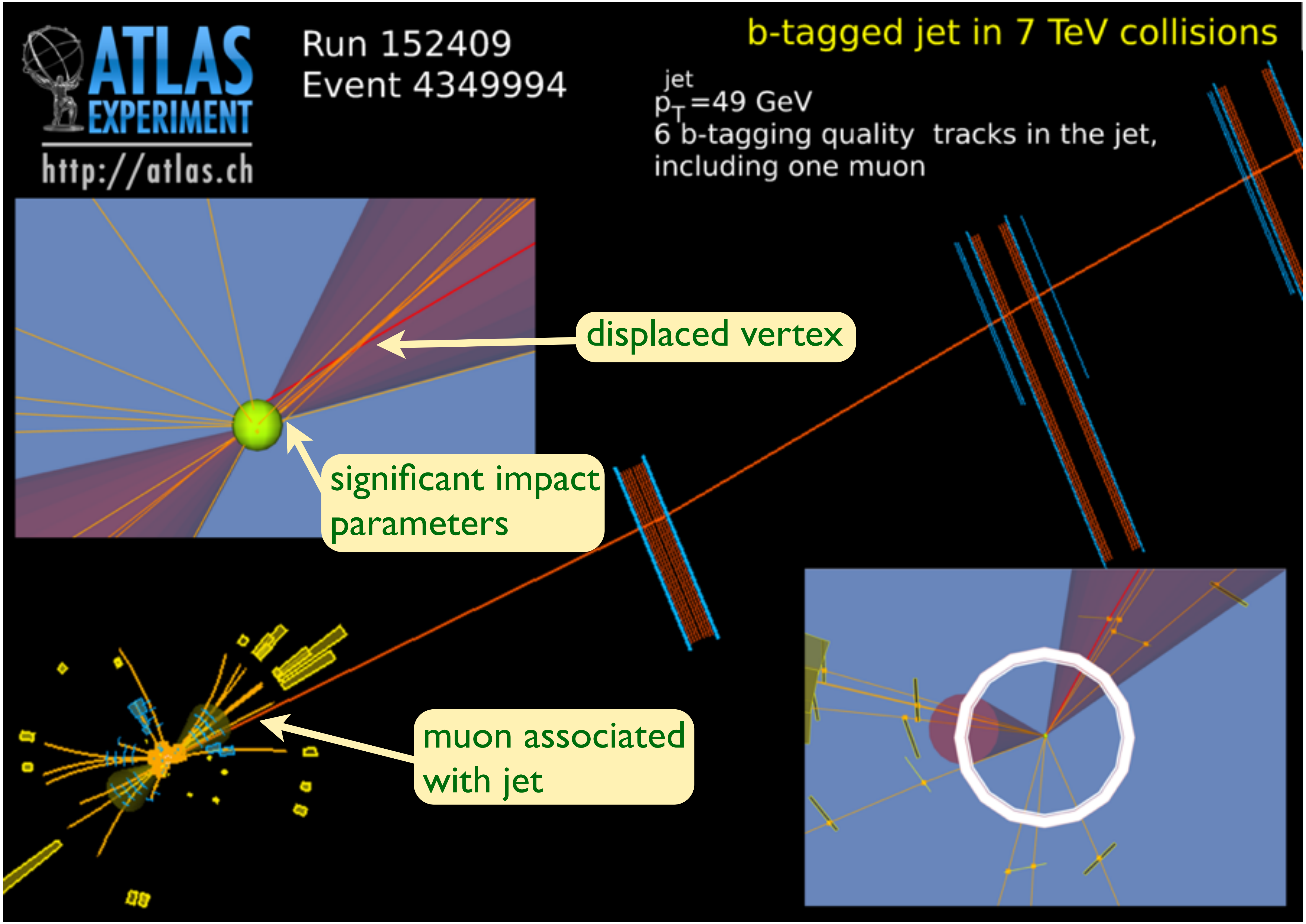
b-Jet tagging in proton-(anti)proton collisions

Frank Filthaut
Radboud Universiteit Nijmegen and Nikhef

Contents:

- tagging algorithms
- tagging performance calibration

Introduction



Physics goals

The largest branching fraction for a M_H=126 GeV Standard Model Higgs boson is H → bb

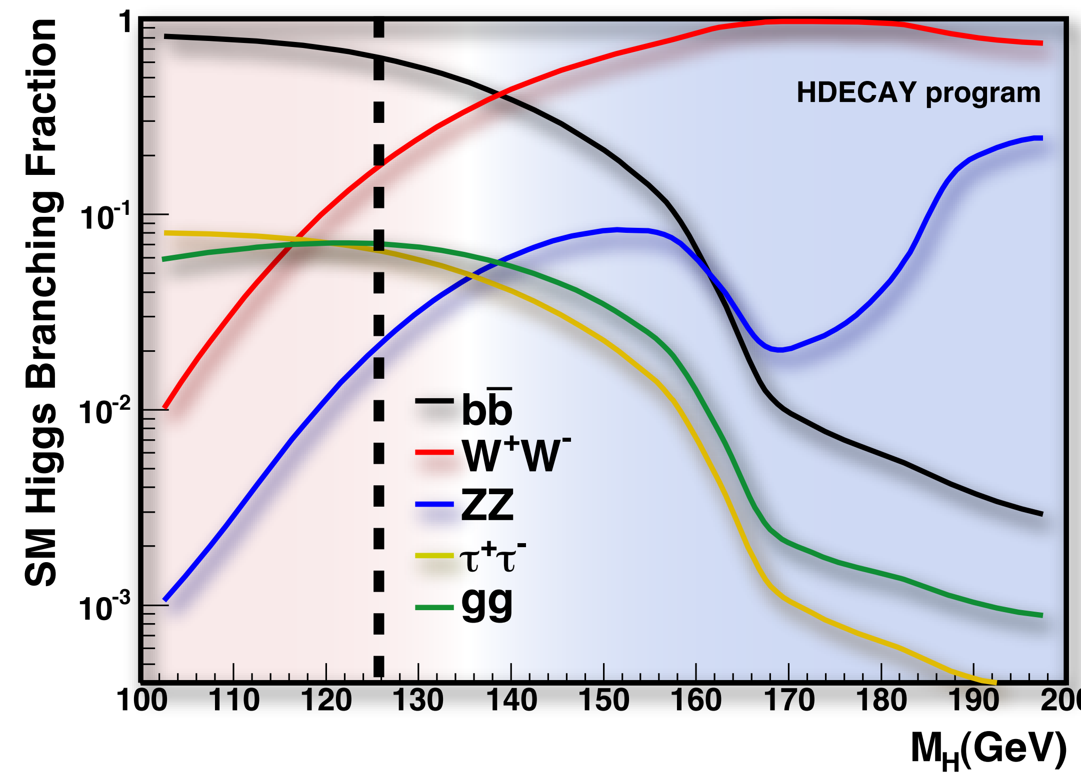
- exploit in the search
- information on coupling to fermions

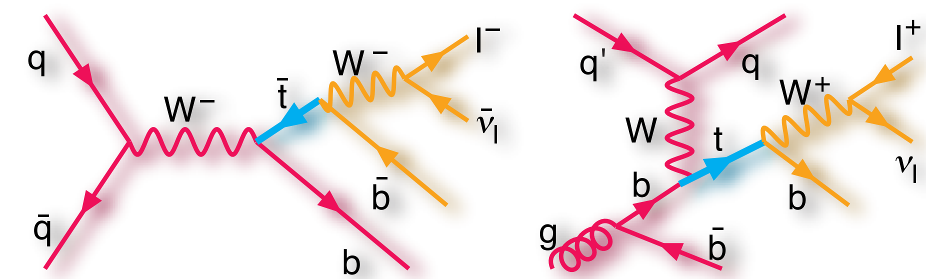
Identifying b-jets is fundamental to measuring $|V_{tb}|$

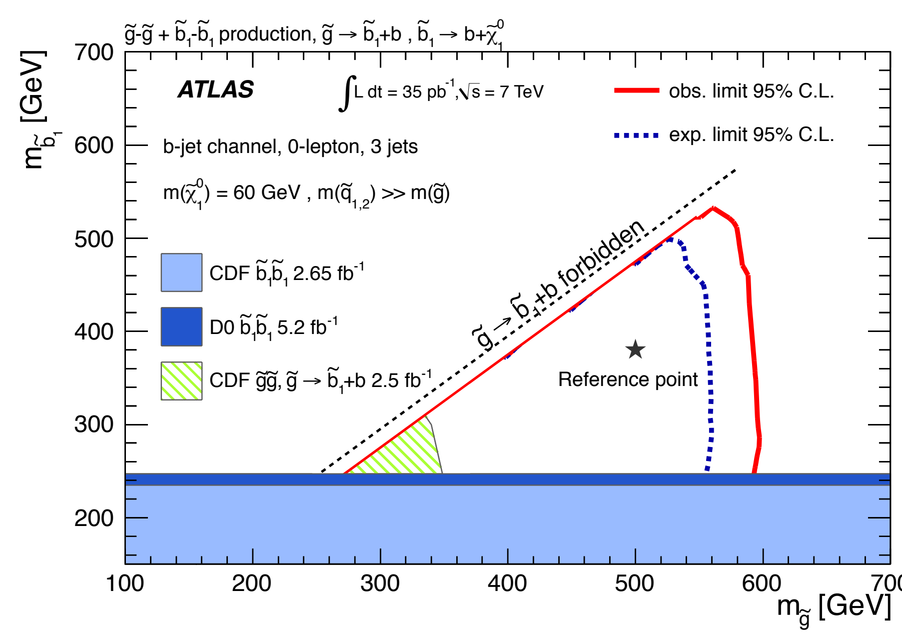
- notably EW production of single top quarks (in association with b quarks)

Many SUSY searches focus on the third quark family being lightest

- exploit in $\tilde{g} \to \tilde{b}b, \tilde{b} \to b\chi^0$



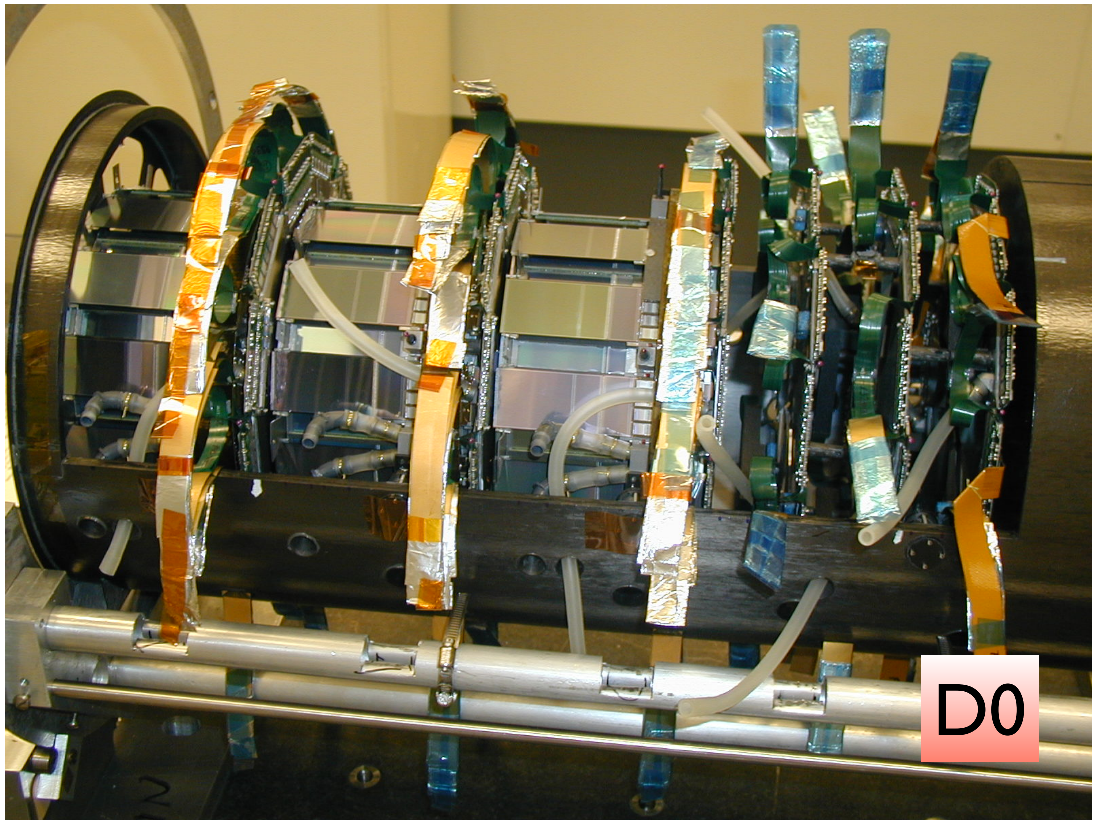


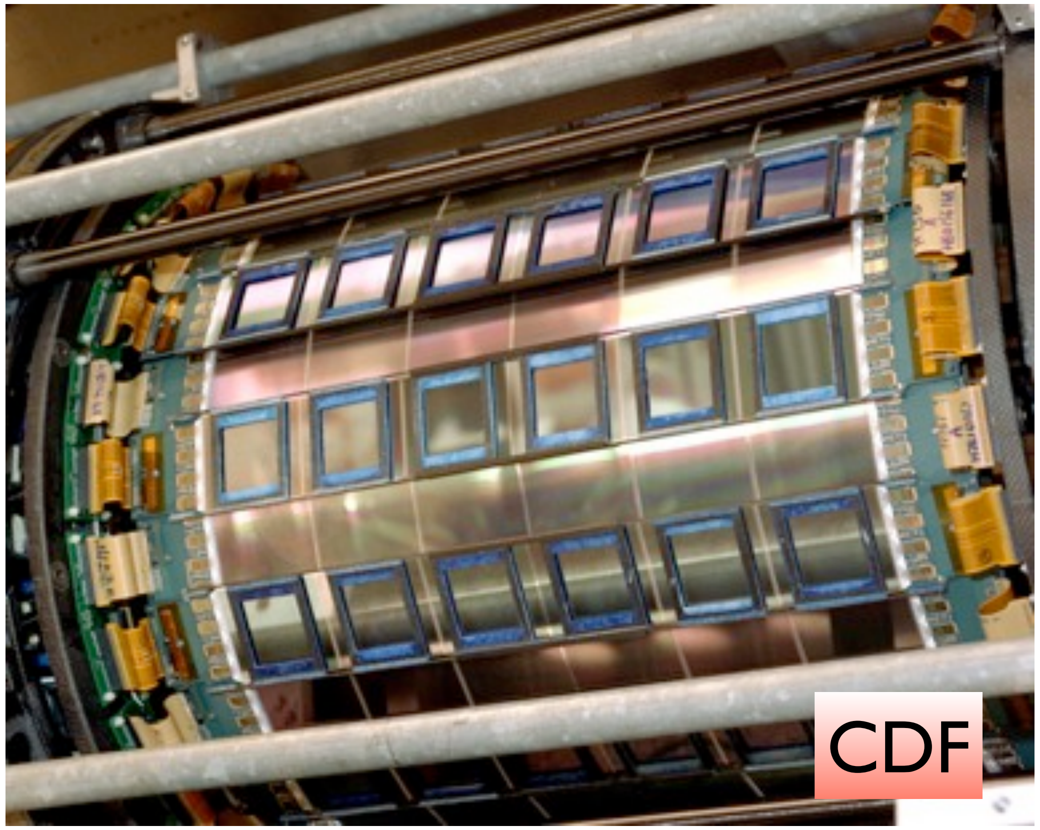


Tracking detectors

Essential: superb and redundant tracking!

