URA thesis award 2009:

Inclusive jet cross section measurement at D0

Mikko Voutilainen

CERN

For work done with:

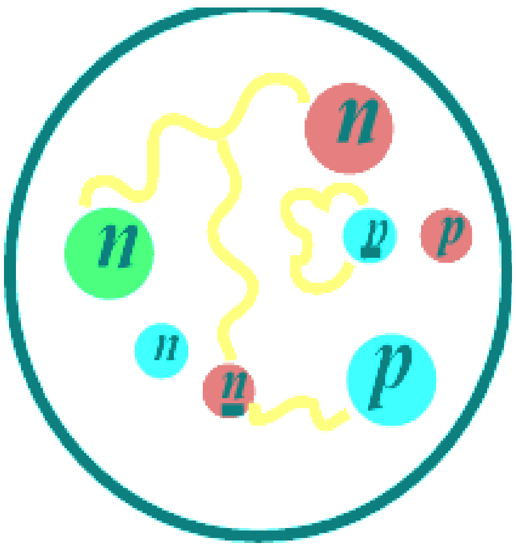
University of Nebraska-Lincoln
Helsinki Institute of Physics
CEA Saclay

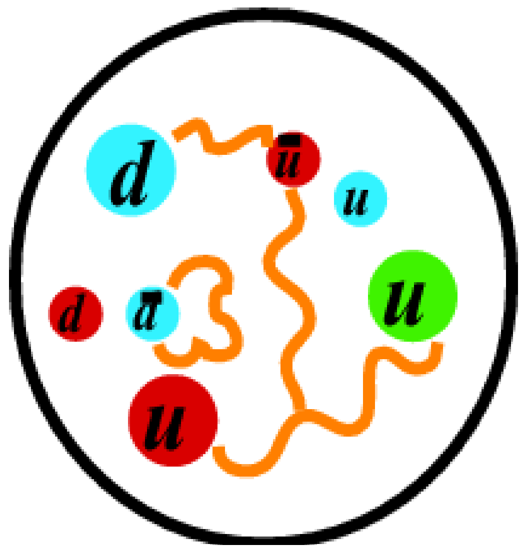
For DSc/PhD from:

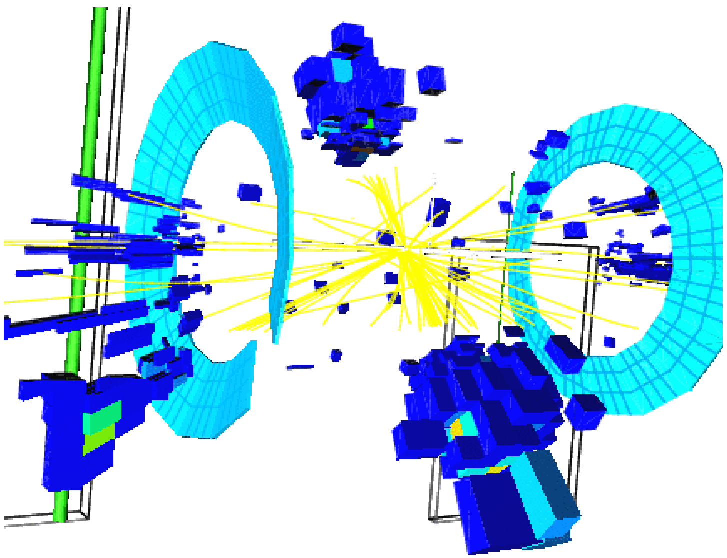
Helsinki University of Technology
Universite de Paris-Sud 11 (Orsay)

Proton structure

- Proton consist of three valence quarks, up-up-down (uud) and antiproton of three antiquarks ($\bar{u}\bar{u}\bar{d}$) bound together by a sea of gluons (g) and virtual quark-antiquark pairs ($u\bar{u}$, $d\bar{d}$, $s\bar{s}$ etc.)



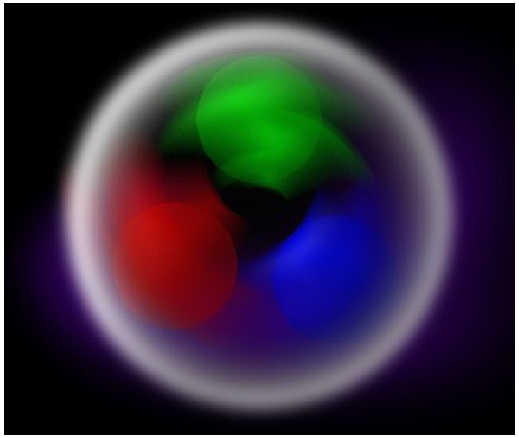


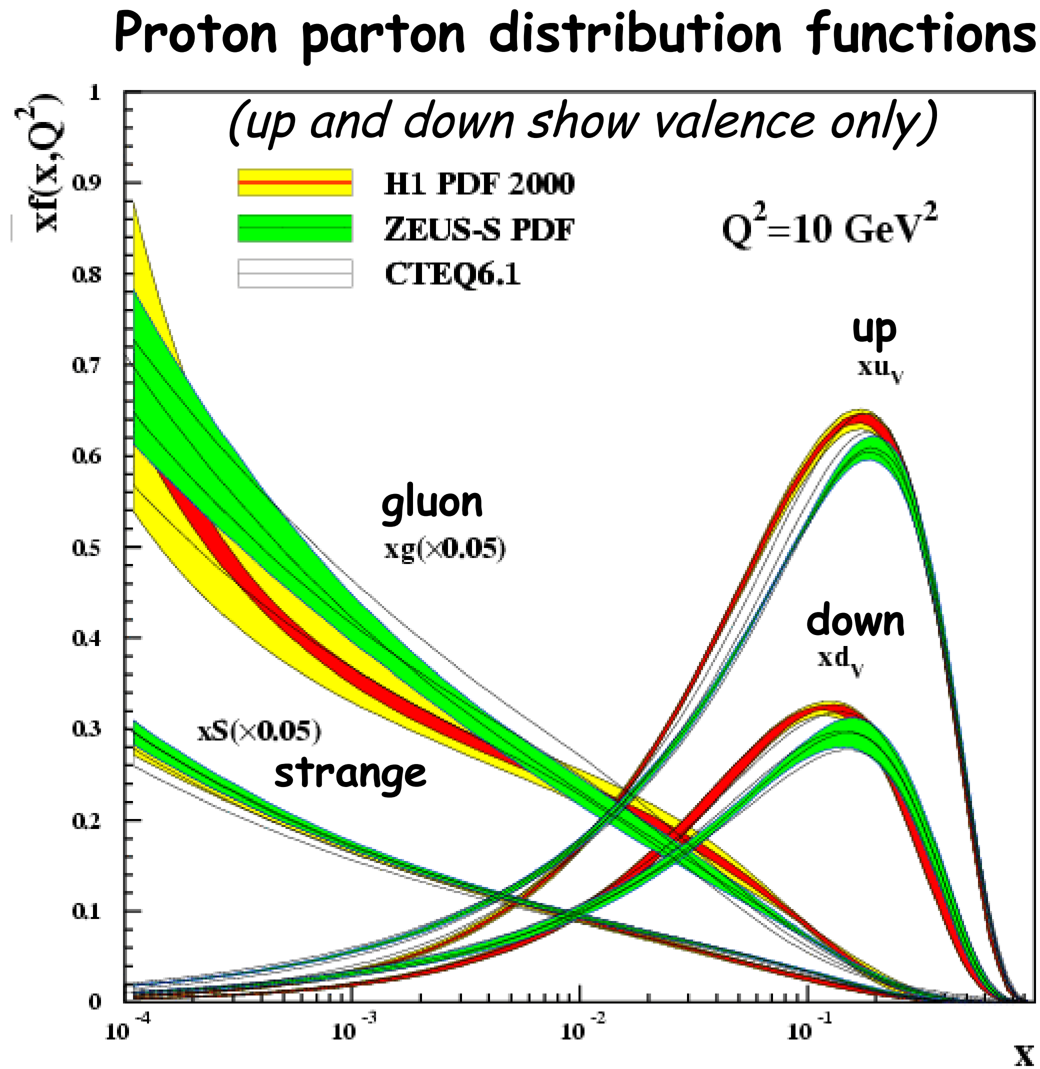


- Higher energies resolve finer detail, and Tevatron's energies are currently the highest available (1.96 TeV)
- Different final states give access to different aspects; inclusive jets look at big picture and can search beyond standard model

Proton Structure

- What can we learn by measuring jets?
  - Distribution of constituents in protons and antiprotons, called partons
  - Nature of the basic interaction between quarks and gluons
  - Find quark substructure?



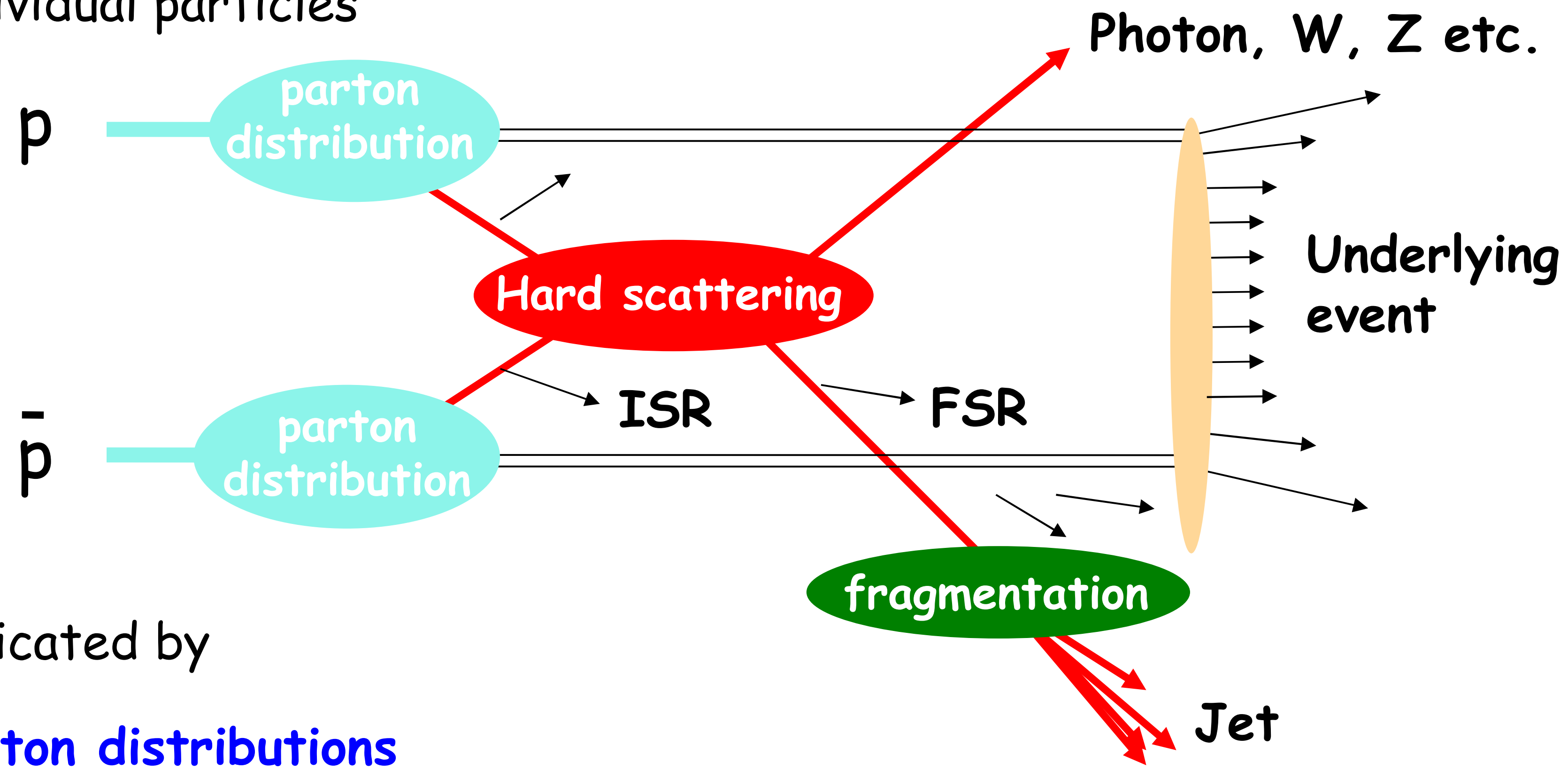


x: fraction of proton momentum carried by individual parton
f(x,Q2): probability of finding parton with momentum fraction x in interval dx

