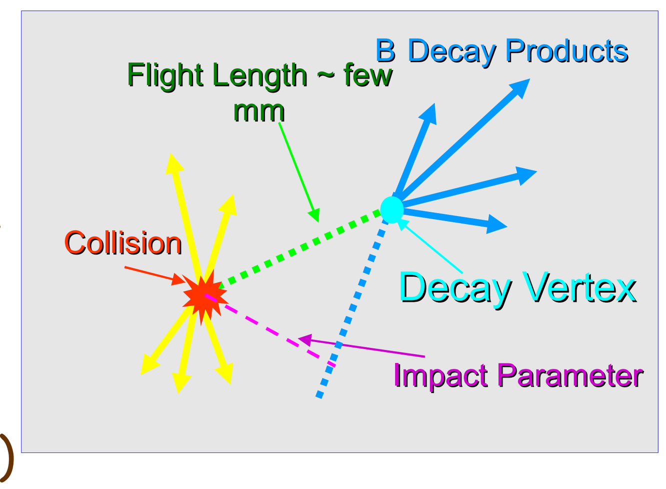
Gennadiy Kukartsev
Brown University
on behalf of
the CMS Collaboration

Strategies for b-tagging calibration using data at CMS

b jet identification (tagging)

Exploit (weak) decay and production properties of b-hadrons:

- Lifetime: secondary (tertiary) decay vertex; displaced tracks with large IP
- High mass (~5.2 GeV) and decay multiplicity (~5 charged tracks)
- Decay kinematics
- Hard b-quark fragmentation function:
  - decay products with larger p_T relative to b hadron flight direction
  - Semi-leptonic decays (per lepton flavour: BR ~ 11%, ~20% incl. cascade)

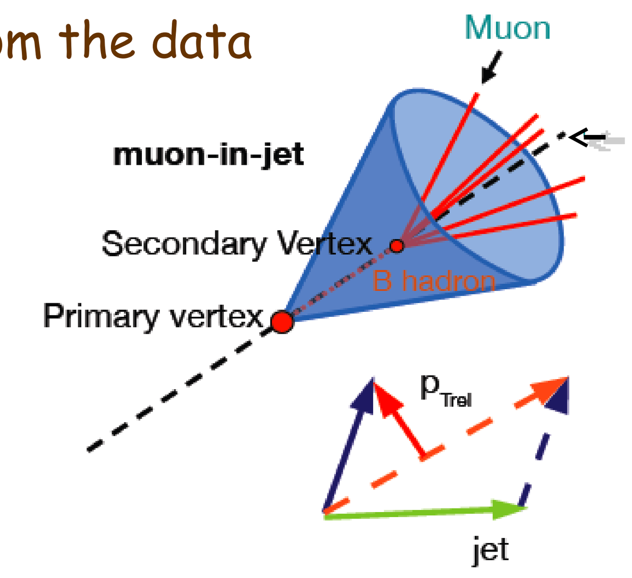


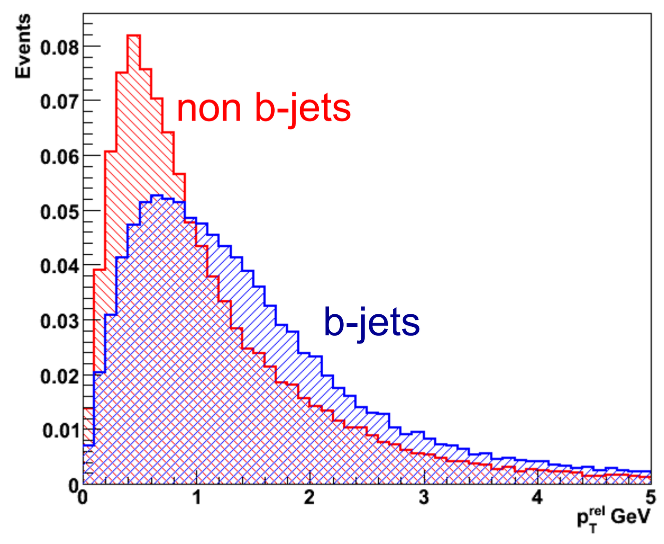
Performance measurements

- Tagging efficiencies and mistag rates from the data
- Calibration of input objects
- Efficiency measurements
  - "System 8" and p_{Trel} methods

$$p_{Trel} = \frac{\vec{p}_\mu \times (\vec{p}_\mu + \vec{p}_{jet})}{\vec{p}_\mu + \vec{p}_{jet}}$$

