Jet Energy Scale & Resolution
-
Heavy Flavour identification

S. Greder

Institut Pluridisciplinaire Hubert Curien, Strasbourg

Outline

- Jet Energy Scale/Resolution
  - JES Reminder, methods & results
  - JER Reminder, methods & results
  - Future, Conclusion
- Heavy-flavour identification
  - Reminder & status
  - Future
  - Conclusion
- General Conclusion

Jet Energy Scale – Reminder

$$E_{\text{jet}}^{\text{ptcl}} = \frac{E_{\text{jet}}^{\text{meas}} - E_{\text{O}}}{R_{\text{jet}} S_{\text{jet}}}$$

***Offset* correction, E_O**

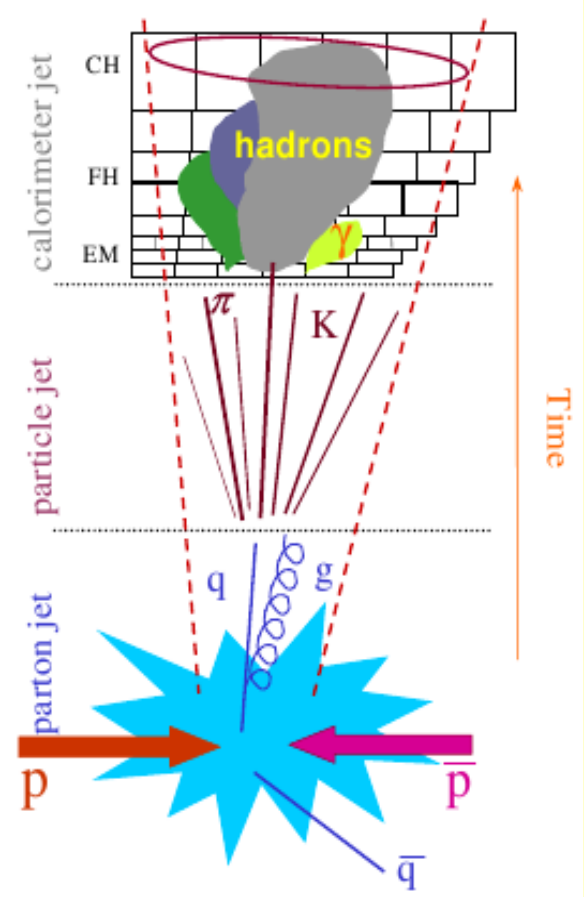
- calorimeter noise, multiple interactions and pile-up
- depends on jet cone R_{cone}, η^{jet}_{det}, instantaneous luminosity

***Shower* correction, S_{jet}**

- fraction of energy deposited outside (inside) the jet cone as a result of the development of showers in the calorimeter and the finite calorimeter cell size
- depends strongly on R_{cone} and, η^{jet}_{det} and mildly on jet energy

***Response* correction, R_{jet}**

- change in the jet energy due to energy response of the calorimeter to different particle jets
- function of R_{cone}, η^{jet}_{det} and jet energy

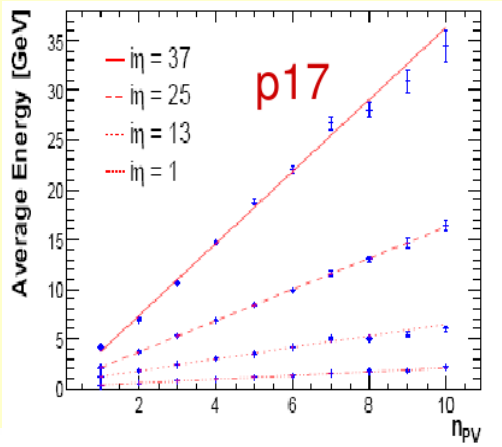


JES – Offset corrections

J. Coss Thesis

Goal:

subtract the energy **not** associated with the high p_T interaction (noise (N), multiple interactions (MI) and pile-up (P))

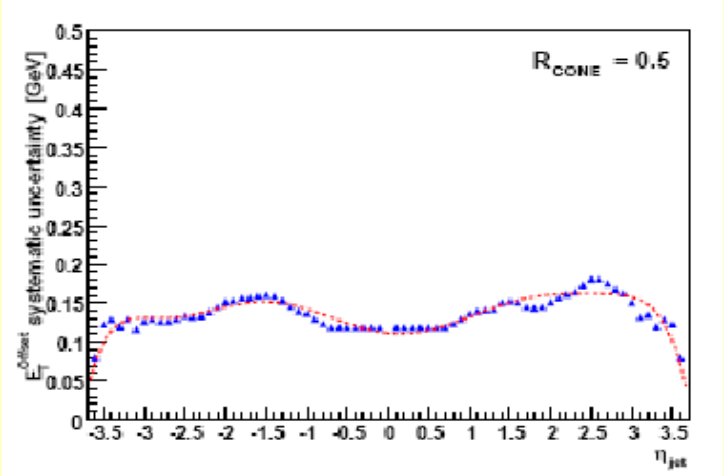


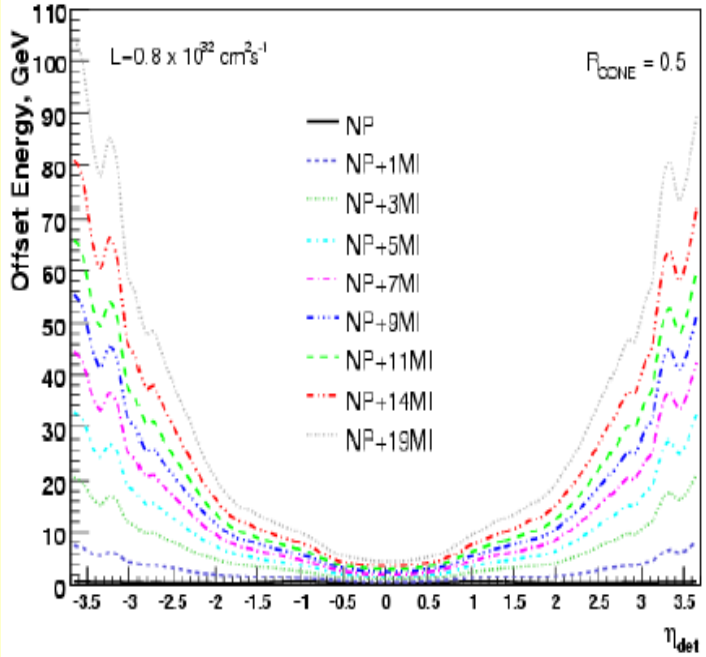
- Average offset energy is estimated for each calorimeter ring i_η, and as a function of n_{PV} and luminosity using Zero-Bias (ZB) and Minimum-Bias (MB) data

$$\mathcal{O}(\mathcal{L}, n_{PV}, \eta) = \mathcal{O}_{ZB}(\mathcal{L}, \eta) + \mathcal{O}_{MB}(\mathcal{L}, n_{PV}, \eta) - \mathcal{O}_{MB}(\mathcal{L}, n_{PV} = 1, \eta)$$

- different impact of zero suppression inside the jet for ZB and MB data (estimated with MC)

- systematic uncertainties are estimated as E_T residuals between data and fits for each i_η ring; **< ~2%**