- CDF, D0 were the first experiments to feature large quantities of precision (silicon) sensors (tracking, not just vertexing)

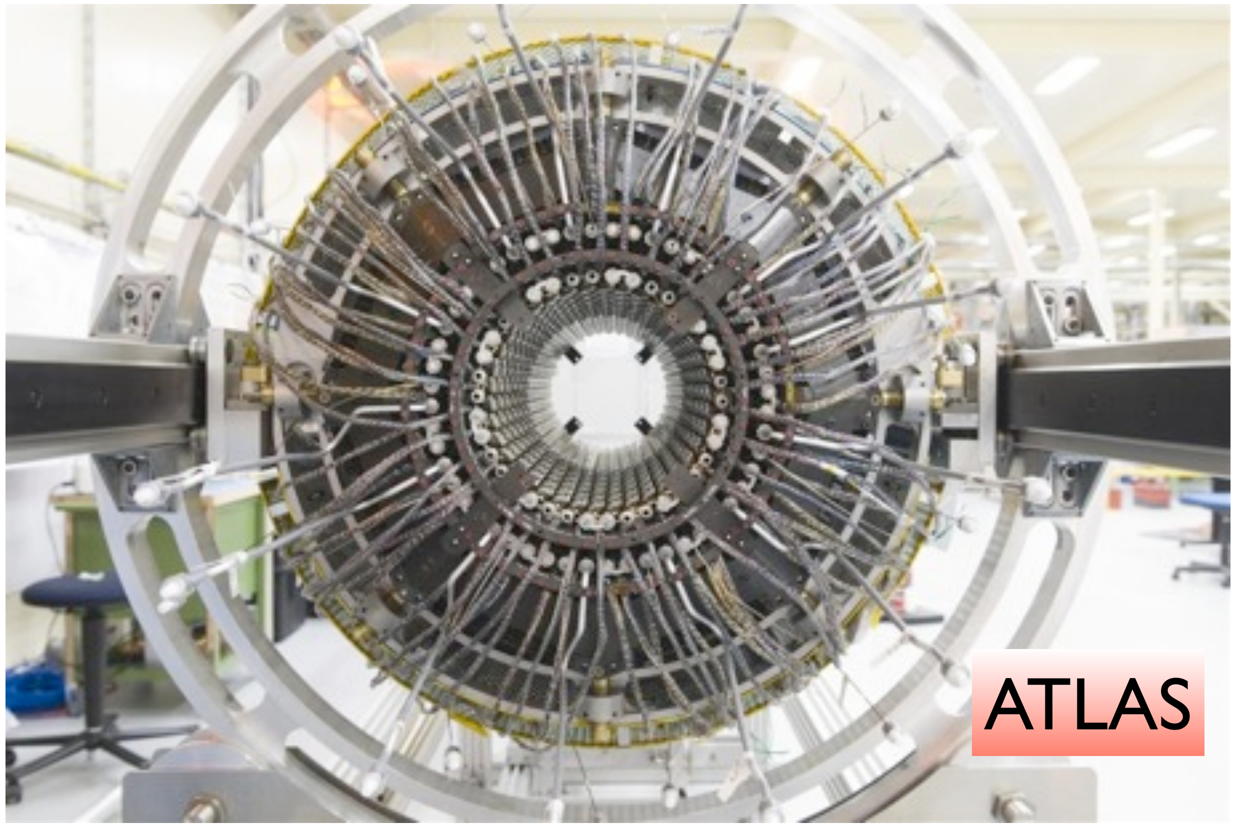




Tracking detectors

Essential: superb and redundant tracking!

- CDF, D0 were the first experiments to feature large quantities of precision (silicon) sensors (tracking, not just vertexing)
- trend continued by ATLAS, CMS: yet more silicon strips; pixels!
- essential to reduce sensitivity to pile-up

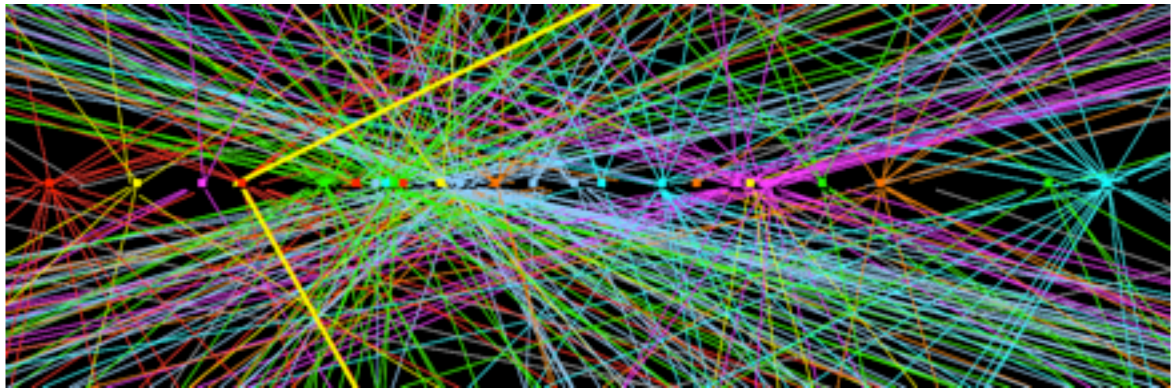


Tagging preliminaries

Primary vertex selection: typically based on tracks' p_T

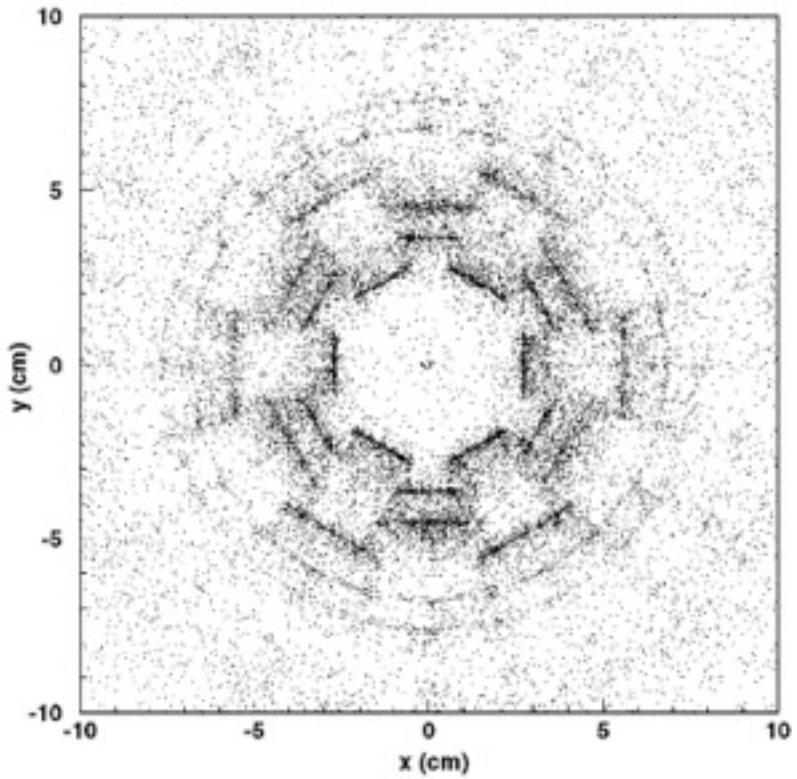
- ATLAS, CMS: select PV with highest Σp_T^2
- will need to revise in view of further increased pile-up starting from 2015

(ATLAS)
$Z \to \mu\mu$ candidate with 25 reconstructed primary vertices



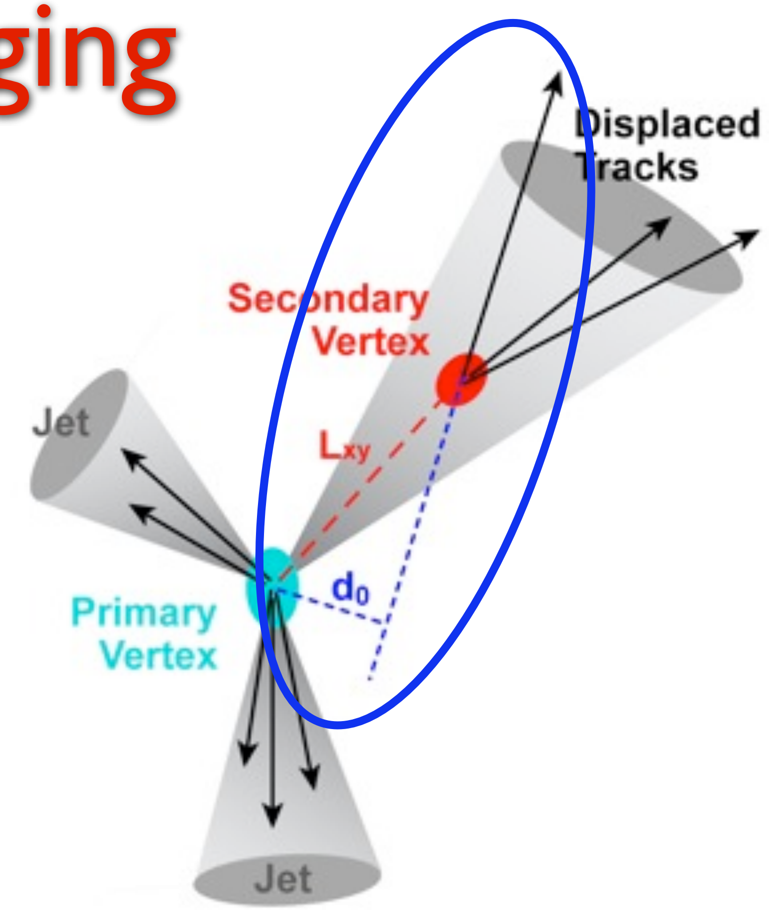
Track selection:

- generally use a fixed cone around calorimeter jet direction
  - but e.g. ATLAS exploits collimation at high jet p_T: smaller cone size
- rejection of tracks from long-lived light hadrons & material interactions
  - attempt to identify these explicitly



