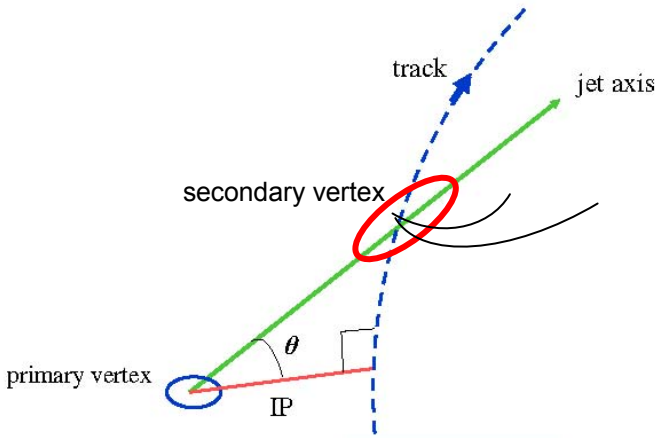
b-tagging from startup to 1 fb^{-1}



- b-hadron properties allow to distinguish b-jets from light (udsg) jets: high mass, long lifetime, large energy fraction, 40% semileptonic decay into soft muon or electron ⇒ secondary vertex and tracks at large impact parameter (IP)
- but mistags due to many processes: V^0 decays (K^0_s and Λ), photon conversion, nuclear interactions, fake tracks
- charm jets contamination is difficult to resolve

b-tagging is at the very end of reconstruction software:

- track reconstruction efficiency and fakes
- track IP resolution
- primary vertex resolution
- jet axis resolution
- soft muon and electron identification

b-tag algorithms

Impact parameter taggers: use 3D IP significance IPs = IP/σ_{IP}

- Track Counting "high efficiency" (2^{nd} highest IPs) or "high purity" (3rd highest IPs)
- Jet Probability: using all tracks IPs or the 4 highest IPs (jet B probability)

Secondary vertex (SV) taggers:

- Simple Secondary Vertex (flight distance $\Delta L/\sigma_{\Delta L}$)
- Combined Secondary Vertex : SSV + IPs and track kinematics if no rec. SV in a Log Likelihood Ratio

Soft lepton taggers: muon or electron

- can use various quantities: IP, p_{Trel} / jet, …, in a combined Likelihood

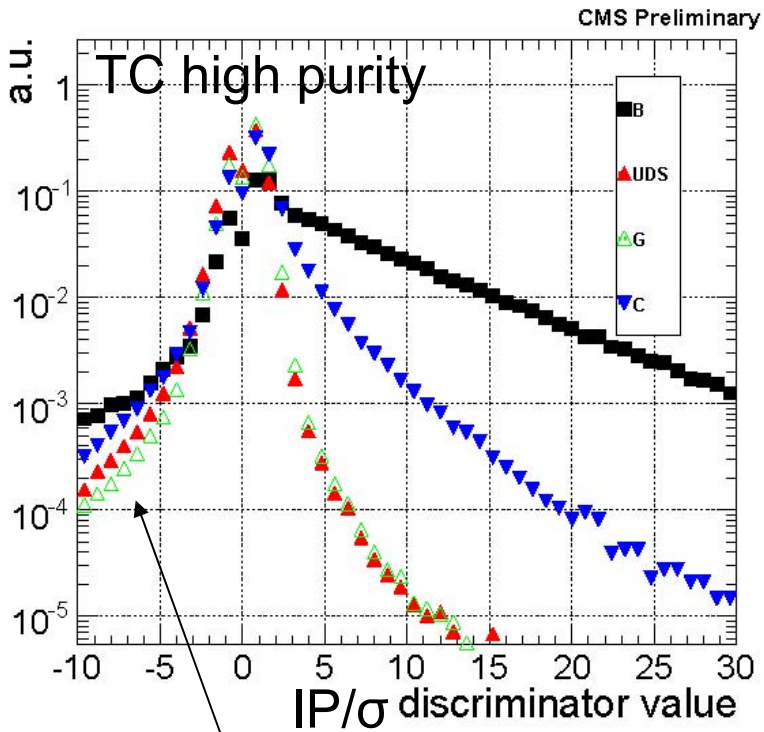
Combined MVA tagger

- include jet probability, secondary vertex, soft leptons discriminants

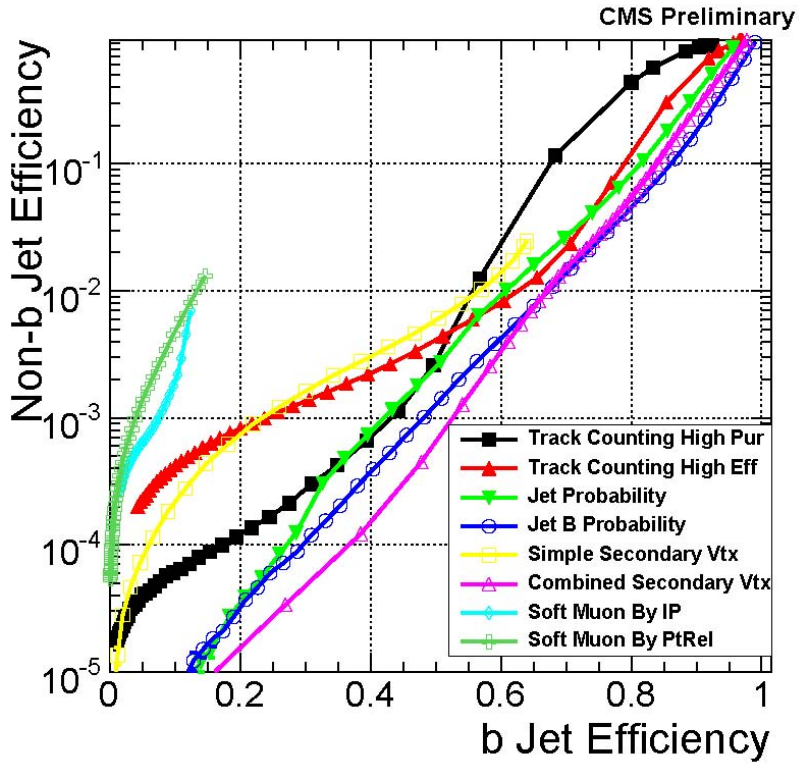
At startup will use the simplest ones: Track counting (maybe Jet Probability)
Simple Secondary Vertex,
Soft muon

expected b-tag performance (1)

- with ideal simulation (BTV-09-01)