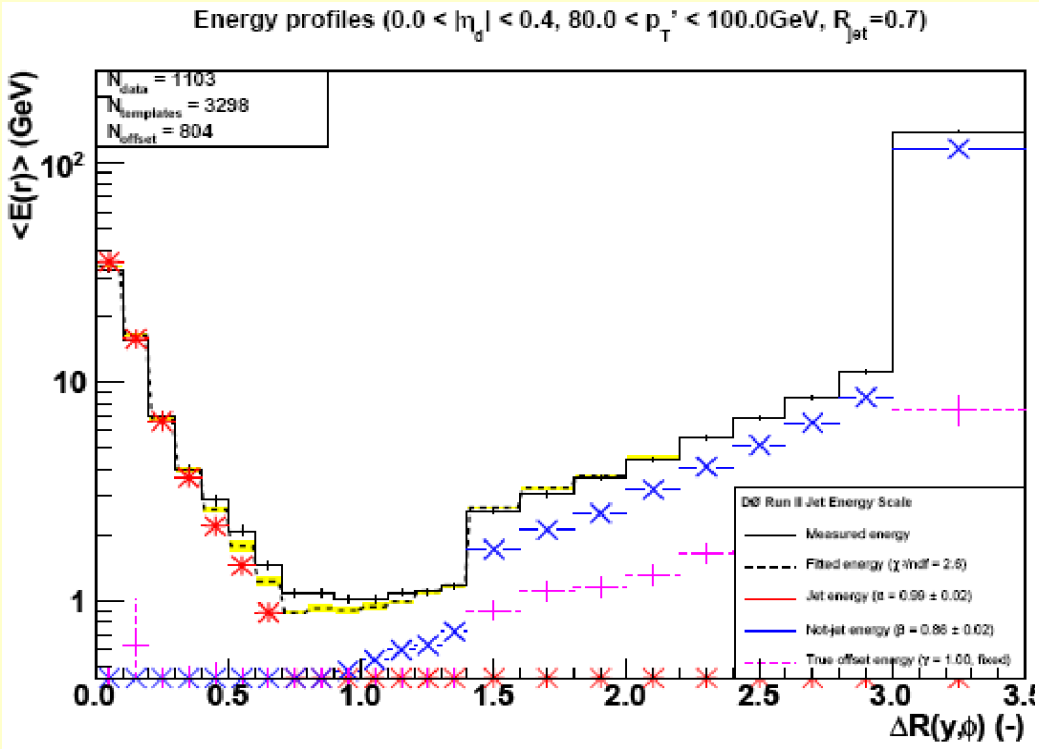


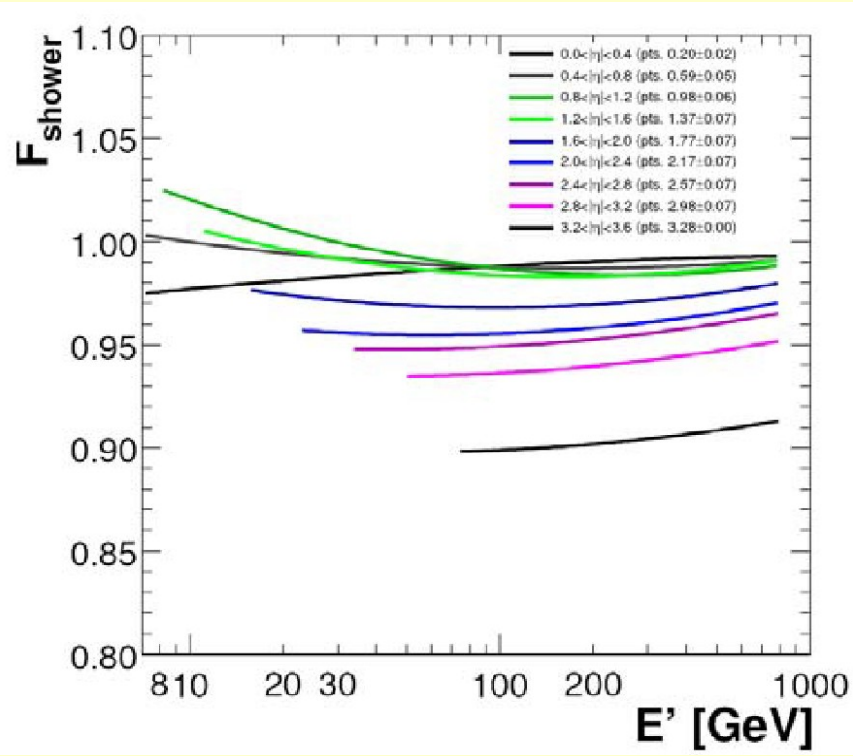
JES – Showering corrections

Goal:

Accounts only for *detector showering (multiple scattering, magnetic. field, etc …),* not physics

- **<u>In MC:</u>** direct evaluation by tracking particles from particle jets in γ +jets events
- **<u>In Data:</u>** measure jet energy profile as a function of distance from the jet axis. Define particle, non-particle and offset profiles.
  - MC templates provide good description of data & method validated on MC.

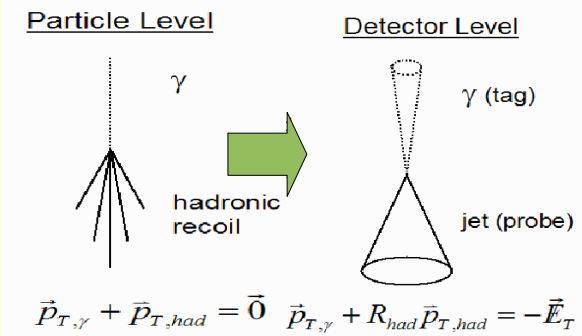




Jet Energy Response Calibration (I)

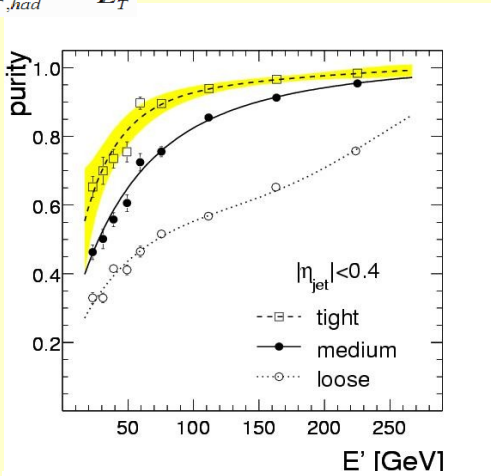
Missing transverse energy Projection Fraction method (MPF)

$$R_{had} = 1 + \frac{\vec{E}_T \cdot \vec{p}_{T,\gamma}}{\vec{p}^{\,2}_{T,\gamma}}$$



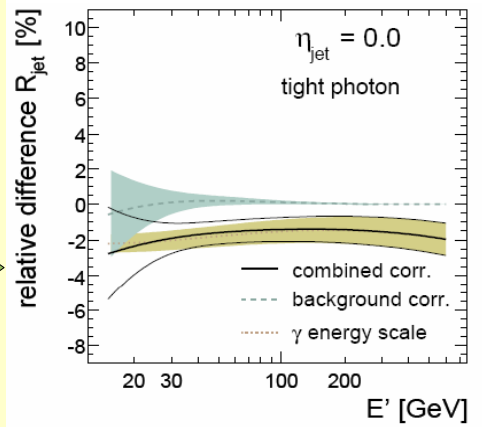
For back-to-back events: $R_{had} \sim R_{jet}$

- insensitive to jet cone and showering effects
- di-jets used to increase stat. explore higher regions in jet p_T spectra



EM Scale

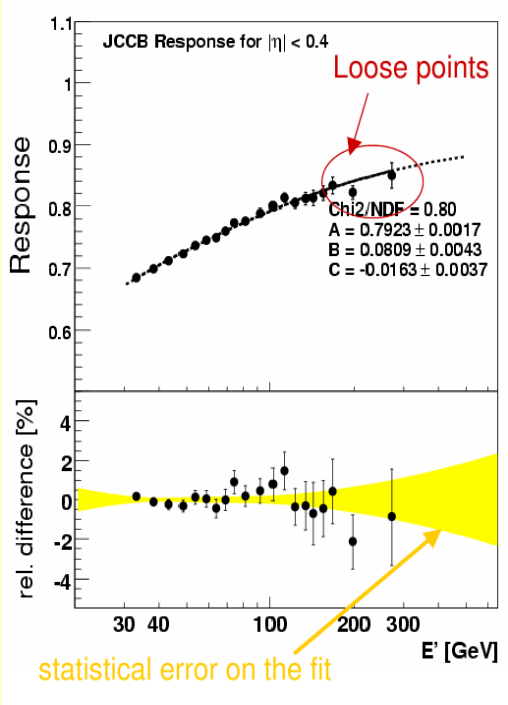
- purity correction: γ+jet altered by background di-jets events (using MC)
- e/γ EM response correction: tuned with detailed MC shower profiling studies (EM scale is calibrated using Z—>e^+e^- data)