  - Using Likelihood ratio and top events
- Measurement of light quark tagging efficiency using multi-jet events and negative impact parameter tags
- Effects of detector misalignment





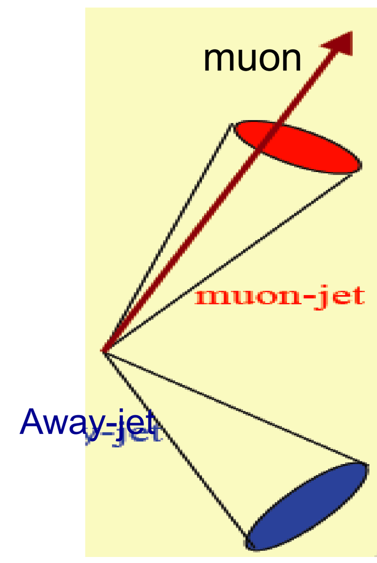
p_{Trel} Method

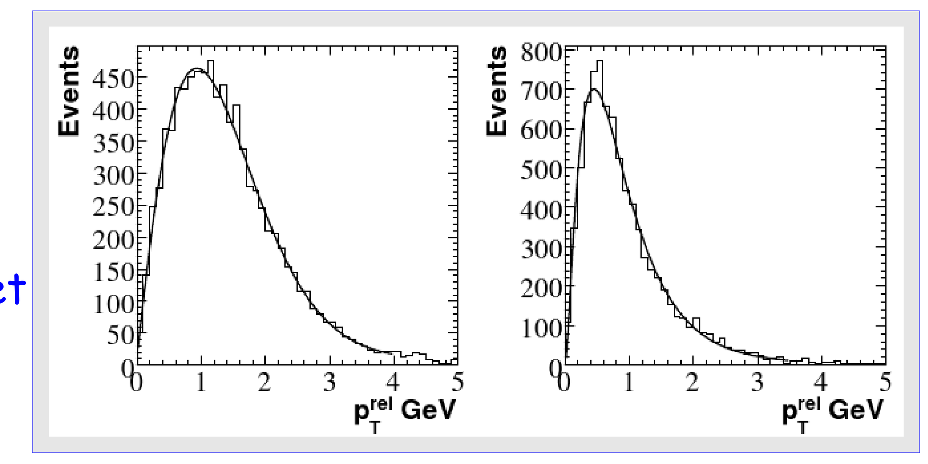
CMS PAS BTV-07-001

- Consider a dataset
  - two reconstructed jets where one jet has a muon
- Subset, where the muon jet is tagged
- Use p_{Trel} template fit to find the amount of
  - b jets with muons (n_b) – from the main set
  - Tagged b jets with muons (n_b^{tag}) – from the subset
- b-tagging efficiency will be

$$\epsilon_b = \frac{n_b^{tag}}{n_b}$$

- This is measured as a function of jet pseudorapidity and transverse momentum
- The main systematical uncertainty comes from the dependence on the Monte Carlo for the p_{Trel} templates

