b-tagging commissioning strategy for ATLAS

Workshop on top physics

LPSC 18 October 2007

Lorenzo Feligioni

Outline

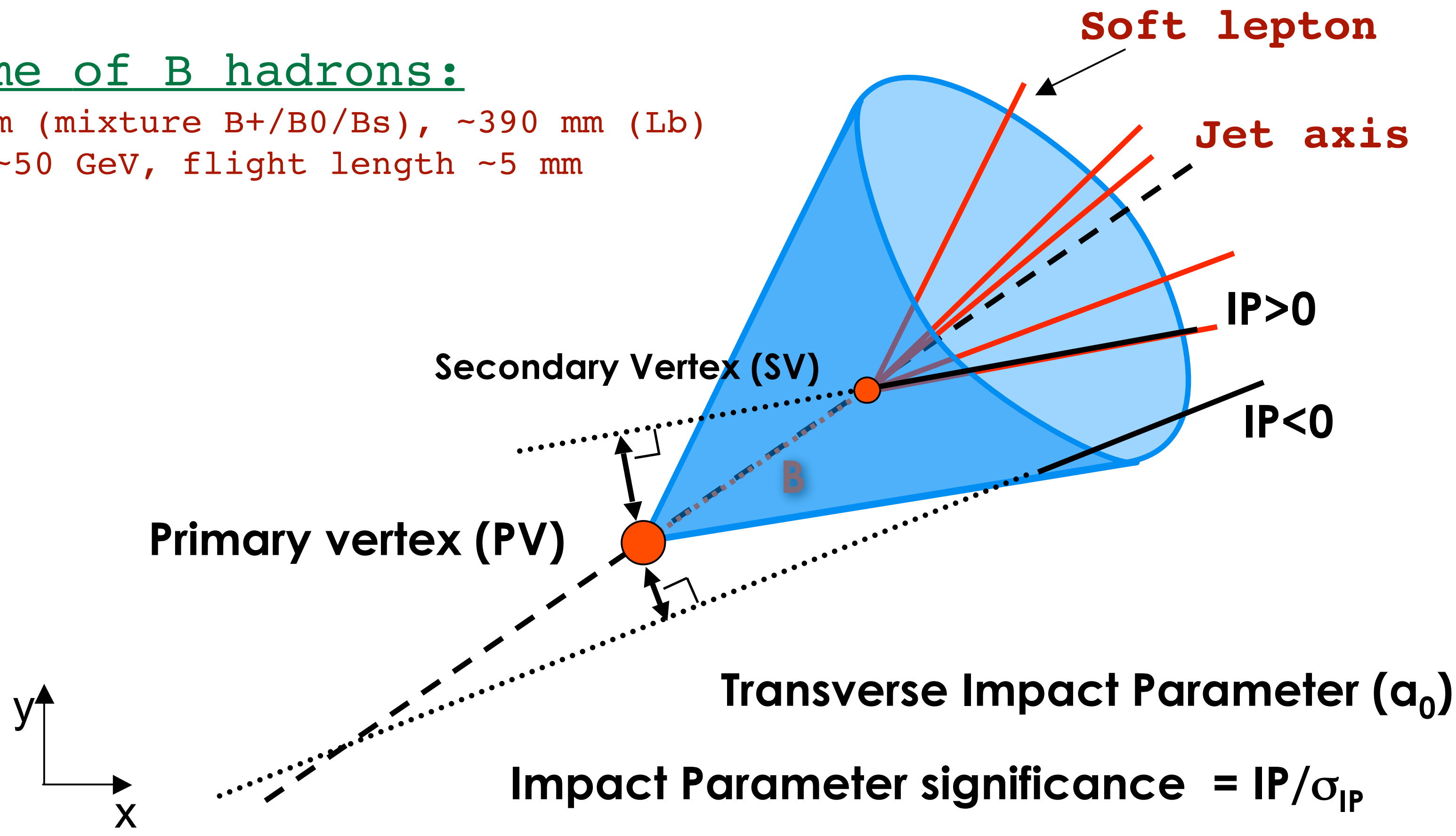
- Introduction to ATLAS tagging algorithms
  - Spatial tagging:
    - Track impact parameter (IP)
    - Secondary vertices (SV)
    - Jet Probability
  - Soft lepton tagging:
    - non isolated low p_T leptons from B(D) decays
- Performances estimation using tt events
  - Event/tag counting method
  - Topological selection
- Calibration using dijet events
  - muon p_T^{rel}
- Conclusions

(IP) and (SV) } combined → *b-tag weight* IP3D+SV1

ATLAS tagging algorithms

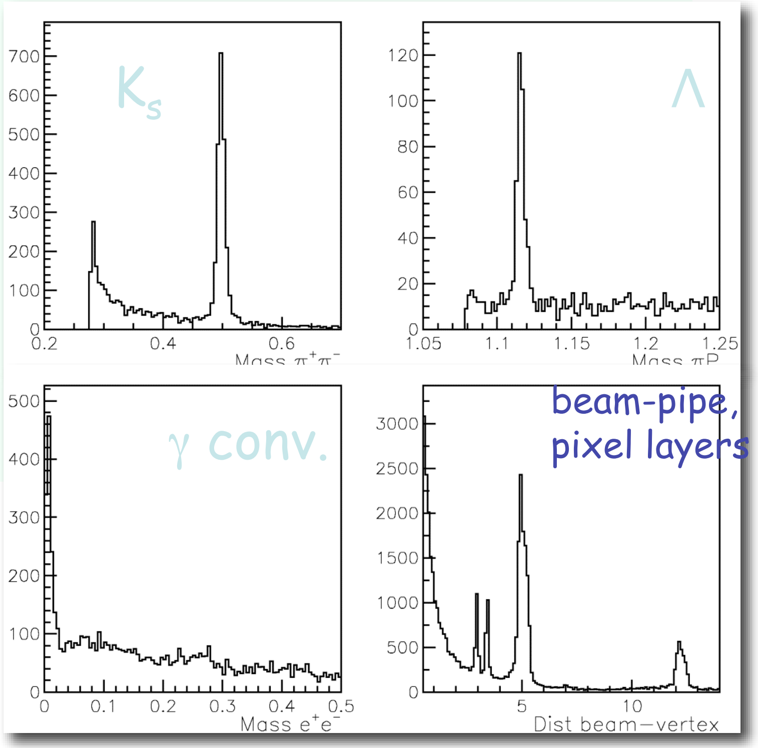
Lifetime of B hadrons:

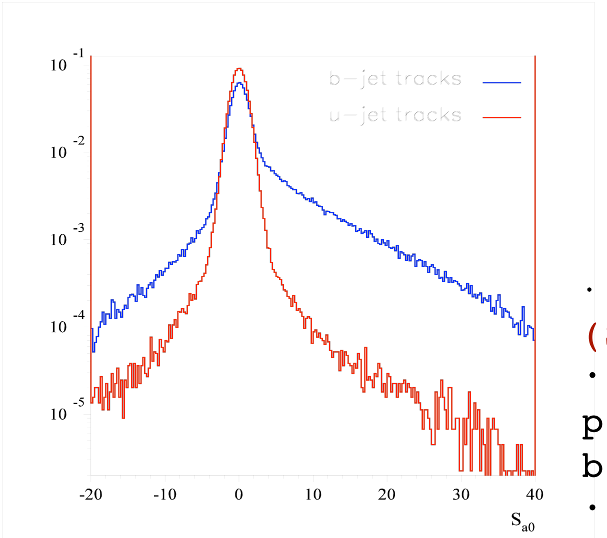
ct ~470 mm (mixture B+/B0/Bs), ~390 mm (Lb)
for E(B) ~50 GeV, flight length ~5 mm