Jet Energy Response Calibration (II)

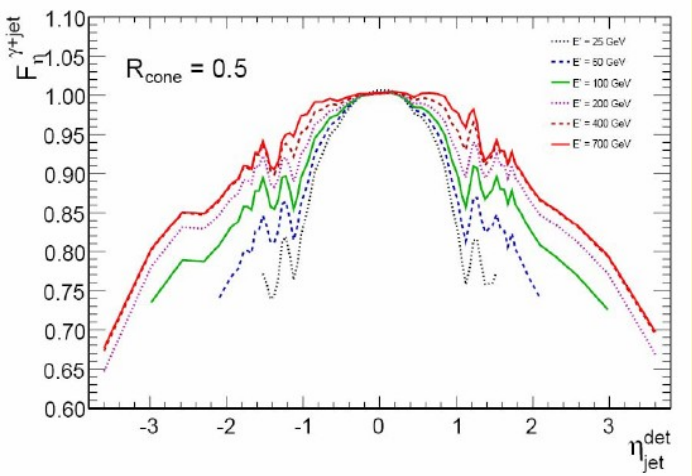
Response in estimated in Central Calorimeter (CC)

- Largest correction in JES calibration
- high p_T region are estimated with fit



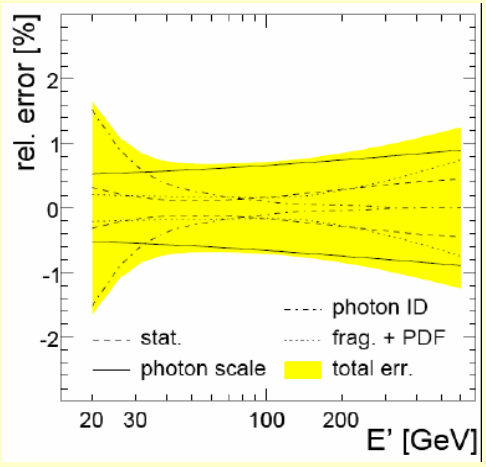
Relative calibration

- calibrate the forward jets w.r.t the central response, F(η)



Main sources of uncertainties

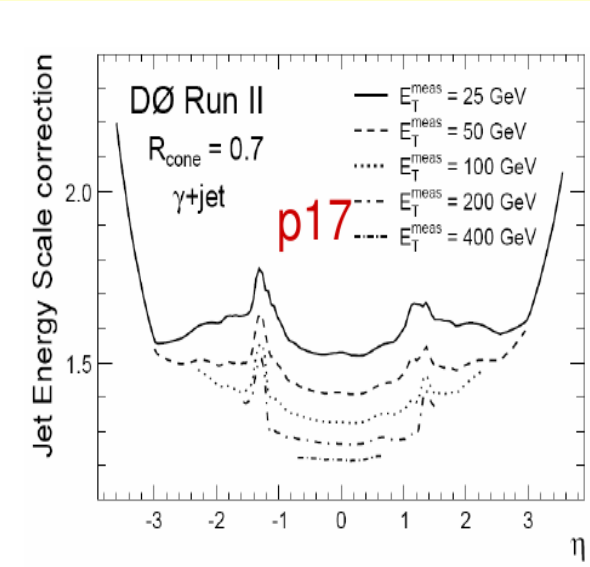
- the statistical uncertainty of the fit
- γ energy scale & purity correction
- high energy extrapolation
- Fragmentation and PDF (@ high p_T)

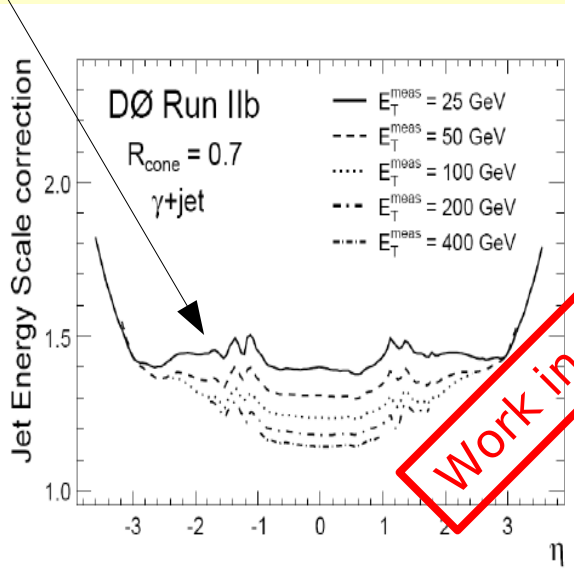


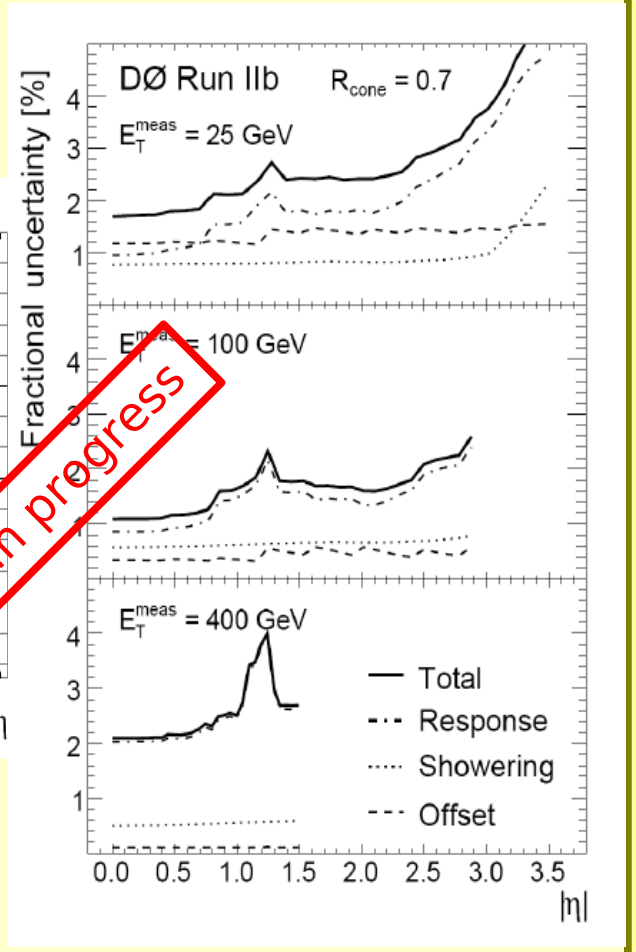
JES – Results

Upudated RunIIb JES calibration in progress

- New ICD calibration







Work in progress

Jet Energy Resolution (JER)

J.L Agram PhD Thesis

True resolution:

$$\frac{\sigma_{p_T}}{p_T} = RMS\left(\frac{p_T^{reco} - p_T^{ptcl}}{p_T^{ptcl}}\right)$$

... can be measured in data using di-jet asymmetry:

Raw asymmetry — *Raw resolution*

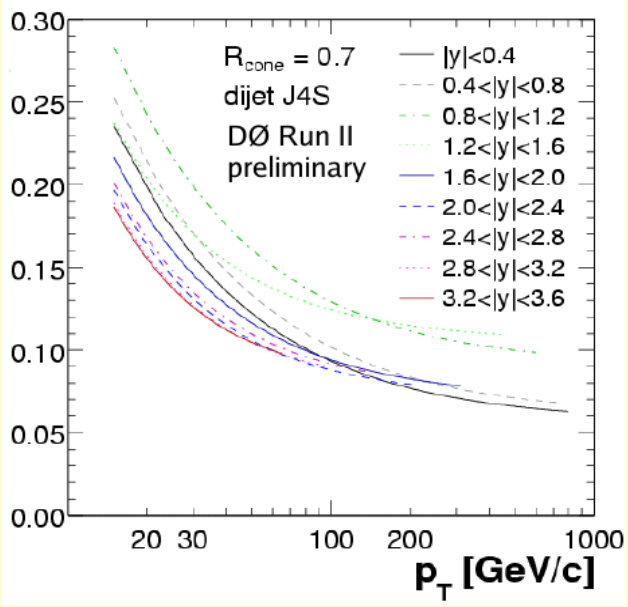
$$A = \frac{p_{T,1} - p_{T,2}}{p_{T,1} + p_{T,2}} \Rightarrow \frac{\sigma_{p_T}}{p_T} = \sqrt{2}\,\sigma_A$$

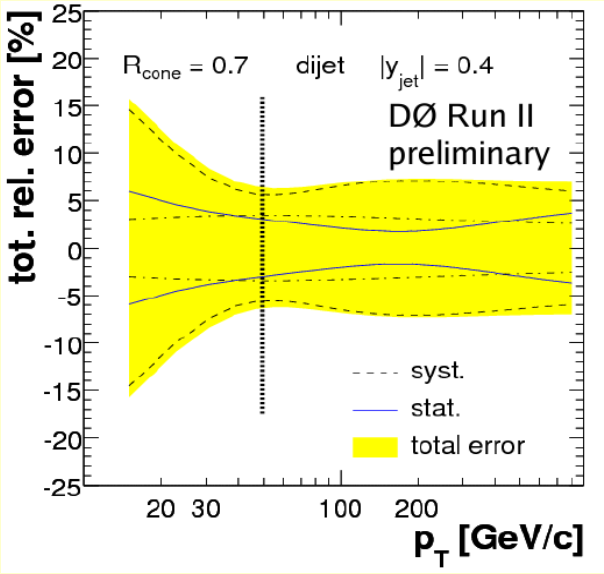
- Corrections: soft radiation (un-reconstructed soft jets) + particle level imbalance (fragmentation., ...)

Resolution measured directly in data using *A* parametrized by:

- Noise, Stochastic and Constant terms:

$$\frac{\sigma_{p_T}}{p_T} = \sqrt{\frac{N^2}{p_T^2} + \frac{S^2}{p_T} + C^2}$$





Jet Smearing, Shifting, Removing (S.S.R)

N. Makovec / C. Ochando Phd. Theses

After standard JES corrections residual differences *between data and simulation* may still remain

Rather than the *absolute* energy scale, it is thus the *relative* energy scale that is of interest for some physics analyses.

The **S.S.R** method proposes to correct, all in a consistent way, three items of the simulated data which are not correctly reproduced:

- Jet energy scale, energy resolution & (identification ∗ reconstruction) efficiency

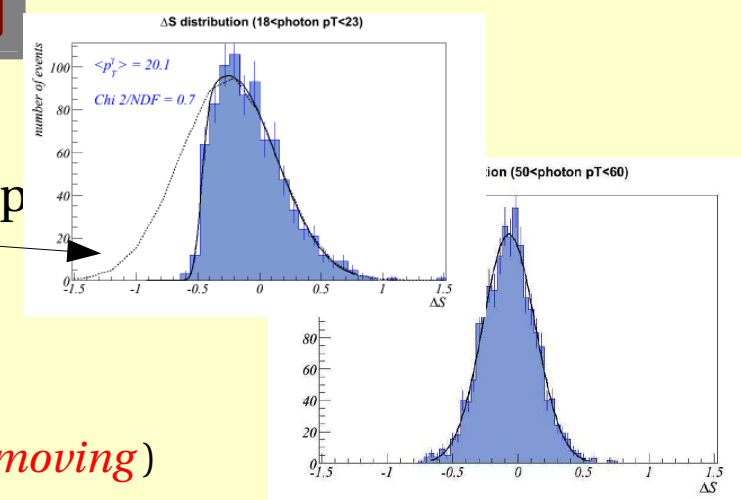
by measuring the *imbalance* variable:

$$\Delta S = \frac{p_T^{\mathrm{jet}} - p_T^{\gamma}}{p_T^{\gamma}}$$

in γ / Z(-->ee)+jet data

ΔS is described by a function:

$F^{\Delta S}$ = Gauss (* Turn-On) in different bins of γ /Z p

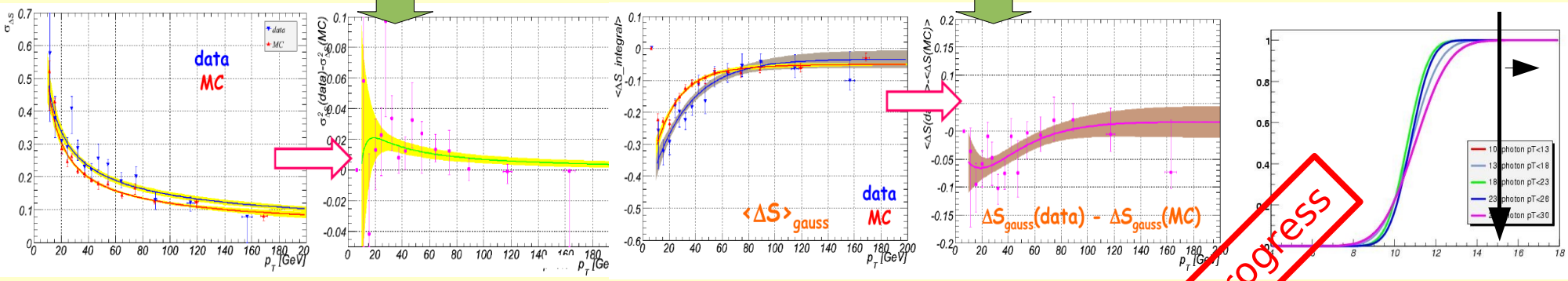


- Mean: energy scale (*shifting*)
- Width: energy resolution (*smearing*)
- Turn-on: p_T threshold bias, i.e (reco.* id.) eff. (*removing*)

S.S.R method

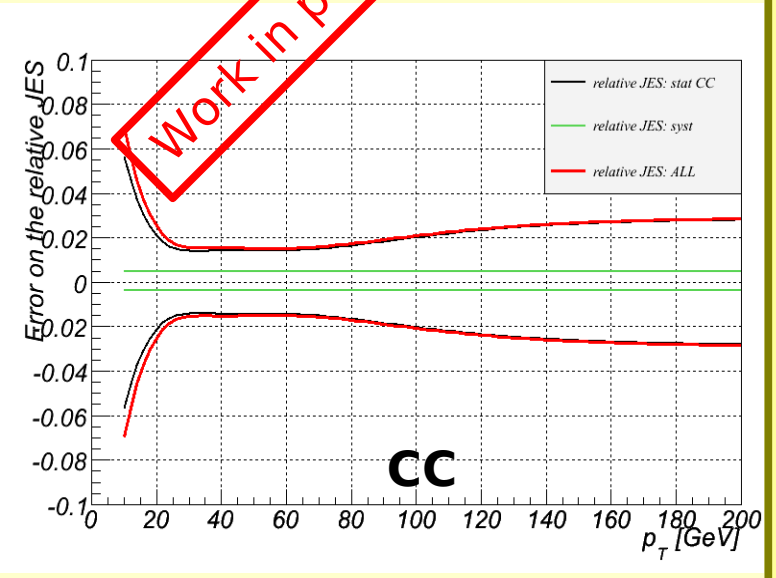
Corrections to be applied to simulation

Δ_{width}(data,MC) = **smearing** ***(resolution)*** Δ_{mean}(data,MC) = **shifting** ***(Energy scale)*** **...** and finally **removing**

