B-Tagging Algorithms for CMS Physics

Gavril Giurgiu
Johns Hopkins University
on behalf of the CMS collaboration

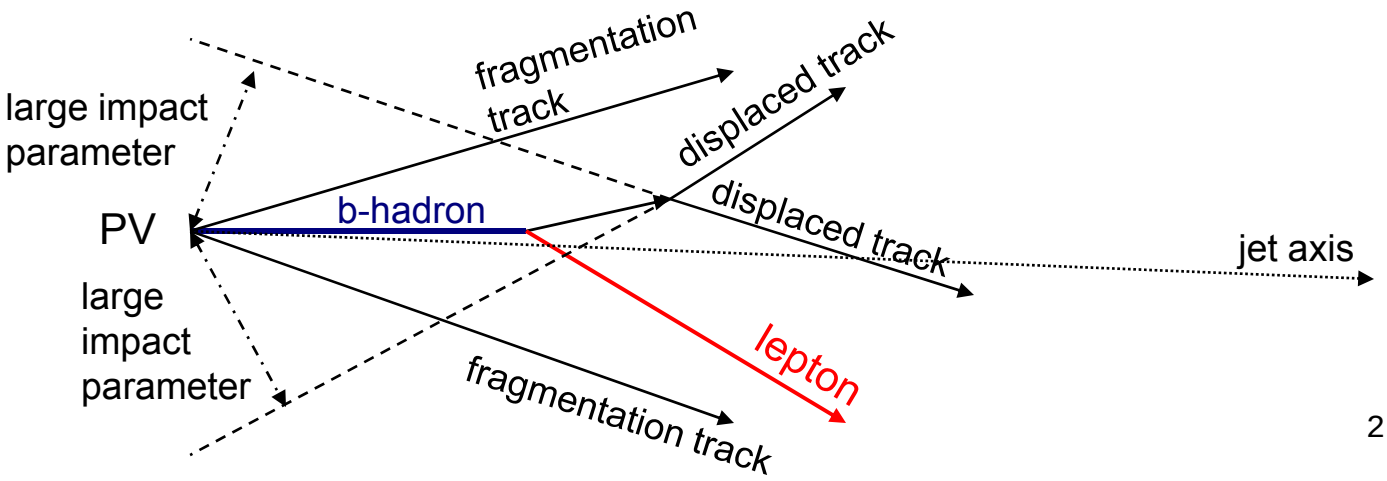
Flavor Physics and CP Violation 2009
Lake Placid, NY
May 27 – June 1, 2009

Introduction

- B tagging algorithms attempt to identify (tag) jets produced by b-quarks

- Identifying b-quarks is crucial for analyses of top quarks, Higgs, SUSY,…

- Separate b-jets from light (u,d,s,g) jets using specific b-hadron properties:
  - large lifetime ~1.5 ps (large decay length: 20 GeV b-hadron decays after ~2 mm)
    - search tracks or vertexes displaced w.r.t. primary vertex
  - large mass ~5 GeV
    - search leptons, from semileptonic B decays, with large transverse momentum w.r.t. jet axis



Impact Parameter

- Large B hadron lifetime → large impact parameter,d_0, of B decay products

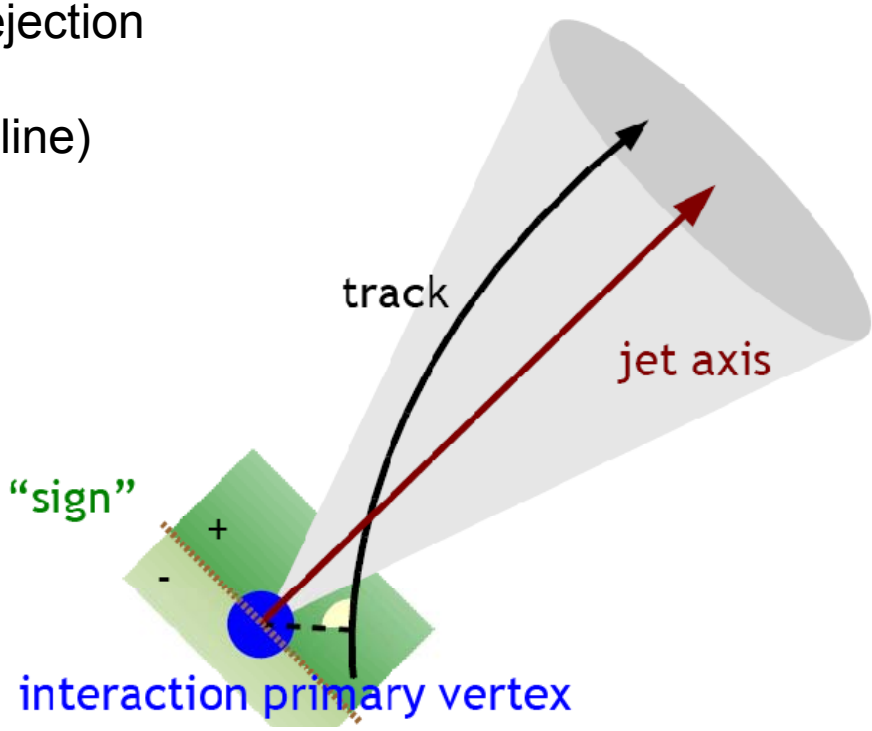
- Search for tracks displaced w.r.t. primary vertex