Track selection and b-tag weight

- Only well reconstructed tracks with p_T>1 GeV are used
- For b-tagging IP is calculated wrt PV
- SV mass used to reject K^0_s/Λ, γ
- b-tag weight:
  - All taggers rely on a comparison between two hypotheses
  - likelihood for the jet coming from a b-quark vs light quark
  - Example: transverse IP sig





· use the impact parameter significance (S = IP/σ) of each track
· compare it to predefined calibration p.d.f. for the b and light hypothesis b(S) and u(S)
· sum over all tracks → jet btag weight

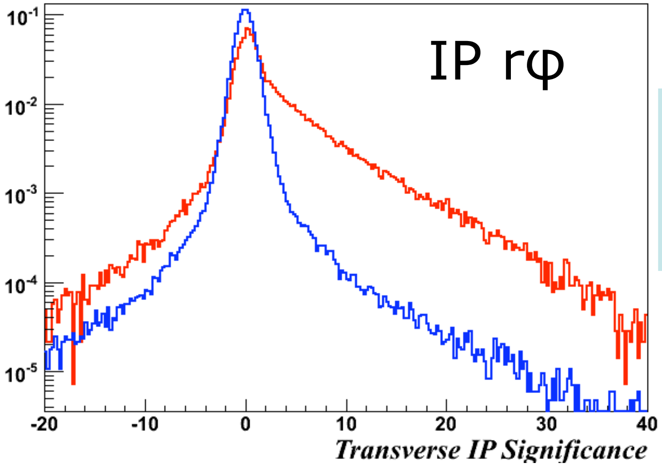
high W ➔ likely a b-jet

$$W_{jet} = \sum_{tracks} \ln\left(\frac{b(S)}{u(S)}\right)$$

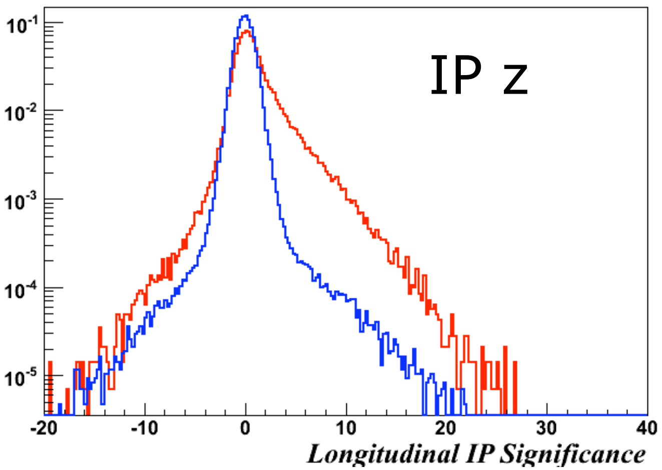
Impact parameter: IP3D

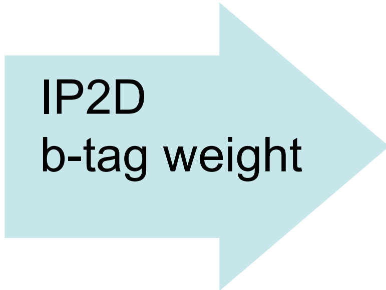
IP3D: combination of transverse & longitudinal normalized IP