Event schematic

Jet is a spray of particles coming from **hard interaction**

- Jets are formed by collisions of partons (quarks and gluons) from individual particles



Complicated by

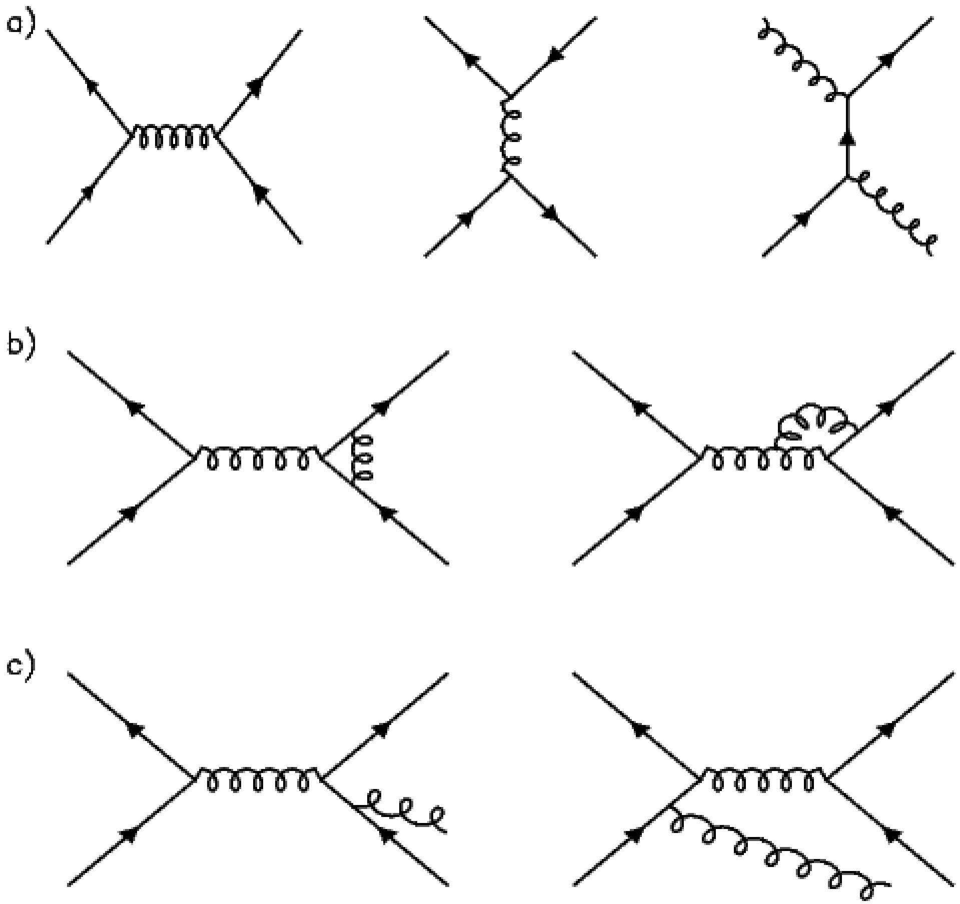
- **Parton distributions**
  - hadron collider is really a broad-band quark and gluon collider
  - both the initial and final states can be colored and radiate gluons
- **Underlying event** from proton remnants

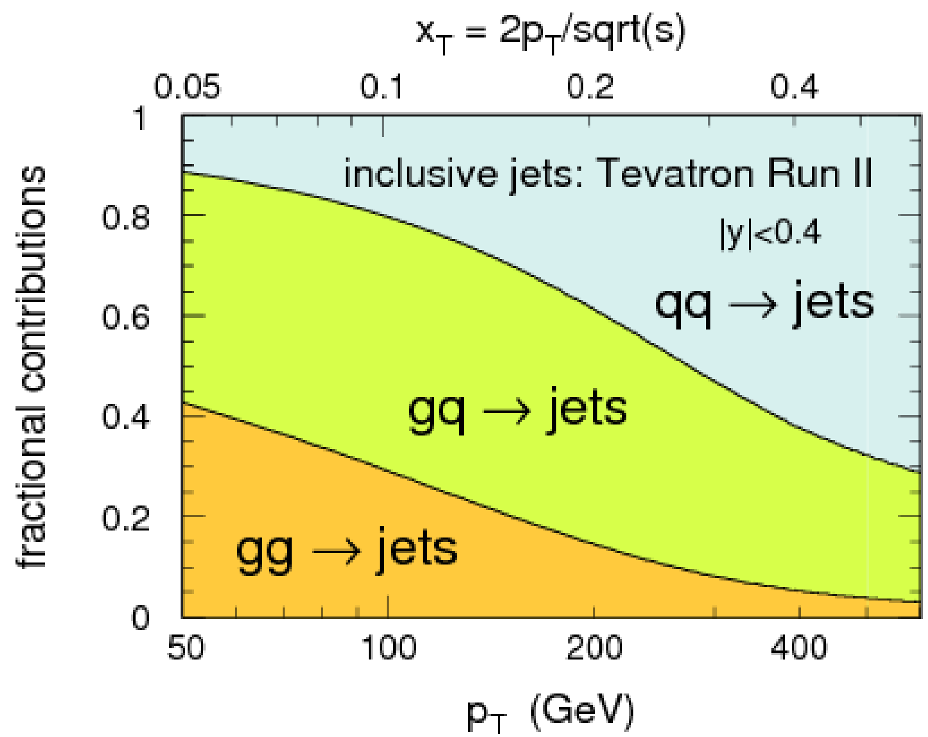
Quantum chromodynamics

- Quantum chromodynamics is calculable using perturbation theory (Feynman diagrams) at high p_T ⇒ **hard scatter**
- Standard is next-to-leading-order (NLO), some NNLO available

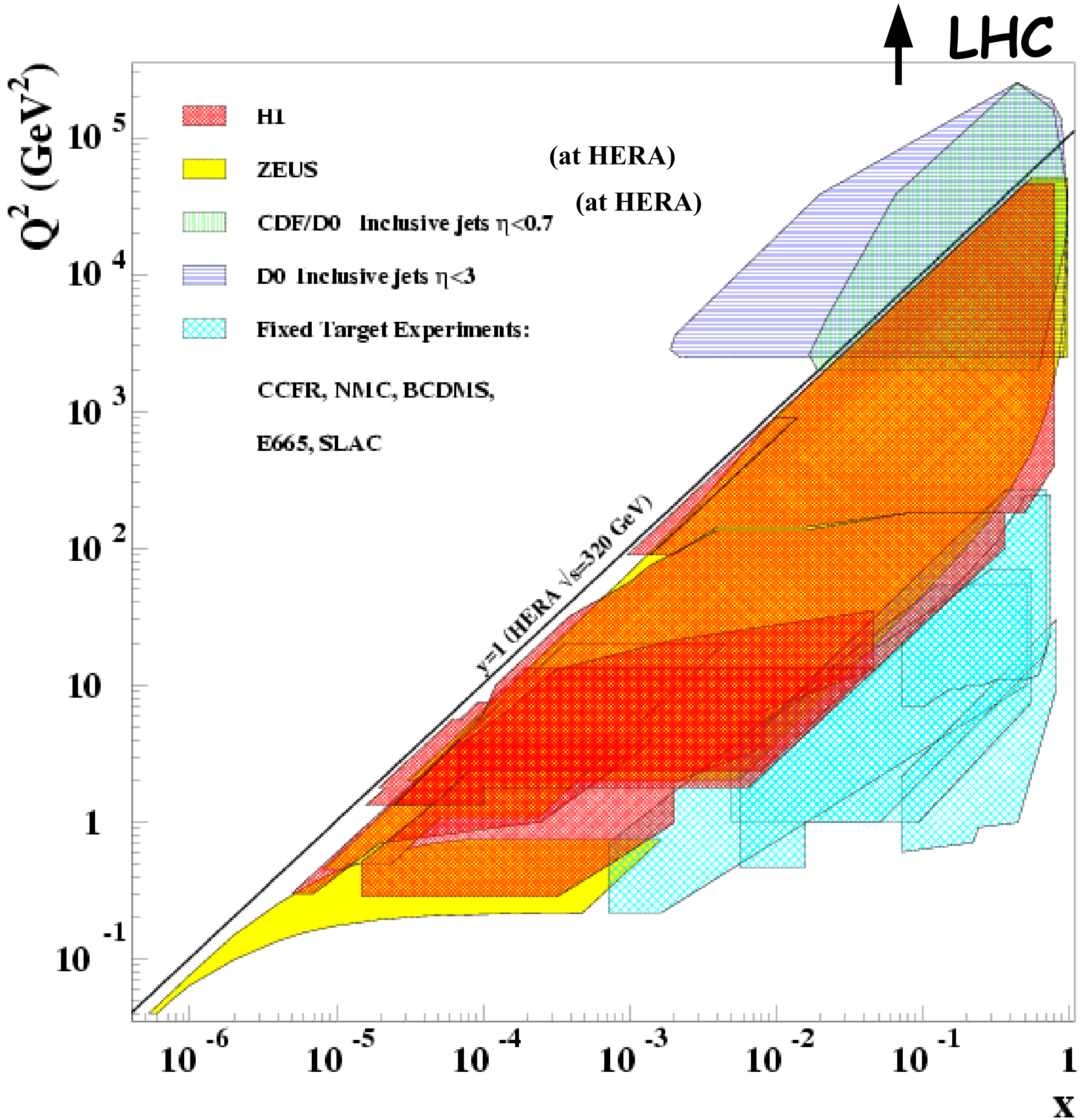
LO

NLO





D0 kinematic range



D0, CDF

HERA was an electron-proton collider in DESY research center in Hamburg, Germany

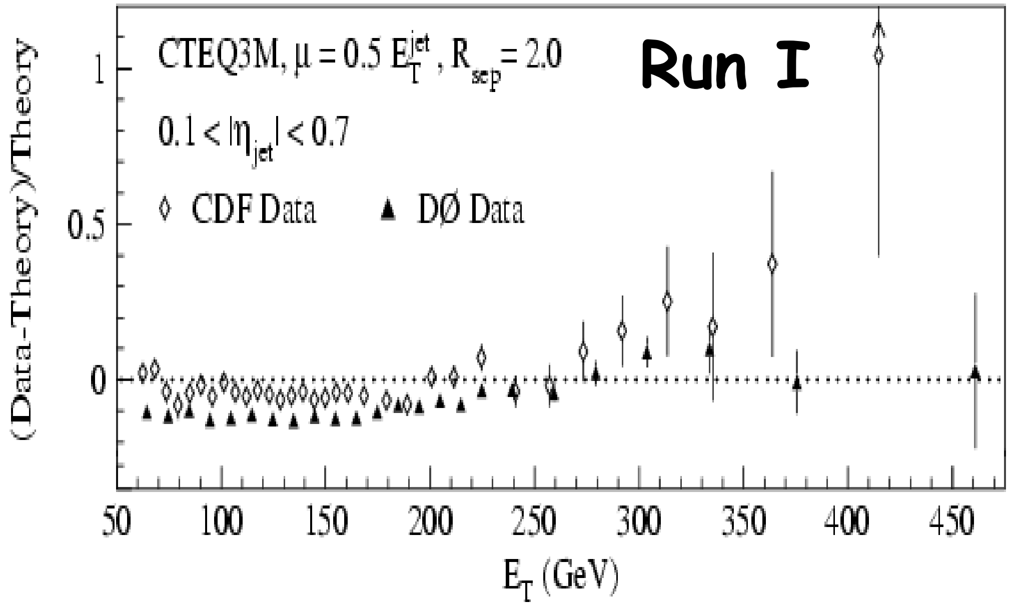
- We are complementary to HERA and fixed target experiments
- Tevatron inclusive jets best to constrain the **high-x gluon distribution**