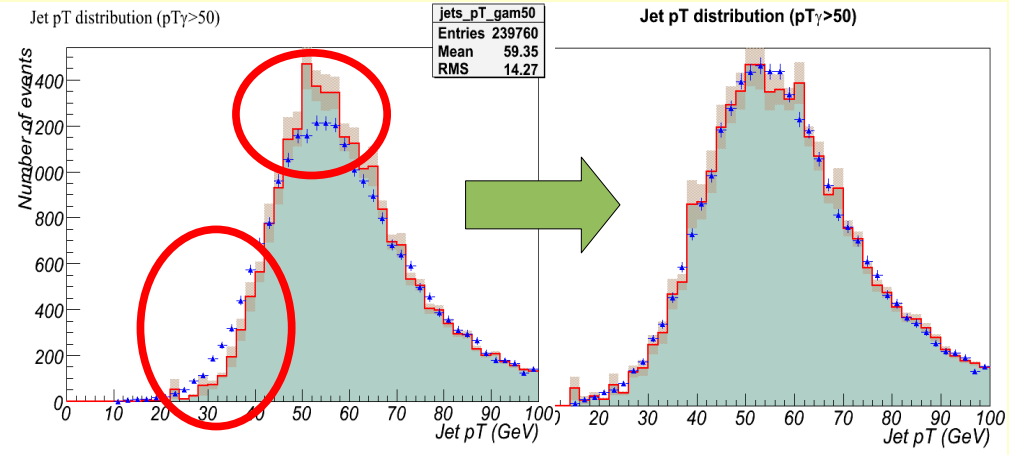


Errors:

- EM energy scale (0,2%) & jet multiplicity (0.3%)
- relative eta calibration (1.5%) (includes also flavour composition differences in γ / Z+jet data)



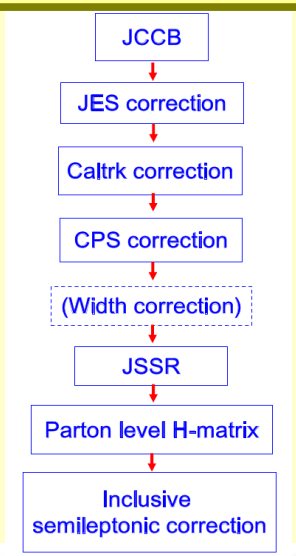
Results:

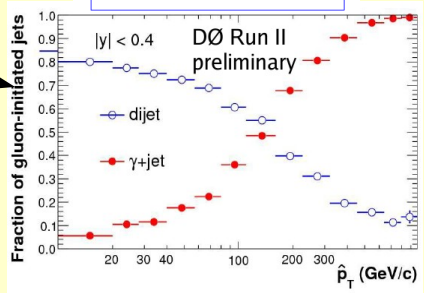


Conclusion/ Future

Room for improvements

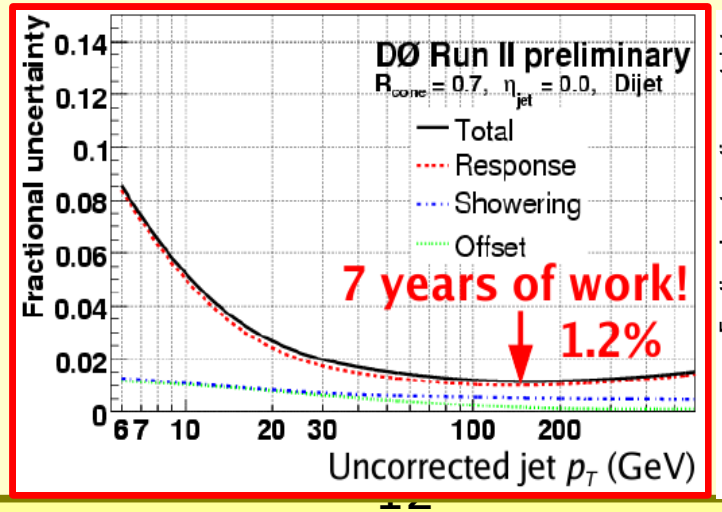
- B-jets dedicated corrections (*crucial for Higgs searches, and if not found to constrain M_{top} vs. M_W)*
  - Look @ jets with <u>semi-leptonic</u> decays and/or displaced vertex
- <u>Energy-flow</u>: track in jets (neutrals ?), preshowers (CPS)
- Particle-jet to parton JES/JER
- **Sample dependance,** γ / Z+jets, di-jets sample have different jet composition, i.e q/g fractions

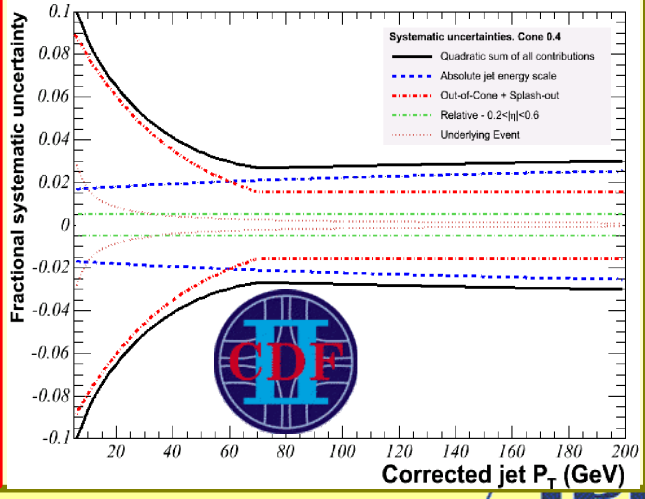




Message for LHC

it's been hard and long to get there ...

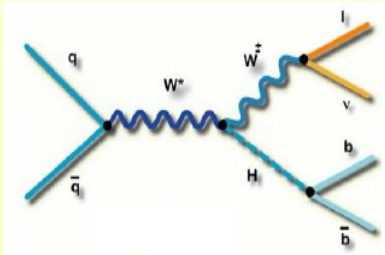


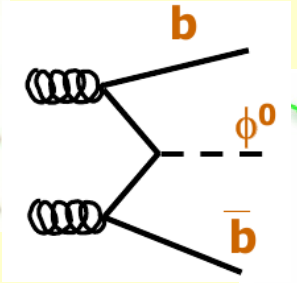


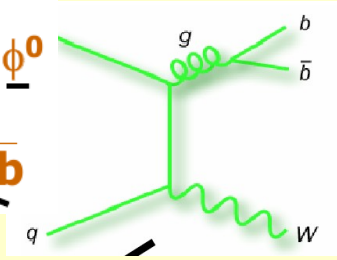
Heavy flavour identification

Physics

- Top physics: x-section, mass, single-top
- Higgs searches: Low-mass, SUSY
- "Backgrounds": W/Z+heavy flavour

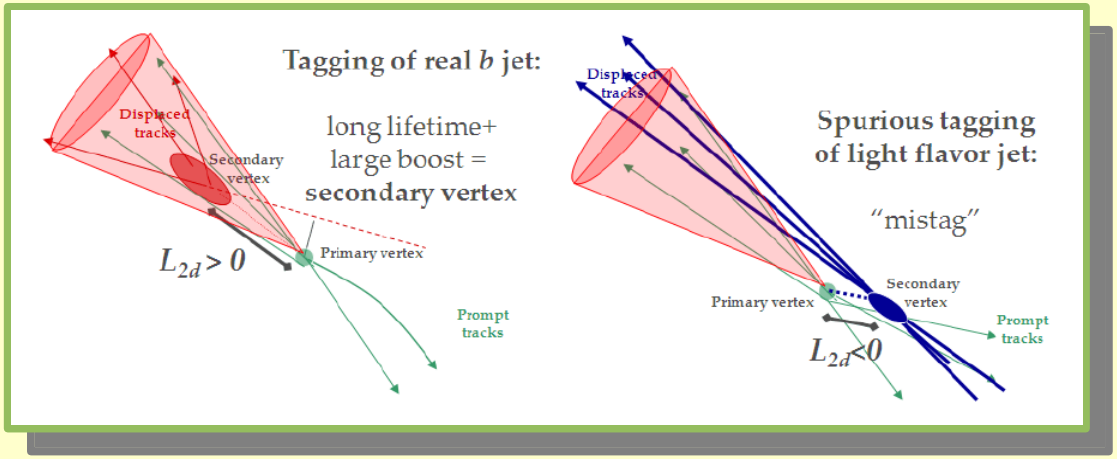






B hadrons properties:

- Mass: ~5 GeV/c^2
- Lifetime: ~1.6 ps
- Semi-leptonic decays



Taggability

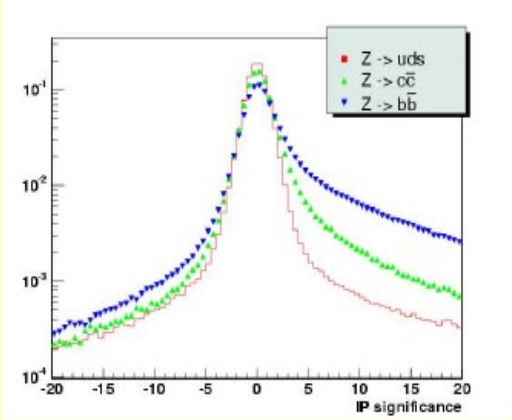
- Only a subset of *all* calorimeter jets is of interest for hf-identification: those with a minimum of good quality tracks attached in cone, i.e ***taggable*** jets (>=2 tracks, with 1 hit in SMT)
- 2-step clustering: along beam axis + 0.5 cone jets (within each z-cluster)
- Require: ΔR(*calo-jet,track-jet*) < 0.5
- HF-identification performance normalized to taggable jets only

Heavy flavour identification

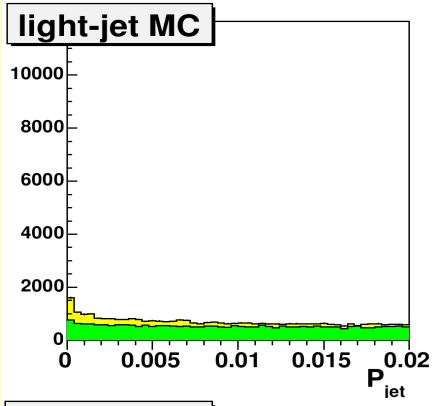
S. Greder , B. Clement & V. Siccardi PhD Theses

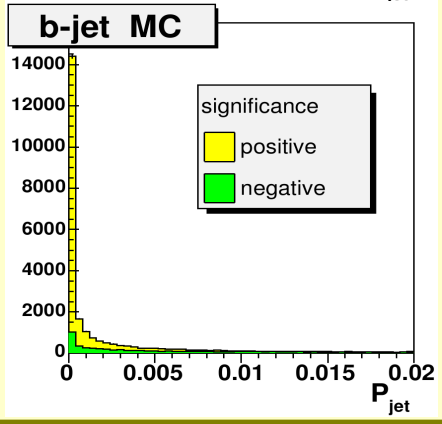
Impact Parameter (IP) based tagger

- Discrete: CSIP, counts tracks with S_{IP} > cut
- Continuous: JLIP, p.d.f from negative S_{IP} resolution function, *R(s)*
  - IP error calibrated in data and simulation for multiple-scattering effects and PV resolution dependence



$$\mathcal{P}_{trk}(\mathcal{S}_{IP}) = \frac{\int_{-50}^{-|\mathcal{S}_{IP}|} \mathcal{R}(s)ds}{\int_{-50}^{0} \mathcal{R}(s)ds} \Rightarrow \mathcal{P}^{\pm}_{jet} = \Pi^{\pm} \times \sum_{j=0}^{N^{\pm}_{trk}-1} \frac{(-\log \Pi^{\pm})^j}{j!} \quad \text{with} \quad \Pi^{\pm} = \prod_{i=1}^{N^{\pm}_{trk}} \mathcal{P}_{trk}(\mathcal{S}^{IP>0}_{IP<0})$$





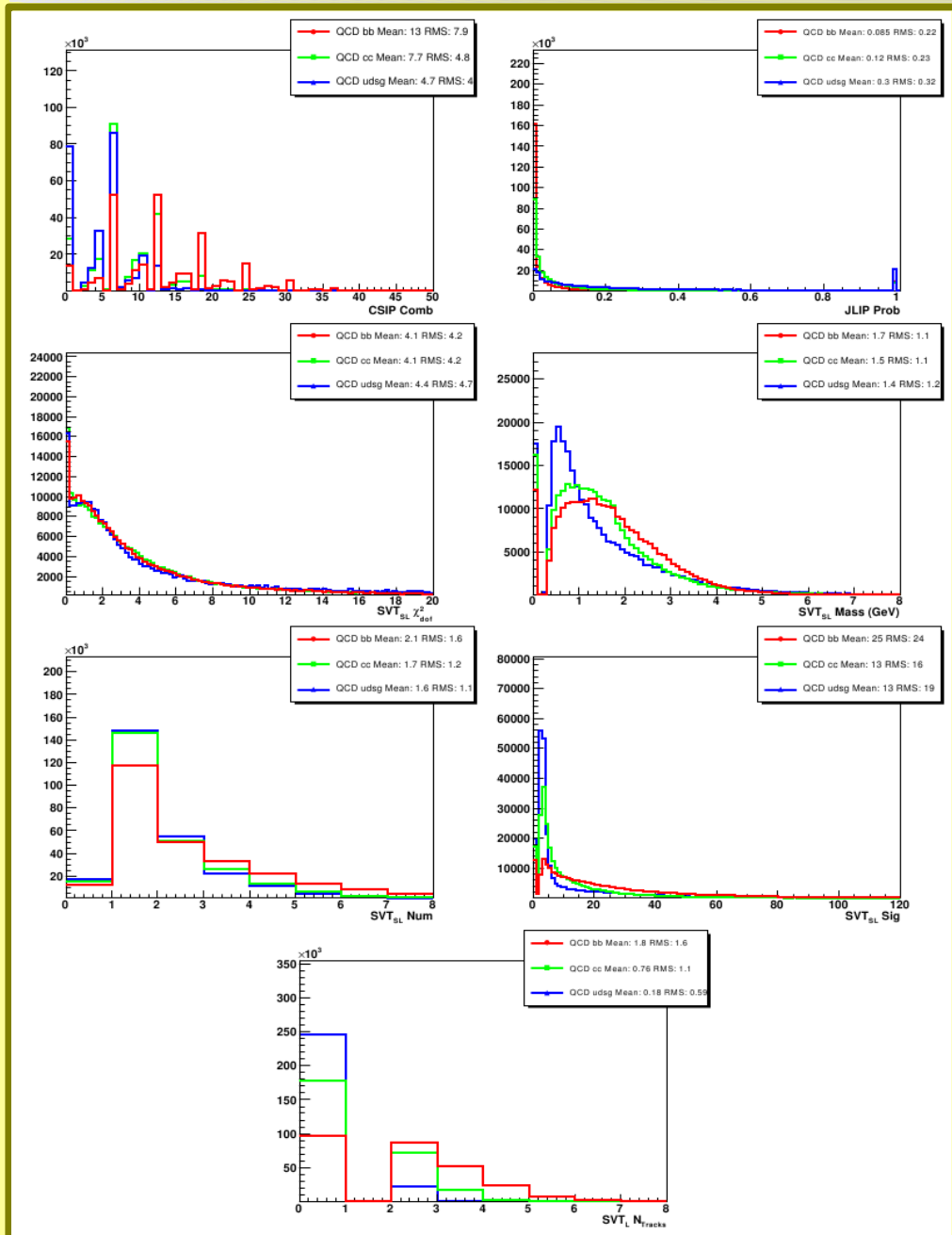
Secondary vertex, SVT

- Kalman-filter based vertex finder
- Track pruning w.r.t χ^2 contribution to vertex

Soft Lepton Tagger, SLT

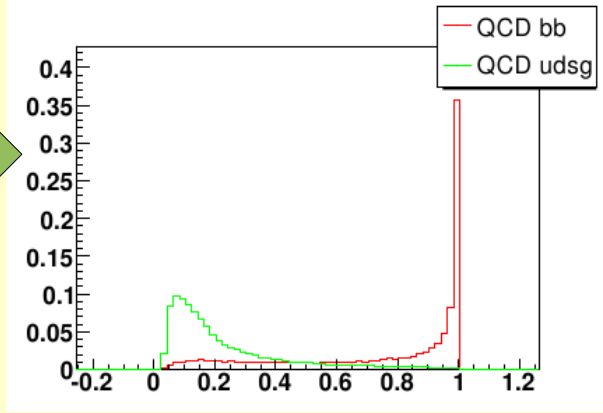
- 40% of low-mass Higgs (H->bb) contain a lepton !