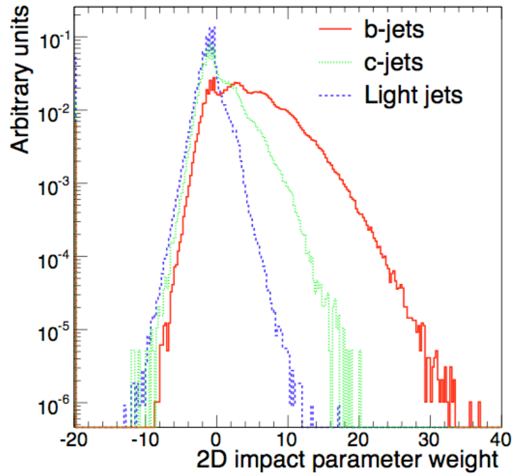
IP Probability Density Functions



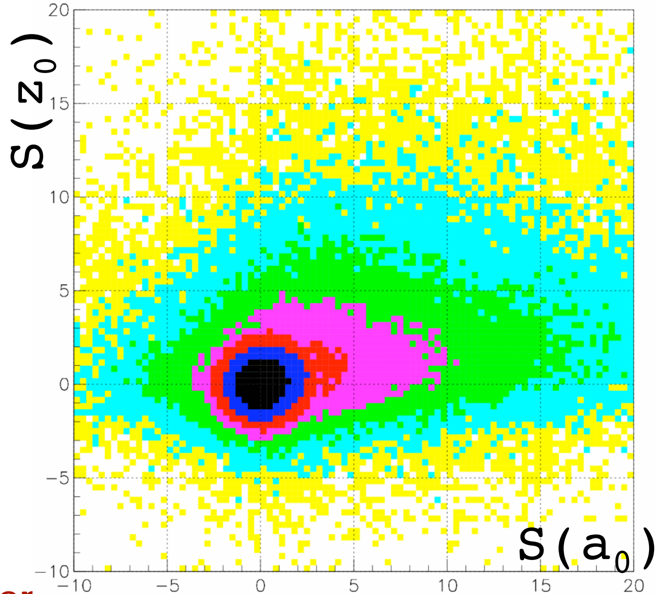
b-jets light jets



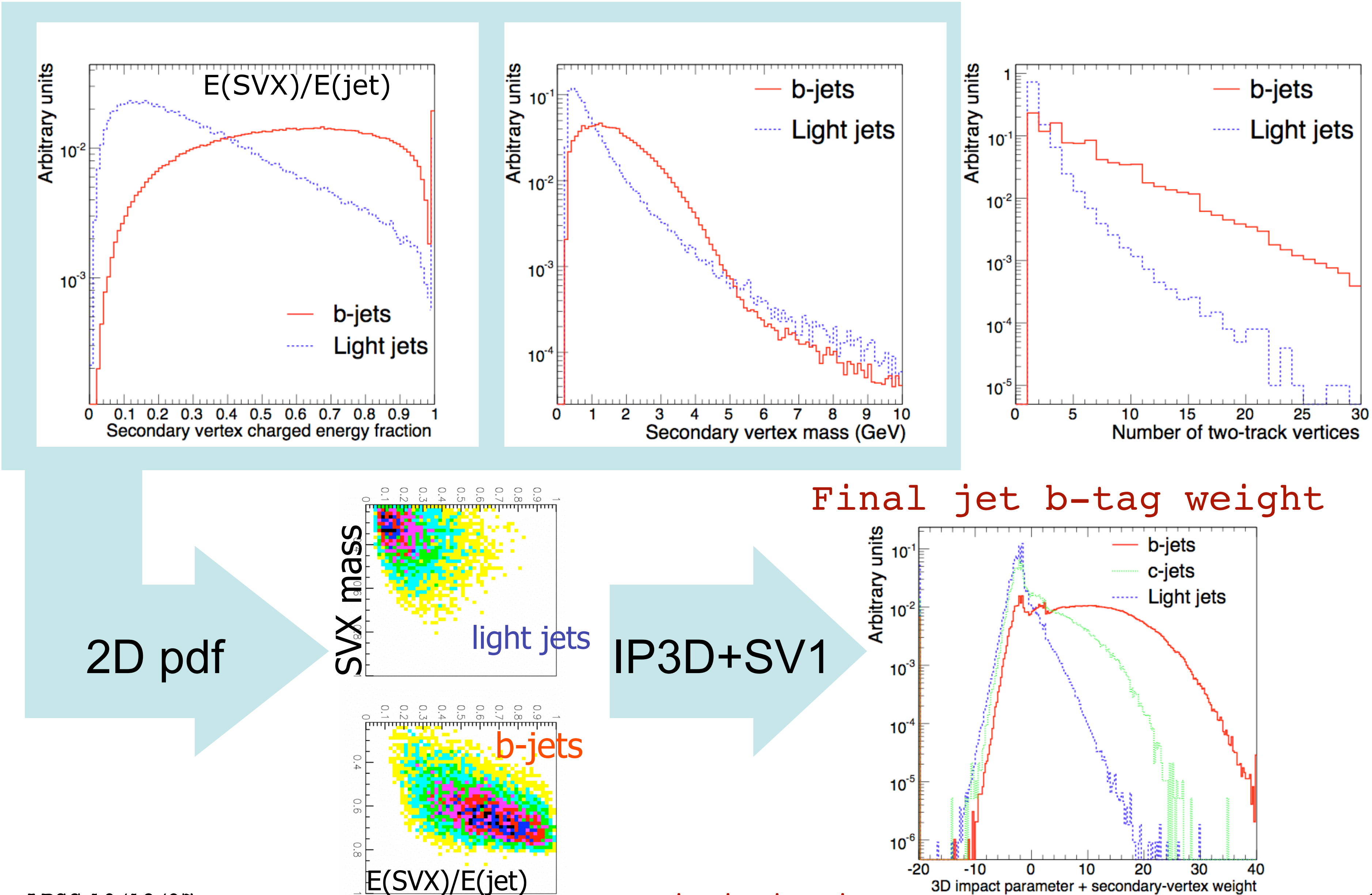




Correlations ⇒ 2D calibration



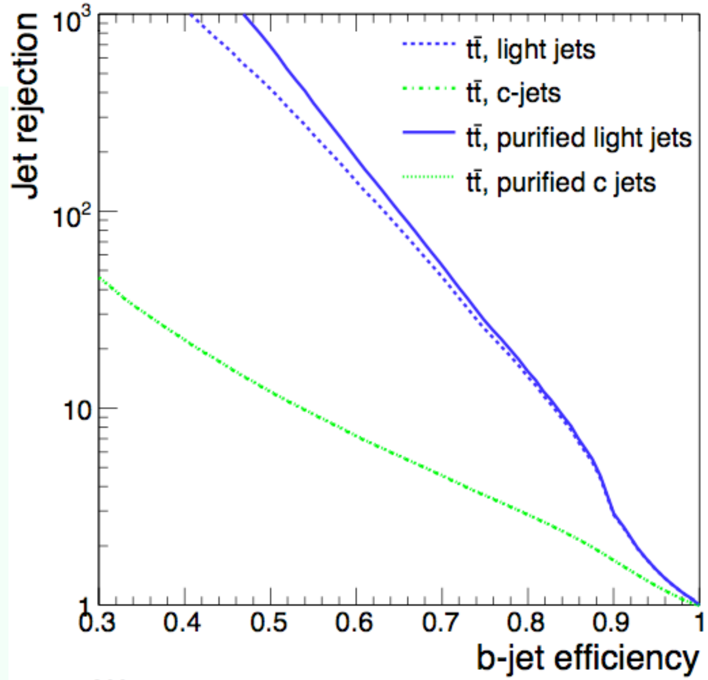
Inclusive Secondary Vertex: SV1

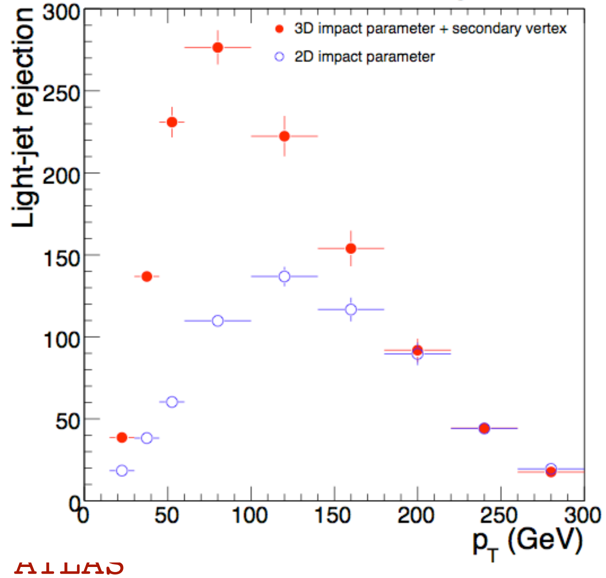


IP3D+SV1: performances

- Different cuts on b-tag weight used to achieve different working points
  - efficiency/light rejection
- Light jet rejection depends on the event environment
  - performances are estimated on purified light jets sample:
    - $\Delta R(\text{light,heavy flavored})>0.8$
- Performances heavily dependent on jet p_T and η

High η: multiple scattering
High p_T: higher track density

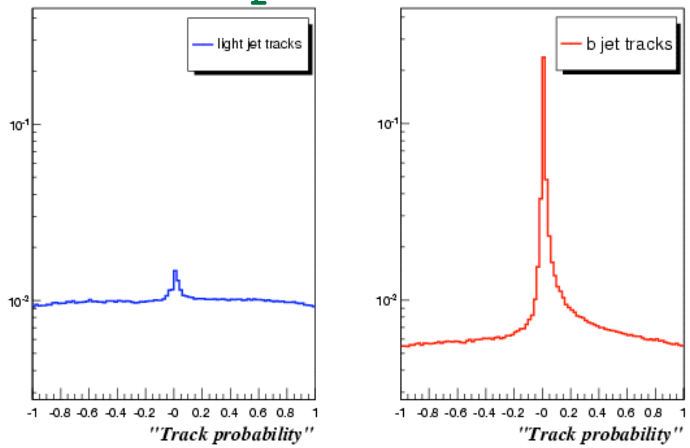




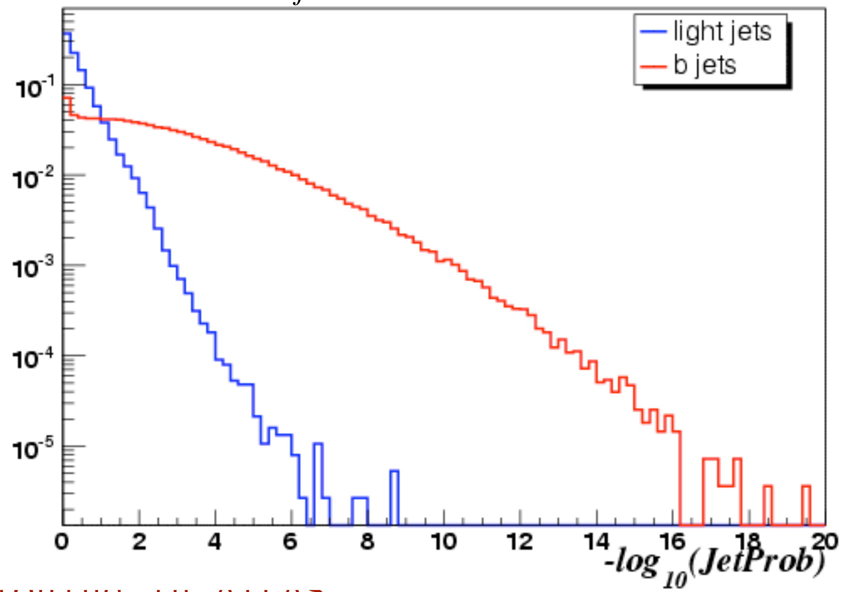
JetProb Tagger

- Tagger à la Aleph/Lep (JLIP in DØ):
  - Compatibility of tracks with PV
  - Fit negative side of signed impact parameter w/ resolution function
  - Use it for the positive signed IP⇒ track probability
  - Combine tracks in a jet probability
- Resolution functions completely derived with first data
- Expect poorer performances than IP3D+SV1
- But easier to understand /calibrate in early data
- Need some development in terms of track categories