- Loose track quality cuts → good fake b rejection

- Use either 2D (transverse plane to beam line)
  or 3D impact parameter

- Use “sign” of the impact parameter:
  - positive if track intersection with jet axis is downstream the PV along jet direction

- Impact parameter significance, $d_0/\sigma d_0$, used as discriminant between signal (true b-jets) and background (fake b-jets)

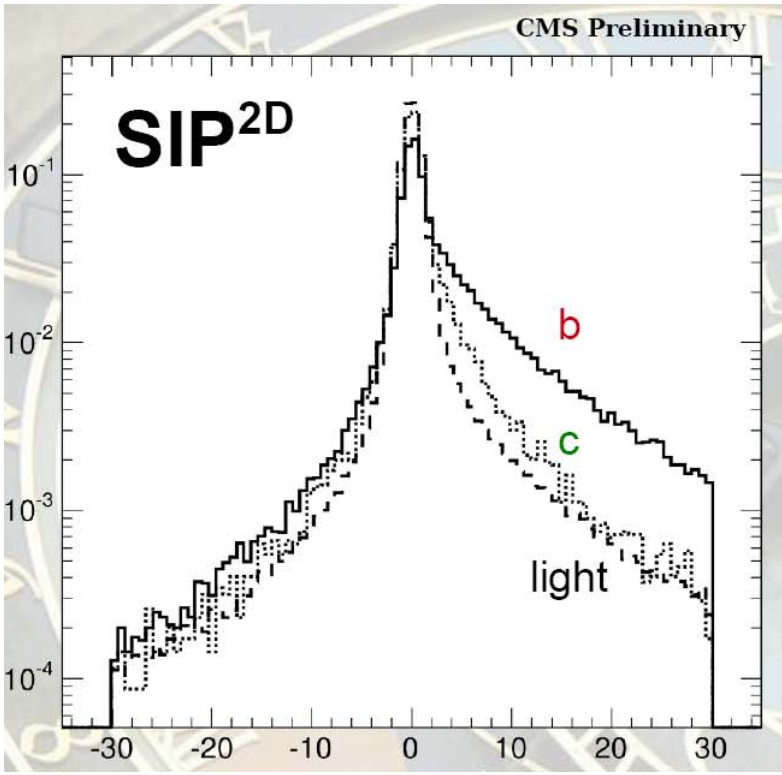


Signed Impact Parameter

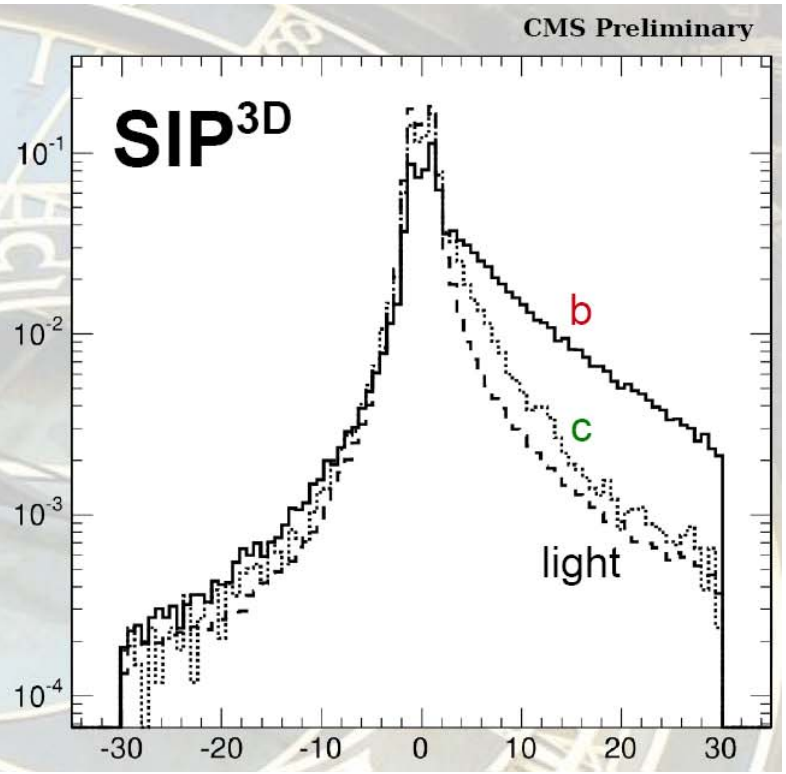
- Different signed impact parameter distributions for bottom, charm or light jets