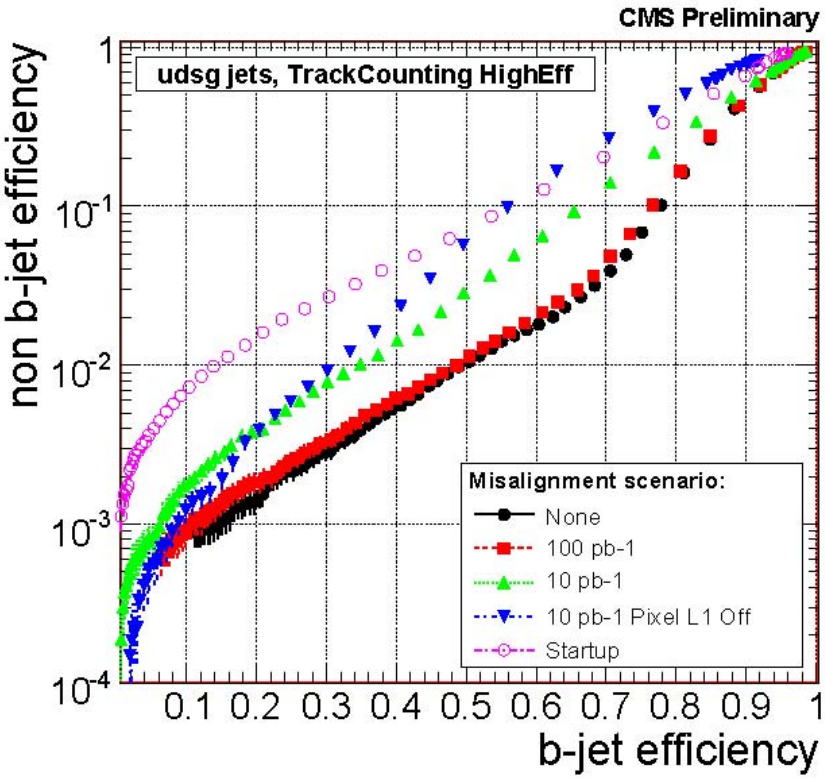


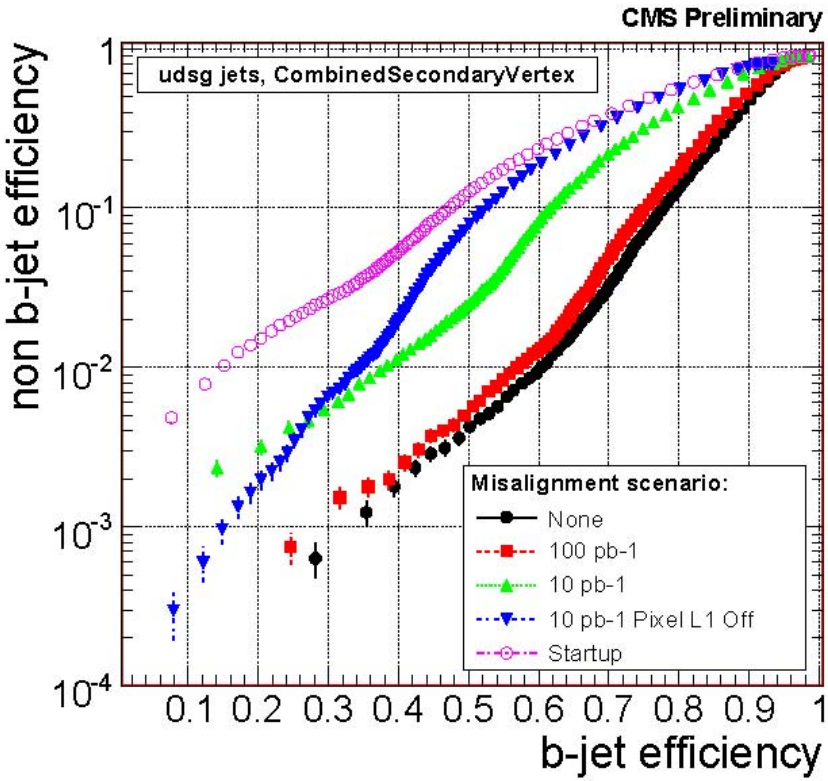
negative IP/σ if tracks "upstream" from the primary vertex



expected b-tag performance (2)

- Sensitivity to alignment (BTV-07-003)