Track probabilities

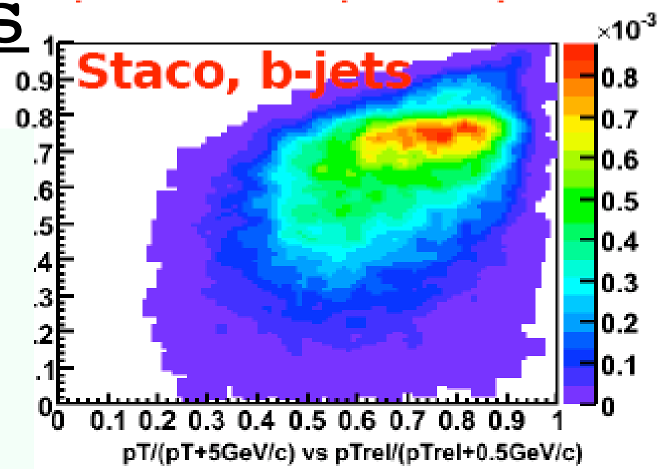


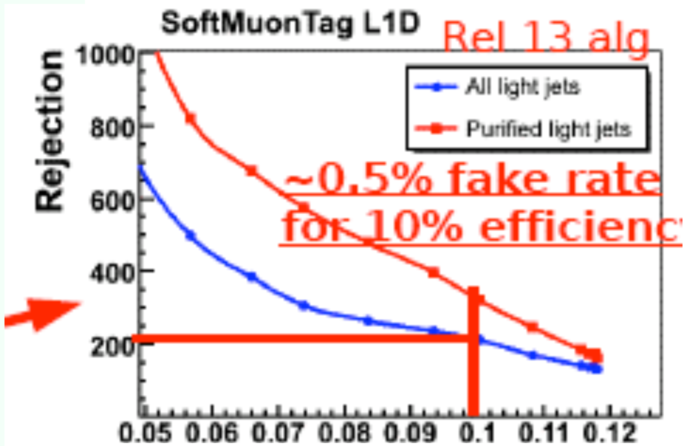
$$P_{jet} = \Pi \cdot \sum_{j=0}^{N_{tr}-1} \frac{(-\ln \Pi)^j}{j!}, \quad \Pi = \prod_{i=1}^{N_{tr}} P_{tr\,i}$$

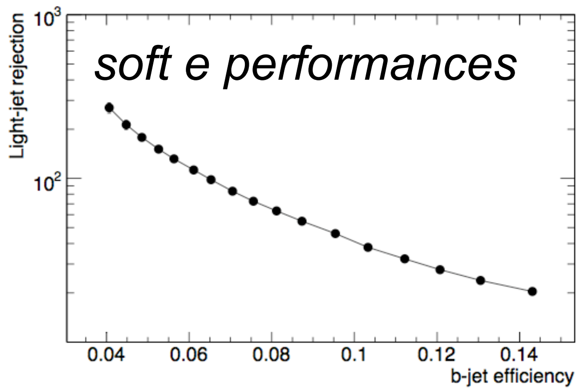


Soft Leptons

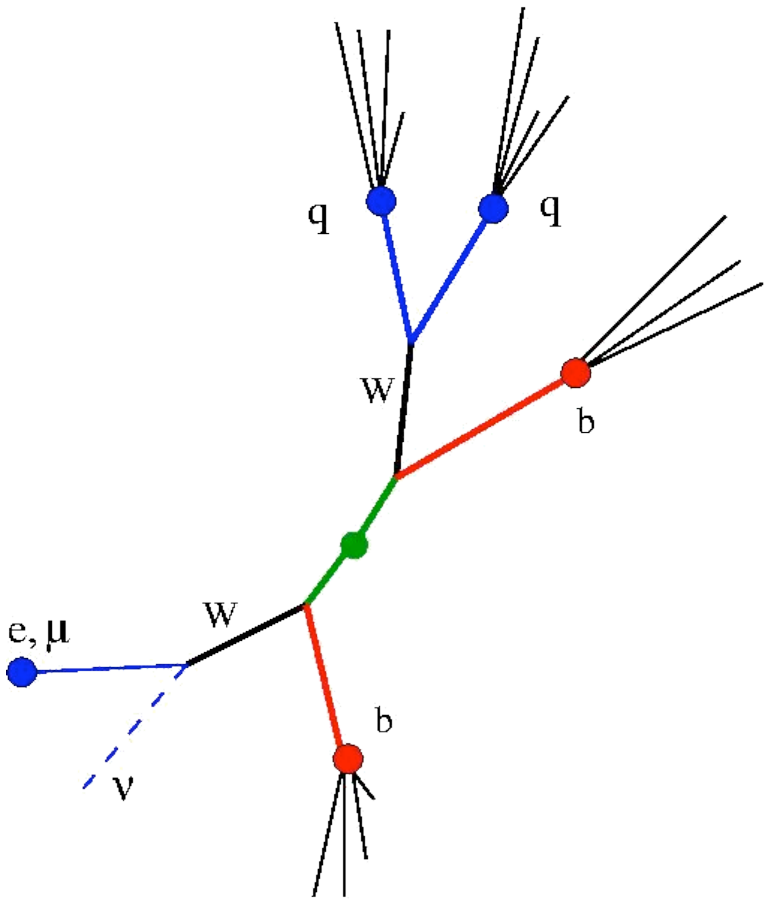
- Br(b→μνX)+Br(b→c→μνX) ~ 20%
- Small correlation w.r.t. SV tag
  - Can be used for cross calibration
- Soft Muon tagging algorithm:
  - currently Staco+Mutag
    - high-efficiency down to 4 GeV p_T
  - 2 steps:
    - association muon-jet (cone ΔR < 0.5)
    - Likelihood ratios (1D, 2D) for b/light hypothesis using pT, pTrel.
  - Checked performance vs pile-up/cavern bckgd with promising results
- Soft Electron:
  - Likelihood ratio ID+shower shape info from the calorimeter
  - 10% efficiency ~3% mistag rate
    - High contamination from γ conversion (40%)