- Jets with transverse momentum of 80 - 120 GeV
  - selected from generic QCD sample

transverse (2D)



three dimensional

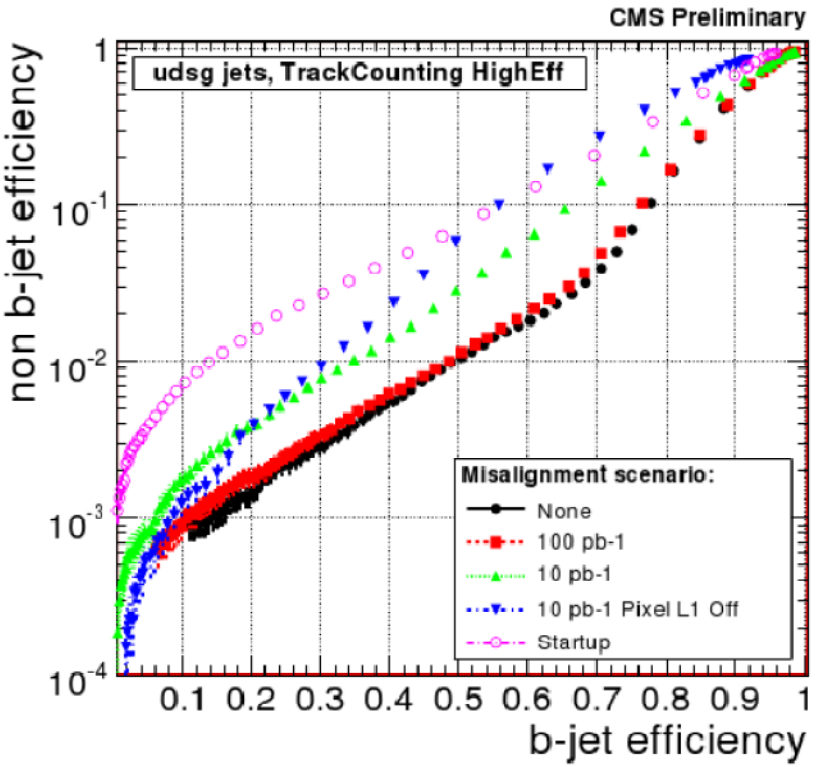


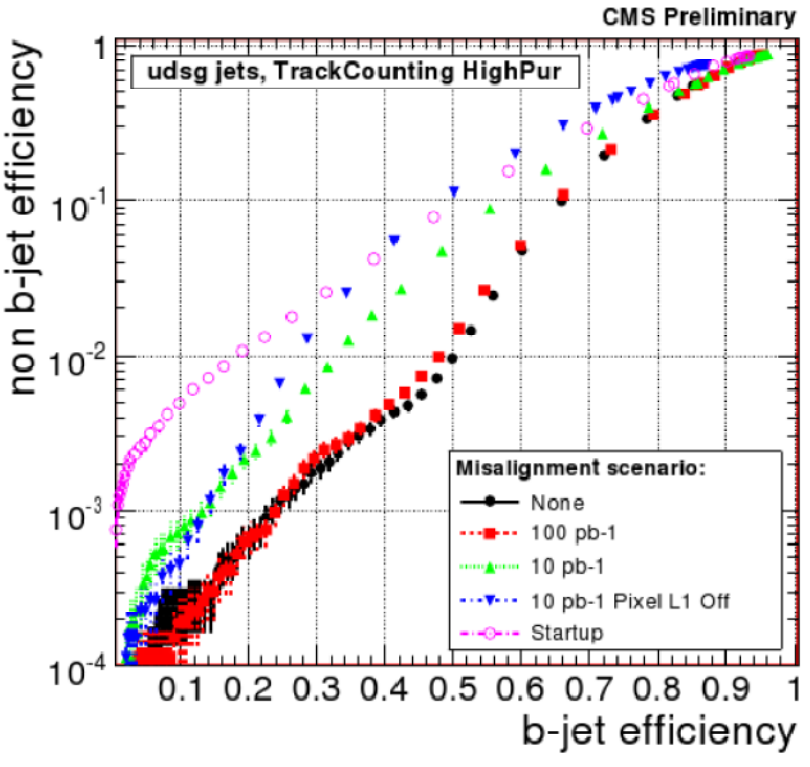
Track Counting B-Tagging

- Use 3D signed impact parameter as discriminator

- Require at least 2 (high efficiency) or 3 (high purity) tracks with
  3D sign IP larger than certain value