γ conversions in the D0 tracker

Impact parameter tagging



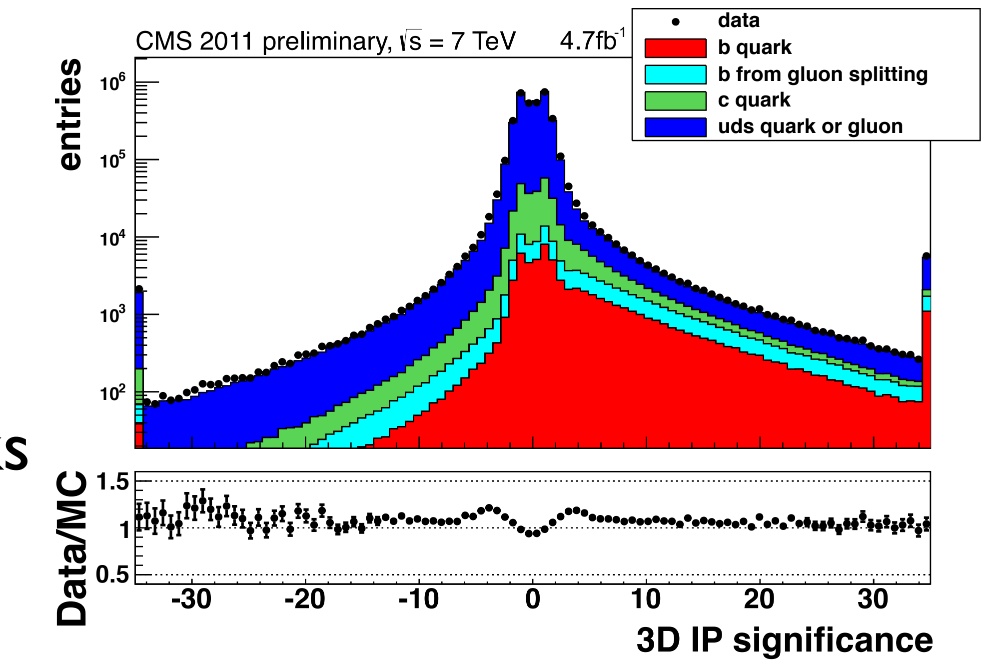
Time-honoured: impact parameter tagging

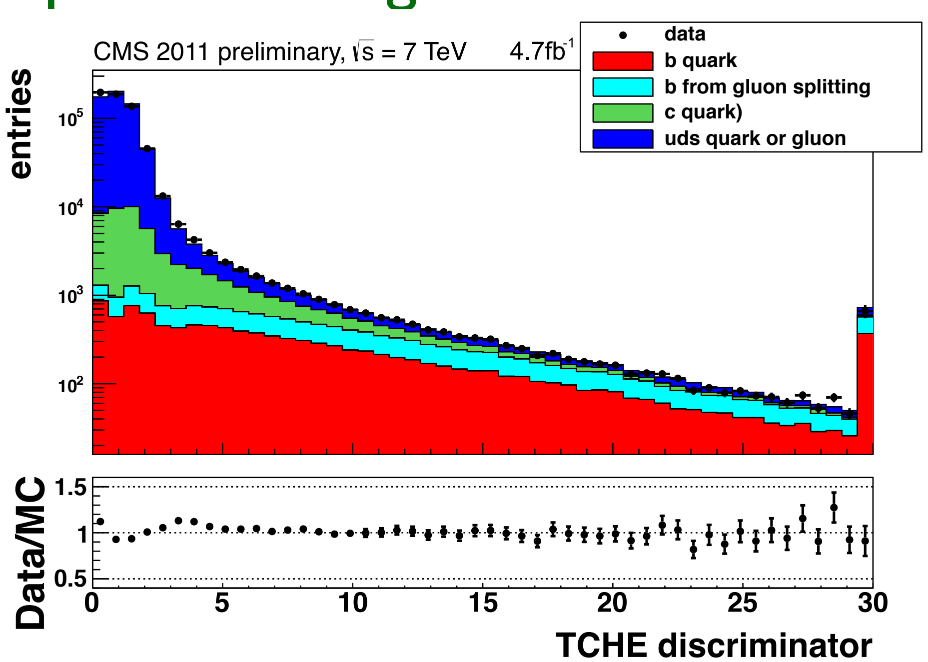
- based on (usually physics-signed) impact parameter significance $S_{d0} \equiv d_0/\sigma(d_0)$
  - 2D or 3D
- Simplest forms: track counting (e.g. number of tracks with significance above a specified minimum)
- nice CMS variant: use 2nd or 3rd highest significance
  - less prone to single outlier tracks
- JetProb: compute joint compatibility of jet's tracks with the PV
  - use of resolution function derived from $d_0 < 0$ tracks: (partly) self-calibrating!

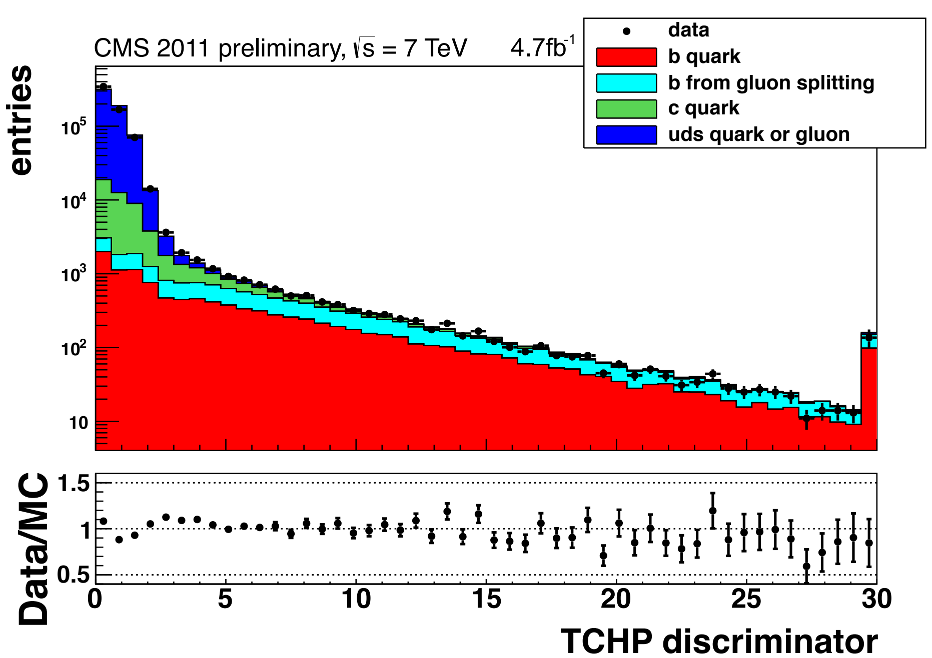
Impact parameter tagging

Time-honoured: impact parameter tagging

- based on (usually physics-signed) impact parameter significance $S_{d0} \equiv d_0/\sigma(d_0)$
  - 2D or 3D
- Simplest forms: track counting (e.g. number of tracks with significance above a specified minimum)
- nice CMS variant: use 2nd or 3rd highest significance
  - less prone to single outlier tracks
- JetP... jet's ...
  - use ... r d_0



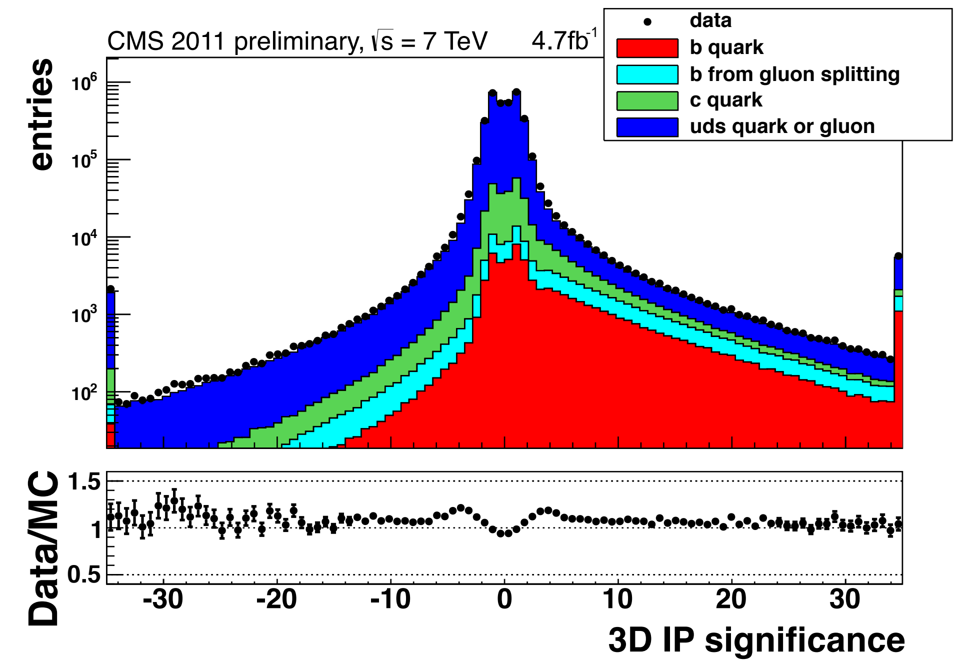




Impact parameter tagging

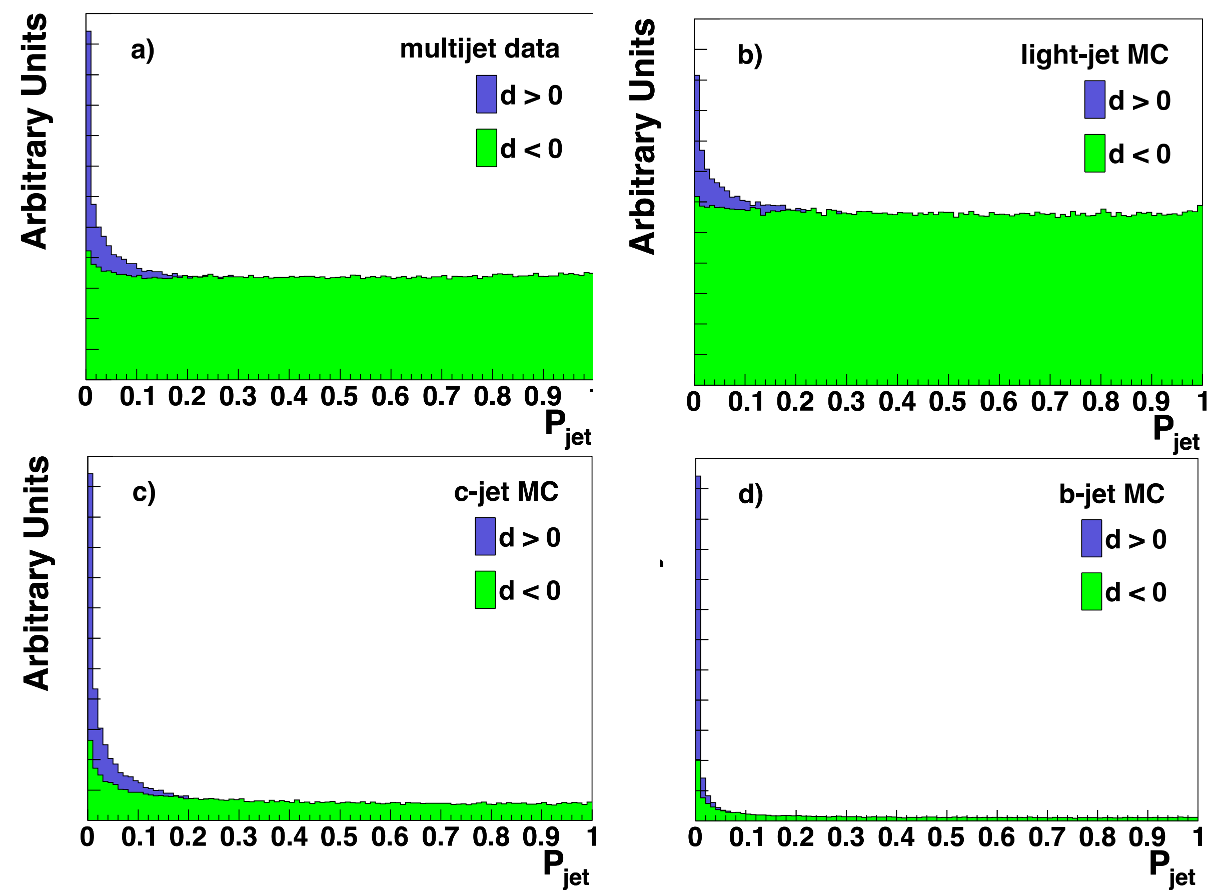
Time-honoured: impact parameter tagging

- based on (usually physics-signed) impact parameter significance $S_{d0} \equiv d_0/\sigma(d_0)$
  - 2D or 3D
- Simplest forms: track counting (e.g. number of tracks with significance above a specified minimum)
- nice CMS variant: use 2nd or 3rd highest significance
  - less prone to single outlier tracks
- JetProb: compute joint compatibility of jet's tracks with the PV
  - use of resolution function derived from $d_0 < 0$ tracks: (partly) self-calibrating!



$$\mathcal{P}_{\mathrm{trk},i} = \int_{-\infty}^{-|S^i_{d_0}|} \mathcal{R}(x)\mathrm{d}x,$$

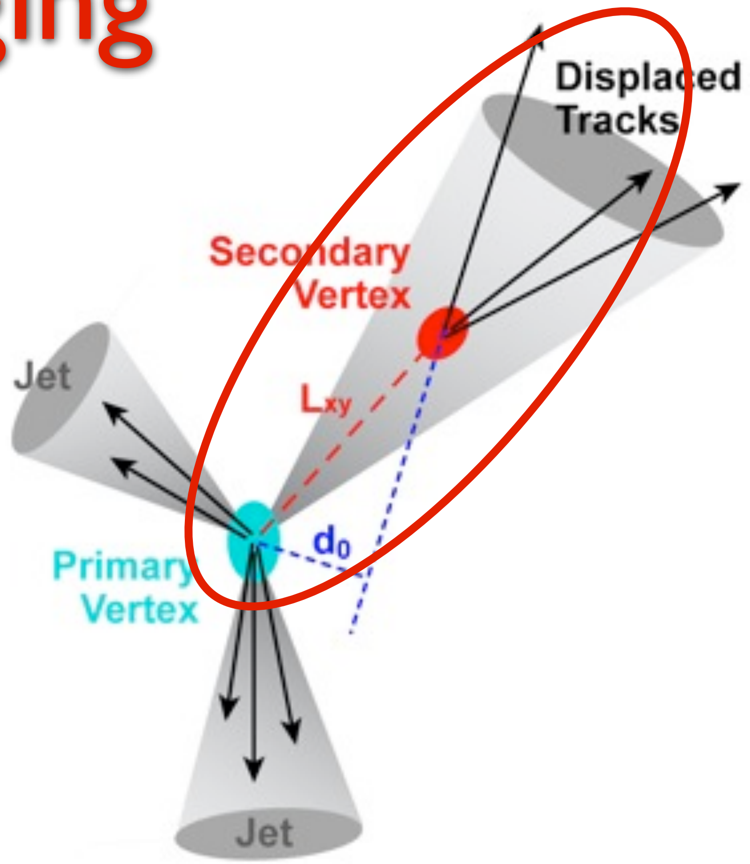
$$\mathcal{P}_{\mathrm{jet}} = \mathcal{P}_0 \sum_{j=0}^{N-1} \frac{(-\ln \mathcal{P}_0)^j}{j!}, \qquad \mathcal{P}_0 = \prod_{i=1}^{N} \mathcal{P}_{\mathrm{trk},i}$$