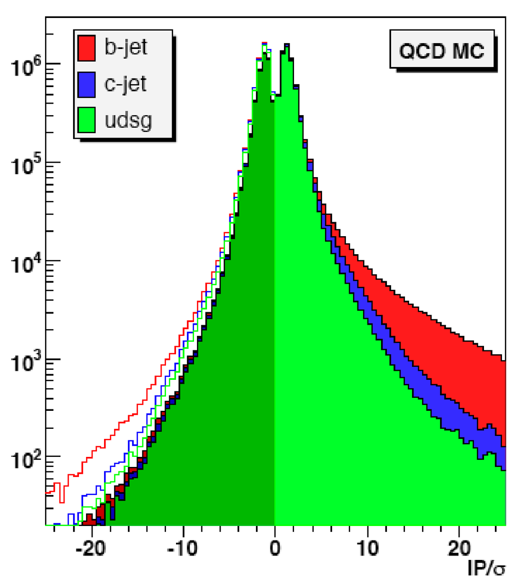


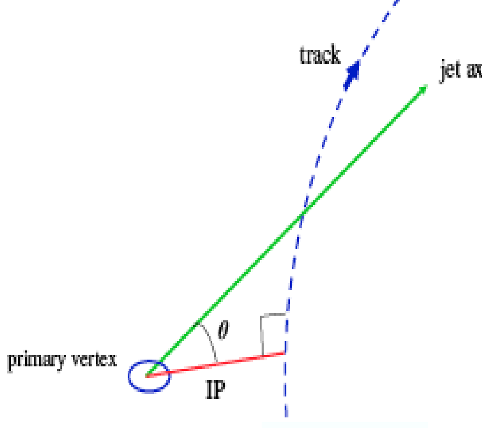


Mis-tag rate measurement using negative tags

CMS PAS BTV-07-002

- The method extracts rejection rates from light quarks looking at tracks with negative impact parameter, and using these distributions to model the mistag rate due to detector effects like resolutions and badly reconstructed tracks
- Uses the lifetime tagger a





The Impact Parameter of a track is positive (negative) if angle θ between the IP direction and the jet axis is smaller (larger) than 90°

- Distribution of negative and positive discriminators should be approximately symmetric for uds and gluon jets.
- However light udsg-jets are also affected by displaced processes such as long lived particles, hadronic interaction with the material and displacements originating from fake and badly reconstructed tracks.

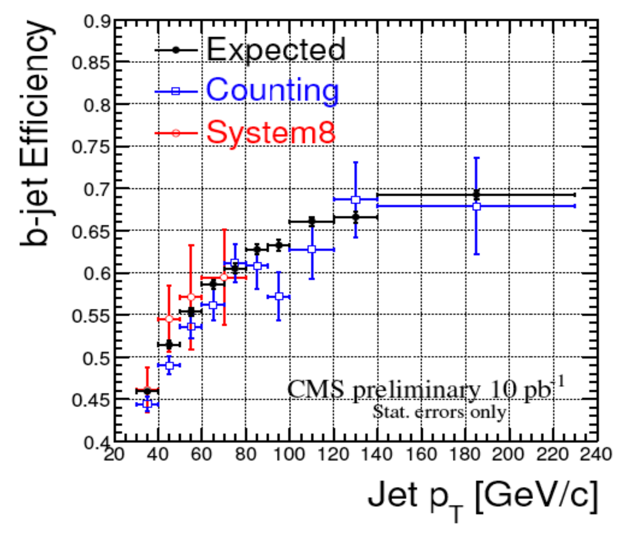
System 8

CMS PAS BTV-07-001

- Only correlation factors between taggers depend on MC
- Two independent taggers
- Two different data sets
  - As in the counting method
- Two jet categories: b and non-b (cl)
- 8 equations, 8 unknowns
- Solve numerically and get b-tagging efficiency as a function of jet pseudorapidity and transverse momentum

$$n = n_b + n_{cl}$$
$$p = p_b + p_{cl}$$
$$n^{tag} = \epsilon_b^{tag} n_b + \epsilon_{cl}^{tag} n_{cl}$$
$$p^{tag} = \beta \epsilon_b^{tag} p_b + \alpha \epsilon_{cl}^{tag} p_{cl}$$
$$n^{\mu} = \epsilon_b^{\mu} n_b + \epsilon_{cl}^{\mu} n_{cl}$$
$$p^{\mu} = \delta \epsilon_b^{\mu} p_b + \gamma \epsilon_{cl}^{\mu} p_{cl}$$
$$n^{tag,\mu} = \kappa_b \epsilon_b^{tag} \epsilon_b^{\mu} n_b + \kappa_{cl} \epsilon_{cl}^{tag} \epsilon_{cl}^{\mu} n_{cl}$$
$$p^{tag,\mu} = \kappa_b \beta \delta \epsilon_b^{tag} \epsilon_b^{\mu} p_b + \kappa_{cl} \alpha \gamma \epsilon_{cl}^{tag} \epsilon_{cl}^{\mu} p_{cl}$$