- Results on full simulation of $t\bar{t}$ decays with different alignment scenarios:



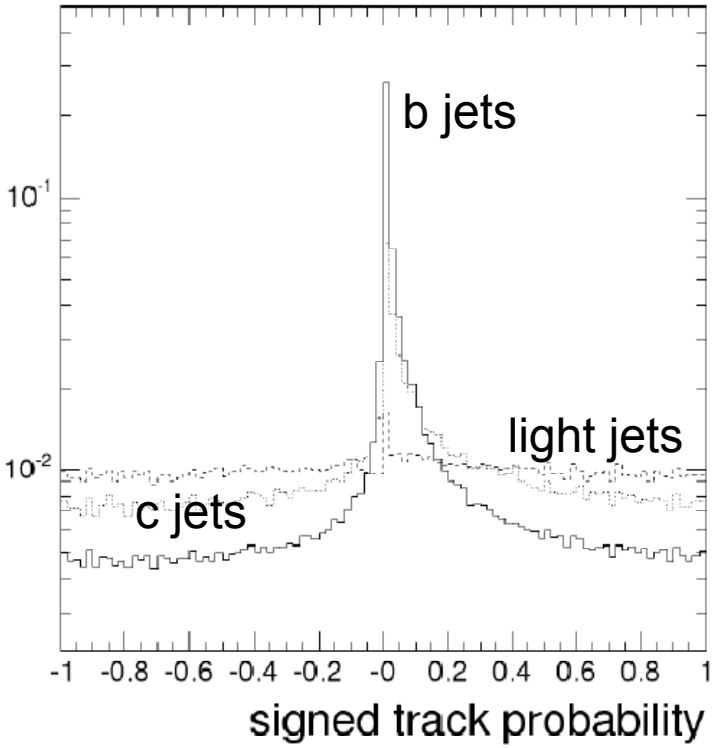


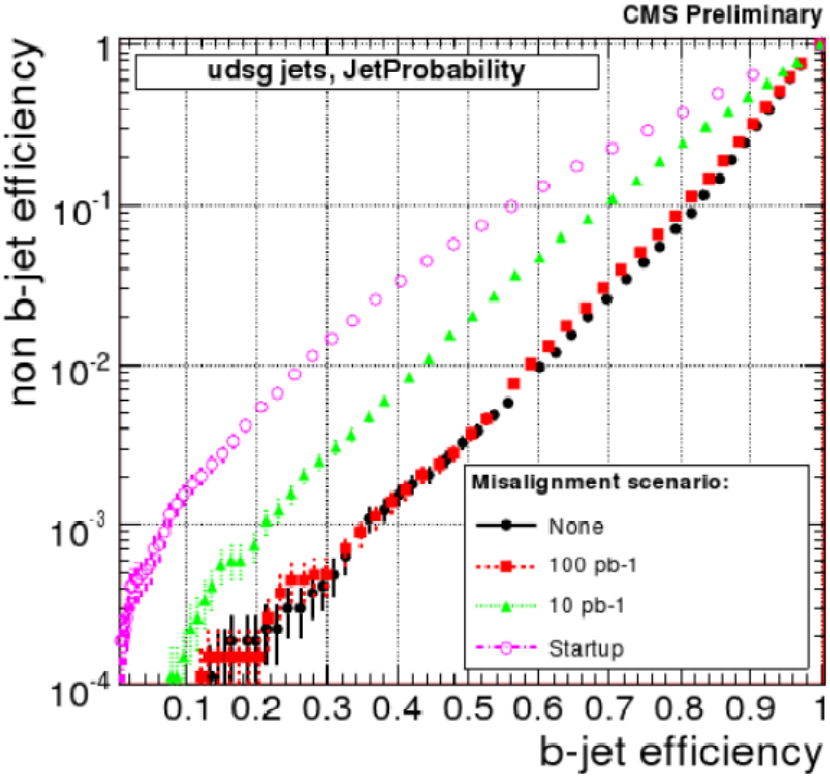
Track Probability B-Tagging

- Determine the probability that each track in jet comes from primary vertex

- Estimate a combined probability that all tracks in jet come from PV