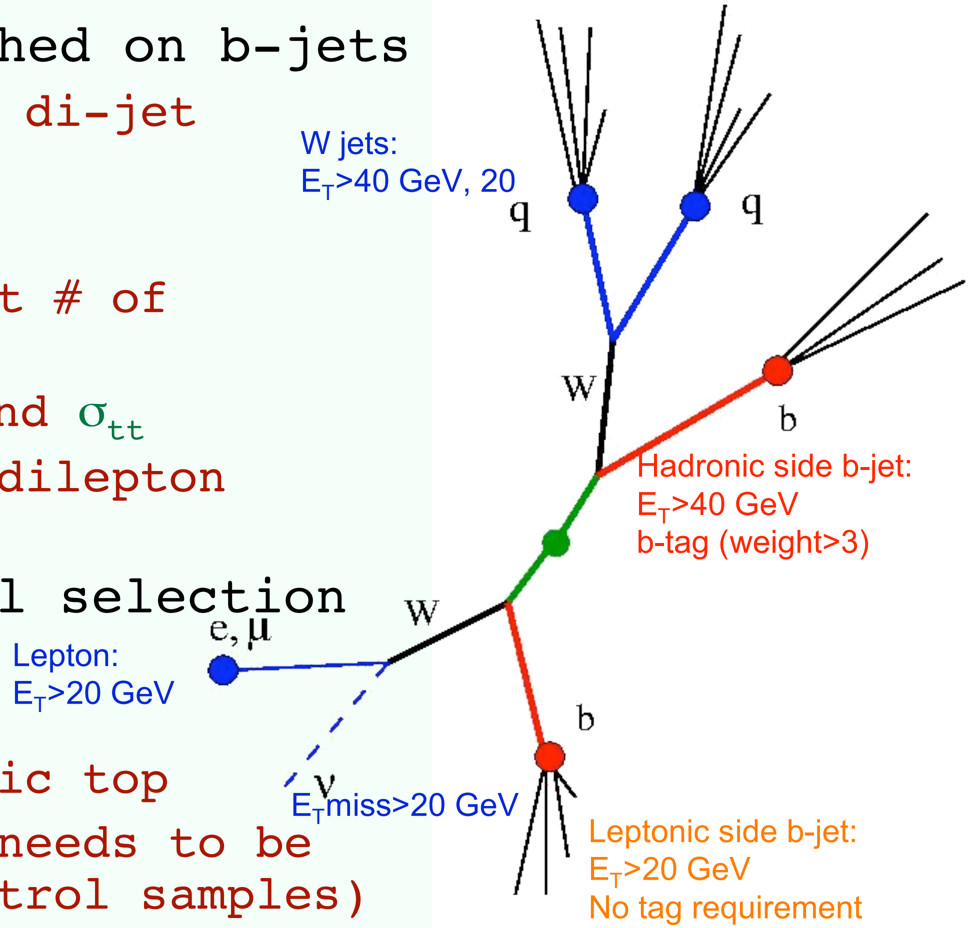


b-tag performances from tt



b-tag performances from tt

- tt at LHC ~800 pb
  - calibration sample
- tt provides a sample enriched on b-jets
  - Different environment than di-jet studies
- Event/tag counting method
  - Count events with different # of tagged b-jets
  - likelihood fit for ε_b, ε_c and σ_{tt}
  - Consider semileptonic and dilepton final state
- Topological and kinematical selection
  - Very energetic events
  - One b-tagged jets used to reconstruct hadronic top
  - Background on a 20% level needs to be subtracted using data (control samples)
    - combinatorial background: misassigned jets
    - Physics background: W+heavy flavors



Event/tag counting method

- $L = Poisson(N_1, < N_1 >) \cdot Poisson(N_2, < N_2 >) \cdot Poisson(N_3, < N_3 >)$
- N_n = Number of observed events with n tags
- $< N_n >$ = Expected number of events with n tags : function of ϵ_B, ϵ_C, ϵ_L, $\sigma_{t\bar{t}}$, etc...

$$< N_n >=$$
$$(L \cdot \sigma_{t\bar{t}} \cdot A_{pre-tag}) \cdot \Sigma_{i,j,k} F_{i,j,k} \Sigma_{combi.} A_i^{i'} \cdot \epsilon_b^{i'} \cdot (1-\epsilon_b)^{i-i'} \cdot A_j^{j'} \cdot \epsilon_c^{j'} \cdot (1-\epsilon_c)^{j-j'} \cdot A_k^{k'} \cdot \epsilon_l^{k'} \cdot (1-\epsilon_l)^{k-k'}$$

i = # b-jets and i' = # tagged b-jets

j = # c-jets and j' = # tagged c-jets

k = # l-jets and k' = # tagged l-jets

$F_{i,j,k}$ = Fraction of events with i b-jets, j c-jets, k l-jets.

$A_i^{i'} = i!/(i'! \cdot (i-i')!)$

$\sigma_{t\bar{t}}$ = production cross-section

$A_{pre-tag}$ = acceptance without b-tagging

L = integrated luminosity

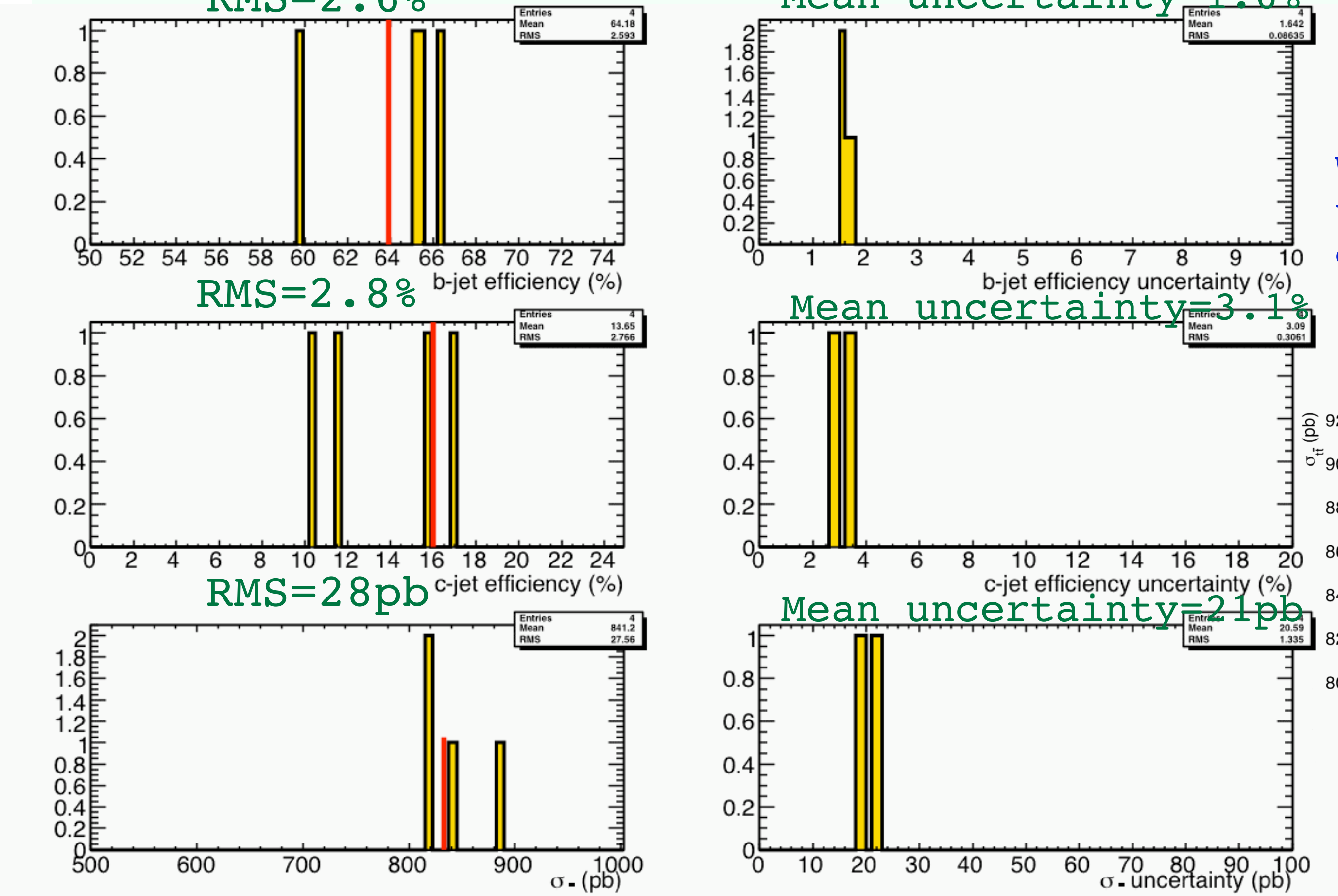
Light quark rejection must be known

an input for the algorithm

BR(t→Wb)=1

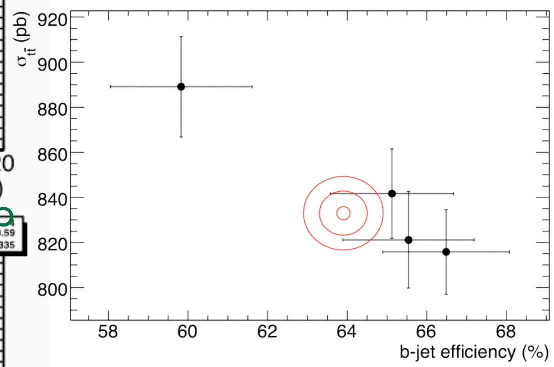
Event/tag counting performances

- Performances on 4 independent MC statistical samples, each equivalent to 100 pb^{-1}