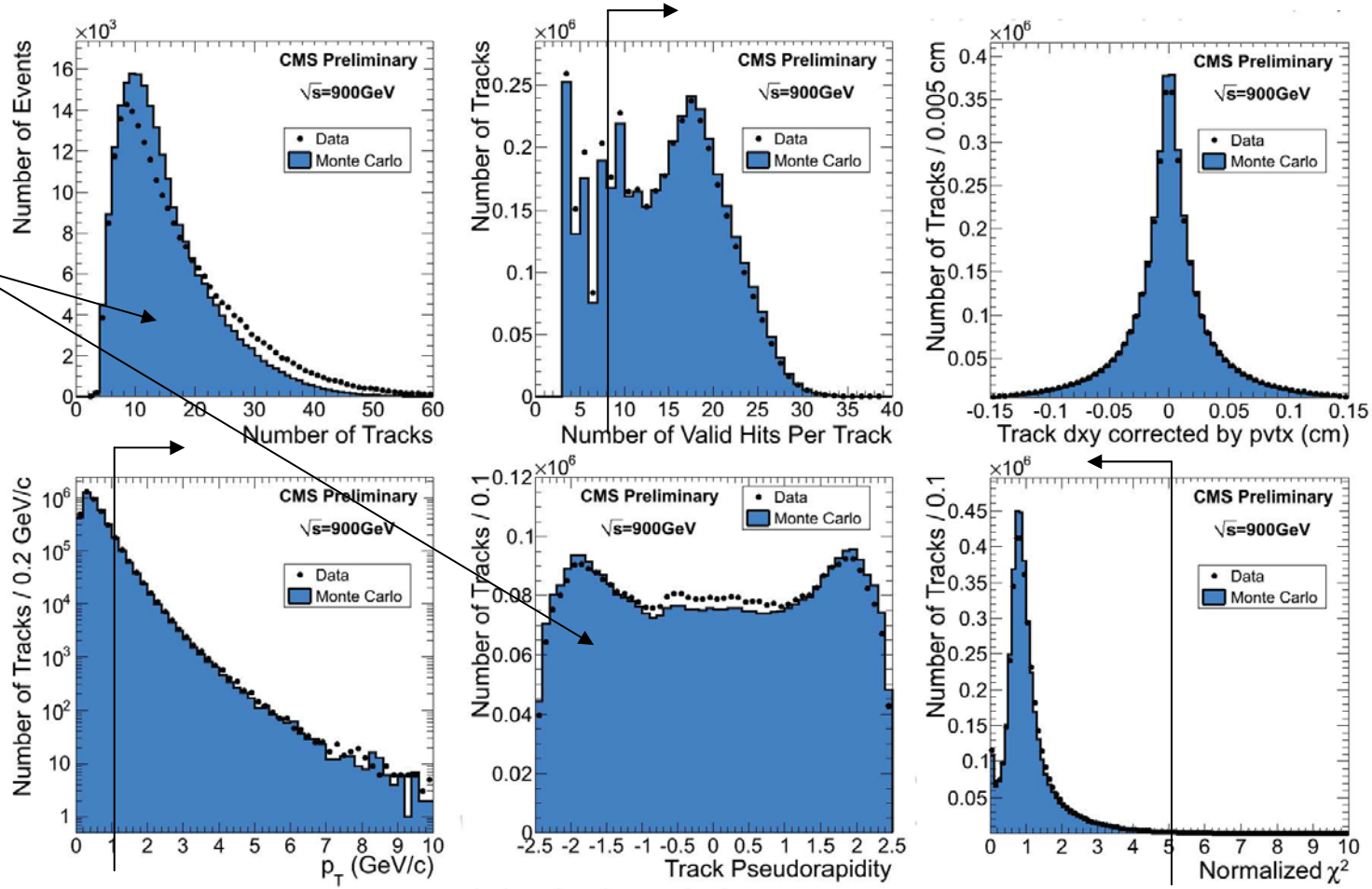
less sensitive for "simple" algorithms

but note that 100 pb^{-1} is already almost ideal !

tracking/btag commissioning

- Striking performance of tracking and vertexing (prelim. TRK-10-001)

imperfect D6T Pythia tuning



primary vertex reconstruction

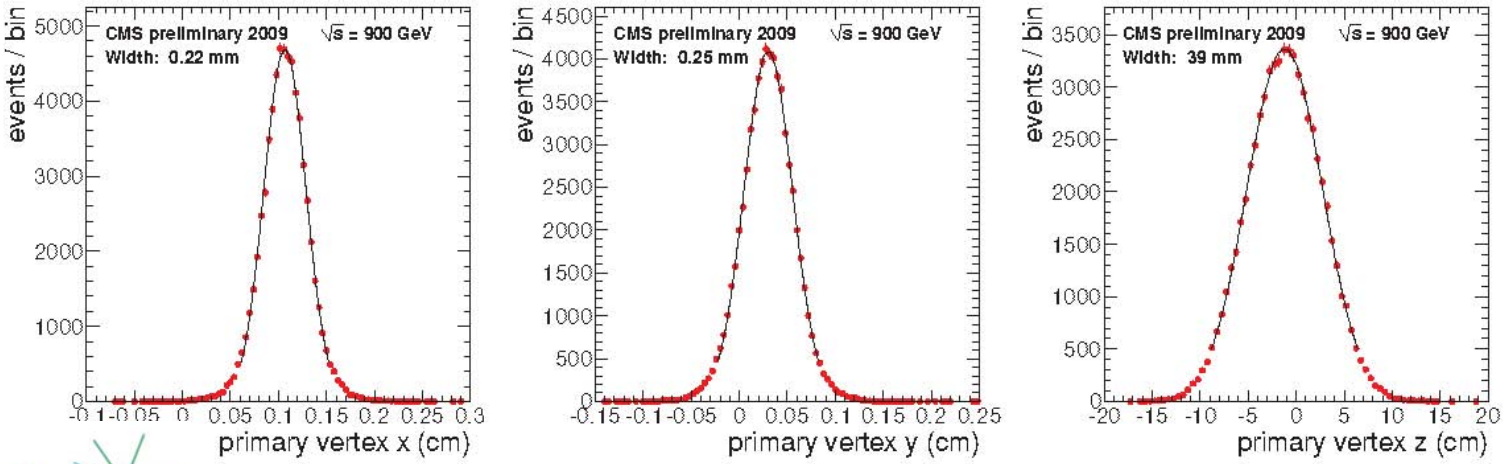


Figure 2: Primary vertex distributions from a single run.

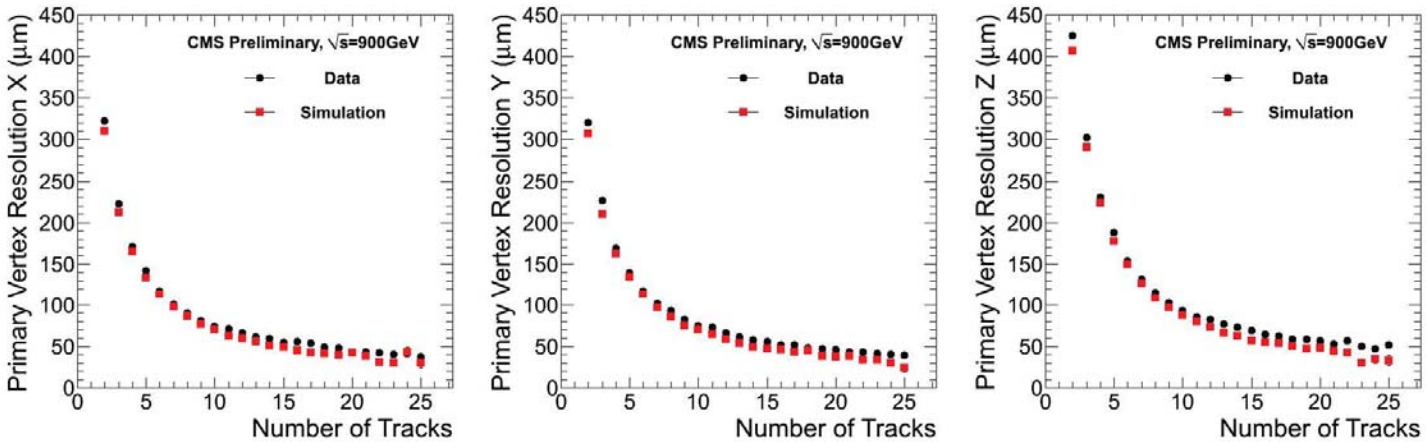


Figure 4: Primary vertex resolution as a function of the number of tracks used in the fitted vertex.

V^0 resonances

important test of (2 tracks) secondary vertex ability and momentun reconstruction