- Use combined probability as discriminant



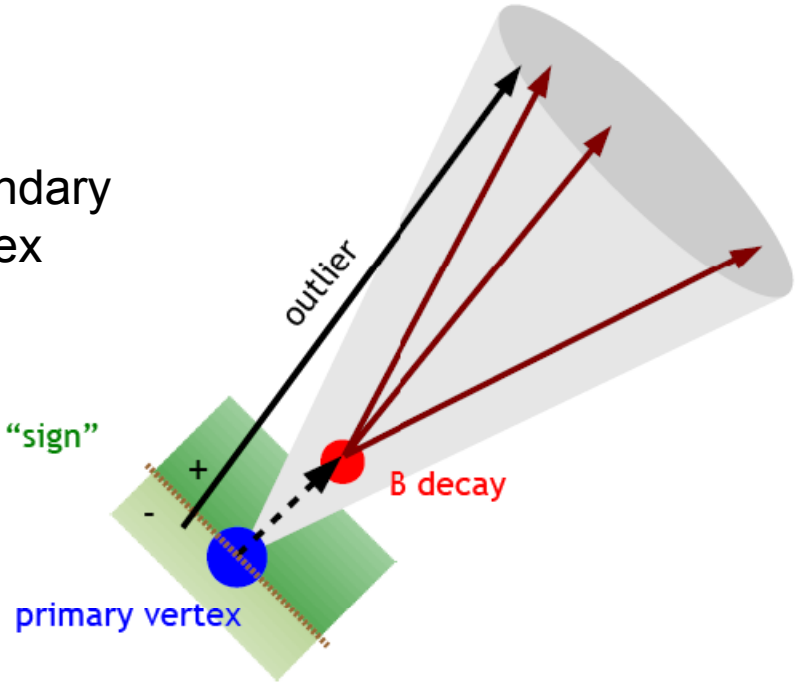


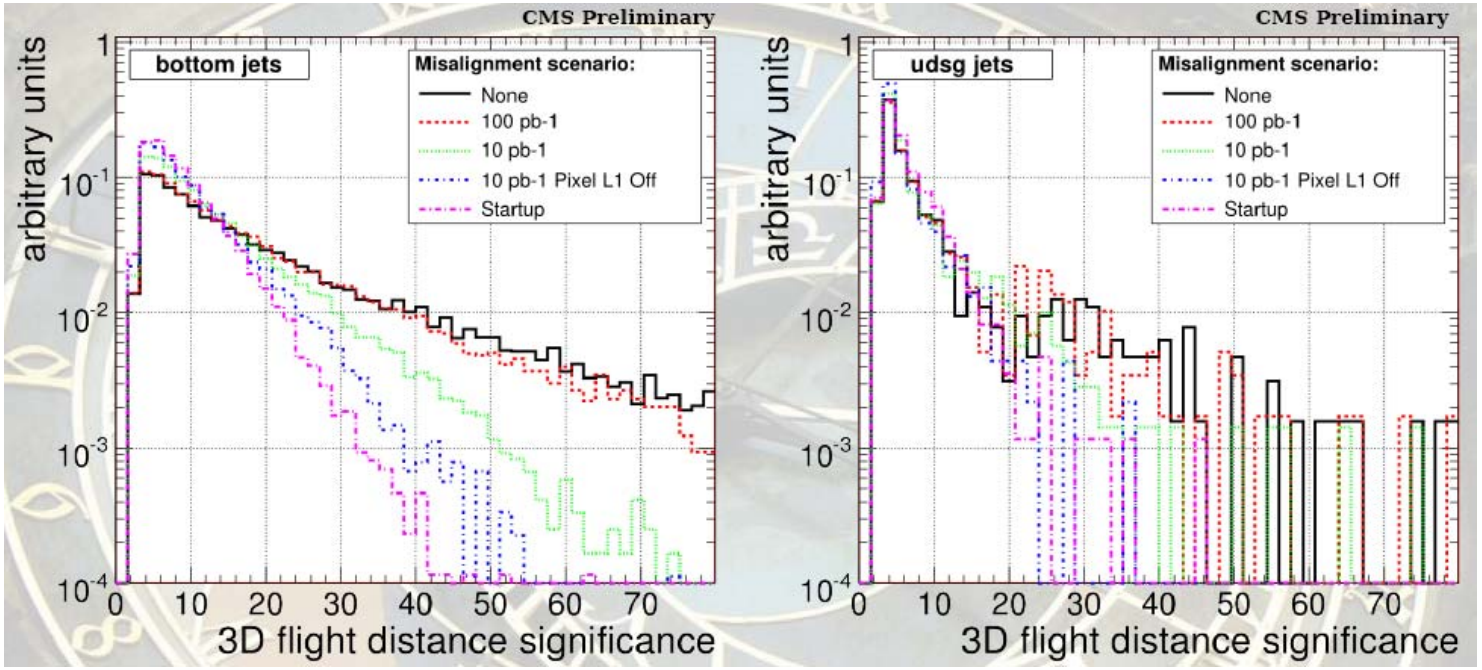
Secondary Vertex

- Large B hadron lifetime → decay vertex (secondary vertex) significantly displaced w.r.t. primary vertex

- Search secondary vertex using all tracks in jet
- Reject M > 6.5 GeV, remove K_s