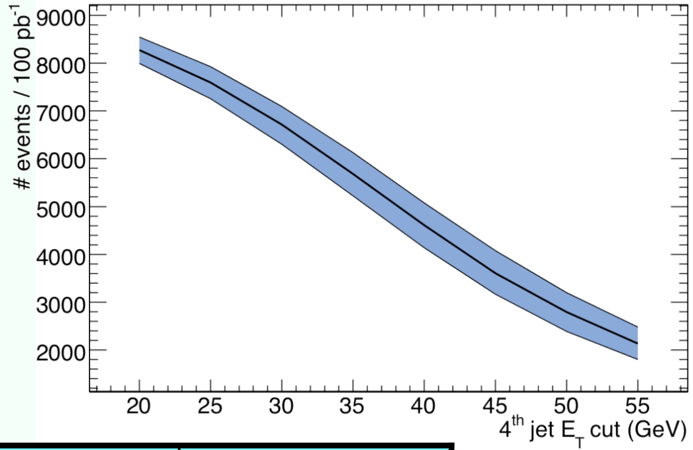
With 100pb-1, statistica
uncertainty on eb
of around 2% absolute

Correlation ε_b vs. σ_{tt}



Event/tag counting systematics

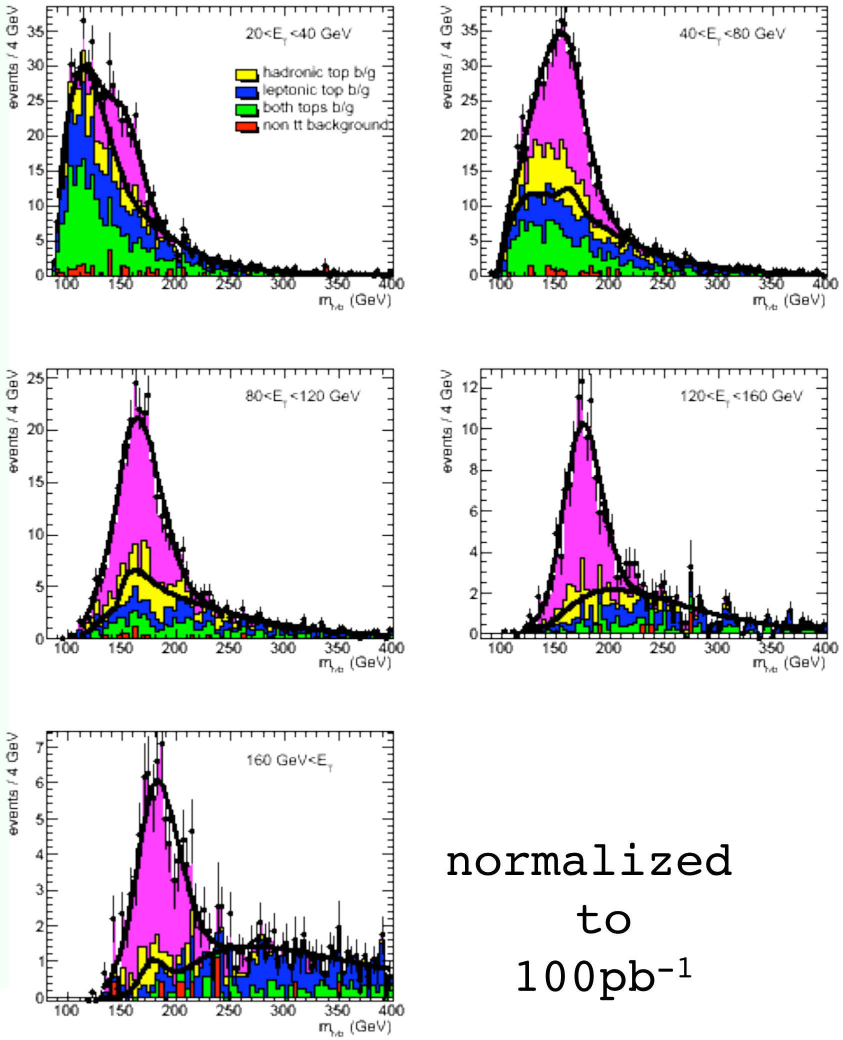
- JES: negligible effect for b-tag efficiency
- Very robust w.r.t to physics background and light quark rejection estimation
  - latter input for the method
- 0.03 absolute uncertainty at ε_b=0.6



Systematic	$\delta\varepsilon_b$ (%)	$\delta\sigma_{tt}$ (%)
Jet energy scale (±5%)	0.3	12
Light jet rejection (±100%)	0.1	0.1
Tag correlation	<0.5	<0.5
Jet labeling	1.4	1.4
MC statistics	0.5	0.5
Background (±100% W/Z+jets)	2.5	4
Total systematic	3	13
Statistical (100 pb^{-1})	3	3

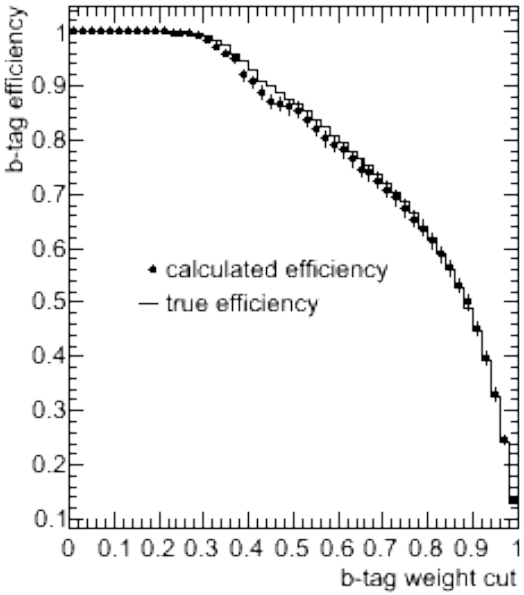
Topological approach using $m_{bl\nu}$ distribution

- Select semileptonic ttbar events
  - Reconstruct m_{bjj} on hadronic side from 'raw' jet energies, cuts on m_W and m_{top}
    - Require b-tag on b-jet, anti-b on W jets
- Use recon mass of leptonic top ($m_{bl\nu}$) to find region enhanced in b-jets
  - No requirement on b-tag of this jet
    - Leptonic top ensures jets are b-flavour
- Have to subtract background from mis-reconstructions
  - Estimate shape from a control sample where hadronic side m_{bjj}>200 GeV, and leptonic top jet is anti b-tagged
  - Estimate flavour composition from signal sample where $m_{bl\nu}$ outside $m_{top} \pm 2\sigma$ (mass sideband region)