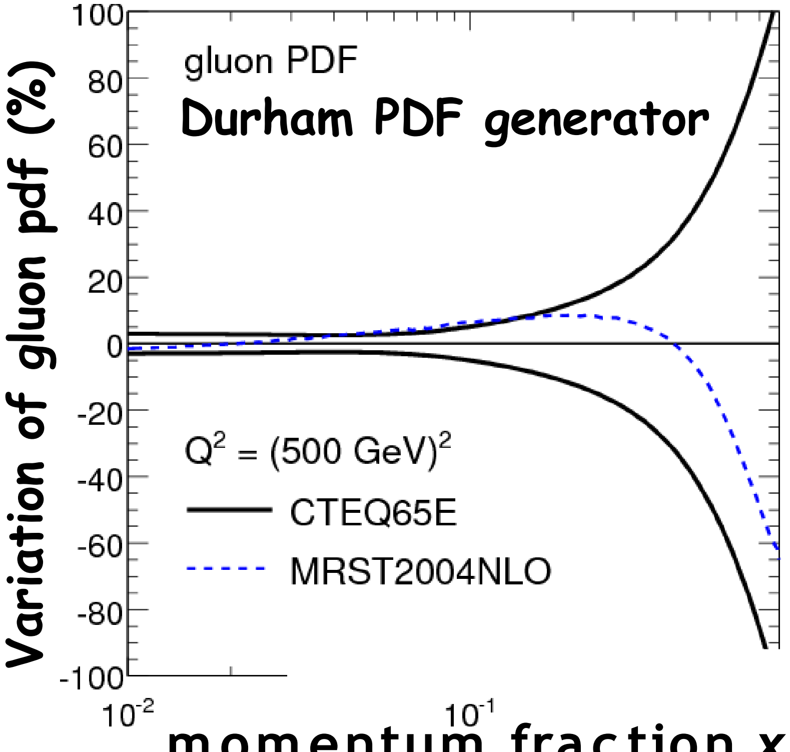
Quark substructure

- Run I legacy: significant freedom for high-*x* gluon PDF
- Once the high-*x* gluon PDF is nailed down we can search for **quark substructure**
- Important measurement to be performed

1) at low rapidities: sensitive to PDF/quark substructure
2) in wide range of rapidity: at high *y*, sensitive to PDF

- One single measurement is sensitive to both effects

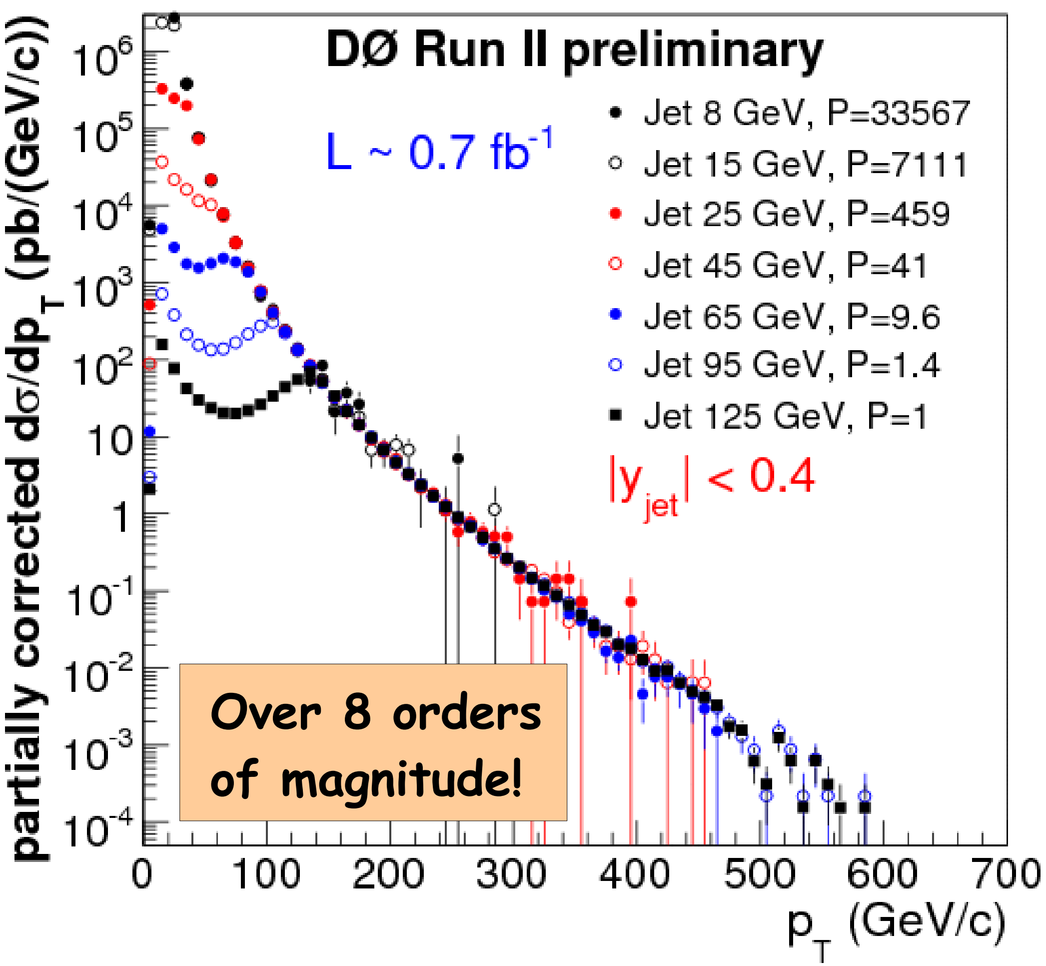




What is so cool about this

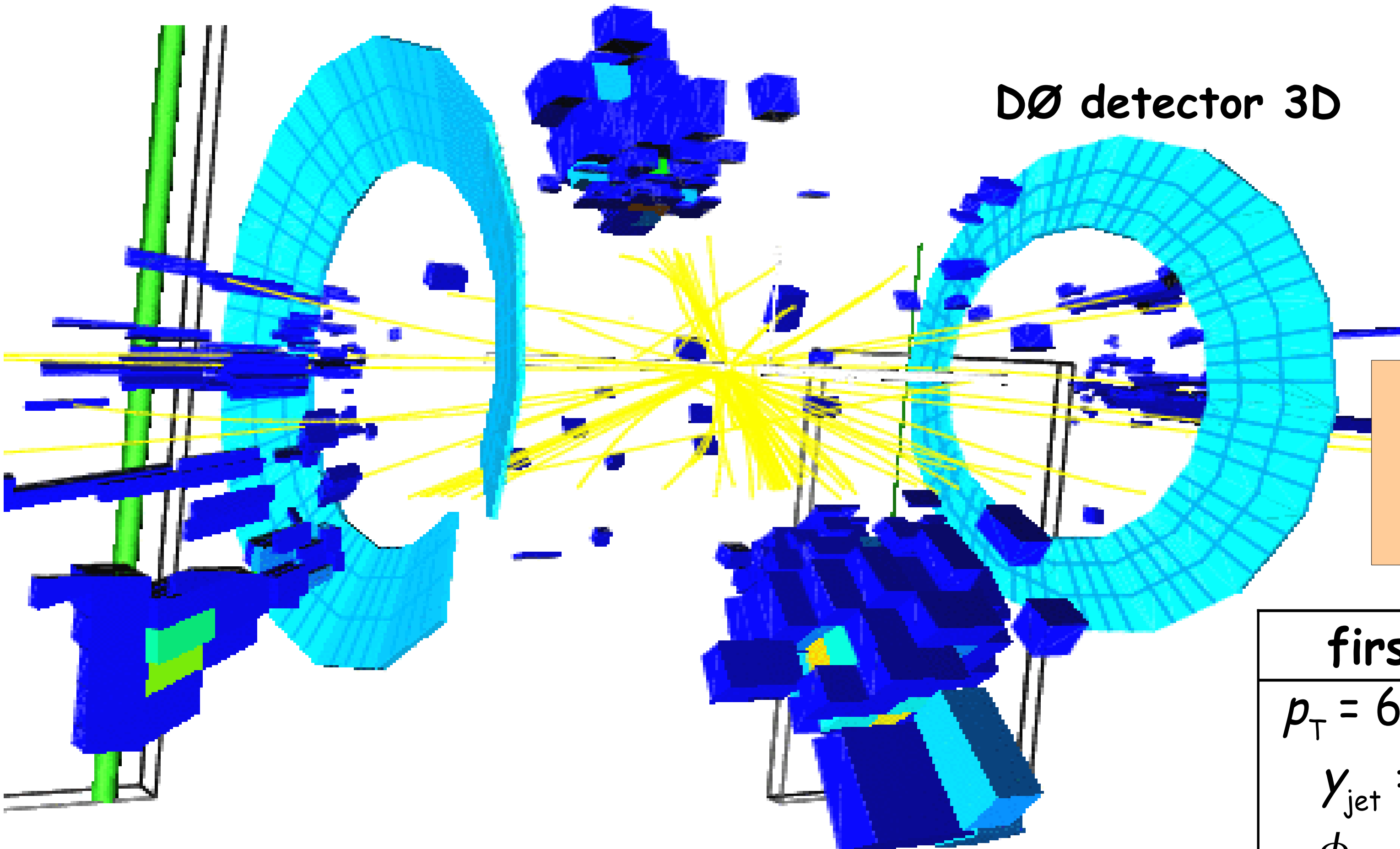
- Inclusive jet cross section spans an impressive eight orders of magnitude
- ...and reaches the **highest energies** ever created in a lab
- Full measurement is a collage of seven different triggers
- Pure data measurement, no theoretical models involved
- Open to new physics at high energy

R_{cone} = 0.7



Highest p_T jet

- 3D view of the event that contains the jet with highest transverse momentum in Run IIa
- Jets are reconstructed offline with iterative cone algorithm (R=0.7)