Left-hand side - observables
Correlation factors (greek) – from MC
All the rest is solved for



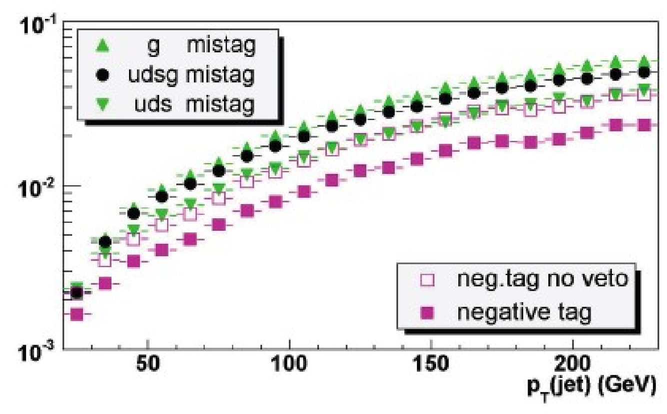
The uncertainty is 15% for 10 pb^{-1}, 10% for 100 pb^{-1} and 5-6% for 1 fb^{-1}.
The p_{Trel} method is dominated by systematic uncertainty and the System8 by statistical.

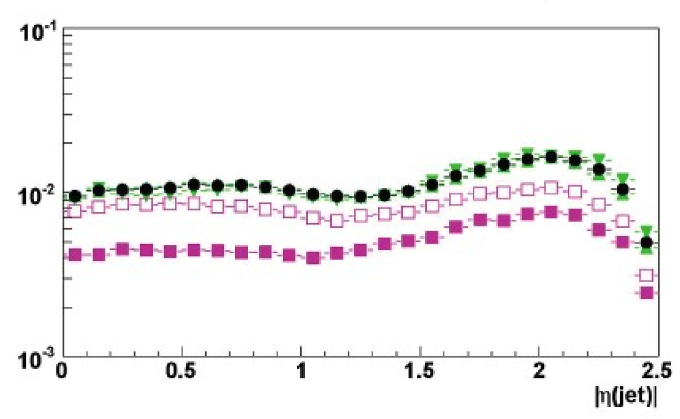
Mistag rate measurement using negative tags

$$\epsilon_{data}^{mistag} = \epsilon_{data}^{-} \cdot R_{light}$$

ϵ_{data}^{-} is the negative tag rate in multi-jet data

$$R_{light} = \epsilon_{MC}^{mistag} / \epsilon_{MC}^{-}$$





- R_{light} is the ratio between the mistag efficiency of *udsg*-jets and the negative tag rate for all jets in simulation
- The evaluation of the mistag efficiency is sensitive to the fractions of c and b quarks in the negative tag jet sample. The *c* and *b* fractions can be significantly reduced by applying a positive tag veto: the negative tag jet is rejected if it has any track with $IP/\sigma_{IP} > 4$

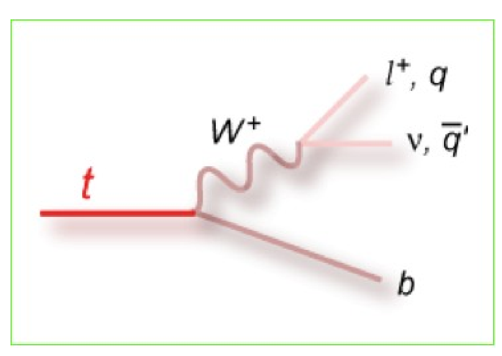
Top Quark based: Likelihood Ratio Method

The method is identical for both semileptonic and fully leptonic decays of ttbar.

- Measure the *b*-jet performance from ttbar events by isolating a jet sample enriched in *b*-jet content using a likelihood ratio

$$\mathcal{L} = \prod_i \frac{f_i(x_i)}{1 - f_i(x_i)}$$

- Build a multivariate classifier for jets.
- Use the classifier to select *b* jets.



Extract the efficiency:

$$\epsilon_b = \frac{1}{x_b}\left[x_{tag} - \epsilon_0 (1 - x_b)\right]$$

fraction of b jets
fraction of tagged jets
mistag rate