normalized to $100pb^{-1}$

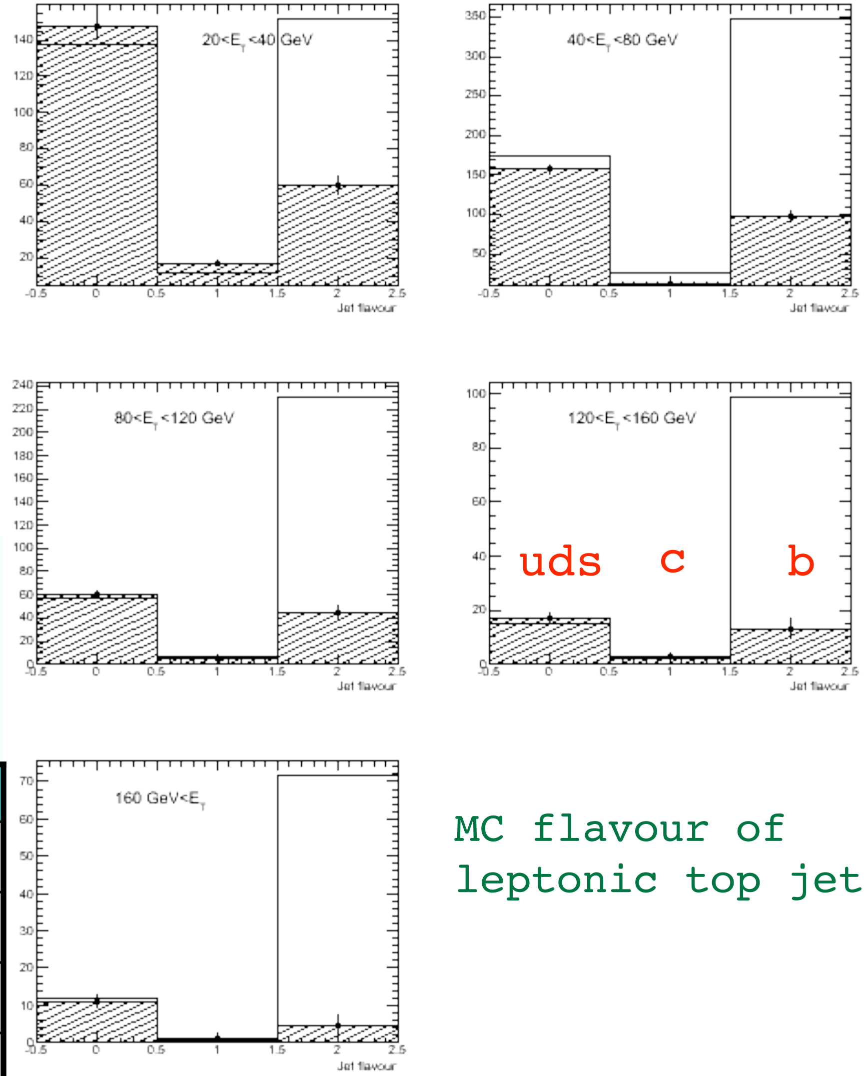
Topological approach: results



Fit performed on L=862 pb-1

Method has large uncertainties below 40 GeV due to very high background

Effect of background estimation



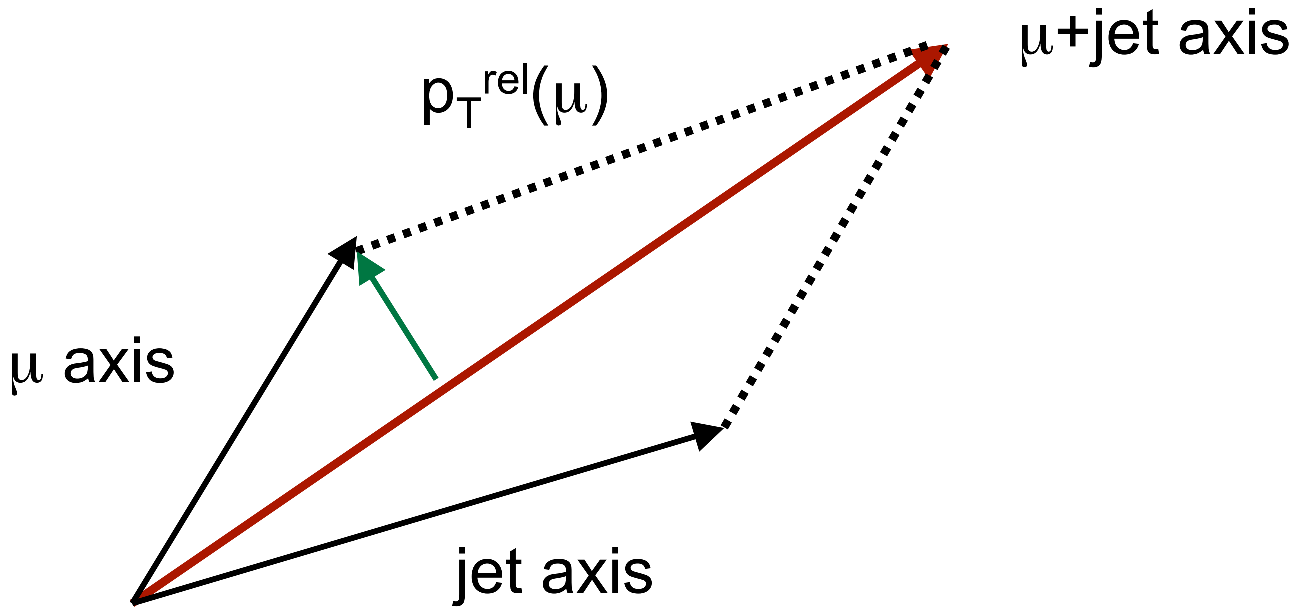
MC flavour of leptonic top jet

- Systematic uncertainties are small:
  - Fit is data-driven, no explicit MC inputs
  - Check $\Delta=\varepsilon_{meas}-\varepsilon_{true}$ for different MC samples

Systematic	MC samples	Sim	δΔ (abs)
ISR/FSR/Q^2	Δ6251-Δ6250	fast	-0.022±0.020
Underlying event	Δ5565-Δ5200	fast	-0.045±0.024
Jet E-scale +5%	5200	full	0.007±?
uds tag *2, *0.5	5200	full	0.001±?
Charm tag ±30%	5200	full	0.004±?

Estimated statistical relative error = 7% for 100 pb^{-1}

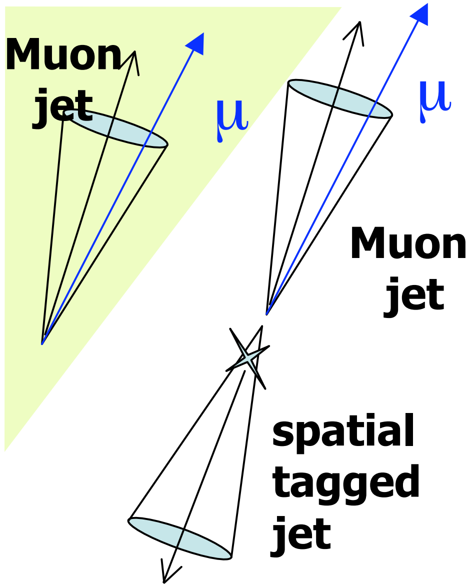
b-tagging performance with di-jet samples



b-tagging efficiency: system 8

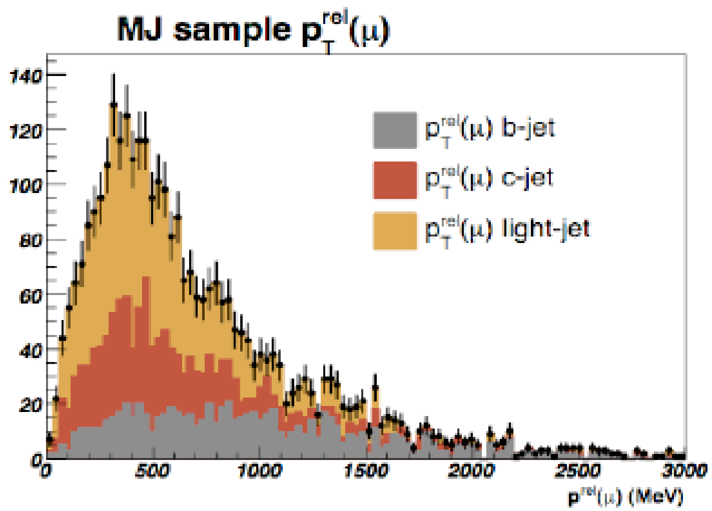
- Non-linear system (System8 à la D0):
  - 2 samples
    - Muon Jets (n)
    - MJ + other tag opposite jet Jet (p)
  - 2 different b/l fractions
  - 2 non-correlated taggers
    - Tracks (impact param./SVT)
    - Soft Muon (pTrel)

  ➔ system can be solved analytically



- Results: (for ~10k MJ)