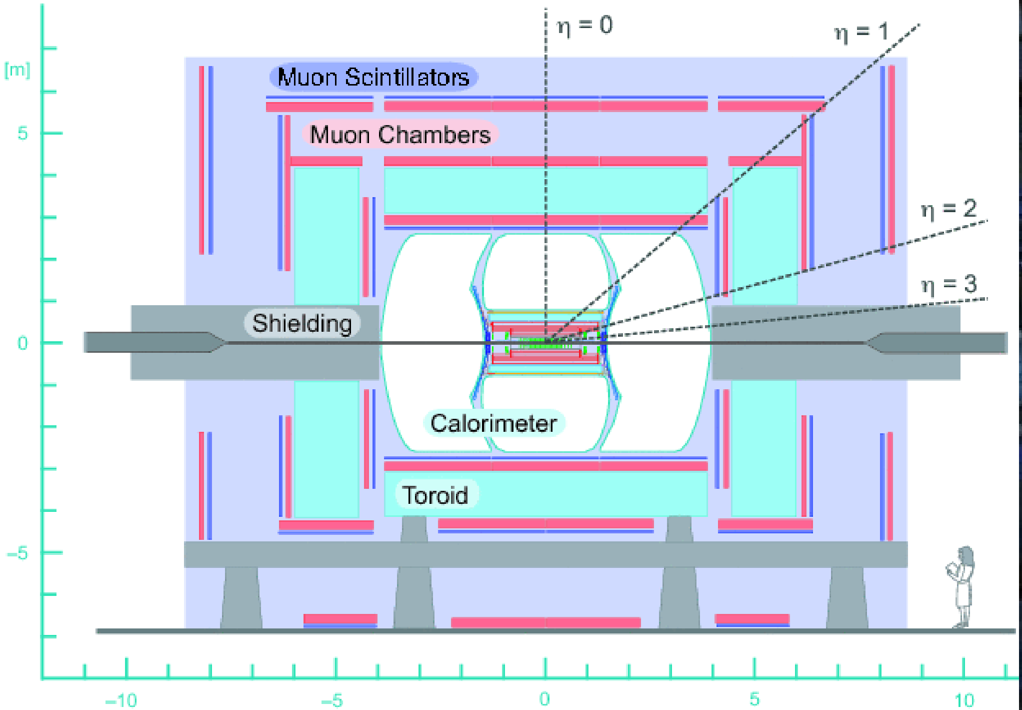


Leading jet has 64% of momentum available in beam proton

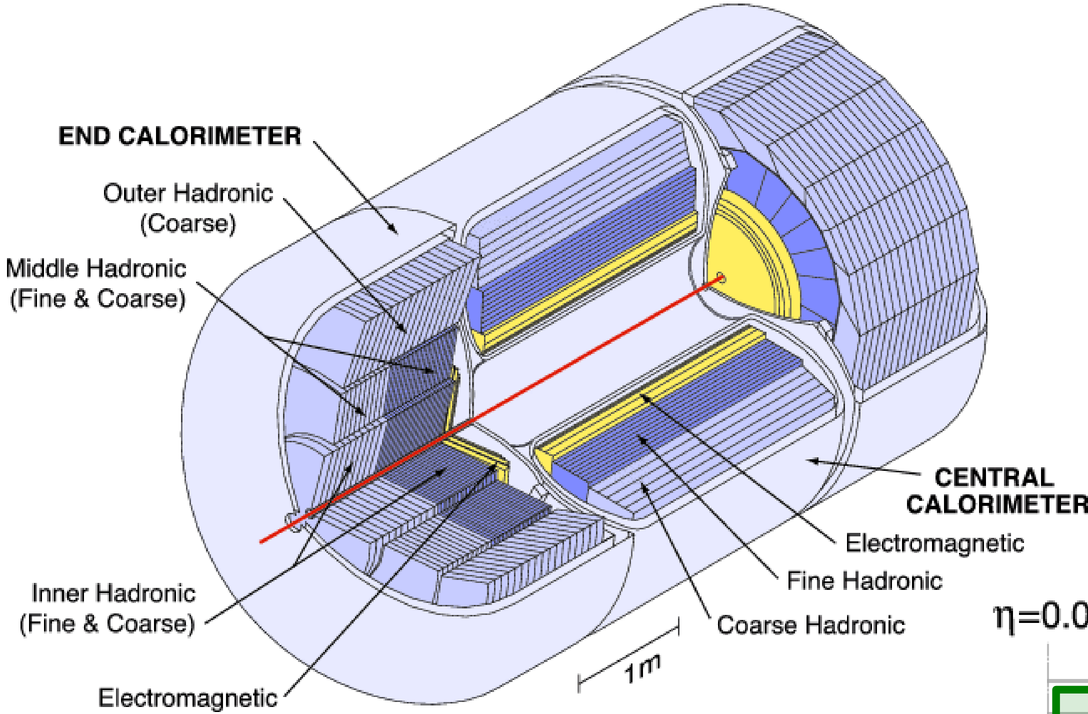
first jet	second jet
p_T = 624 GeV	p_T = 594 GeV
y_{jet} = 0.14	y_{jet} = -0.17
ϕ_{jet} = 2.10	ϕ_{jet} = 5.27
M_{jj} = 1.22 TeV	

D0 experiment

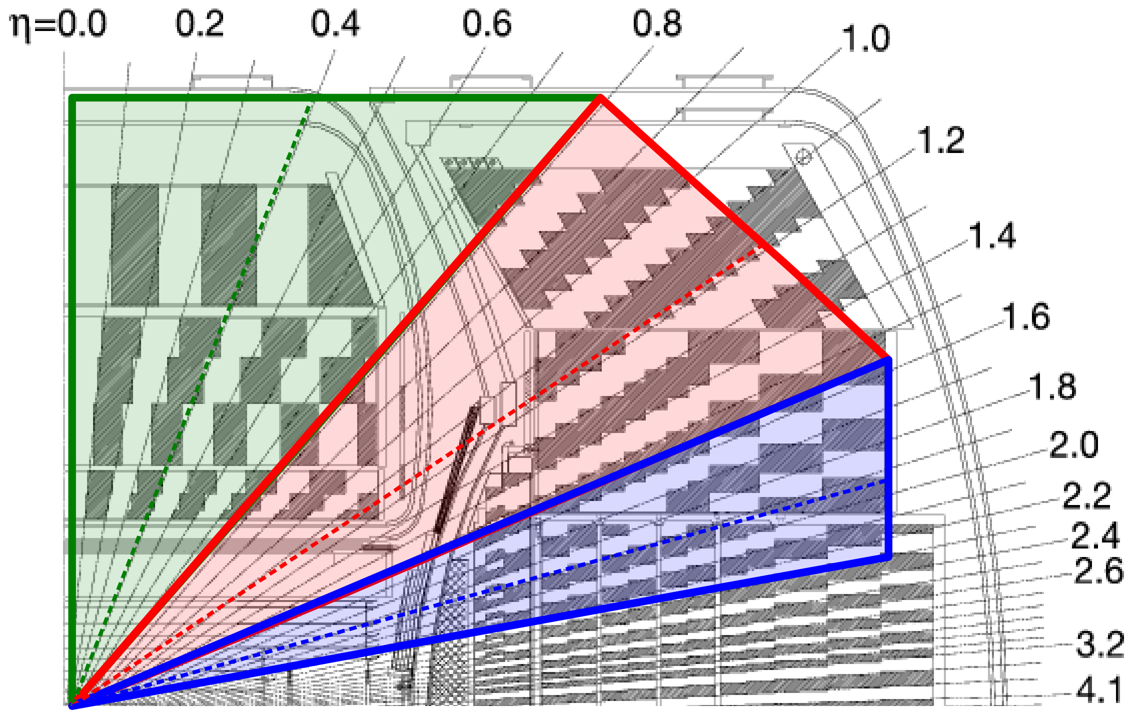
- Three main systems:
  - Tracker (silicon and scintillating fibre)
  - Calorimeter (lAr/U, some scintillator)
  - Muon chambers and scintillators
- First two used in this measurement



D0 calorimeter



- Calorimeter is the most important detector for jet measurements
- Liquid-Argon/Uranium calorimeter:
  - Stable response, good resolution
  - Partially compensating ($e/\pi \sim 1$)
- Gaps covered with scintillator tiles



- Calorimeter structure divides the measurement in three regions:
  - **Central calorimeter** (easiest)
  - **Intercryostat region** (challenging)
  - **End caps** (fine segmentation)

Counting experiment

This is basically a counting experiment:

Double differential cross section

Number of jets

Jet Energy Scale!

$$\frac{d^2\sigma}{dp_T\,dy} = \frac{N}{\epsilon \cdot L \cdot \Delta p_T \Delta y} \cdot C_{smear} \quad \text{versus } p_T$$

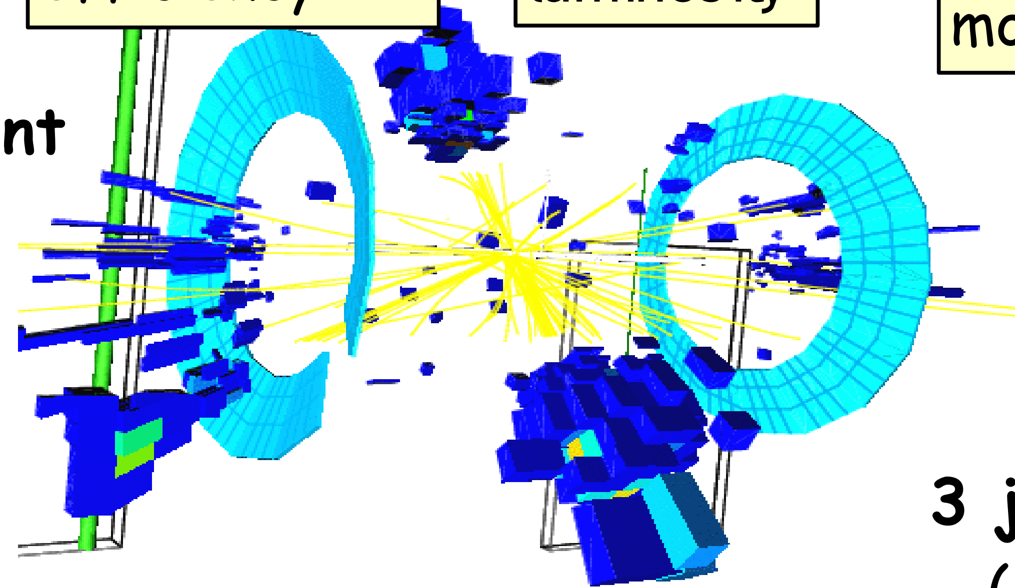
Jet and event efficiency

Integrated luminosity

Bins of transverse momentum and rapidity

Events can move in and out of p_T bins due to **calorimeter energy resolution**

1 event



3 jets

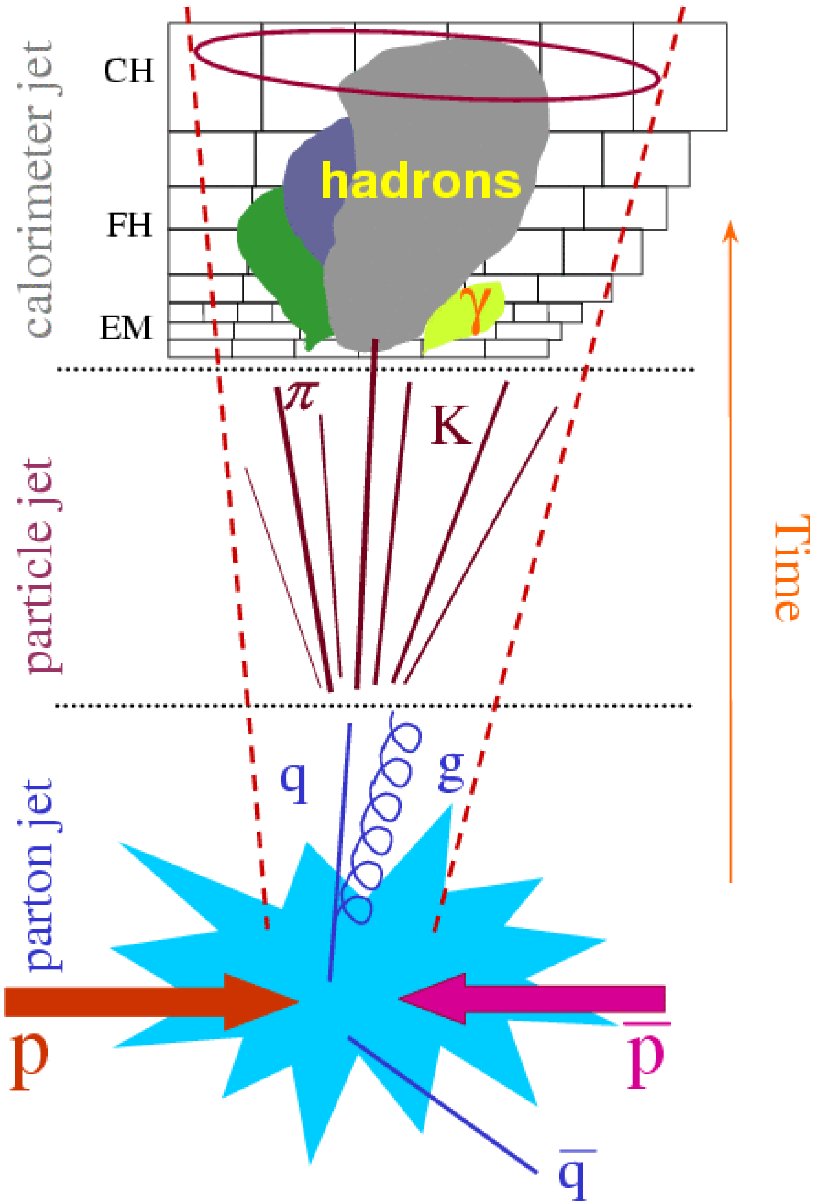
(in different bins)

Jet energy scale

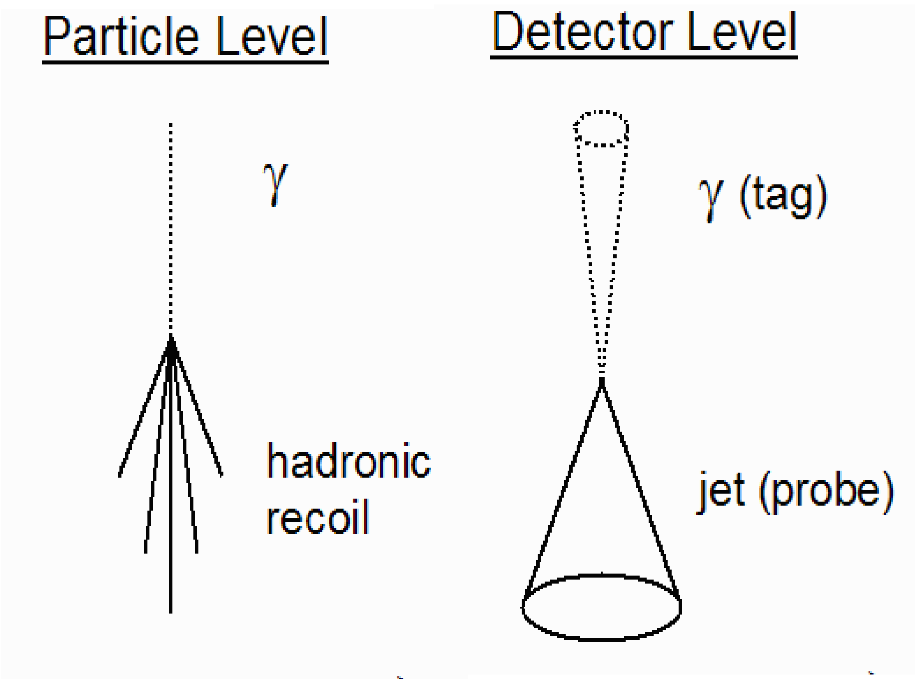
- Jet Energy Scale returns the measured calorimeter jet energy to the **particle level**