- Use decay distance significance
  d/σ_d as discriminant

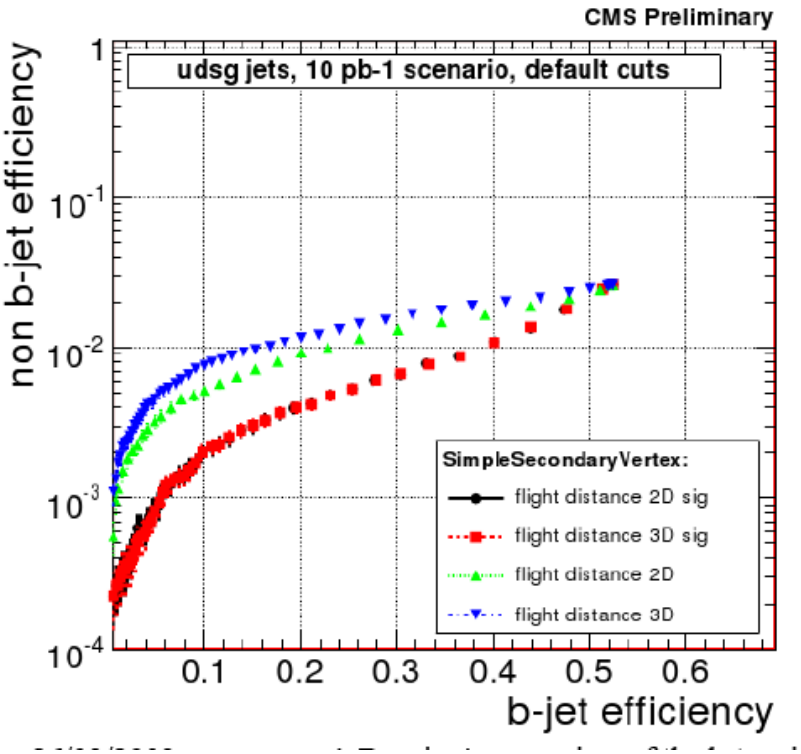


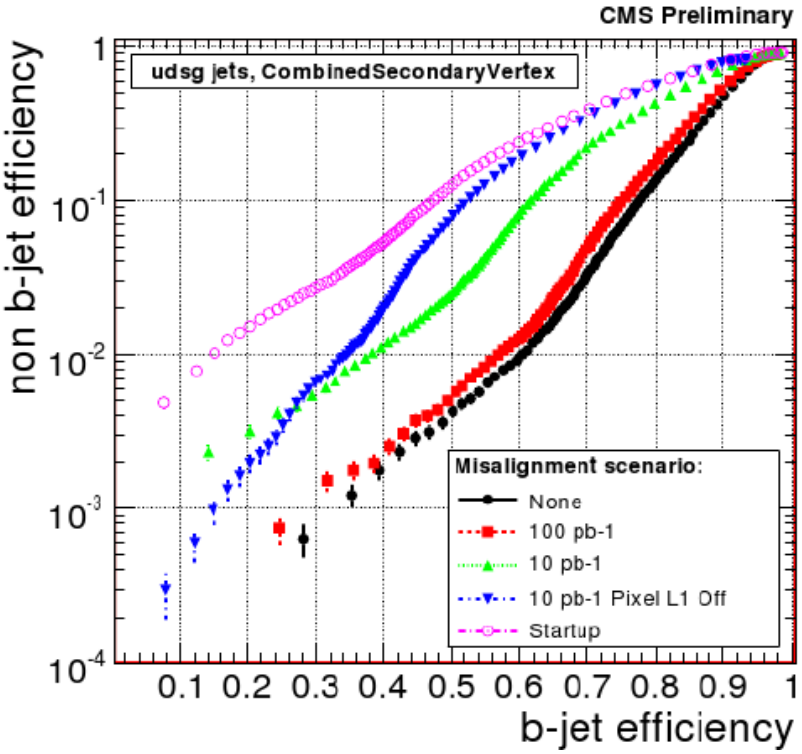


Secondary Vertex B-Tagging

- Search for secondary vertex
  - use either 2D or 3D decay distance
  - use impact parameter and vertex mass as well