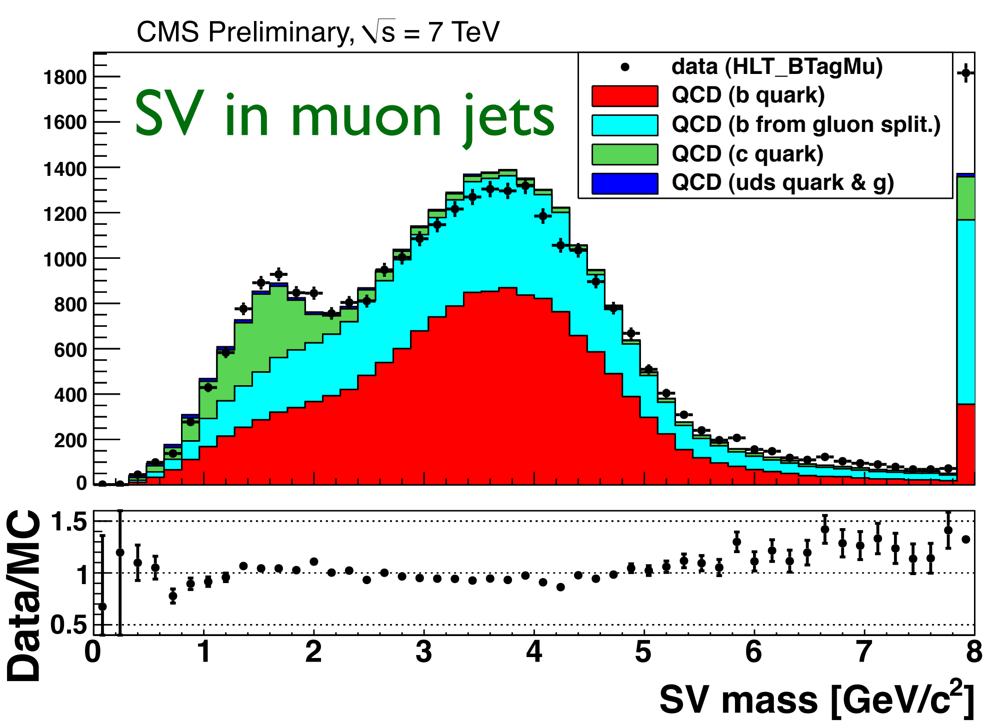


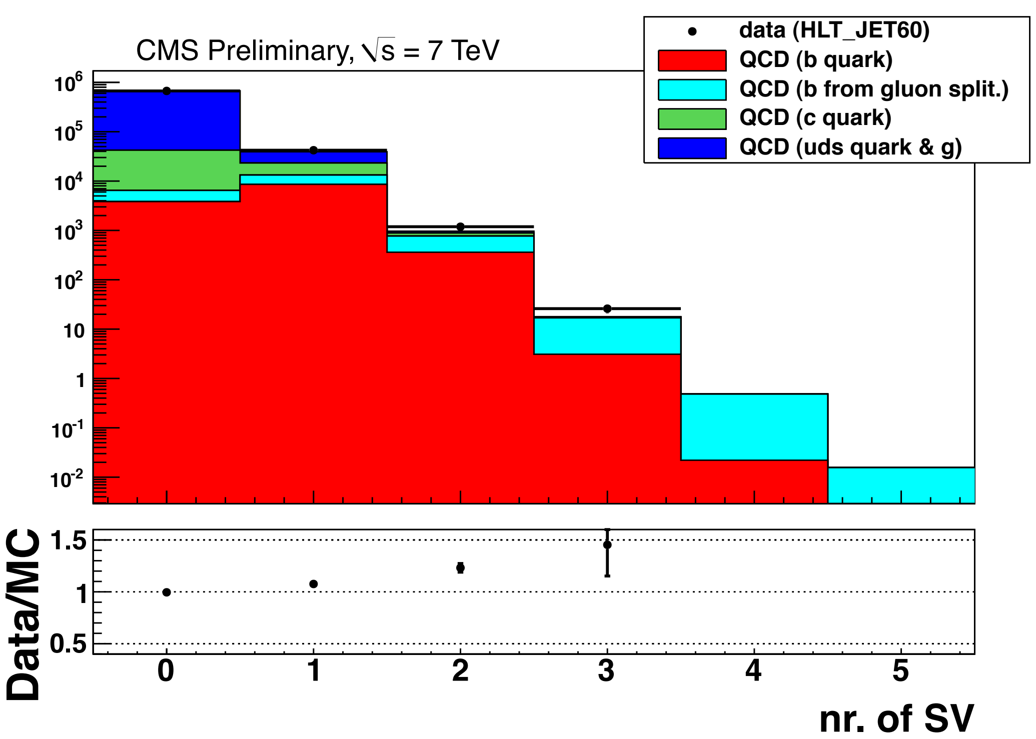
Secondary vertex tagging

Time-honoured: allow for significantly higher purity than impact parameter tagging

- typically use”build-up” algorithms: start with 2-track vertex candidates, attempt to add other tracks
- cleaning stages to remove tracks originating from the primary vertex, fake tracks, ...
- access to interesting kinematic information!
  - e.g. charm; g → bb

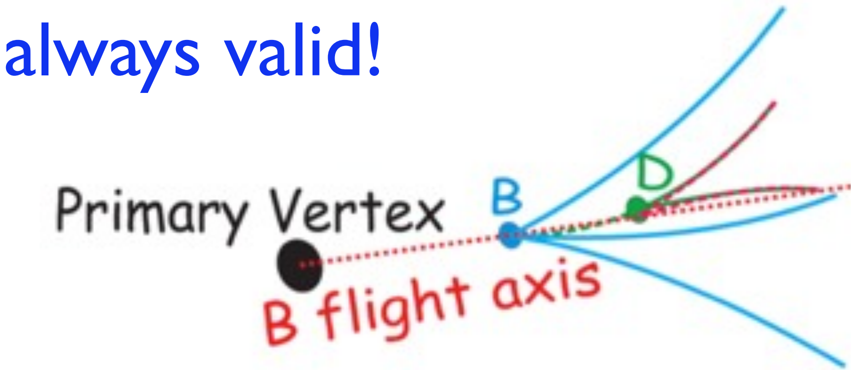




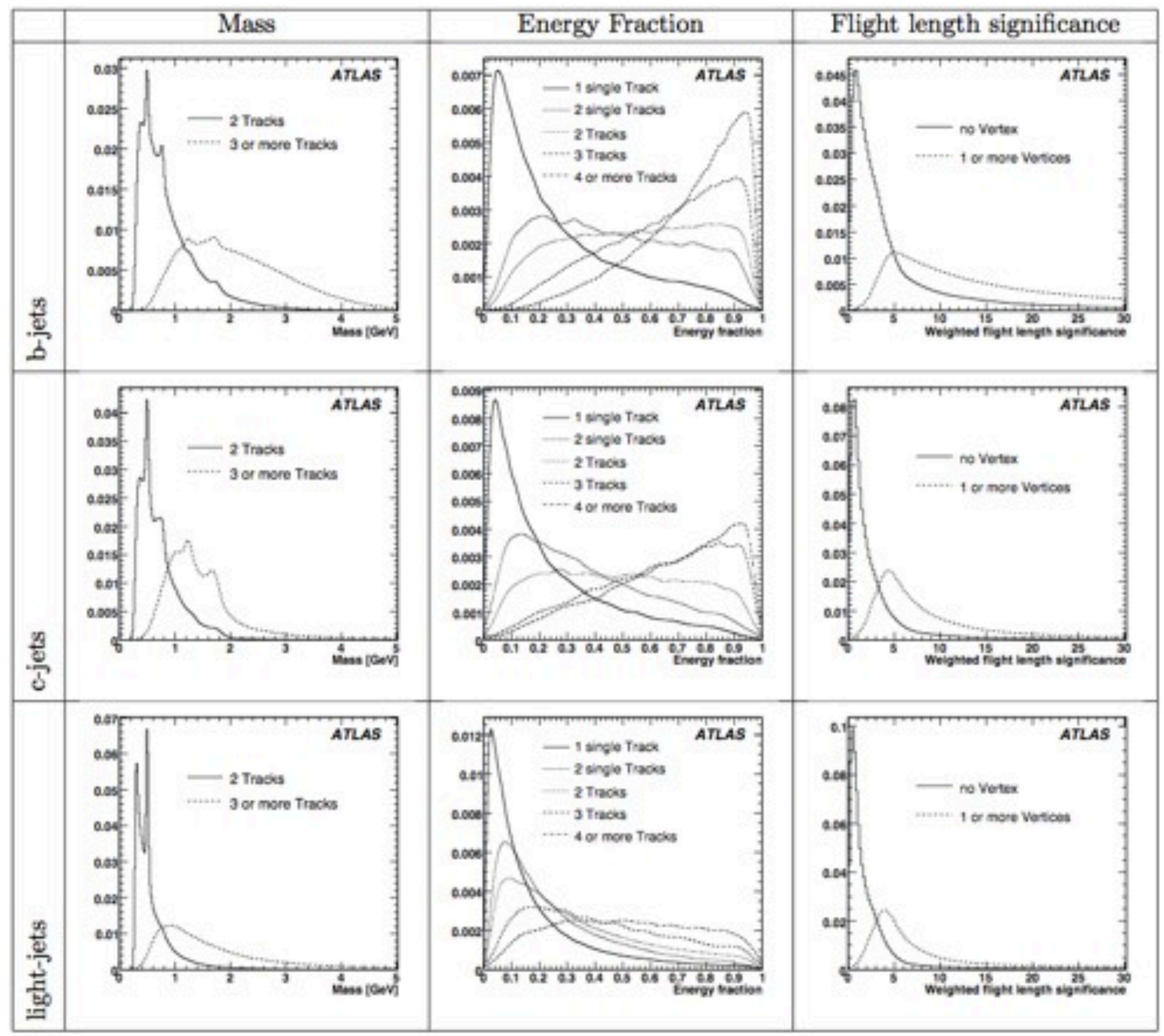


A more unified approach

The assumption of a single secondary vertex is not always valid!



- b → c → l decays: also charmed hadrons live long!
- JetFitter approach used in ATLAS (similar to earlier SLD algorithm): assume vertices are aligned
  - allows for significant efficiency increase



After reconstruction:

- categorise by track and vertex multiplicity
- construct likelihood ratio from discriminating variables

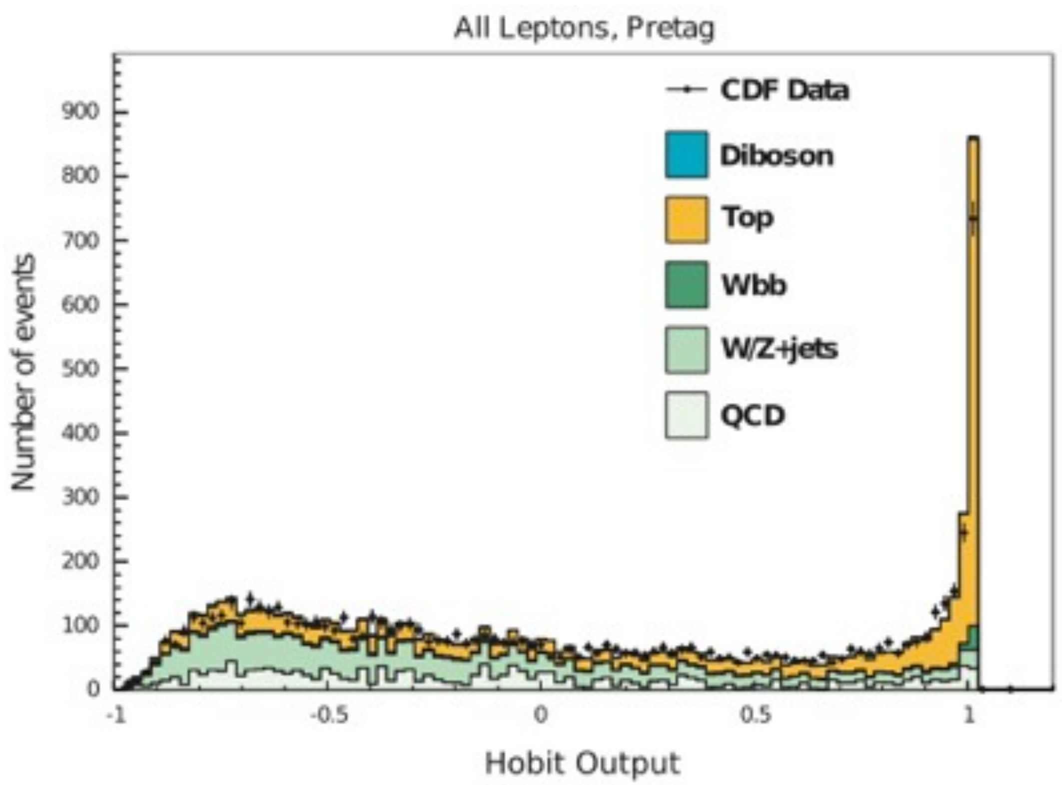
$$\frac{\mathcal{P}_b}{\mathcal{P}_b + \mathcal{P}_c + \mathcal{P}_{\text{light}}}$$