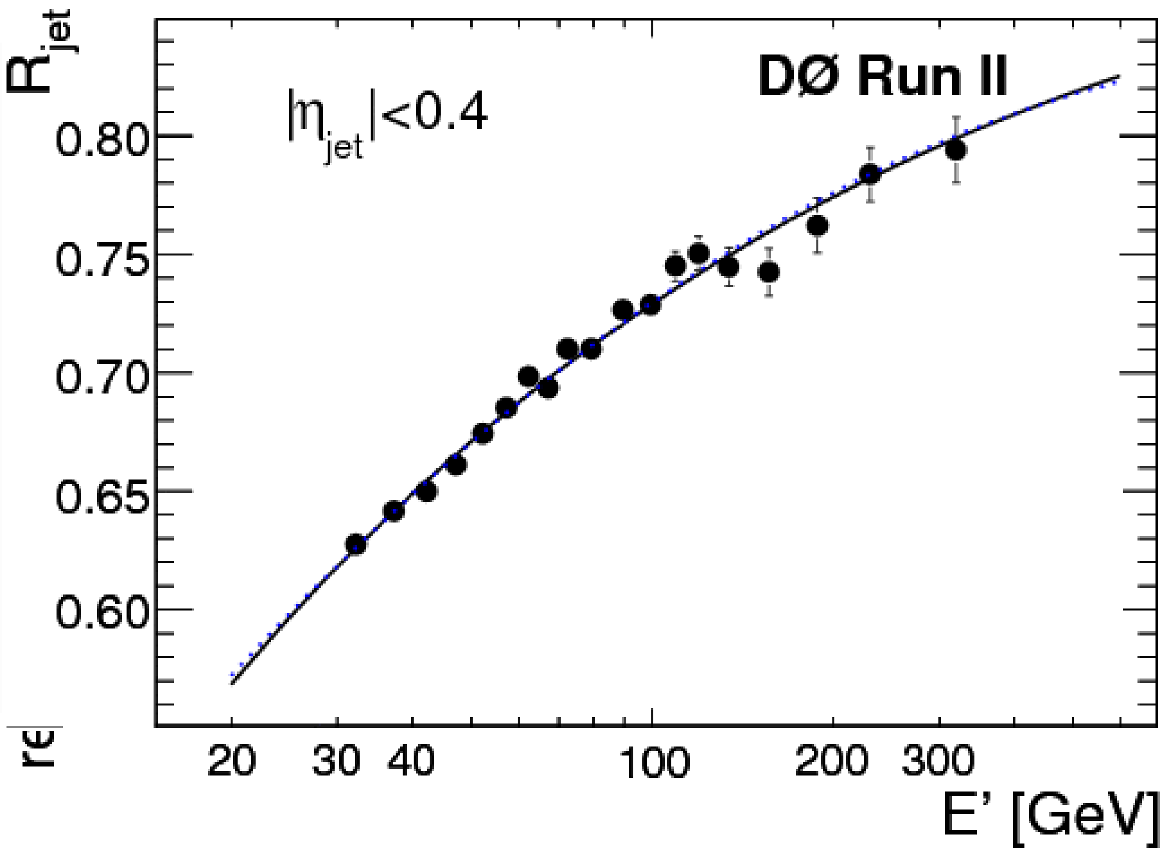
$$E_{ptcl} = \frac{E_{cal} - Offset}{(F_{\eta} \cdot R) \cdot S} \cdot k_{bias}$$

- Offset is energy not associated to the hard scatter: noise, pile-up, **multiple interactions**
- Response is the fraction of particle jet **energy deposited** in the calorimeter by the particles
- Detector showering accounts for **energy flow** in and out of the calorimeter jet due to detector effects (finite calorimeter tower and hadron shower size, magnetic field)
- Method biases corrected using tuned MC



JES: Response calibration

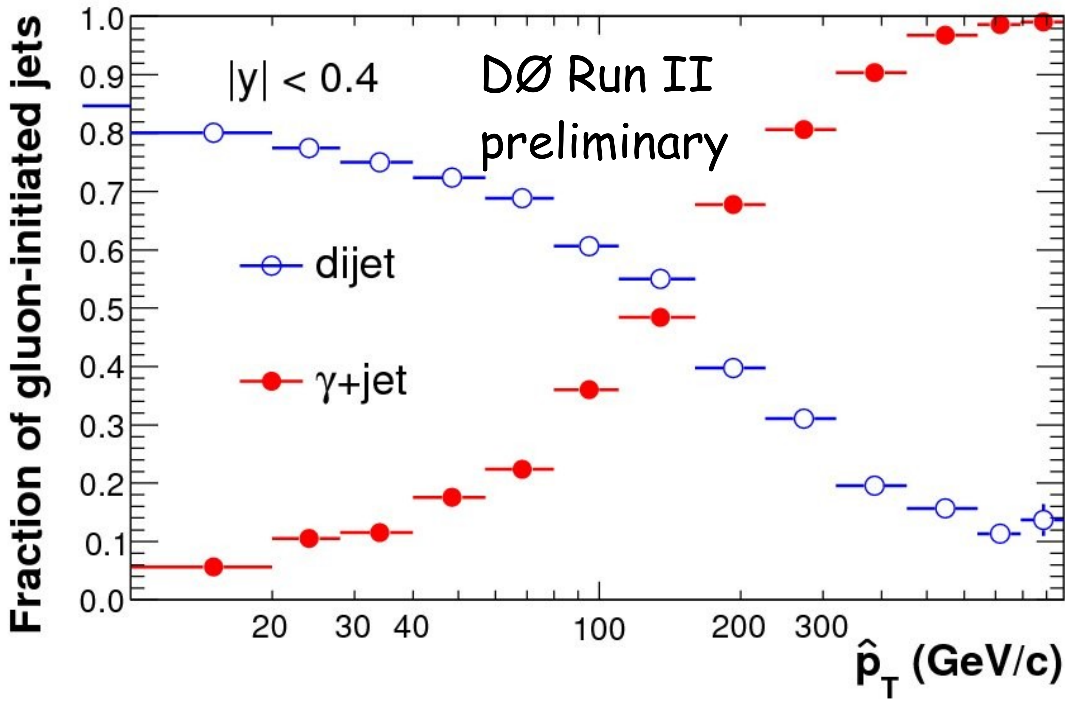


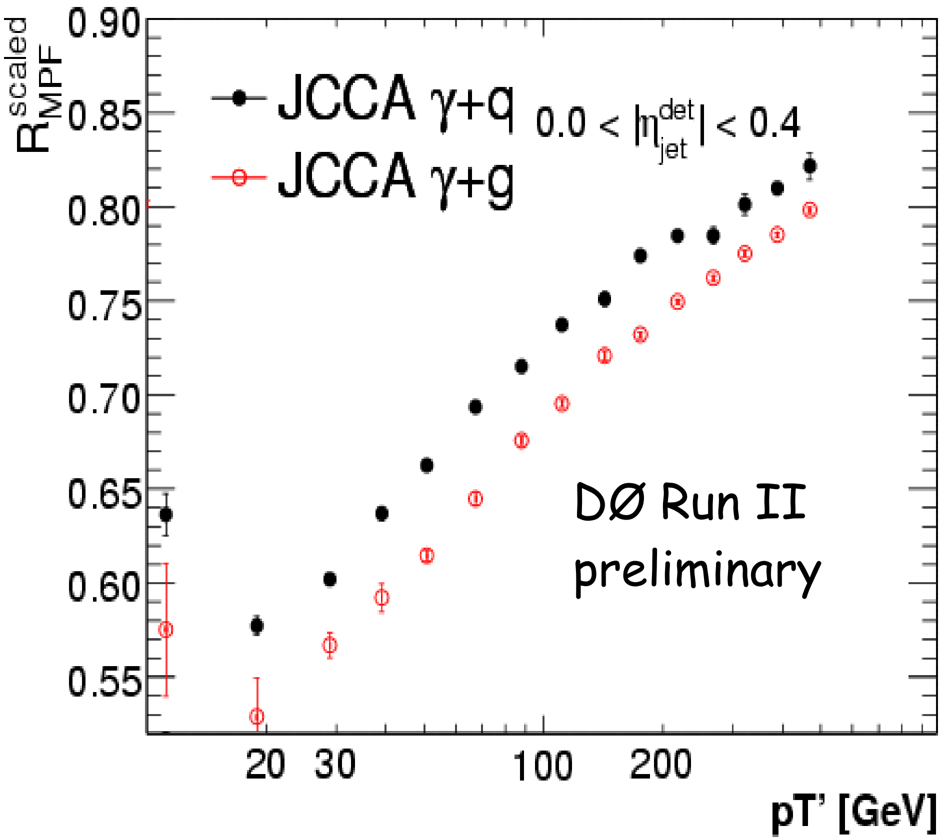


- Response calibration based on transverse momentum conservation
- Photon/central jet and recoil balanced in p_T at parton/particle level
- Calibration through missing-E_T insensitive to the jet cone and showering effects

JES: Sample dependence

- Jets in γ+jet (absolute JES) and dijet (analysis) samples have very different quark/gluon composition
- Gluons fragment to softer particles than quarks ⇒ lower response