- Combined above in likelihood discriminant



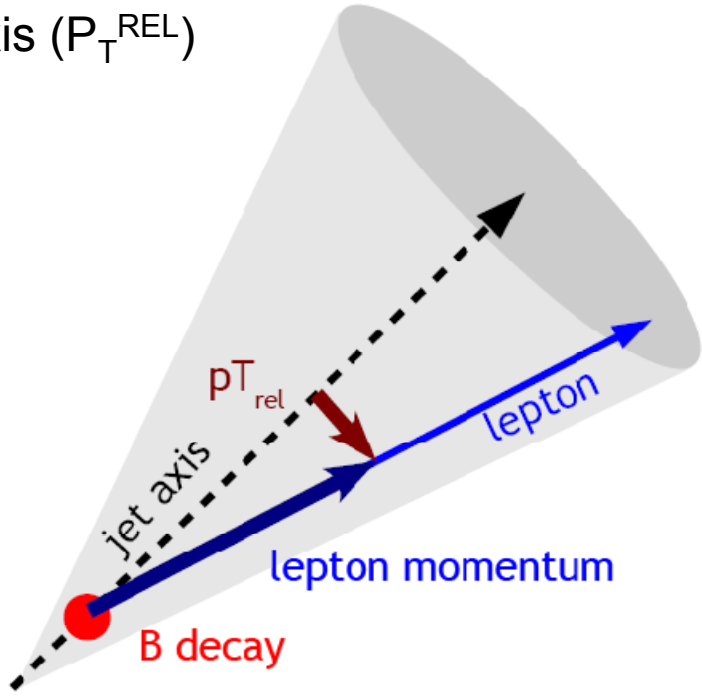


“Soft” Leptons

- Exploit large semileptonic branching fraction of B decays (Br(b → lepton) ≈ 10%) along with the large b-hadron mass (~5 GeV)

- Leptons from b decays are characterized by:
  - large impact parameter w.r.t. PV
  - large transverse momentum w.r.t. jet axis (P_T^{REL})
  - large angular distance w.r.t. jet axis

- Lepton (muon or electron) quality selection improves b-tagging purity



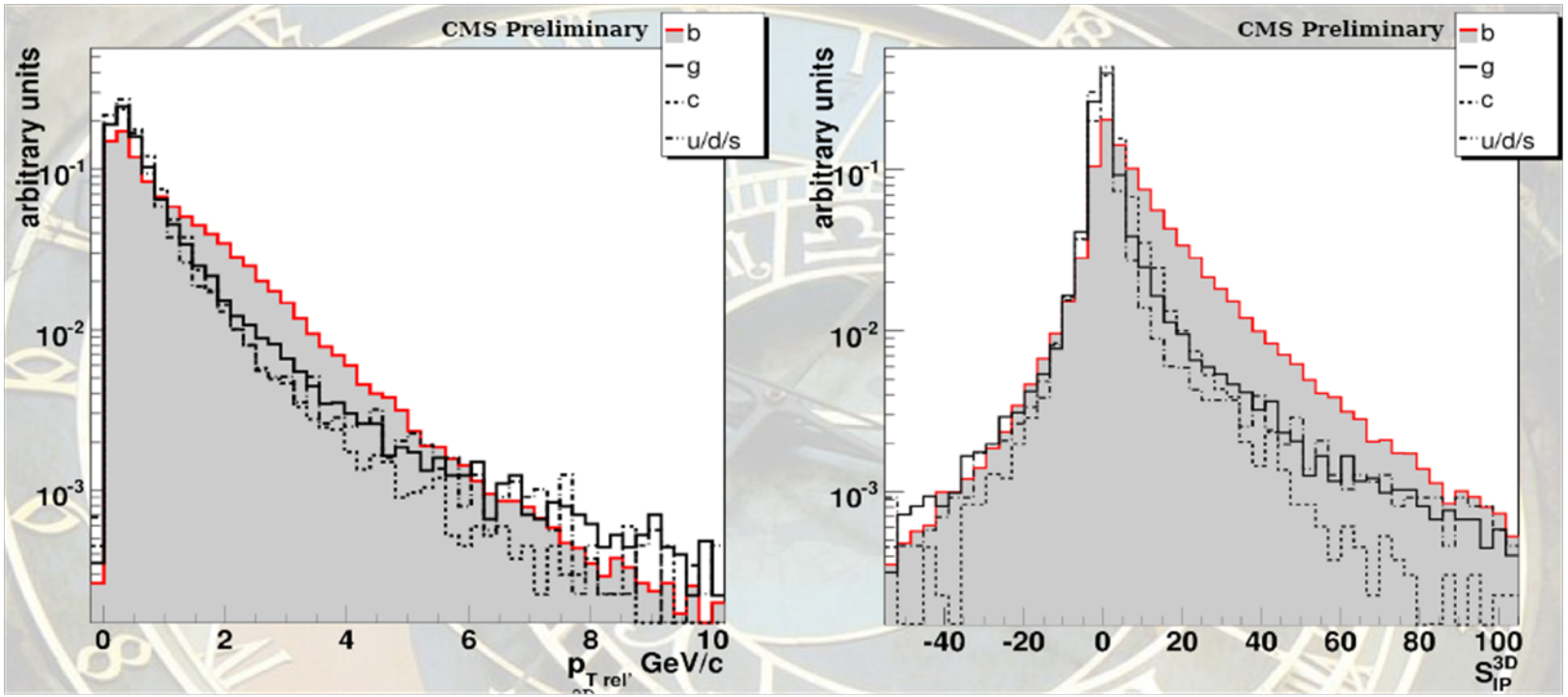
“Soft” Leptons

- Study both P_T^{REL} and signed IP as discriminants between gluon, u,s,d, charm and b-jets