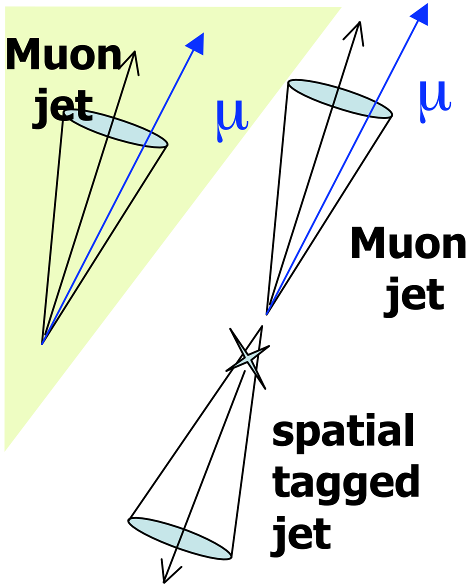
	True ϵ	S8 result
SMT (p_T^{rel} > 1 GeV)	48%	(45±4)%
IP3D+SV1 (w>4)	75%	(80±10)%



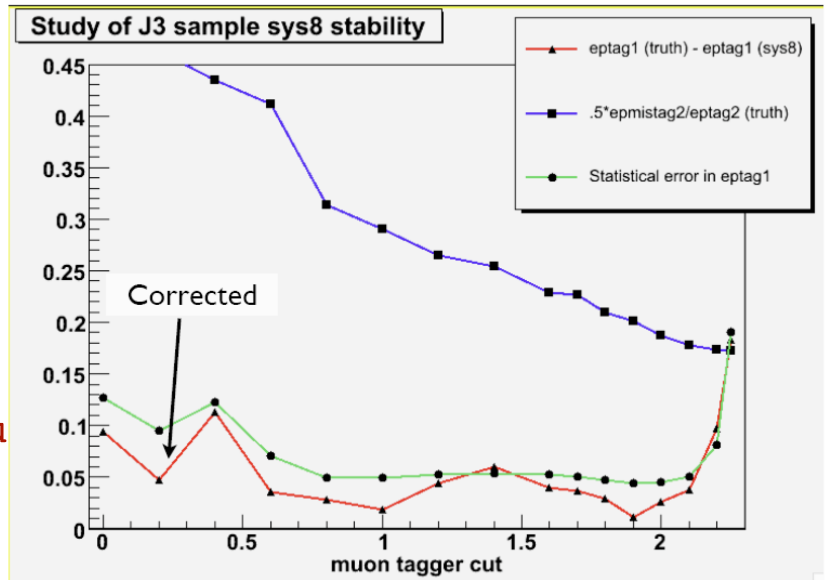
➔ definition of Muon+Jet trigger in progress

b-tagging efficiency: system 8

- Non-linear system (System8 à la D0):
  - 2 samples
    - Muon Jets (n)
    - MJ + other tag opposite jet Jet (p)
  - 2 different b/l fractions
  - 2 non-correlated taggers
    - Tracks (impact param./SVT)
    - Soft Muon (pTrel)

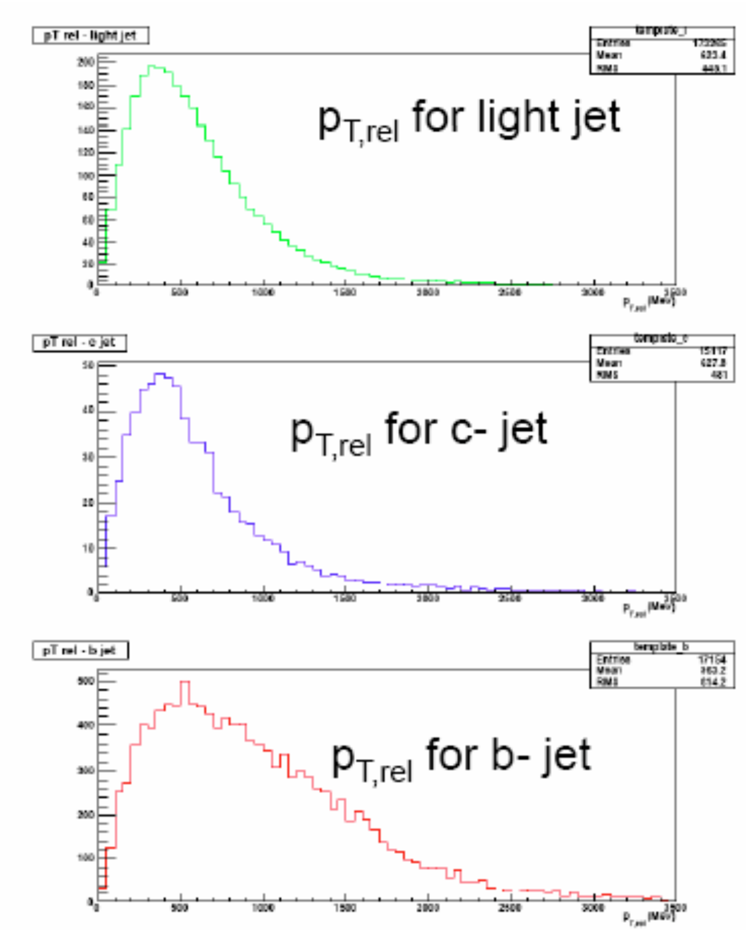


- Muon tagger included in recent studies:
  - Allows cross calibration of spatial and soft muon tagger
- Stability study:
  - Stable as a function of the soft mu
    weight cut applied

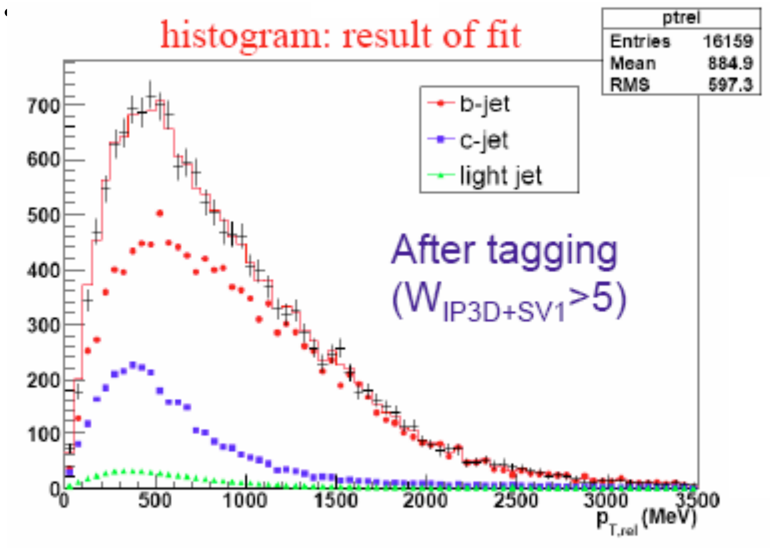


b-tagging efficiency: templates

Use the p_T^{Rel} of a muon fit those to data to determine tag rates.



Use MINUIT to fit



Use half the sample to derive the templates, the other half to test.
Vary the b-fraction (f_b) by reducing the # of light jets and test repeatedly.

$$f_b^{true} = 44\% = f_b^{fitted}$$

Still investigating source of systematics

Conclusions

- ATLAS developed a large variety of spatial b-tagging algorithms
  - More performant (IP3S+SV1):
    - Work on defining negative tags for mistag estimation
  - More robust (JetProb)
    - Easier to calibrate, good for first analyses
    - Need development: implementation of track categories
- Calibration with tt
  - Already in good shape, different approaches, interesting ideas
  - Need to carefully study systematics and effect of physics background.
  - No clear strategy to validate likelihood shapes yet
- Dijet:
  - Work just started look promising
  - Trigger study will tell how much integrated luminosity needed
  - Mistag rate methods under development