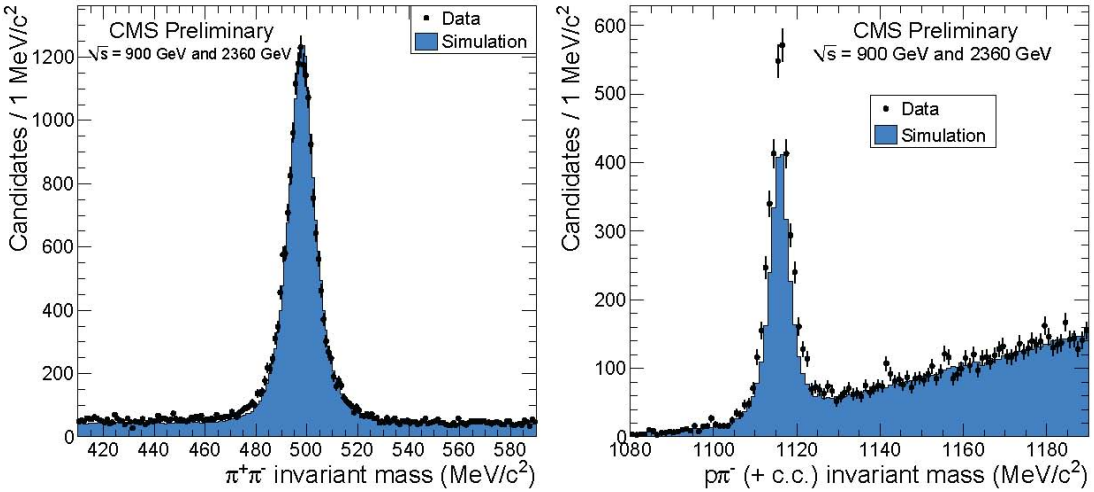
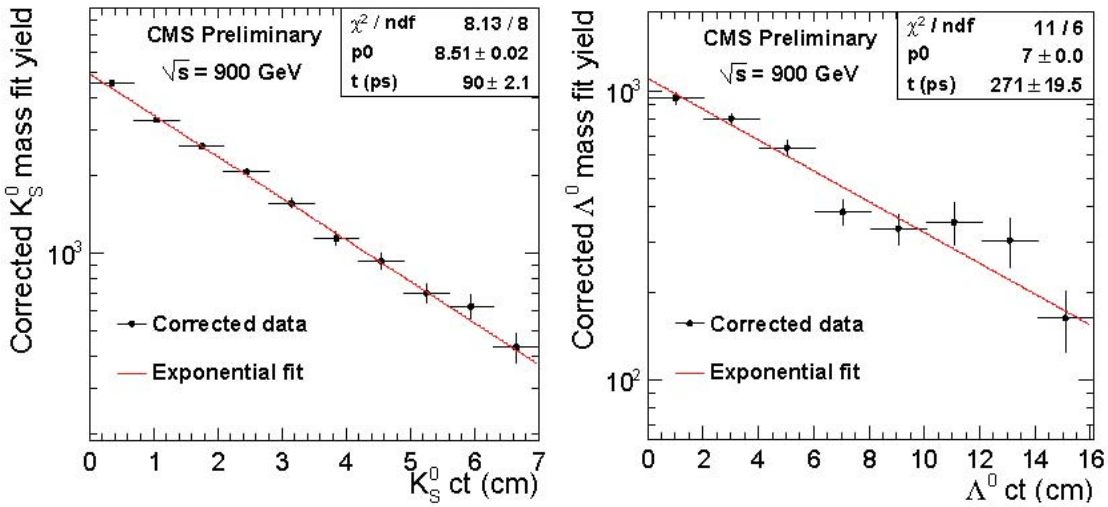


Figure 12: Yield and signal-to-background comparison of K^0_S and Λ^0 between data and simulation. Both simulation histograms are scaled by the ratio of K^0_S yields.

the observed V^0 multiplicity enters in the mistag systematics : *data/MC discrepancies to be studied at 7 TeV*

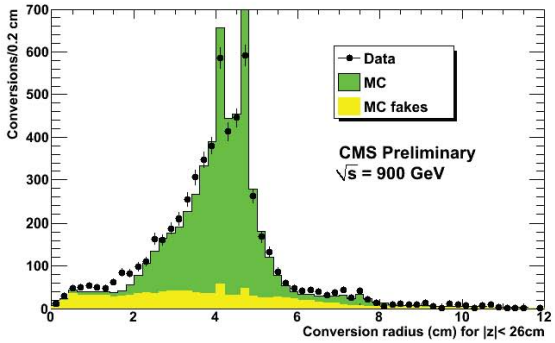


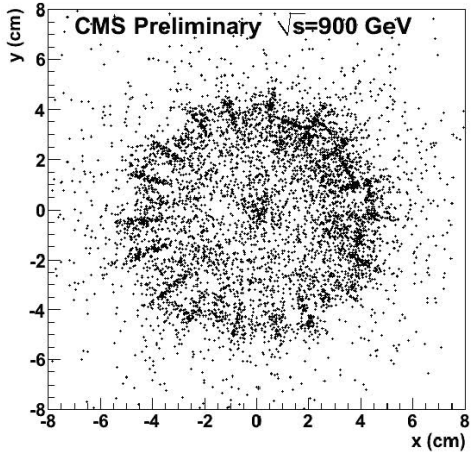
nice decay proper time reconstruction !

secondary interactions

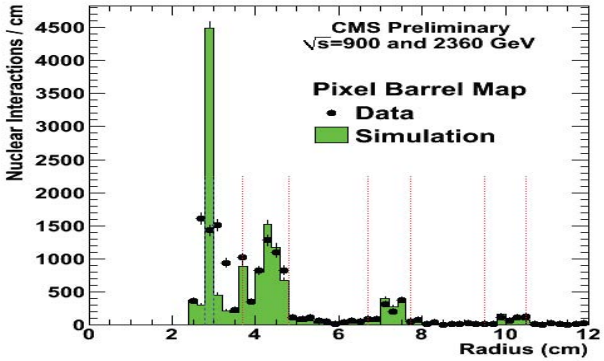
These are also an input to the mistag systematics:

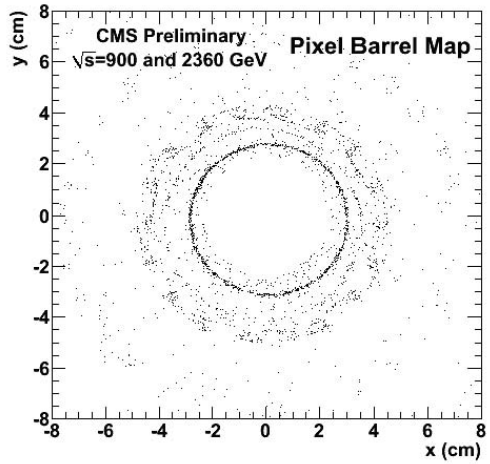
- photon conversions:





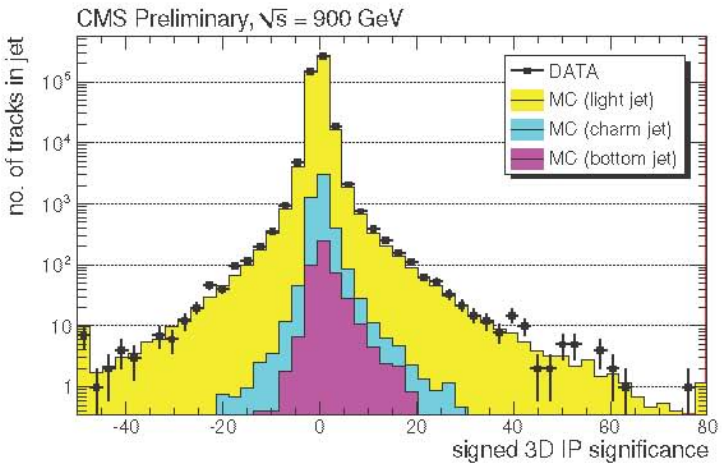
- nuclear interactions:

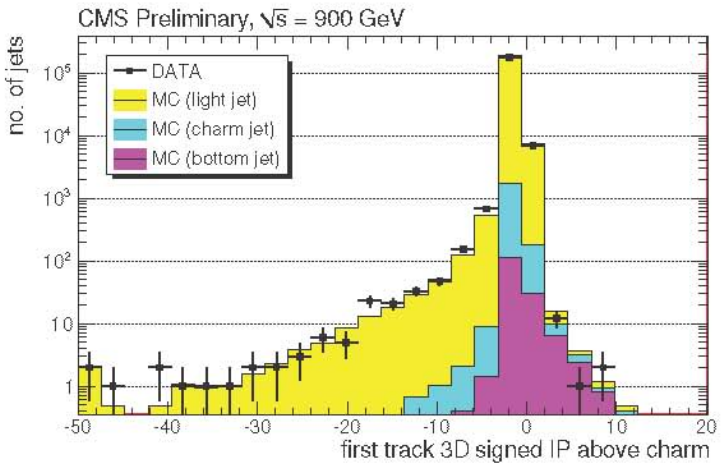


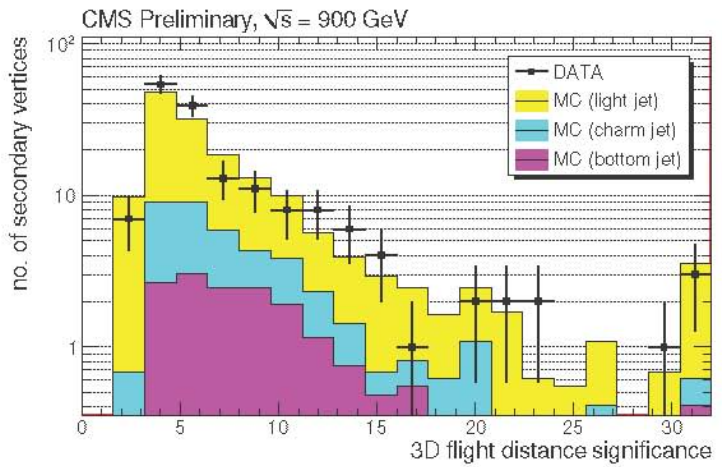


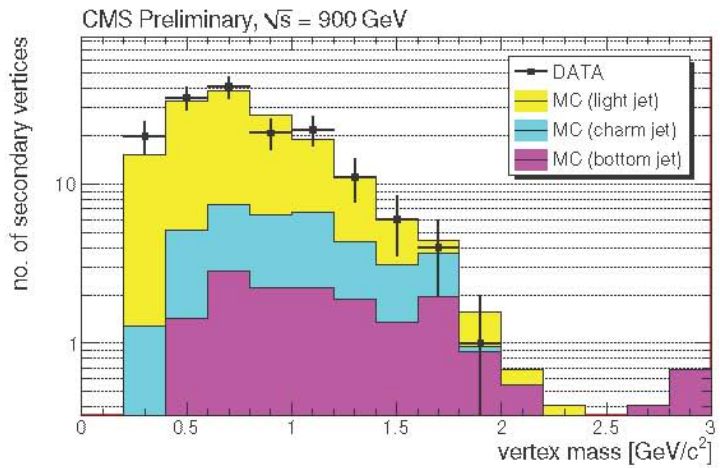
already a good agreement at ± 20% :
to be extended to 2nd and 3rd pixel layers with more statistics