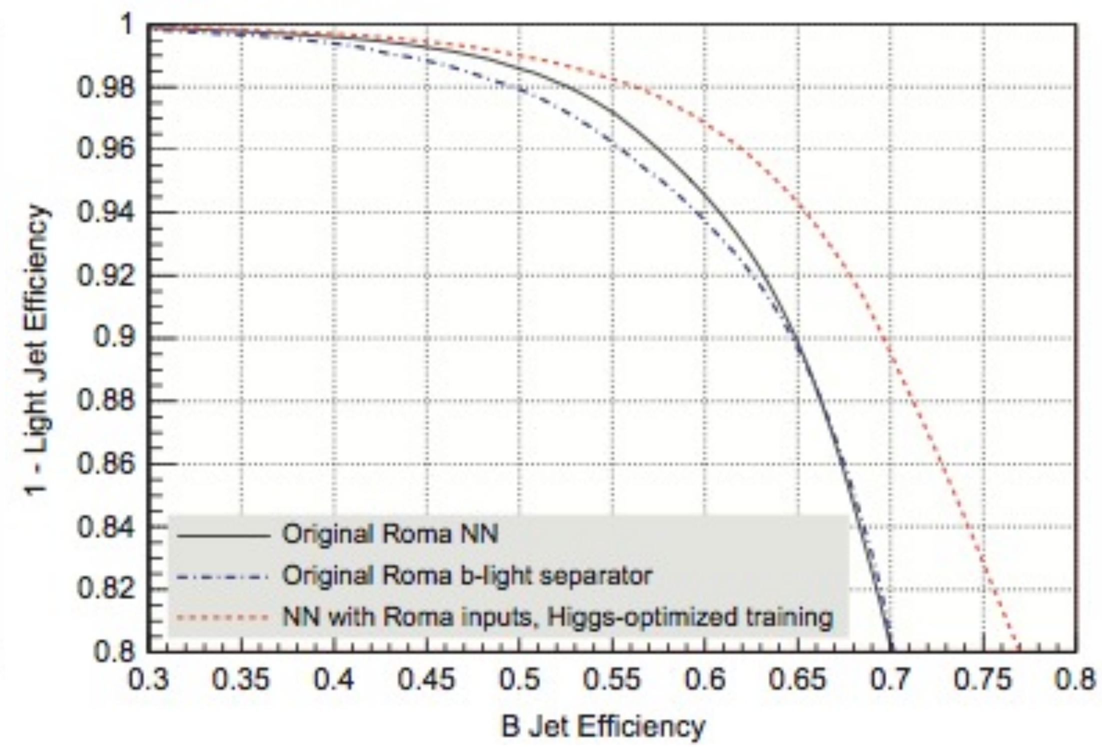
Similar features used in CMS SV reconstruction

Multivariate tagging

Recent example: CDF's Higgs-Optimized b-Identification Tagger (HOBIT)

- combine 25 variables sensitive to b ↔ non-b differences in a neural network
  - secondary vertex related, impact parameter related, soft muons
  - including outputs of earlier neural networks
- significant improvements, can be optimised for specific physics use cases



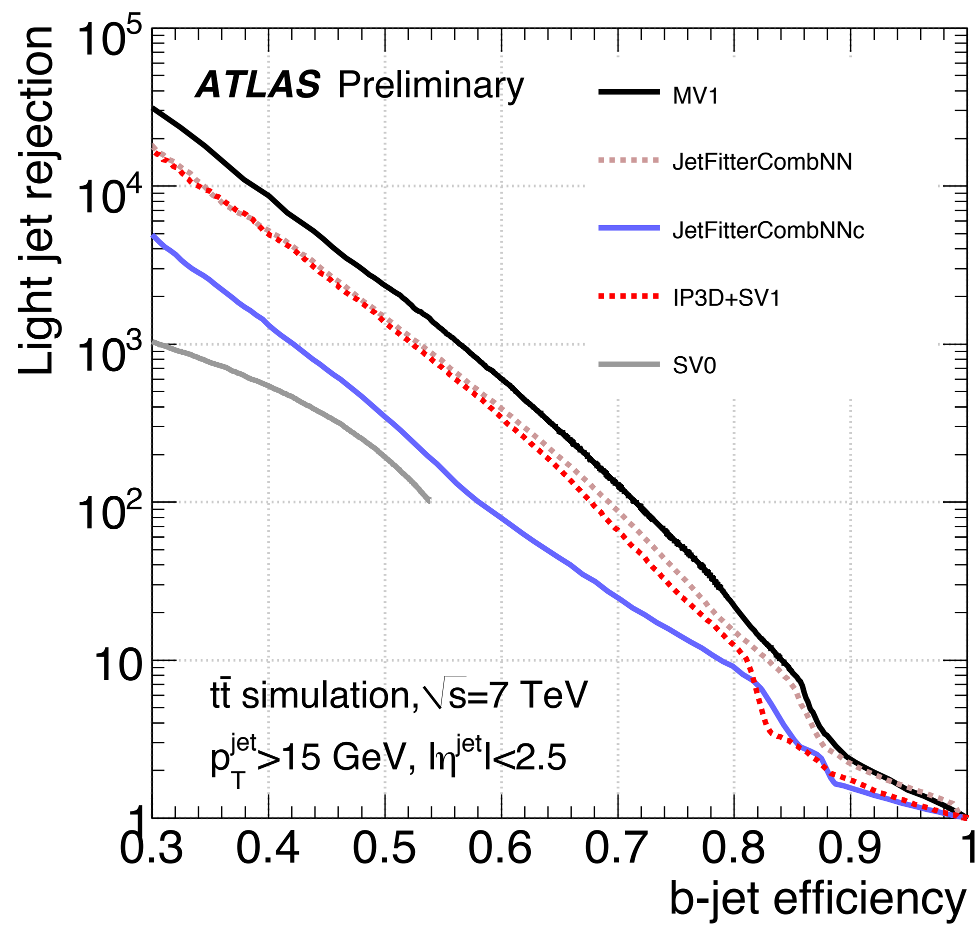


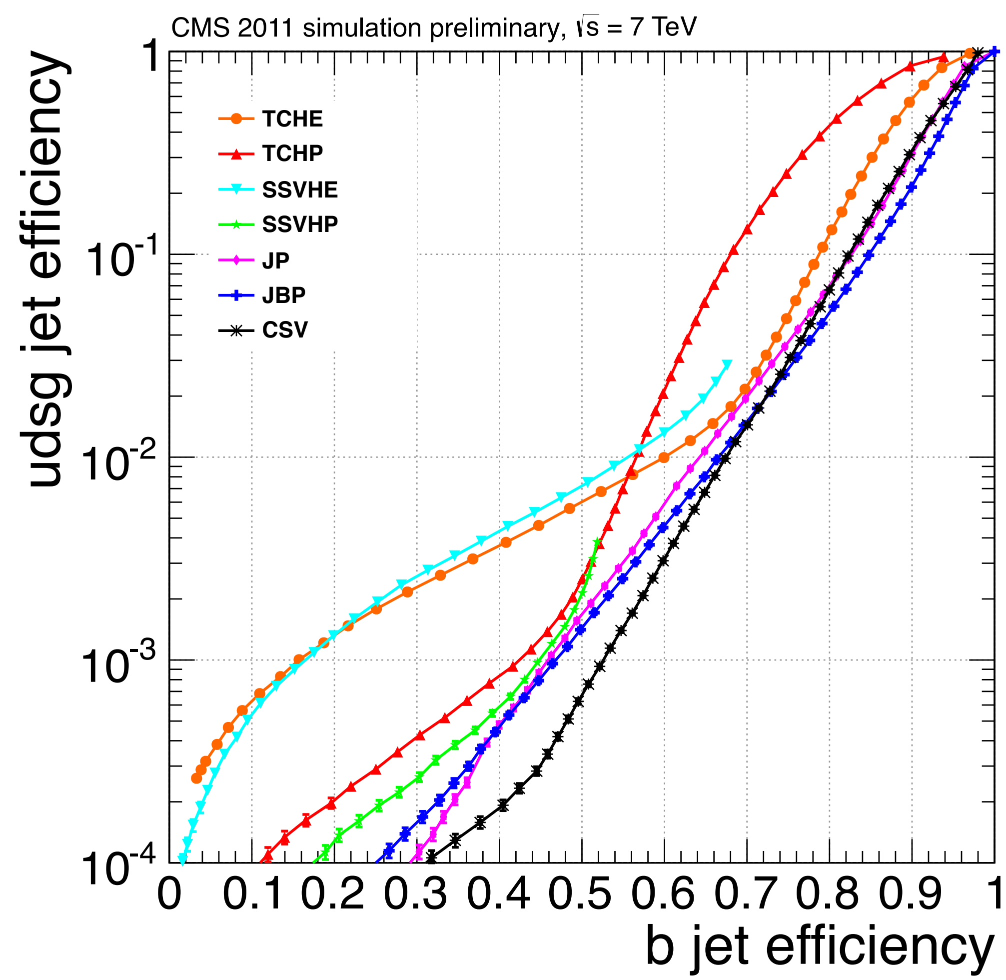
Similar taggers in D0, ATLAS

- versatility exploited in ATLAS by optimisation for charm (instead of light-jet) rejection

Performance estimates

Impressive range of algorithms, but MC estimates should be checked!





Calibrations of efficiency for b / c / light-flavour jets to satisfy given cuts on tagging outputs

- result typically reported as data/MC efficiency scale factors

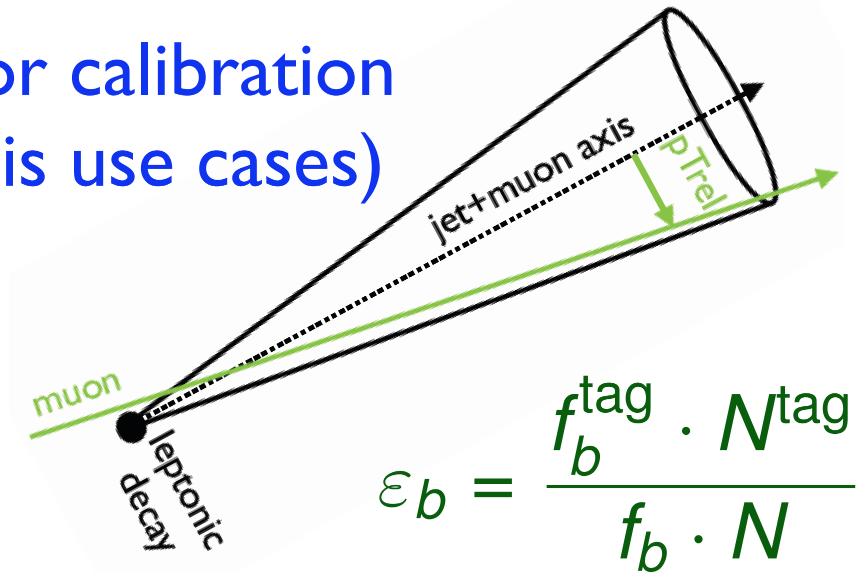
Muon-based calibration (1)

Most focus in b-jet tagging is on lifetime tagging

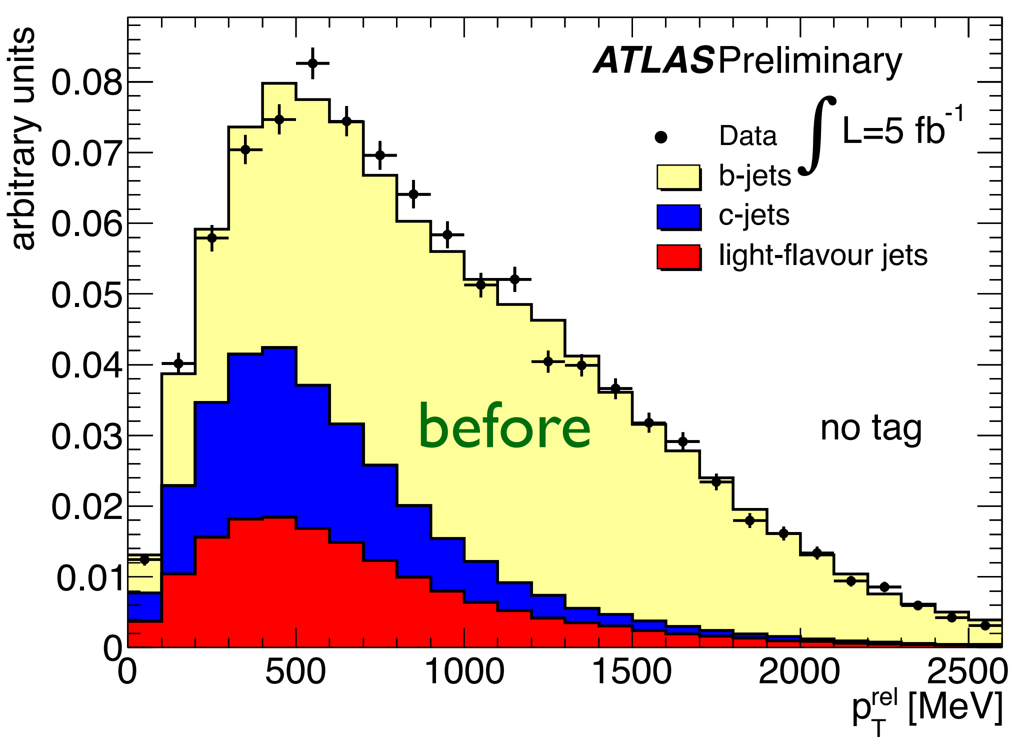
- semileptonic decays (especially involving muons) suffer from low branching fractions; but useful in calibrations (see later) and for specific physics analyses