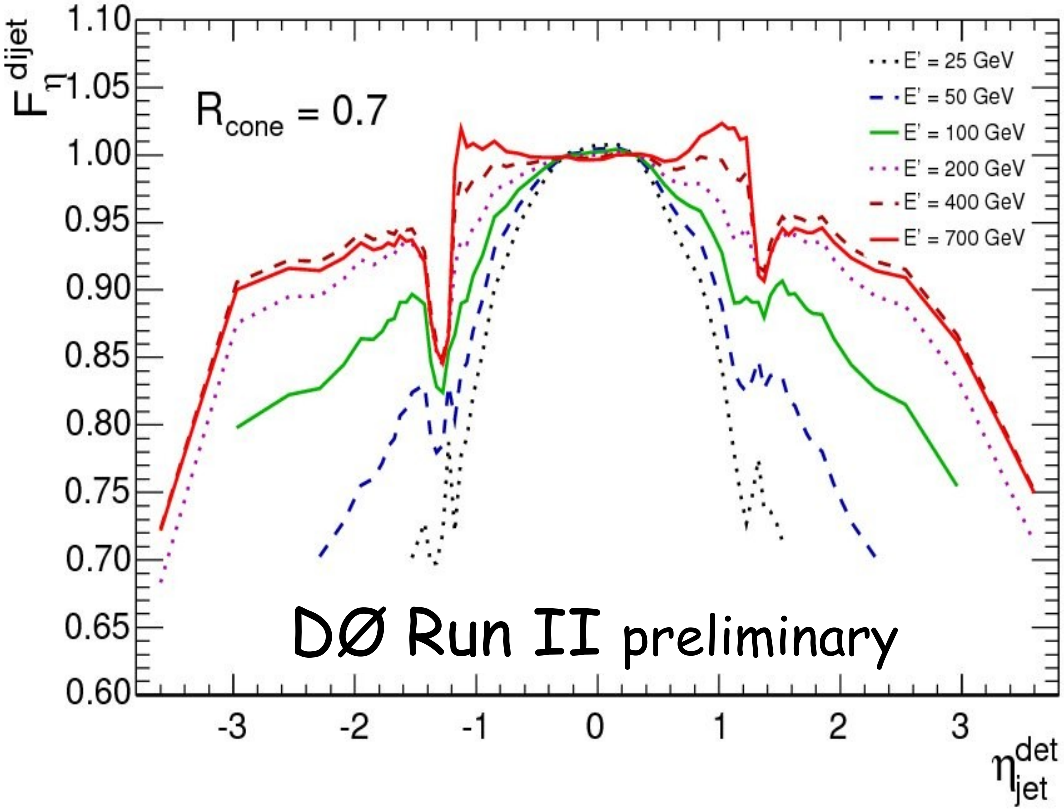


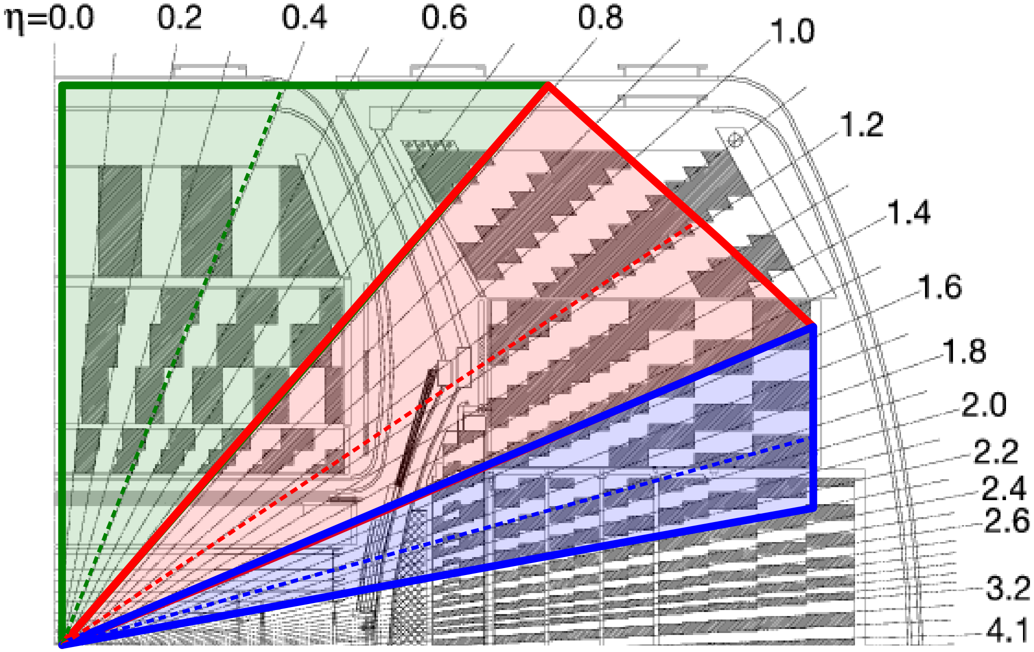


- Knowledge of single pion response is essential to predict the response differences correctly
- Monte Carlo single pion response tuned using γ+jet data

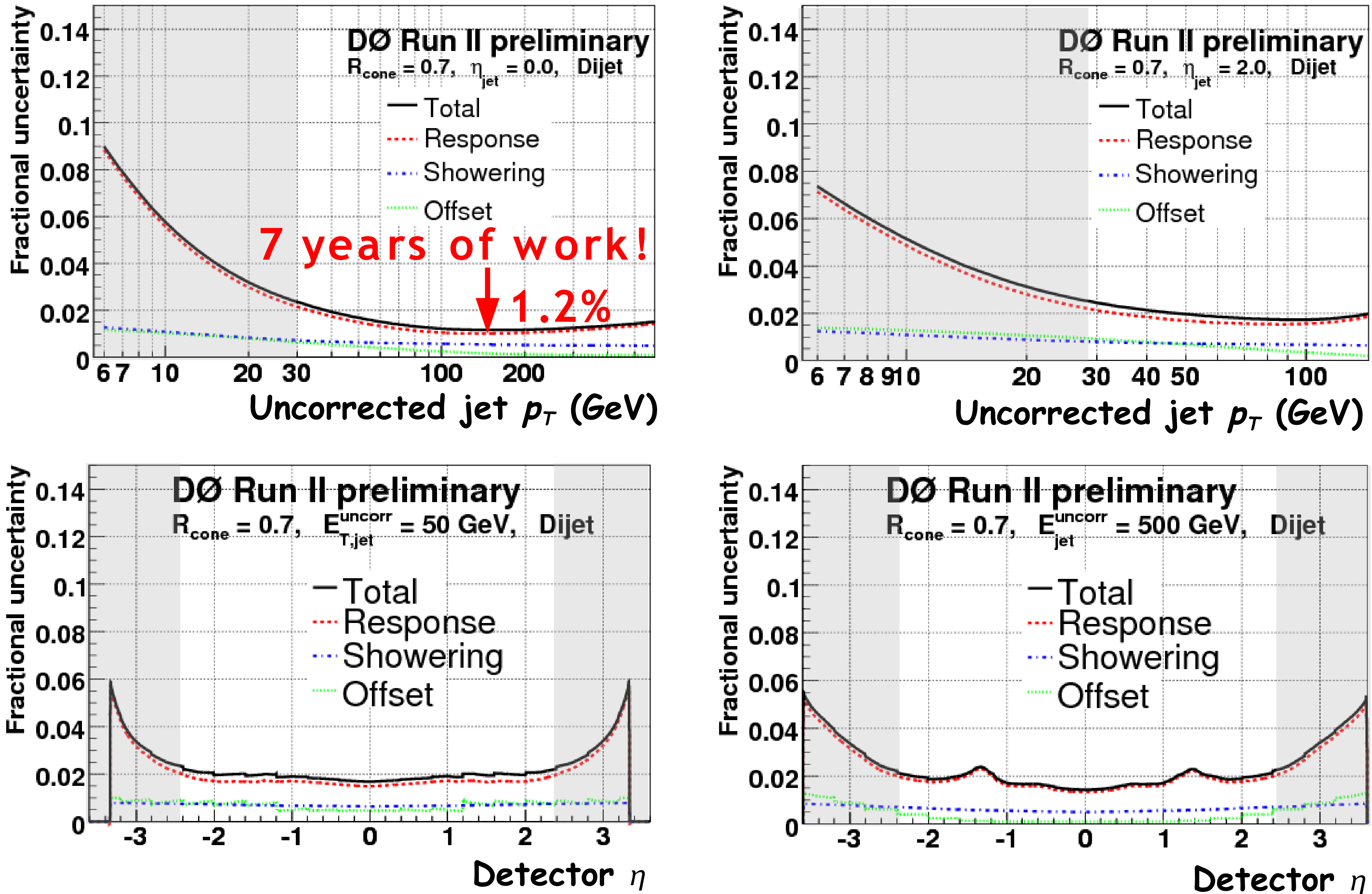
JES: η-dependence

- Response depends on calo region (central, intercryostat, end cap)
- Probe jet balanced against a photon or a central jet
- Simultaneous fit to **dijet** and γ+jet samples taking into account sample differences ⇒ relative JES for analysis **directly from data**



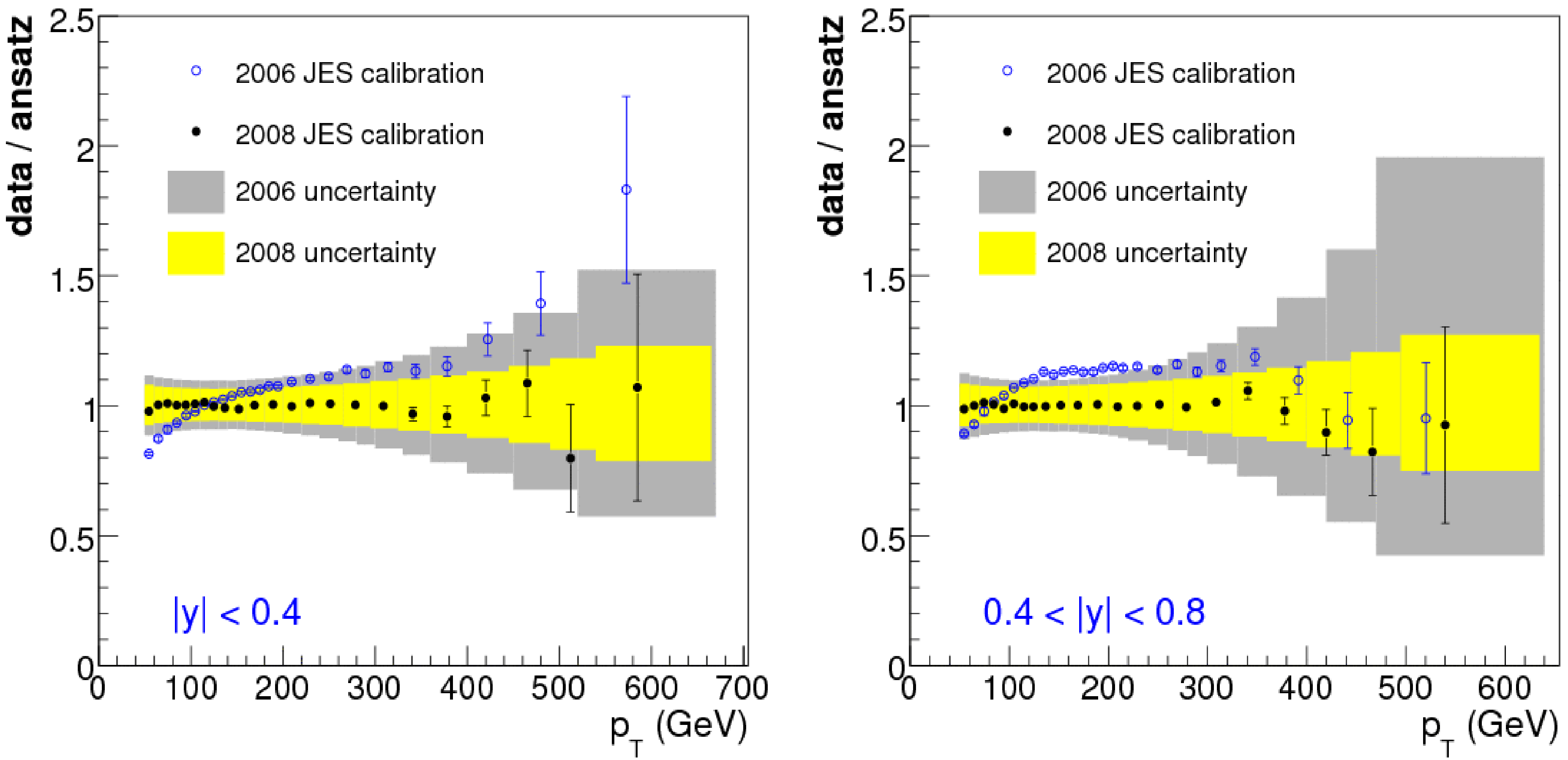


Dijet JES uncertainty



Improvement since 2006

- Uncertainties have improved by up to factor two and more in the central region since preliminary JES (2006)
- Forward regions not published before, but improvement over factor 10