- Results with QCD and tt samples and full detector simulation:

P_T^{REL}

signed 3D IP

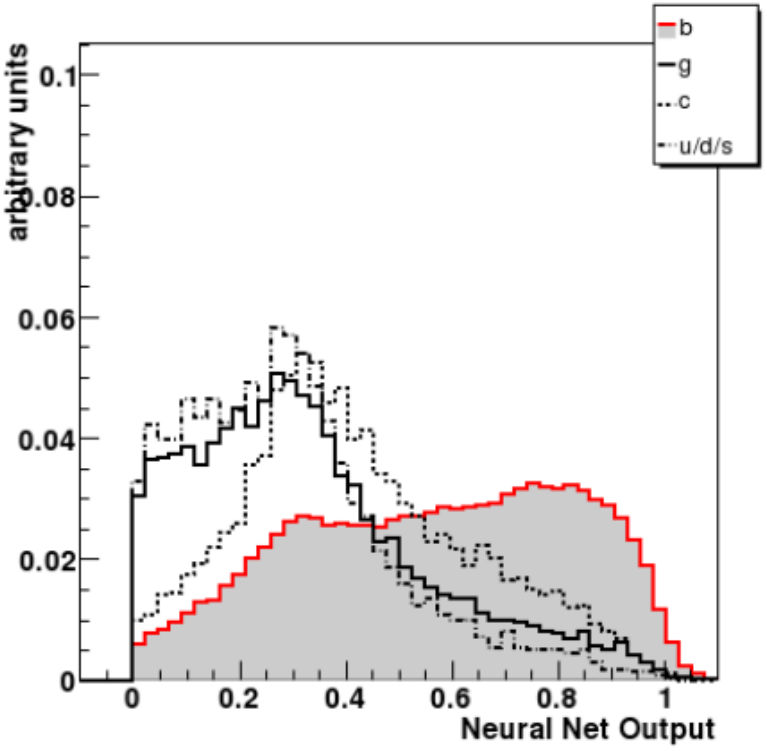


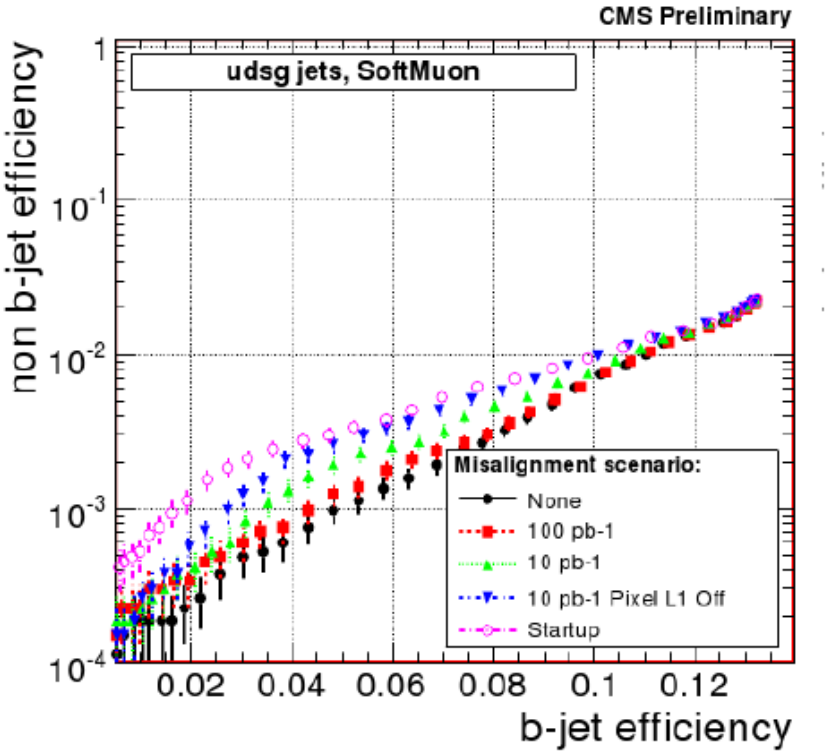
“Soft” Leptons B-tagging

- Search for muons and leptons in each jet

- Combine discriminating variables in neural network to discriminate light from heavy flavor jets

- Efficiency limitted by semileptonic b branching fraction





Conclusions

- Various B-tagging algorithms exploit characteristics of b hadrons:
    - lifetime, mass, semileptonic decays

- Different algorithms are available at CMS

- Use simple and more robust b-taggers at the beginning of data taking
    - track counting, displaced vertex, lepton tagging

- Use multivariate b-tagging techniques latter as data is better understood
    to take full advantage of available information