900 GeV data b-tag studies



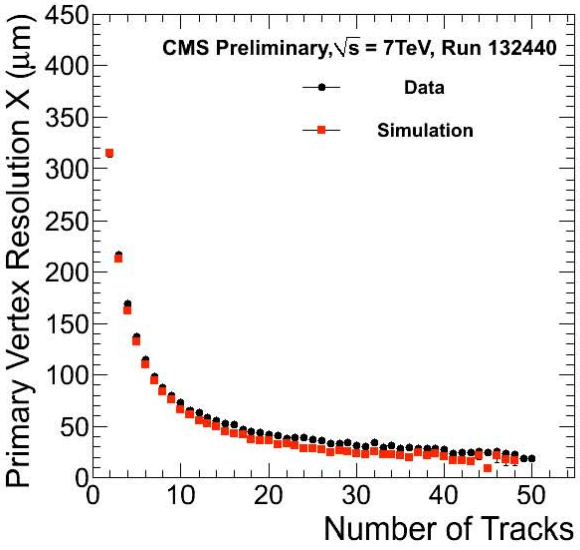


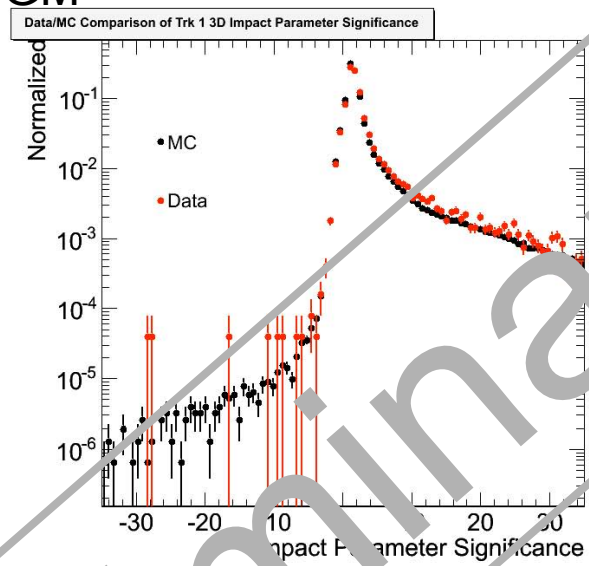


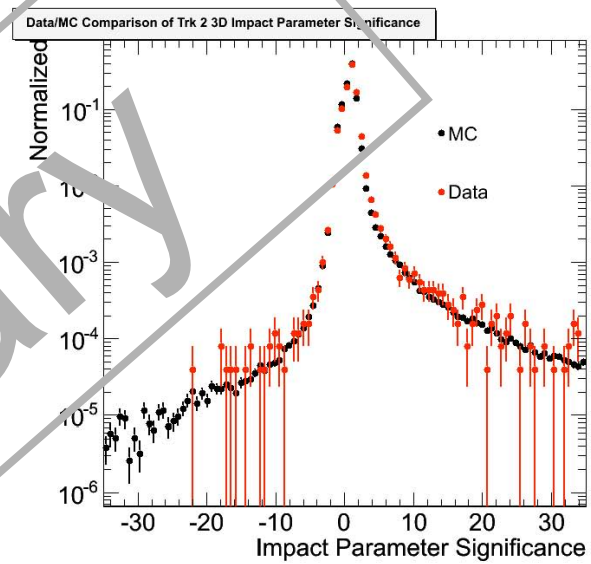


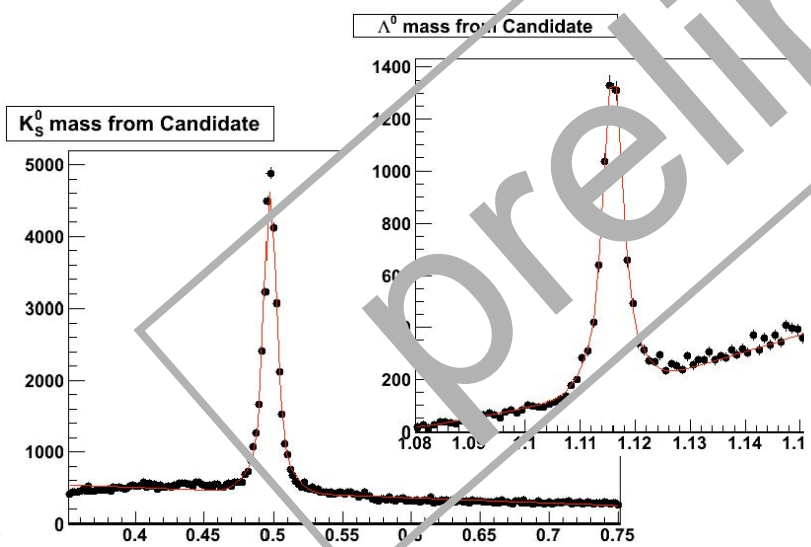
and 7 TeV data preliminary…

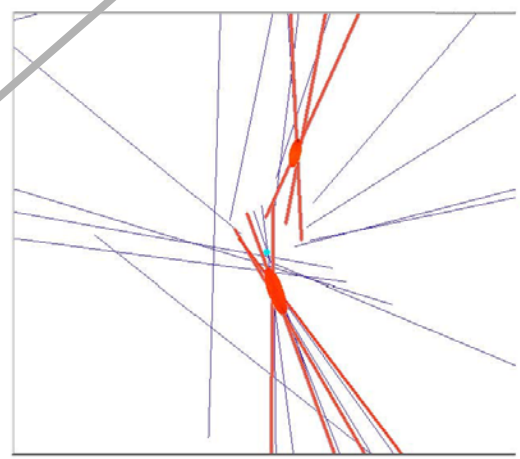
from Yanian Gao, March 31 GM

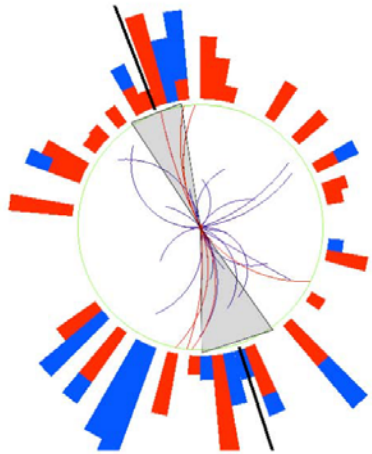






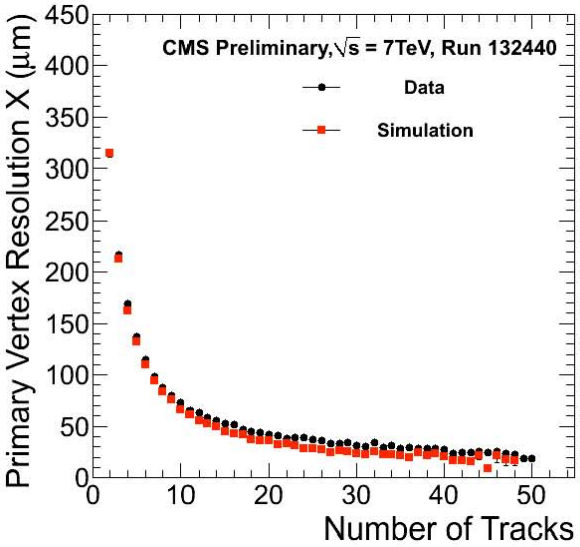


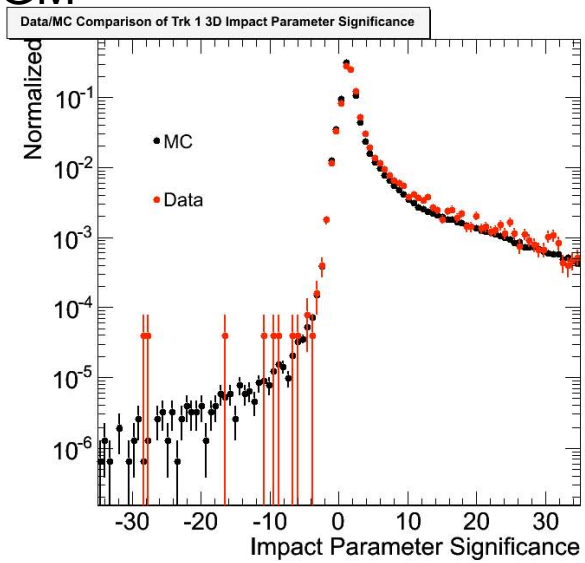


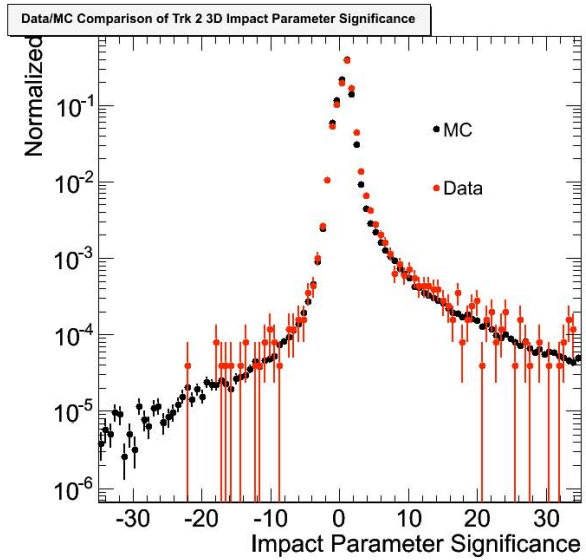


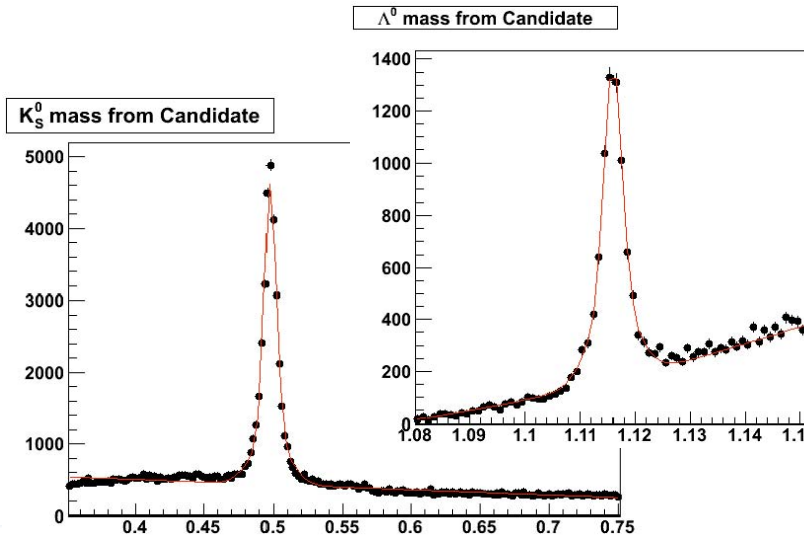
and 7 TeV data preliminary…

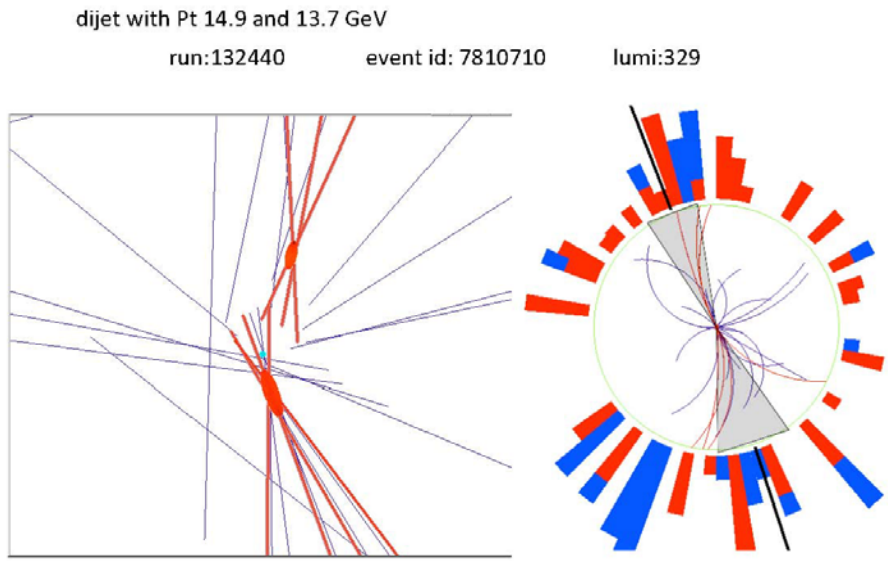
from Yanian Gao, March 31 GM











How to measure the b-tag and mistag efficiencies ?

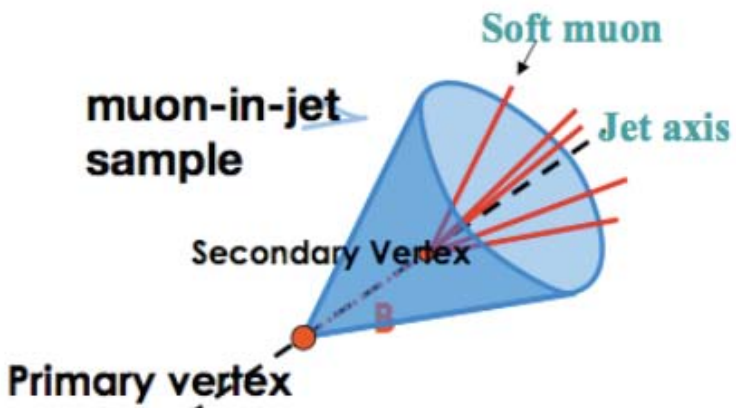
- For a given tagger, adjust the cut on the discriminant in order to get 10% (loose), 1% (medium), 0.1% (tight) mistags (according to QCD MC)