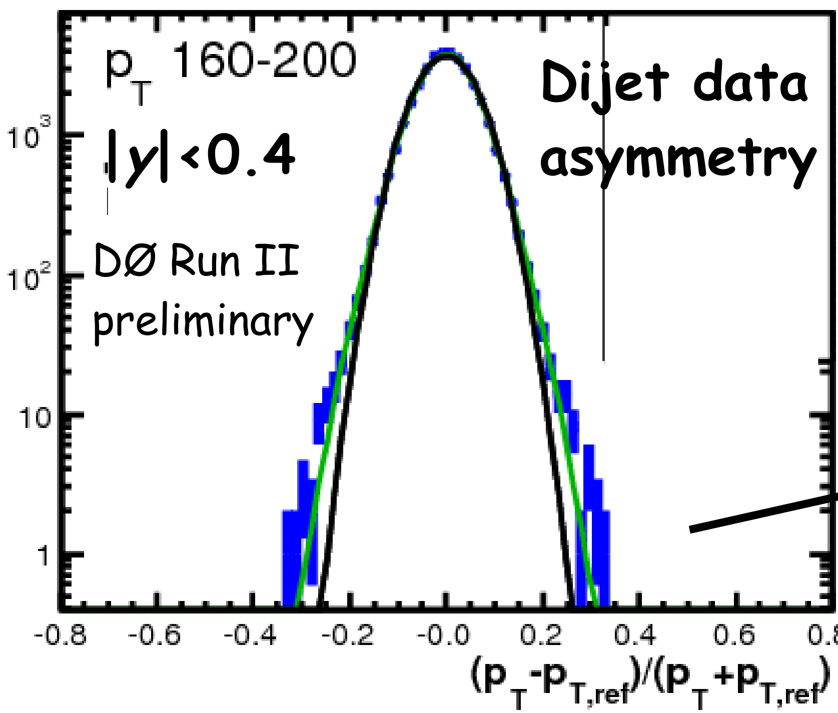


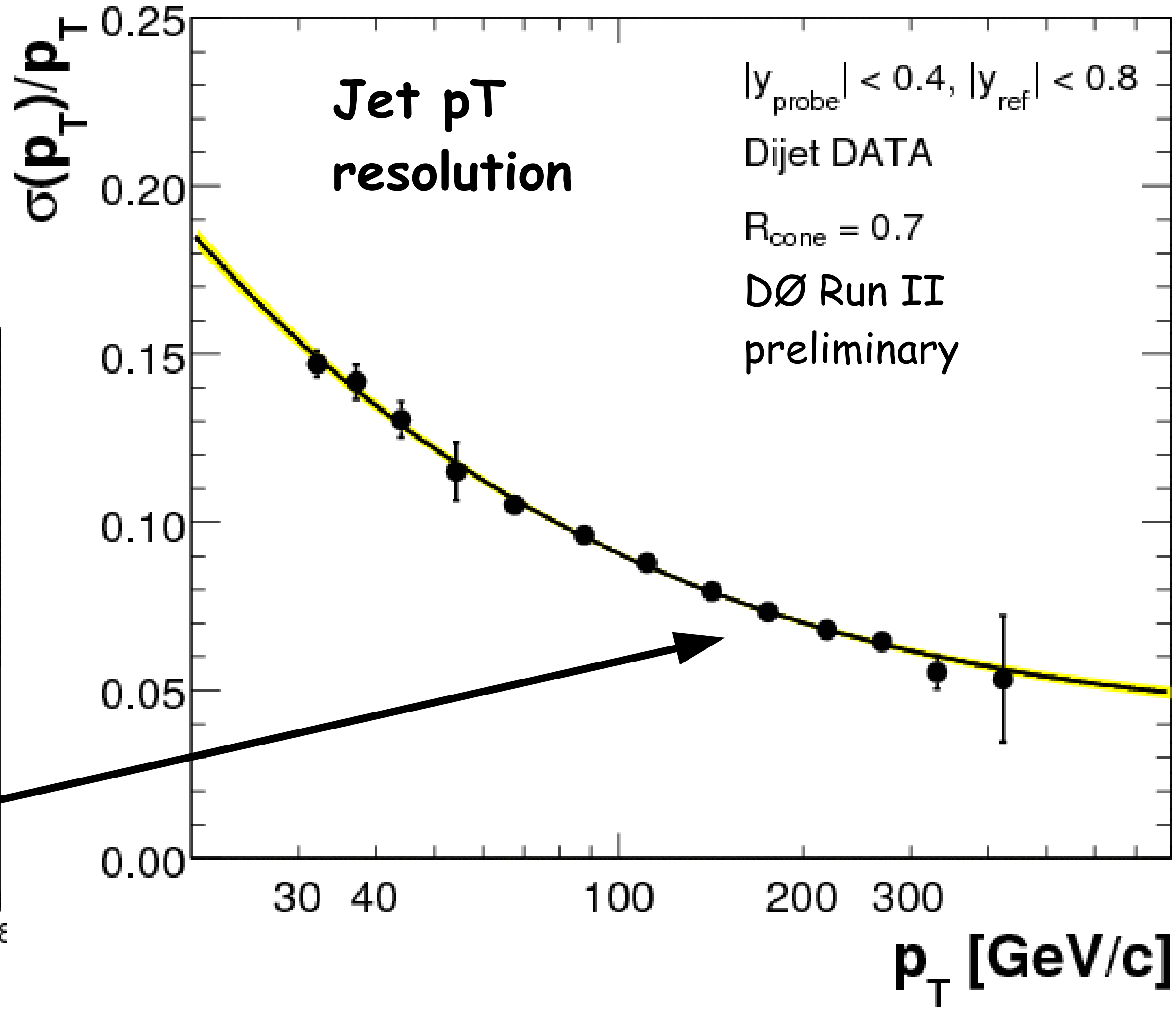
Jet p_T Resolution

- Jet p_T resolution was measured directly **on data using dijets**
- Basic variable dijet asymmetry A, which is corrected for **soft radiation** (third jet below reconstruction threshold) and **particle level imbalance**

$$A = \frac{p_{T,2} - p_{T,1}}{p_{T,2} + p_{T,1}},$$

$$\frac{\sigma_{p_T}}{p_T} = \sqrt{2} \cdot RMS(A)$$

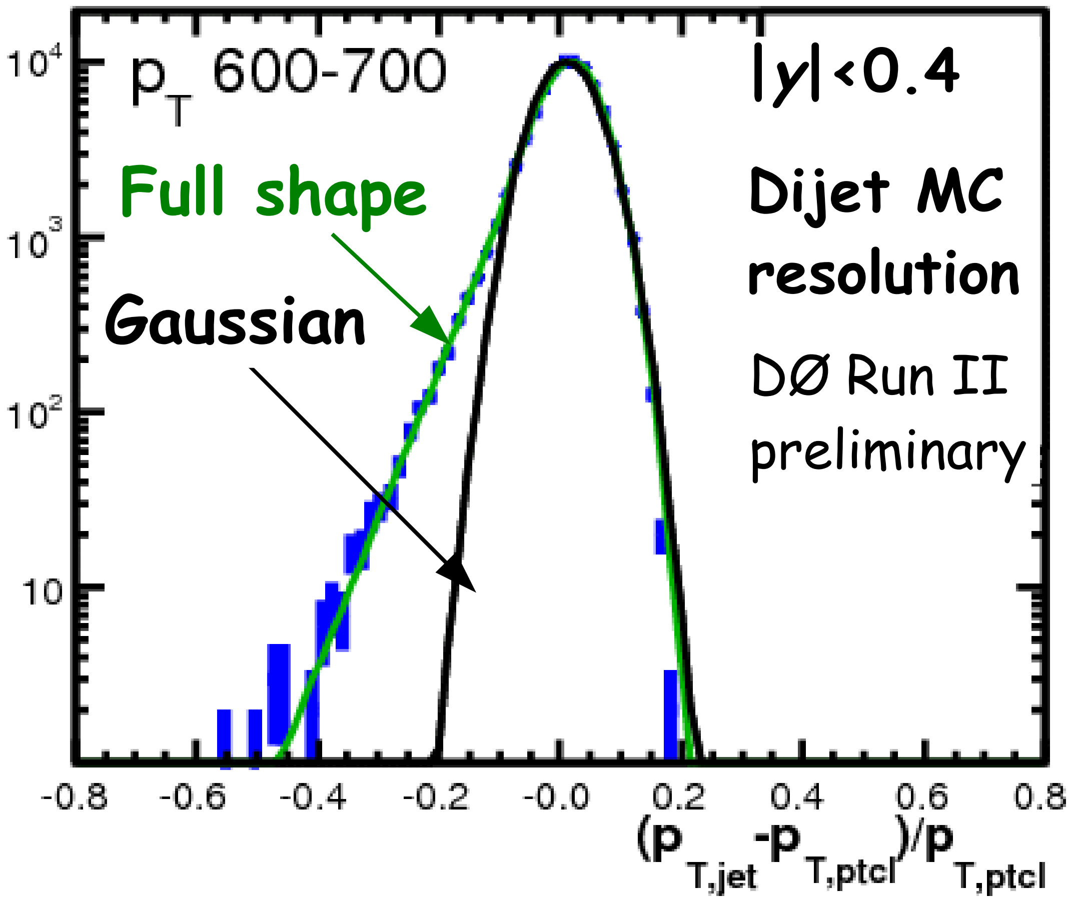


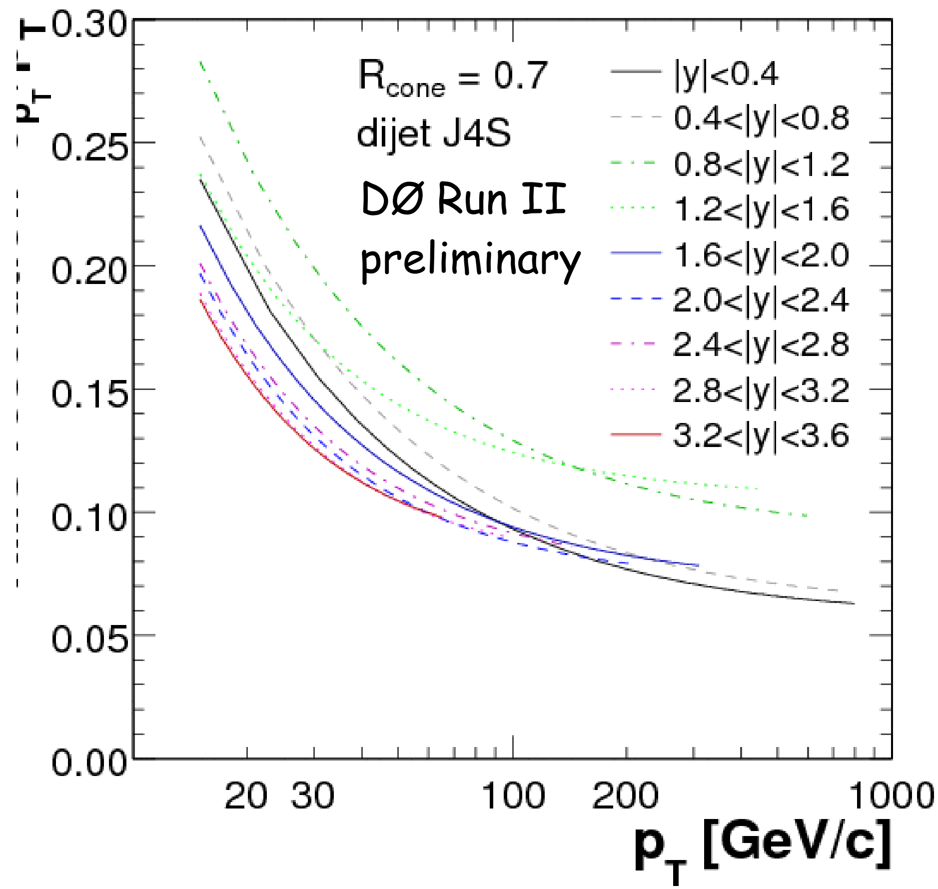


Jet p_T resolution: details

- Parametrized by Noise, Stochastic and Constant terms in every |*y*| bin
- Smearing shape from MC truth: non-Gaussian tails explicitly accounted for
- Especially high p_T punch-through sizable effect

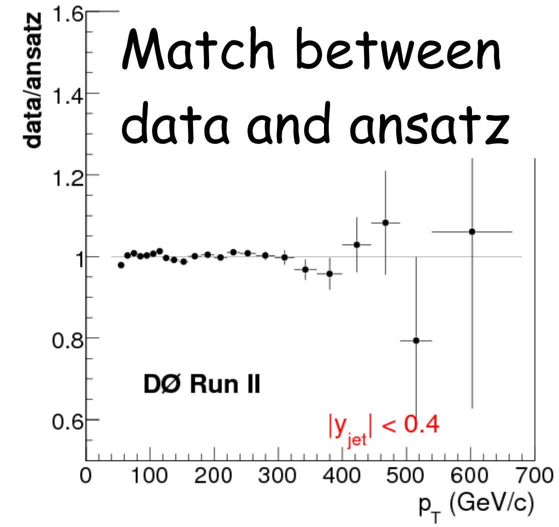
$$\frac{\sigma_{p_T}}{p_T} = \sqrt{\frac{N^2}{p_T^2} + \frac{S^2}{p_T} + C^2}$$