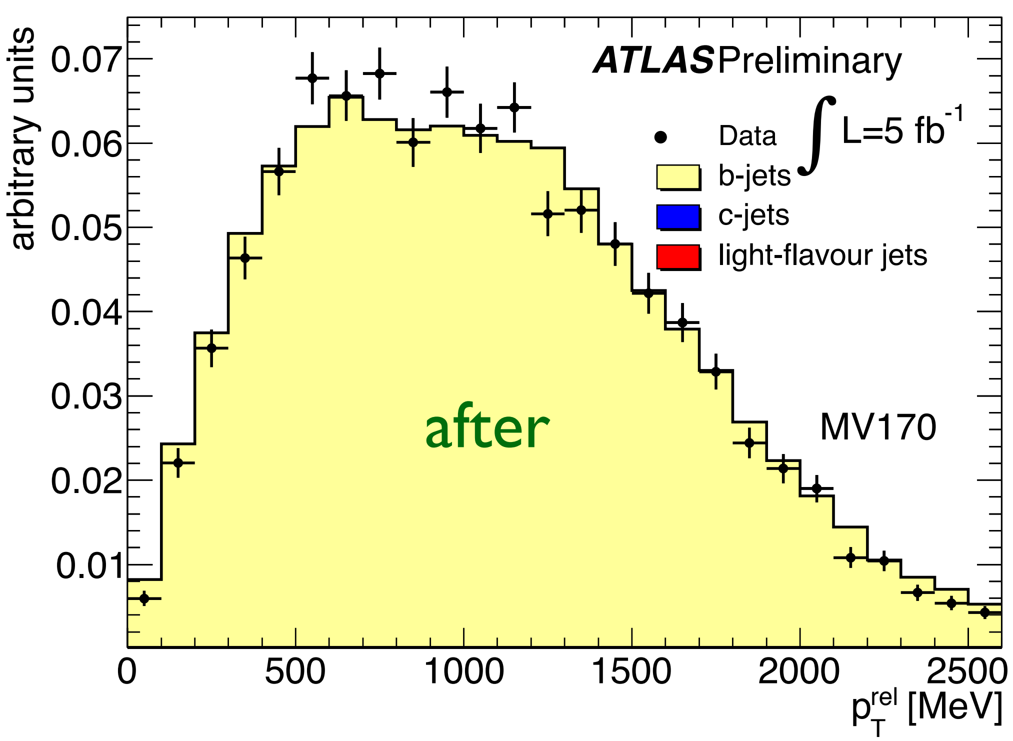
But jets with associated muons are very useful for calibration purposes (in addition to providing specific analysis use cases)

- most commonly used: $p_{T,rel}$ (good at low p_T)
  - also muon impact parameter (better at high p_T)



$$\varepsilon_b = \frac{f_b^{tag} \cdot N^{tag}}{f_b \cdot N}$$

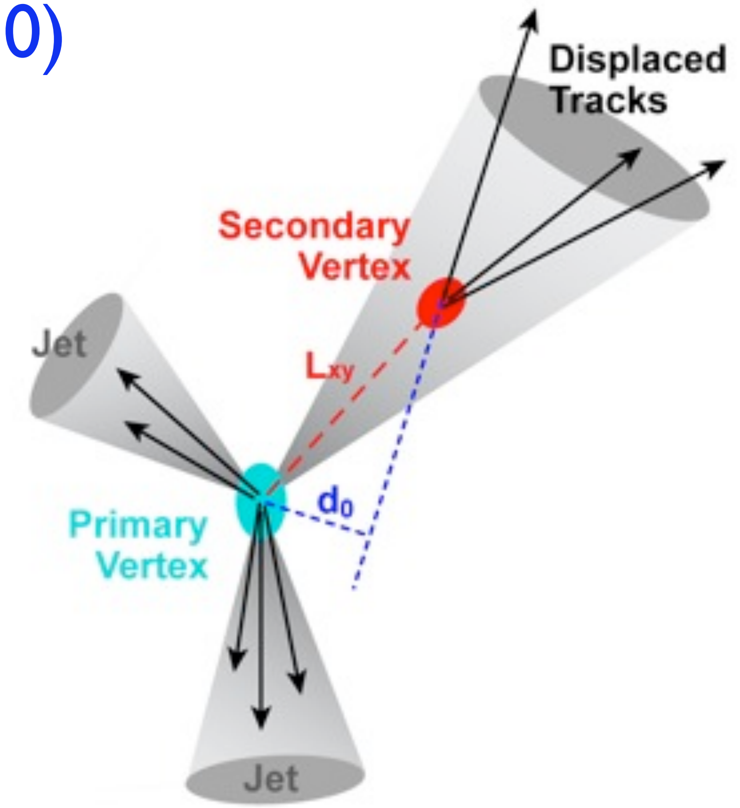




Muon-based calibration (2)

Reduced dependence on MC: System8 (pioneered by D0)

- use 3 (essentially uncorrelated) cuts each discriminating between b-jets and c/l-jets: the tagger under consideration (LT), a $p_{T,rel}$ cut (MT), and an away-jet tag ($n \rightarrow p$): $2^3 = 8$ combinations ➠ solve for 8 unknowns
  - need to lump together charm and light-flavour jets
  - correlations (mostly inferred from MC) are typically small. Exception: charm purification by away-jet tag (difficult to estimate but extracted b-jet efficiency not very sensitive)



$$
\begin{aligned}
n &= n_b + n_{cl} \\
p &= p_b + p_{cl} \\
n^{LT} &= \varepsilon_b^{LT} n_b + \varepsilon_{cl}^{LT} n_{cl} \\
p^{LT} &= \alpha_6 \varepsilon_b^{LT} p_b + \alpha_4 \varepsilon_{cl}^{LT} p_{cl} \\
n^{MT} &= \varepsilon_b^{MT} n_b + \varepsilon_{cl}^{MT} n_{cl} \\
p^{MT} &= \alpha_5 \varepsilon_b^{MT} p_b + \alpha_3 \varepsilon_{cl}^{MT} p_{cl} \\
n^{LT,MT} &= \alpha_1 \varepsilon_b^{LT} \varepsilon_b^{MT} n_b + \alpha_2 \varepsilon_{cl}^{LT} \varepsilon_{cl}^{MT} n_{cl} \\
p^{LT,MT} &= \alpha_7 \alpha_6 \alpha_5 \varepsilon_b^{LT} \varepsilon_b^{MT} p_b + \alpha_8 \alpha_4 \alpha_3 \varepsilon_{cl}^{LT} \varepsilon_{cl}^{MT} p_{cl}
\end{aligned}
$$

$$
\begin{aligned}
\alpha_1 &= \varepsilon_b^{LT,MT,n} / (\varepsilon_b^{LT,n} \varepsilon_b^{MT,n}) & \alpha_2 &= \varepsilon_{cl}^{LT,MT,n} / (\varepsilon_{cl}^{LT,n} \varepsilon_{cl}^{MT,n}) \\
\alpha_5 &= \varepsilon_b^{MT,p} / \varepsilon_b^{MT,n} & \alpha_3 &= \varepsilon_{cl}^{MT,p} / \varepsilon_{cl}^{MT,n} \\
\alpha_6 &= \varepsilon_b^{LT,p} / \varepsilon_b^{LT,n} & \alpha_4 &= \varepsilon_{cl}^{LT,p} / \varepsilon_{cl}^{LT,n} \\
\alpha_7 &= \varepsilon_b^{LT,MT,p} / (\varepsilon_b^{LT,p} \varepsilon_b^{MT,p}) & \alpha_8 &= \varepsilon_{cl}^{LT,MT,p} / (\varepsilon_{cl}^{LT,p} \varepsilon_{cl}^{MT,p})
\end{aligned}
$$

Both results need to be extrapolated to inclusive b jets

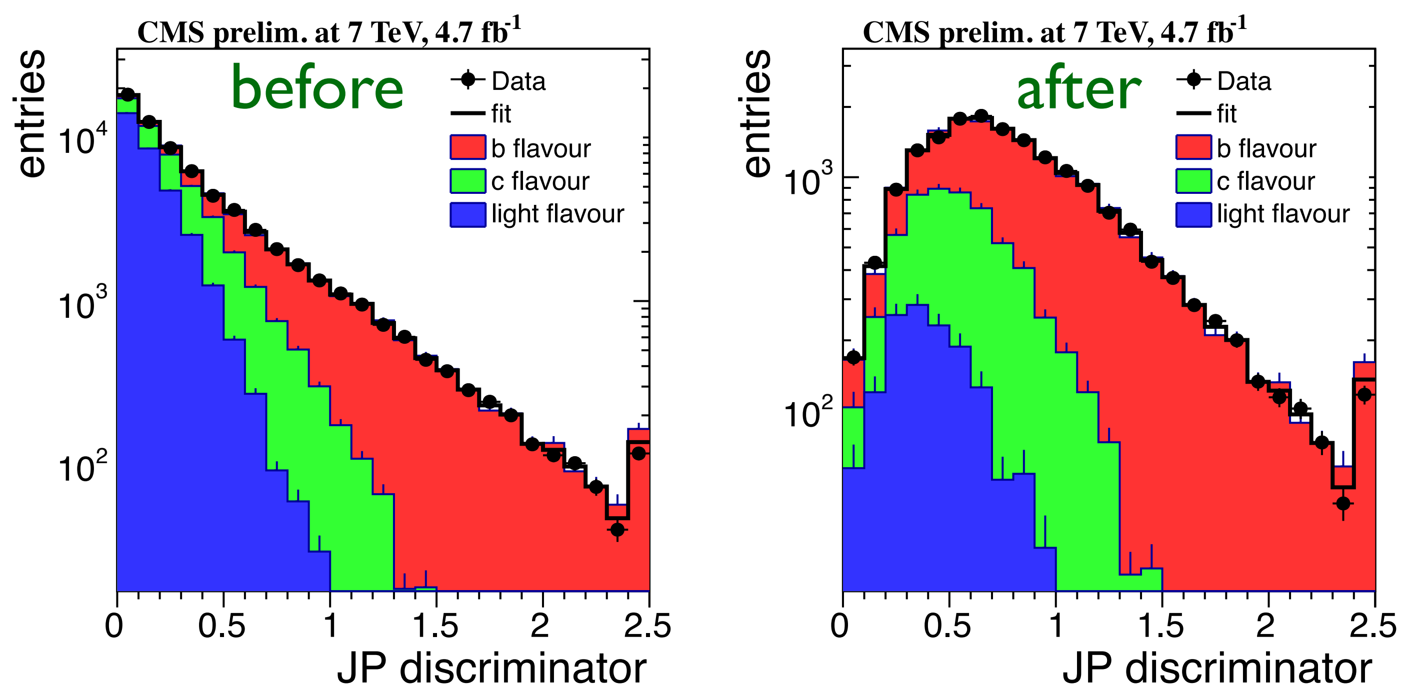
- assume same data/MC scale factors hold

Reference lifetime tagger

Basic problem with muon-based calibration: $p_{T,rel}$ spectra for b-jets and lighter flavours become similar at high p_T ➠ not useful for p_T > 200 GeV

CMS are using an alternative method: JetProb template fitting

- templates still partly based on MC but under improved control:
  - self-calibrating feature for light-flavour jets
  - JetProb calibration



tt-based calibration

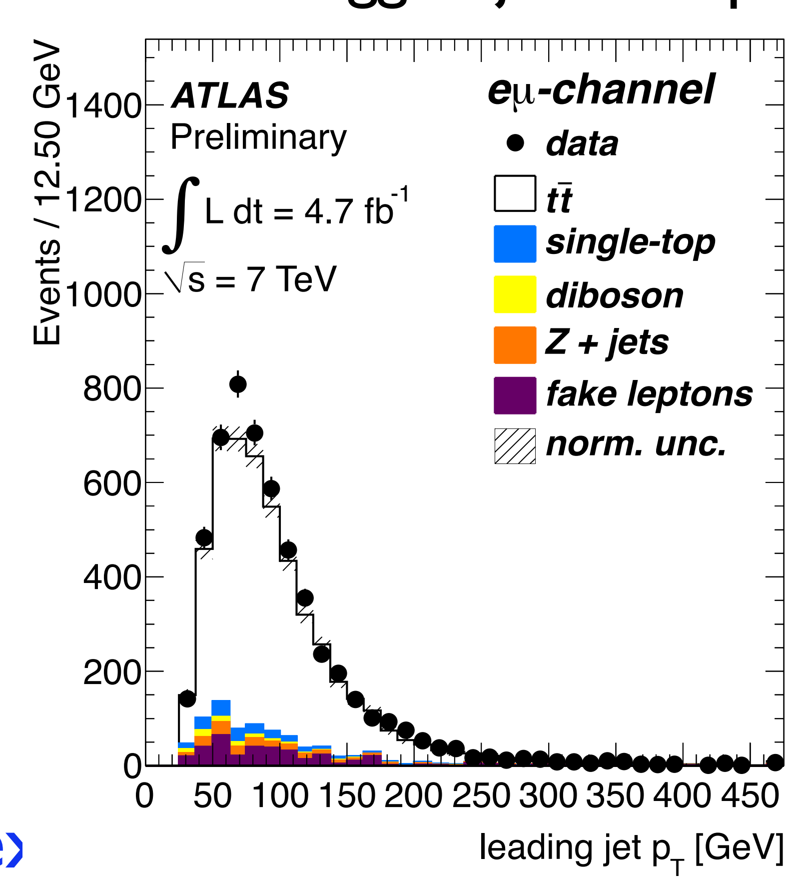
For the LHC experiments only!