All in one: Neural Network tagger

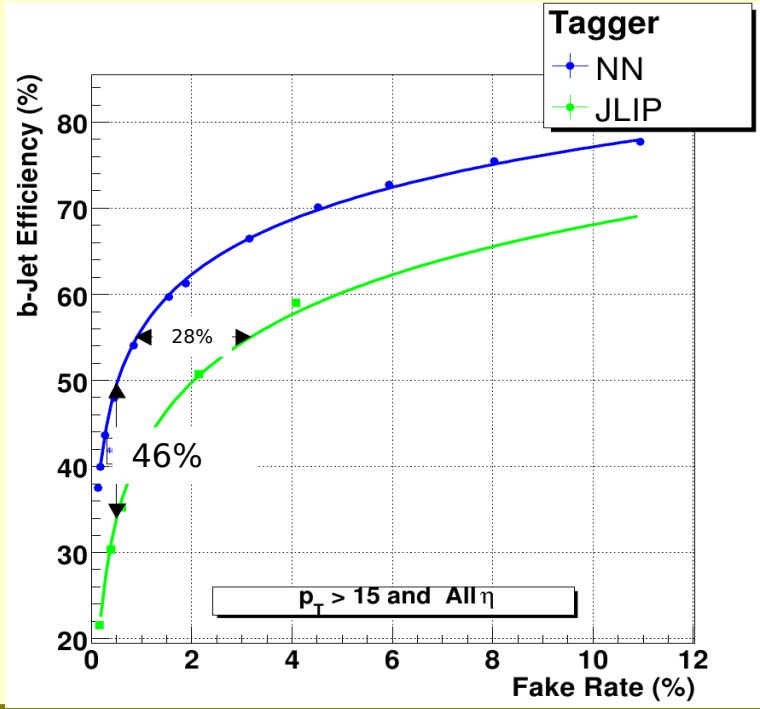


Optimized selection of inputs:

CSIP, JLIP & 4 SVT properties



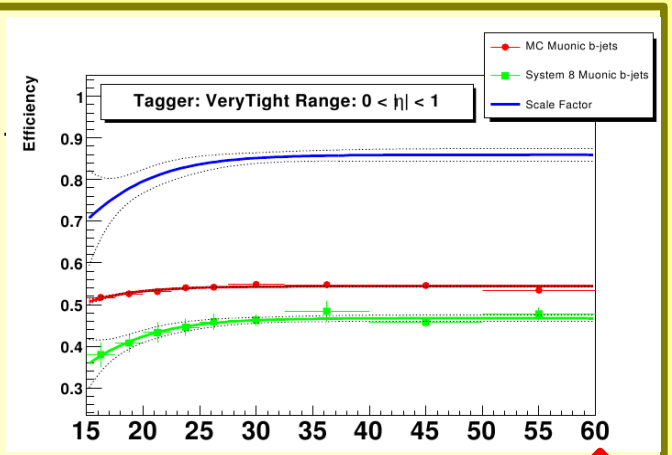
... can lead to significant improvement:



Performance (preliminary)

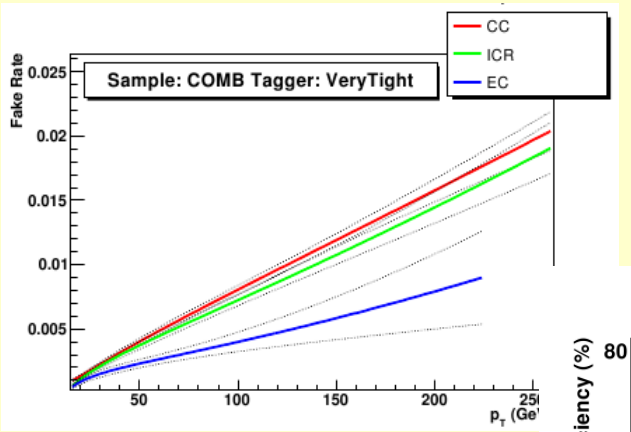
Measured in data

- Original method to extract signal efficiency *directly* in b-enriched data (*muon-in jet*)
- Minor corrections (*correlations*) from MC
- V^0 removal



Fake rate estimated from negative tags corrected for:

- HF contamination
- negative/positive asymmetry

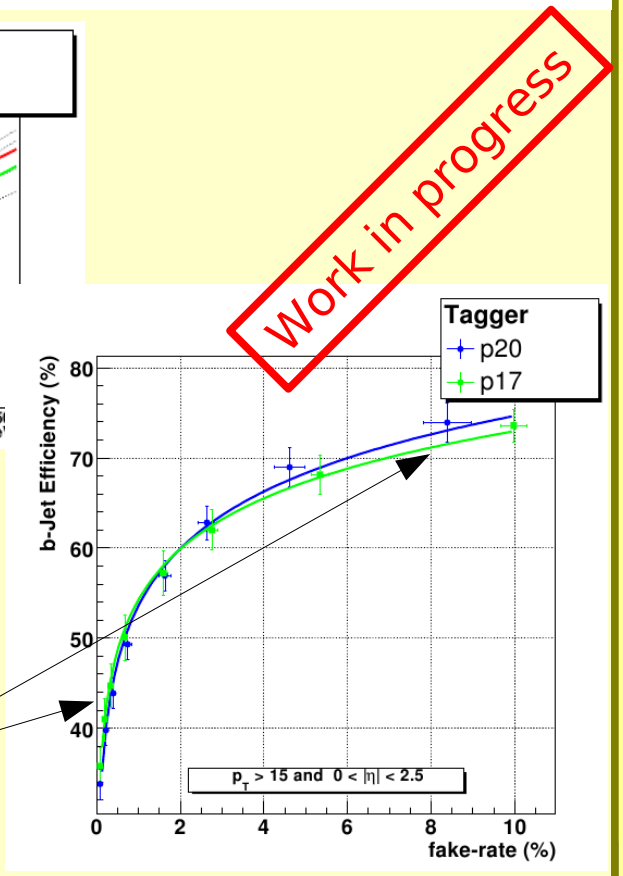


Systematics

- MC samples dependence
- Taggers correlations
- Total: **~3-4%**

Scale Factors applied to simulated jets SF(p_T,η,flavour)

- Defined for different performance operating points



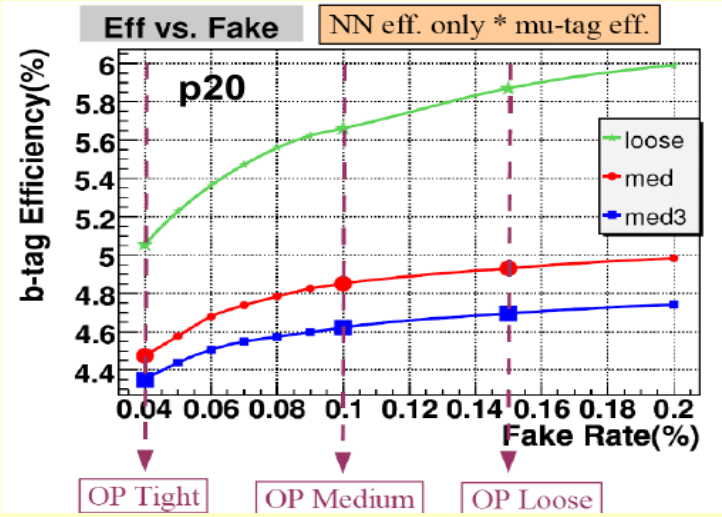
Future (I)