  cf. https://twiki.cern.ch/twiki/bin/view/CMS/BtagOctober09ExerciseUsePayload

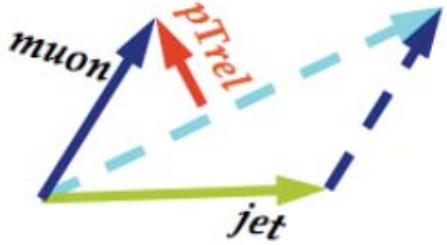
- For each tagger and each operating point,
  apply data driven methods (inspired from Tevatron) to evaluate
  the b-tag efficiency with b-enriched sample
  the mistag rate with multijet events
  and/or
  the b-tag efficiency and mistag rate data/MC scale factor
  parametrised according to
  the jet p_T and η (at least)

- This will be the job of the coming month(s) !

b-tag eff. from muon p_{Trel}

- Use jets with a soft muon (cf. BTV-07-001)





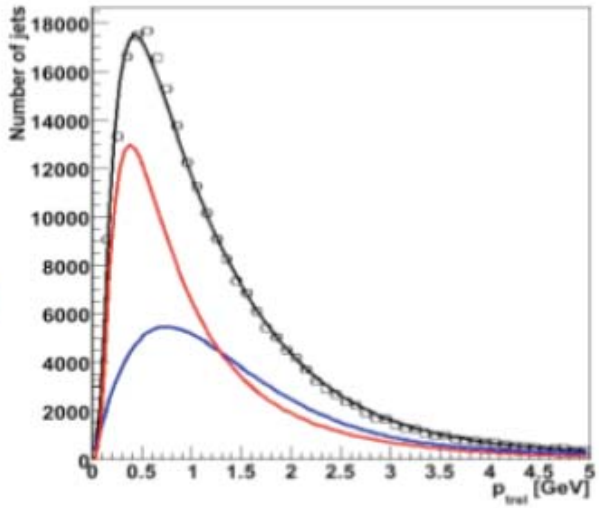
$$p_T^{rel} = \frac{\| \bar{\mu} \times (\bar{\mu} + \bar{J}) \|}{\| \bar{\mu} + \bar{J} \|}$$

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

Number of b-jets in the muon-in-jet-tagged sample

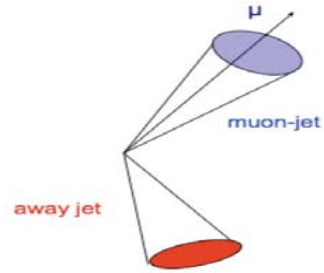
Number of b-jets in the muon-in-jet sample