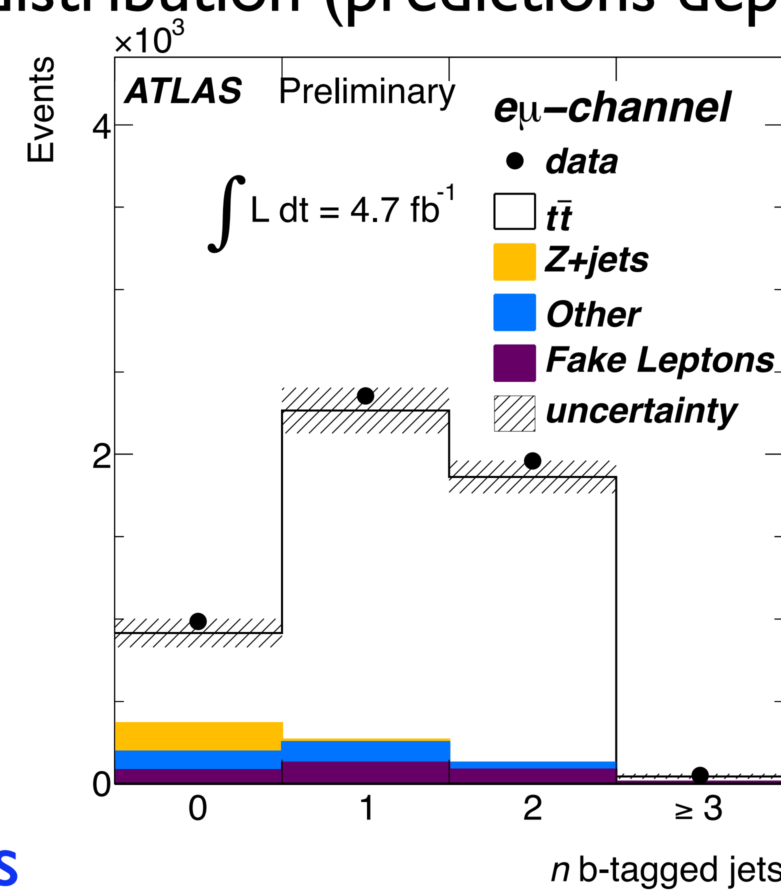
- greatly increased tt production cross section (from ~ 7 pb to ~ 200 pb)
- relatively clean source of b jets (assuming B(t→Wb)=1)

Exploited using a variety of methods, both in dilepton & lepton+jets events

dilepton sample:

- either count fraction of tagged jets
- or fit to tagged jet multiplicity distribution (predictions depend on ε_b)





No ex … jets … ties

tt-based calibration

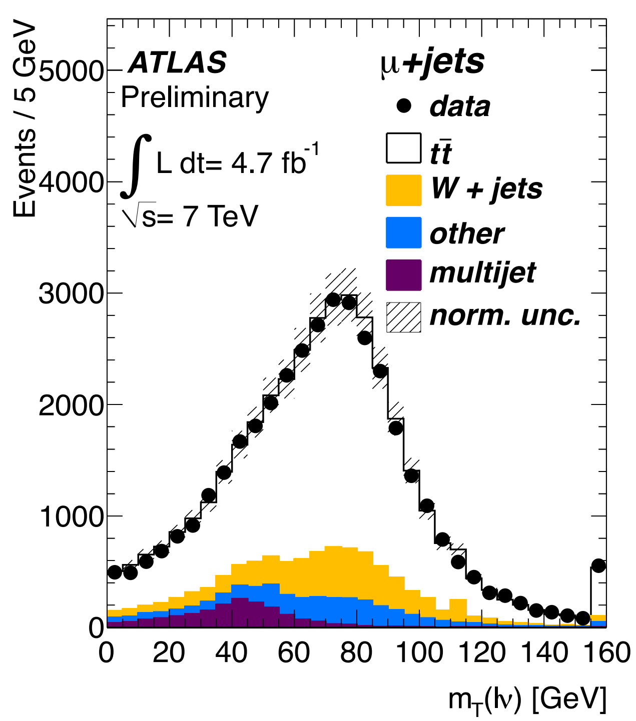
For the LHC experiments only!

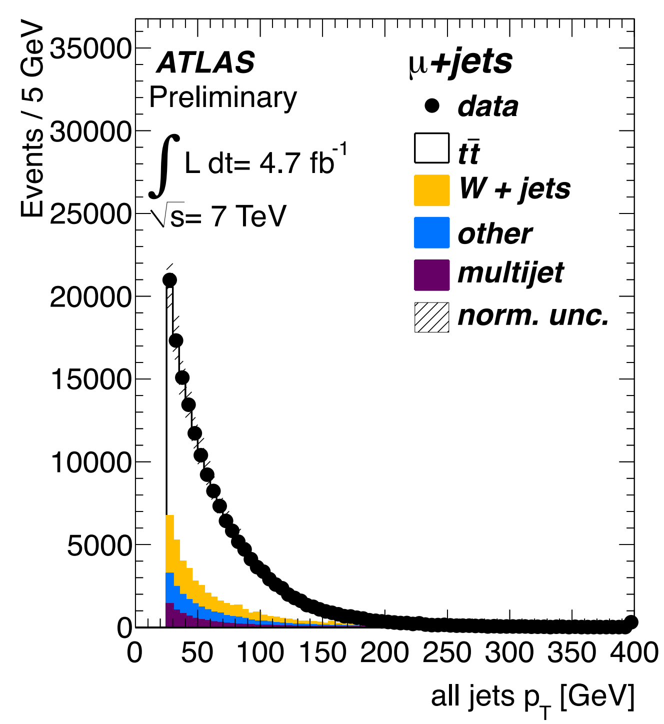
- greatly increased tt production cross section (from ~ 7 pb to ~ 200 pb)
- relatively clean source of b jets (assuming B(t→Wb)=1)

Exploited using a variety of methods, both in dilepton & lepton+jets events

lepton+jets sample:

- enhance tt content using kinematic selection, count fraction of tagged jets
- or apply kinematic fit, subtract wrong combinations & count





No extrapolation to inclusive b jets; but modelling uncertainties

tt-based calibration

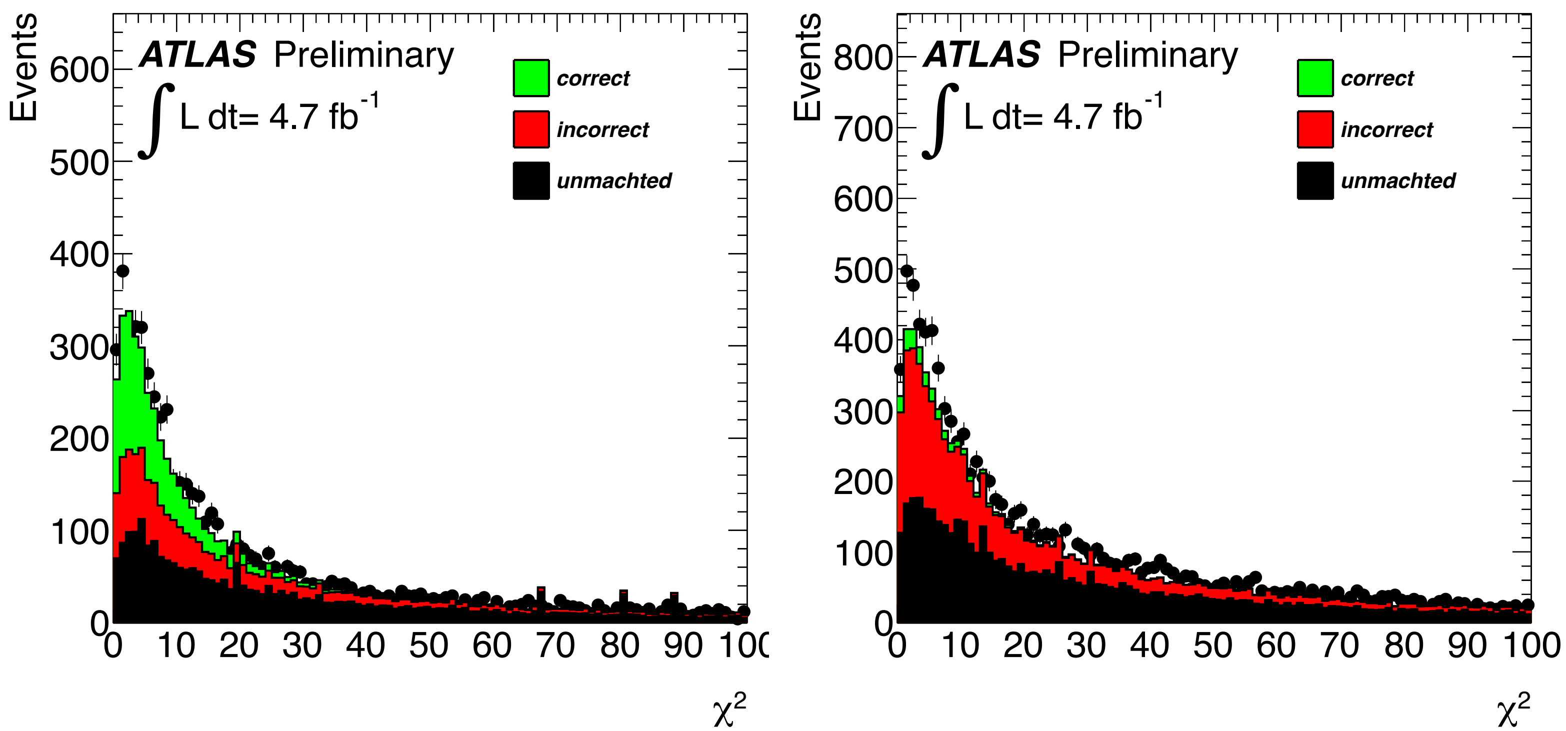
For the LHC experiments only!

- greatly increased tt production cross section (from ~ 7 pb to ~ 200 pb)
- relatively clean source of b jets (assuming B(t→Wb)=1)

Exploited using a variety of methods, both in dilepton & lepton+jets events

lepton+jets sample:

- enhance tt content using kinematic selection, count fraction of tagged jets
- or apply kinematic fit, subtract wrong combinations & count

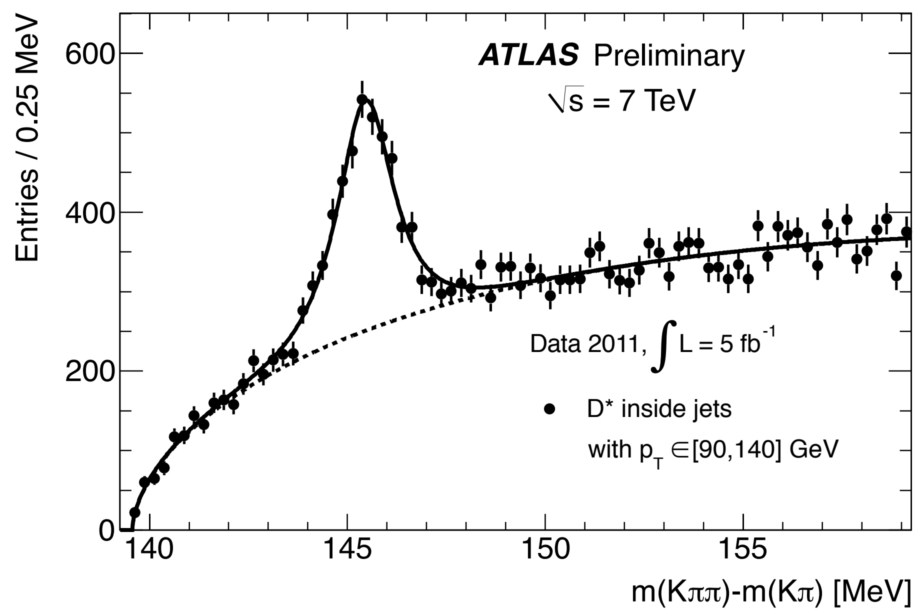


No extrapolation to inclusive b jets; but modelling uncertainties

D^*-based calibration

ATLAS: use $D^{*+} \rightarrow D^0(\rightarrow K^- \pi^+)\pi^+$ decays to remove light-flavour jets

- vertexing but without track impact parameter cuts / fit χ^2 cuts to minimise biases on tagging measurements
  - minimum $E(D^*)/E(\text{jet})$: suppress combinatorics + $b \rightarrow D^*X$ decays
- estimate remaining b-jet component using pseudo-$c\tau$ fit (before tagging)
- subtract (small) remaining) b-jet contribution using (corrected) simulation



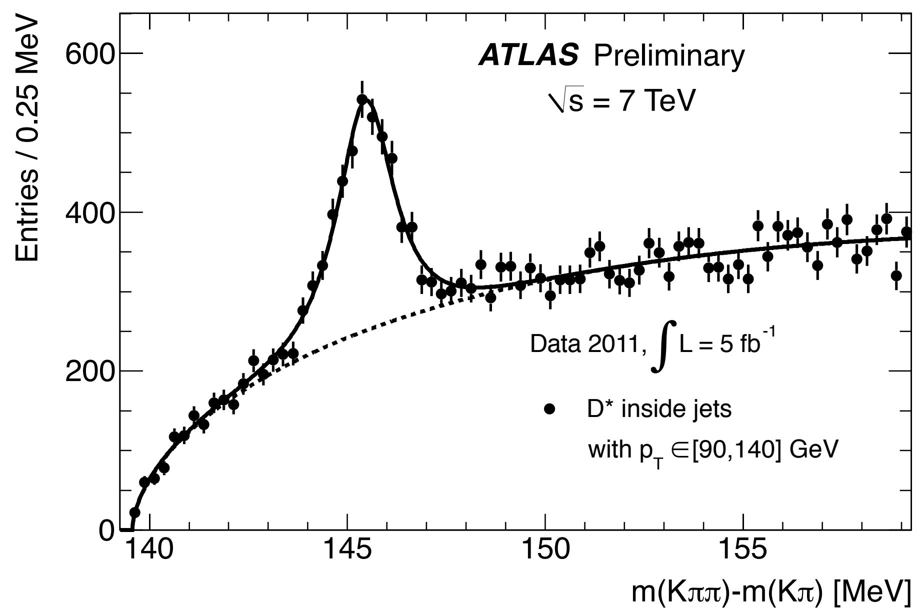
Results need to be extrapolated to inclusive charm jets