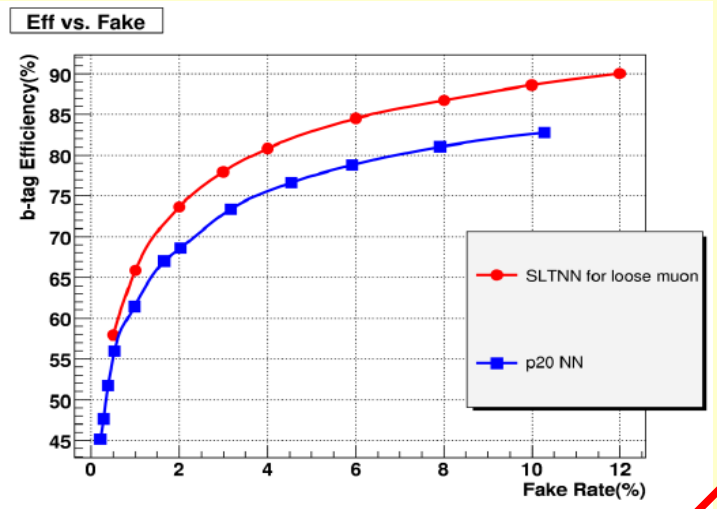
D. Jamin, F. Beaudette PhD Theses

Improved tagging in dedicated topologies

SLTNN with muons

- SLT variables (p_T^{rel}, χ^2, ΔR(jet), ...) can be combined with lifetime variables in a dedicated NN to improve identification performance for semi-leptonic b decays
- Up to **10%** relative increase of signal efficiency @ same fake rate level





SLT with electrons

- Reconstruction of (low-p_T) electrons in jets is more challenging
- b->eX ~25% identification efficiency for 1% fake rate

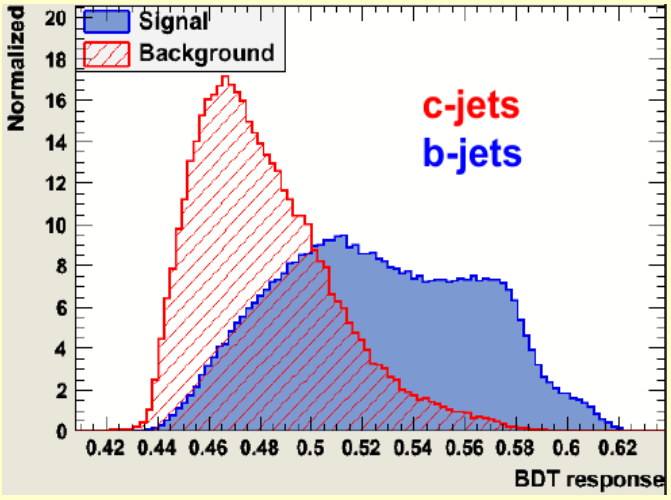
Work in progress

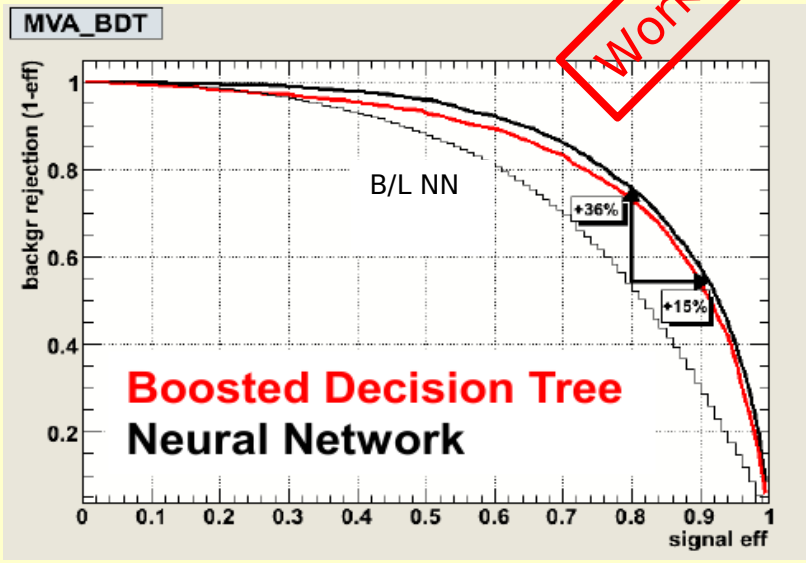
Future (II)

Testing different Multivariate (MVA) techniques allows to efficiently exploit differences of fragmentation properties between b and c hadrons

- charged multiplicity, transverse neutral missing charged momentum, collimation, ...

- **Boosted Decision Trees** trained against NN (b vs.light) selected candidates jets





- Tune physics analysis on *(NN,BDT)* 2D plane

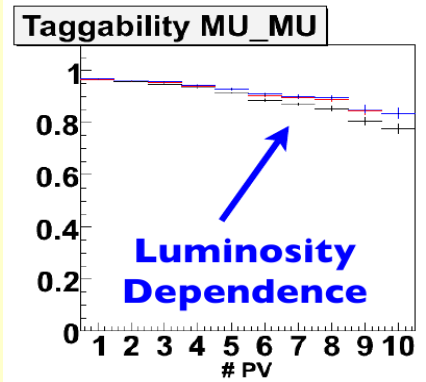
Conclusion

Very good performance despite of *complex* and *busy* hadronic environment:

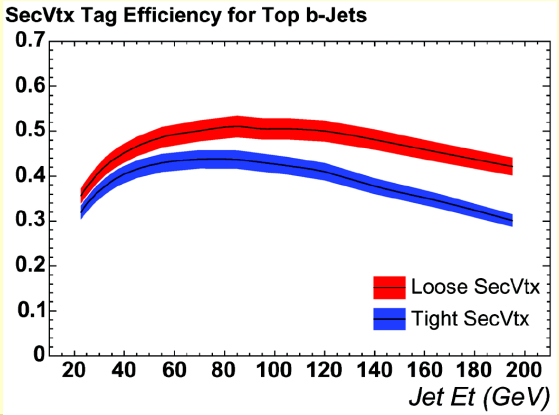
- Advanced multi-variate tools are an asset to keep high signal efficiency/low fake rates
  - And simplify procedures: reduce complexity to 1 variable
- Neural Networks (*b vs. l, SLT*), Boosted Decision Trees (*b vs. c*), ...

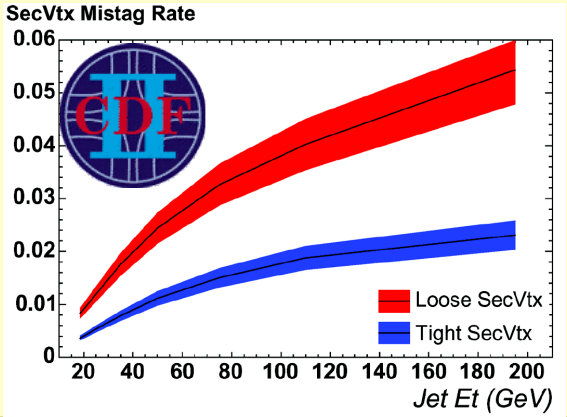
Increasing luminosity must be *carefully* handled

- Performance normalization sensitive to the number of multiple interaction



DØ: 60% sig. eff./ 2% fakes, **CDF:** 50% sig. eff./ 1.8% fakes





General conclusion

Precise and exhaustive understanding of jet energy scale / resolution and hf-identification is crucial for many physics analyses

- Part of top-3 main systematics for top mass, single-top, low-mass Higgs searches, …

Challenging environment stimulates for developing new original ideas

but investing a lot of effort therein … well, is worth it :