Obtained by fitting



(from Francisco Yumiceva, Top.com, March 5)

b-tag eff. from SystemD



- use 2 (almost) independent b-taggers:
  lifetime tagger under study and the muon p_{Trel} cut
- use 2 samples: muon-in-jet + away jet without or with b-tagging
- allow 8 non-linear equations with 8 unknowns (⊃ b-tag efficiency)
- Need the MC to estimate some correlation parameters (⇒ systematics)

$$n = n_b + n_{cl}$$ muon-in-jet + away-jet

$$p = p_b + p_{cl}$$ muon-in-jet + tagged-away-jet

$$n^{\text{tag}} = \varepsilon_b^{\text{tag}} n_b + \varepsilon_{cl}^{\text{tag}} n_{cl}$$

$$p^{\text{tag}} = \beta\, \varepsilon_b^{\text{tag}} p_b + \alpha\, \varepsilon_{cl}^{\text{tag}} p_{cl}$$ apply "probe" tagger

$$n^{\text{pTrel}} = \varepsilon_b^{\text{pTrel}} n_b + \varepsilon_{cl}^{\text{pTrel}} n_{cl}$$

$$p^{\text{pTrel}} = \delta\, \varepsilon_b^{\text{pTrel}} p_b + \gamma\, \varepsilon_{cl}^{\text{pTrel}} p_{cl}$$ apply "tag" tagger

$$n^{\text{tag,pTrel}} = \kappa_b\, \varepsilon_b^{\text{tag}} \varepsilon_b^{\text{pTrel}} n_b + \kappa_{cl}\, \varepsilon_{cl}^{\text{tag}} \varepsilon_{cl}^{\text{pTrel}} n_{cl}$$

$$p^{\text{tag,pTrel}} = \kappa_b\, \beta\, \delta\, \varepsilon_b^{\text{tag}} \varepsilon_b^{\text{pTrel}} p_b + \kappa_{cl}\, \alpha\, \gamma\, \varepsilon_{cl}^{\text{tag}} \varepsilon_{cl}^{\text{pTrel}} p_{cl}$$ apply "tag" and "probe" taggers

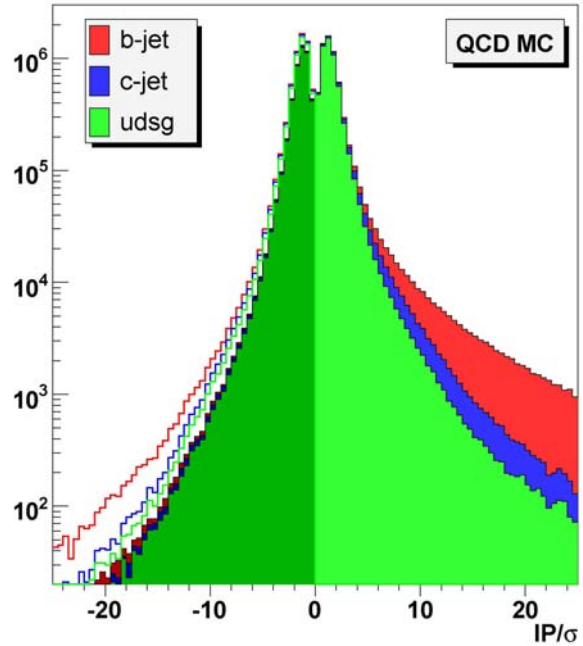
mistag rate from negative tags

- a negative tag can be defined for all taggers (cf. BTV-07-002)
- should be symetric for light udsg-jets … apart V^0, 2^{ndary} interactions, fake tracks
- systematics also account for the remaining:

  b- and c-jet contributions

  difference between uds- and g-jets

  data-MC difference in sign flip of the neg. tagger



- the mistag rate is computed from the negative tag rate as:

$\varepsilon_{data}^{mistag} = \varepsilon_{data}^{neg.tag} \times R_{light}$ where $R_{light} = \varepsilon_{MC}^{mistag}(udsg) / \varepsilon_{MC}^{neg.tag}(udsg+c+b)$

- a veto can be applied on the negative tagged jet if there is a track with a large positive IPs (>4) in this jet
  $\Rightarrow$ reduce a lot the b and c contamination on the negative side

systematics

- b-tag eff. with SystemD (BTV-07-001):

operating point	Loose			Medium			Tight		
Luminosity (pb^{-1})	10	100	1000	10	100	1000	10	100	1000
Systematics (%)									
β	5.8	5.8	2.9	6.3	6.3	3.2	5.7	5.7	2.9
α	0.4	0.4	0.2	0.4	0.4	0.2	0.4	0.4	0.2
κ_b	3.4	3.4	1.7	3.6	3.6	1.8	3.3	3.3	1.7
κ_{cl}	0.2	0.2	0.1	0.2	0.2	0.1	0.2	0.2	0.1
p_{Trel}	2.8	2.8	2.8	2.9	2.9	2.9	3.0	3.0	3.0
statistics MC (%)	2.3	2.3	2.3	2.6	2.6	2.6	2.7	2.7	2.7
statistics data (%)	7.2	2.3	0.7	8.4	2.6	0.8	8.7	2.7	0.9
Total error (%)	10.5	8.0	6.4	11.8	8.6	5.4	11.6	8.2	5.3

- udsg mistag from negative tags (BTV-07-002):

operating point	Loose			Medium			Tight		
Luminosity (pb^{-1})	10	100	1000	10	100	1000	10	100	1000
Systematics (%)									
b fraction	1.4	1.4	0.6	0.8	0.8	0.3	1.2	1.2	0.5
c fraction	0.8	0.8	0.3	0.7	0.7	0.3	1.3	1.3	0.5
g fraction	0.8	0.8	0.4	1.4	1.4	0.7	2.3	2.3	1.2
V^0 fraction	1.4	1.4	0.7	3.6	3.6	1.8	4.6	4.6	2.3
other displaced processes	1.4	1.4	0.7	3.6	3.6	1.8	4.6	4.6	2.3
IP sign flip	0.7	0.3	0.2	4.5	1.9	1.4	24.0	10.2	7.6
statistics MC	0.1	0.1	0.1	0.4	0.4	0.4	1.2	1.2	1.2
statistics data	0.4	0.1	—	1.6	0.5	0.2	5.5	1.7	0.6
sampling	2.0	2.0	2.0	5.0	5.0	5.0	13.0	13.0	13.0
Total syst.	3.4	3.4	2.4	8.8	7.6	5.9	28.7	18.1	15.5

plan for 2010 and beyond

- Will benefit from much higher statistics and high p_T jets
- Repeat and extend the data/MC comparisons of tracking and b-tagging observables: track IP, PV and SV resolutions
- Use soft muons (maybe electrons) within jets
- Optimize track, jet, vertex selections
- Measure the b-tag efficiency and mistag rate from the data
- Will need feedback from users on how to apply the b-tag efficiencies or scale factors to physics analyses
- Close collaboration with the Tracking and JETMET POGs, B, QCD and TOP PAGs