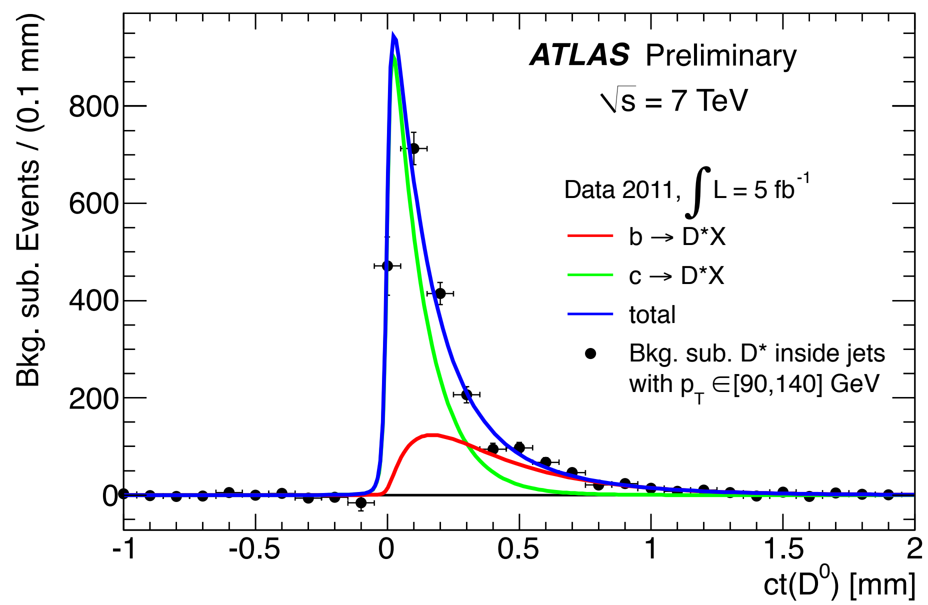
- additional systematics from charm fragmentation & decay modelling

D^*-based calibration

ATLAS: use $D^{*+} \to D^0(\to K^- \pi^+)\pi^+$ decays to remove light-flavour jets

- vertexing but without track impact parameter cuts / fit χ^2 cuts to minimise biases on tagging measurements
  - minimum $E(D^*)/E(\text{jet})$: suppress combinatorics + $b \to D^*X$ decays
- estimate remaining b-jet component using pseudo-$c\tau$ fit (before tagging)
- subtract (small) remaining) b-jet contribution using (corrected) simulation





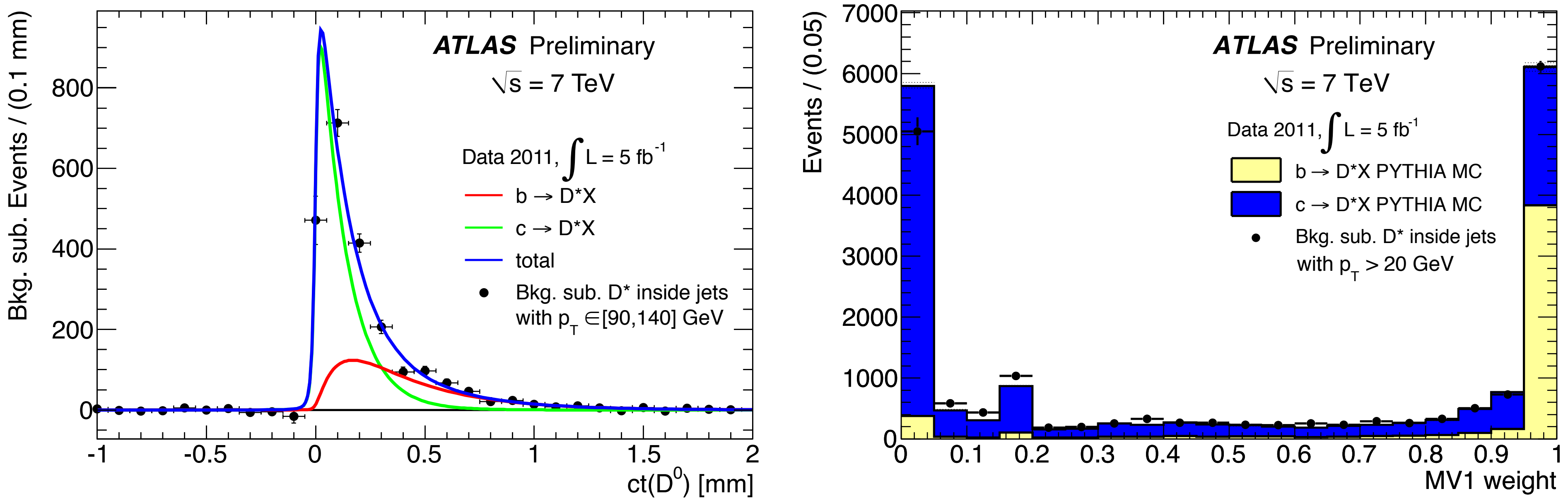
Results need to be extrapolated to inclusive charm jets

- additional systematics from charm fragmentation & decay modelling

D^*-based calibration

ATLAS: use $D^{*+} \rightarrow D^0(\rightarrow K^- \pi^+)\pi^+$ decays to remove light-flavour jets

- vertexing but without track impact parameter cuts / fit χ^2 cuts to minimise biases on tagging measurements
  - minimum $E(D^*)/E(jet)$: suppress combinatorics + $b \rightarrow D^*X$ decays
- estimate remaining b-jet component using pseudo-cτ fit (before tagging)
- subtract (small) remaining b-jet contribution using (corrected) simulation



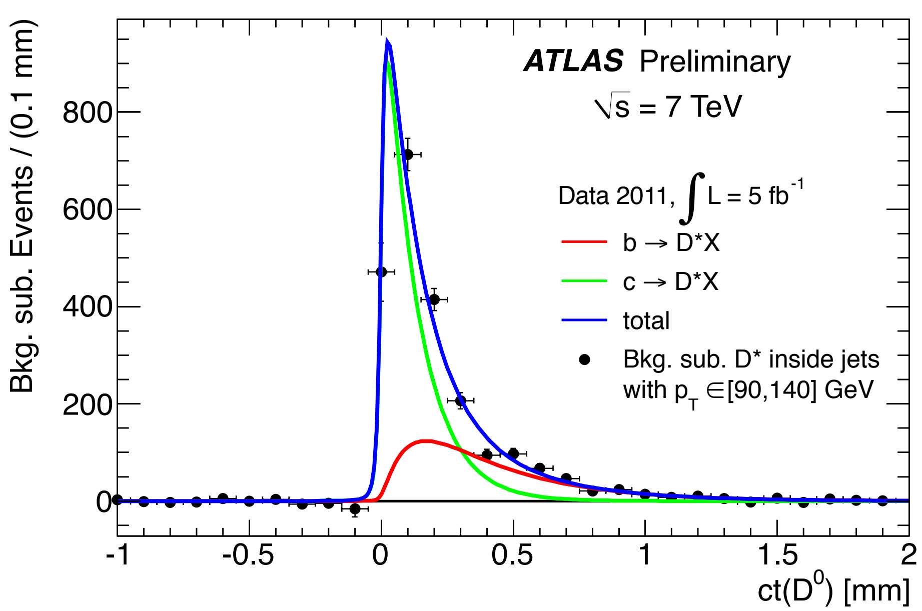
Results need to be extrapolated to inclusive charm jets

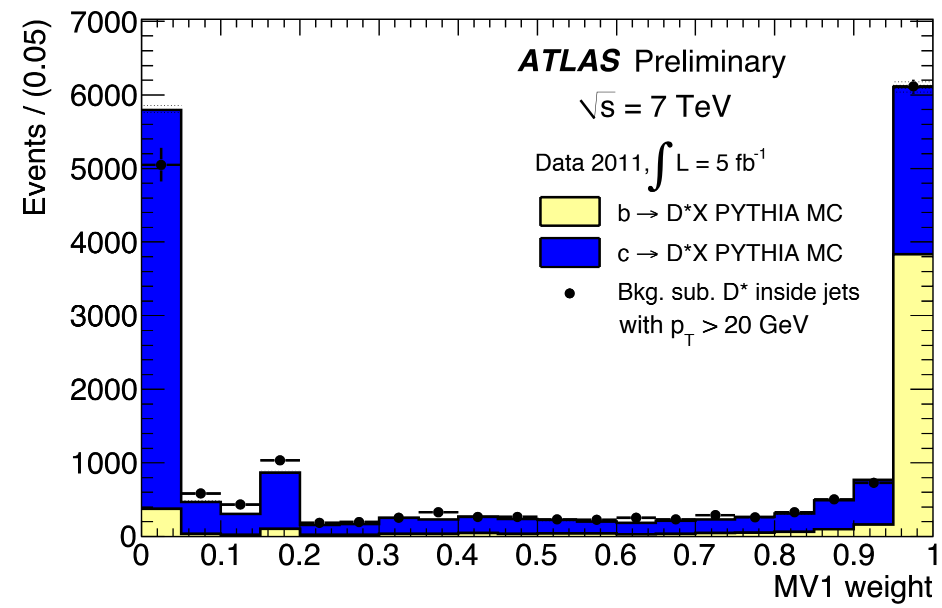
- additional systematics from charm fragmentation & decay modelling

D^*-based calibration

ATLAS: use $D^{*+} \to D^0(\to K^-\pi^+)\pi^+$ decays to remove light-flavour jets

- vertexing but without track impact parameter cuts / fit χ^2 cuts to minimise biases on tagging measurements
  - minimum $E(D^*)/E(\text{jet})$: suppress combinatorics + $b \to D^*X$ decays
- estimate remaining b-jet component using pseudo-cτ fit (before tagging)
- subtract (small) remaining b-jet contribution using (corrected) simulation





Results need to be extrapolated to inclusive charm jets

- additional systematics from charm fragmentation & decay modelling

Mis-tag rate calibration

Negative-tag calibration: exploit symmetry of track resolution functions

- ... and the availability of d < 0 tracks, L < 0 secondary vertices

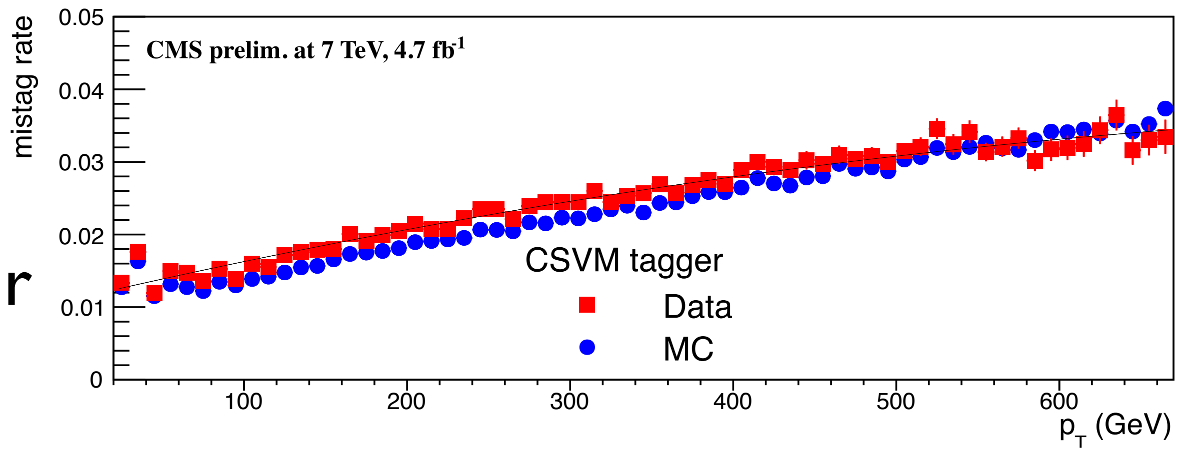
Determine negative-tag rate in data, correct for sources of asymmetry (MC)

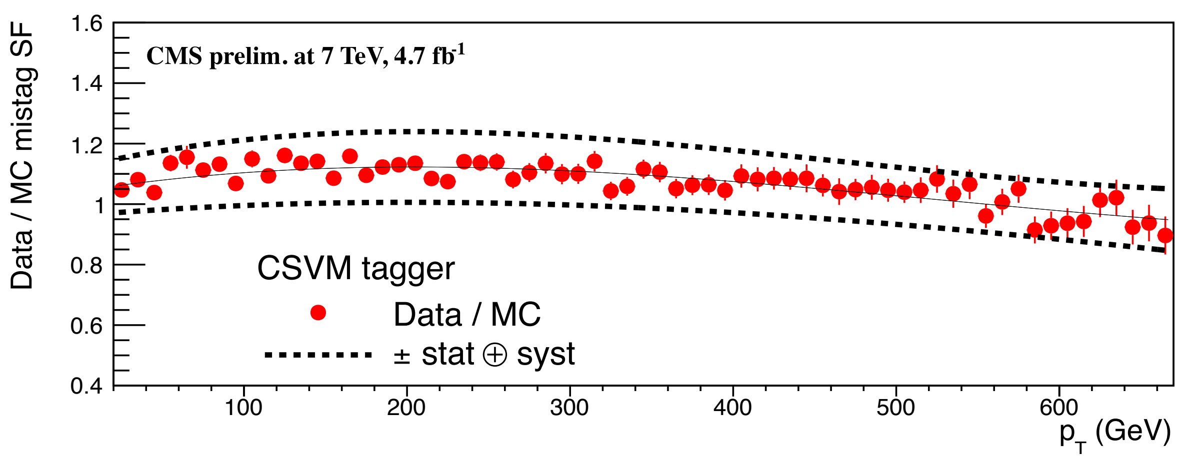
- heavy-flavour contamination of negative-tag data
- long-lived hadrons in light-flavour jets

$$\epsilon = \epsilon^{-} \cdot R_{hf}^{-} \cdot R_{ll}$$

Substantial MC dependence

- but can use e.g. measured heavy-flavour fractions, strange particle production rates; quality of detector simulation
- CMS example: good agreement over a large p_T range





Summary & Outlook

Recent years have seen significant advances in b-tagging:

- Tevatron experiments have continued the improvements of their b-tagging algorithms
- The LHC experiments benefit greatly from detector advancements, very good detector simulations, and much higher (tt) cross sections

As a result, powerful and robust tagging algorithms exist, and their performances are characterised by a wide variety of calibration methods

The (LHC) work isn't done

- the next challenge will be to accommodate the further luminosity increase after the upcoming shutdown