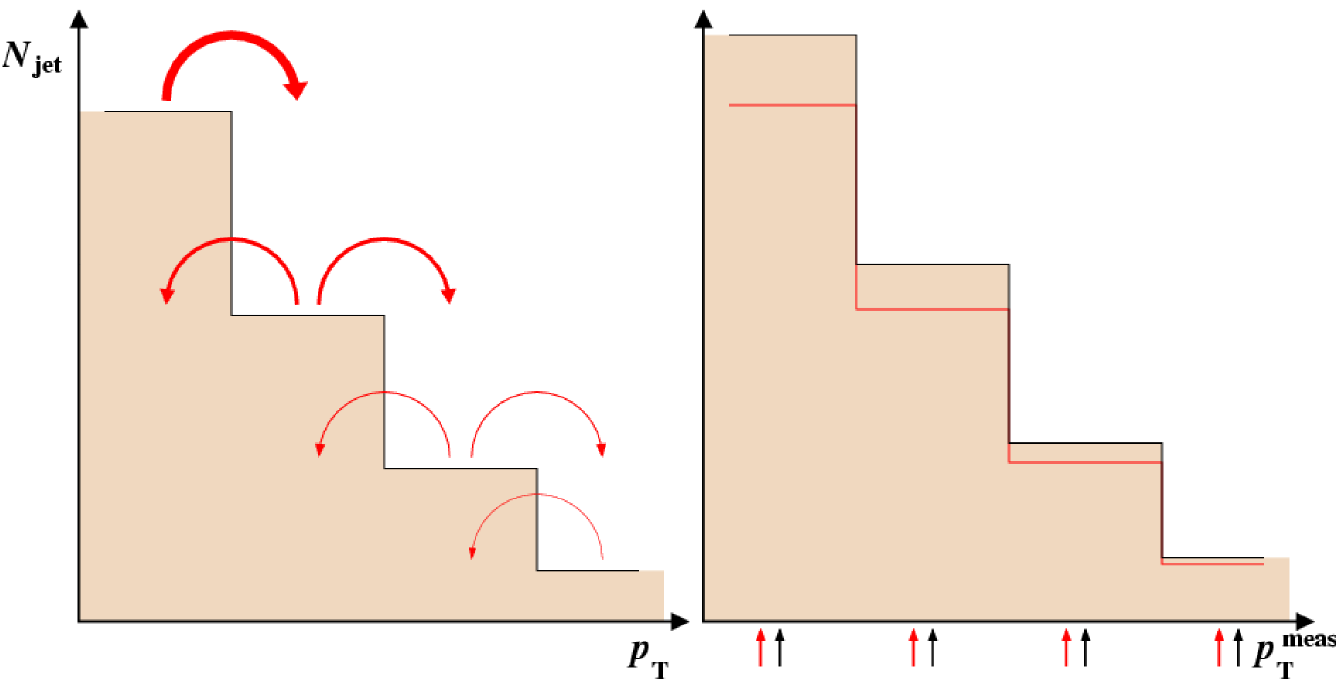


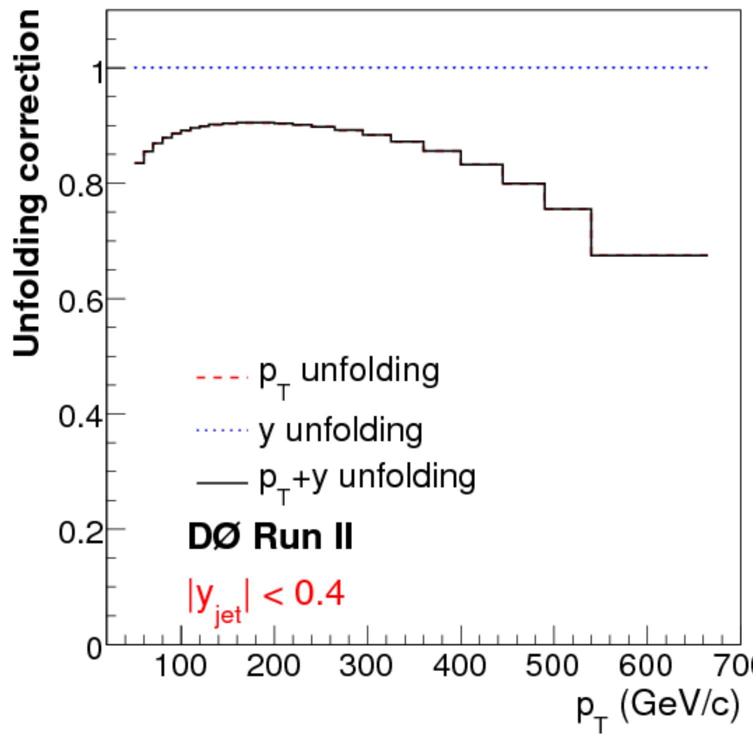


Jet p_T resolution unfolding

- Jets can move in and out of p_T bins due resolution
- We know how jets migrate; guess where they were (**ansatz**) and calculate where they would then go
- Iterate the initial guess a few times until match with data; take ratio of before and after



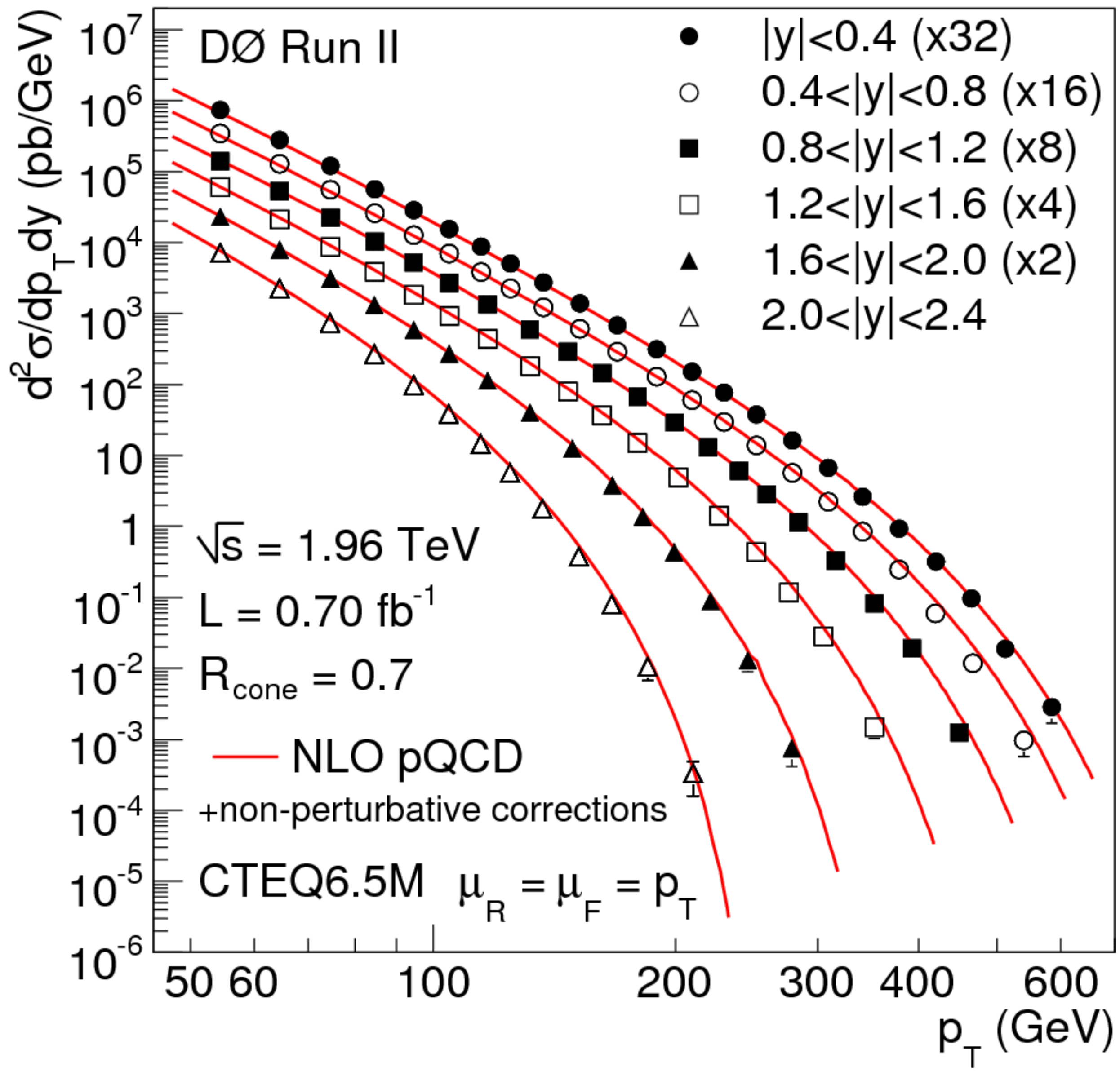




Final results

- Largest data set from Run II with the widest rapidity coverage (|y|<2.4) and smallest uncertainties to date
- D0 and CDF Run II data have superseded Run I data in MSTW08 PDFs
- Jet spectrum presented at **particle level** with **midpoint cone** (R_{cone} = 0.7)
- Compared to next-to-leading order (NLO) theory with CTEQ6.5M PDFs and non-perturbative corrections from Pythia

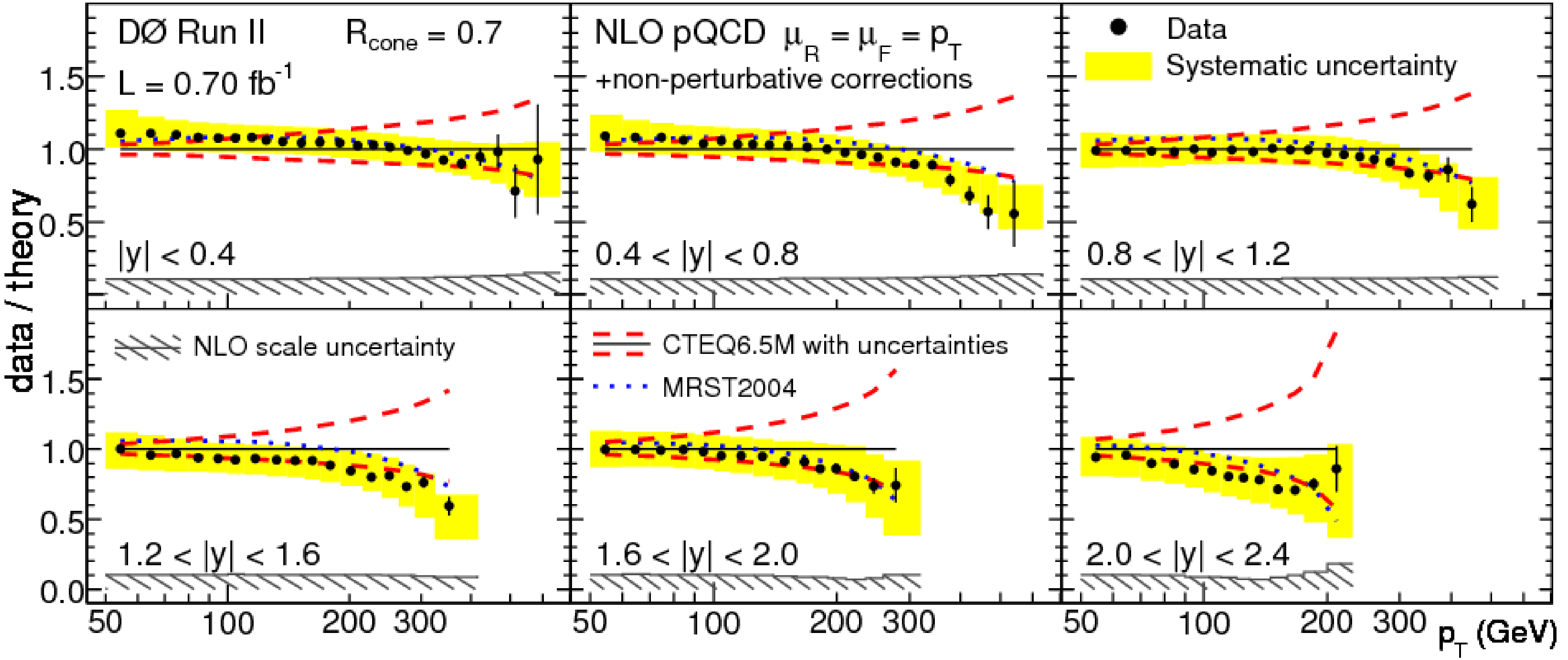
PRL 101, 062001 (2008)



Final results

- Good agreement between data and theory at all rapidities; MRST2004 PDFs and low end of CTEQ6.5 PDF uncertainty favored (soft gluon)
- Scale uncertainty in next-to-leading order (NLO) theory comparable to experimental uncertainty at low p_T

PRL 101, 062001 (2008)



MSTW08: DØ and CDF comparison

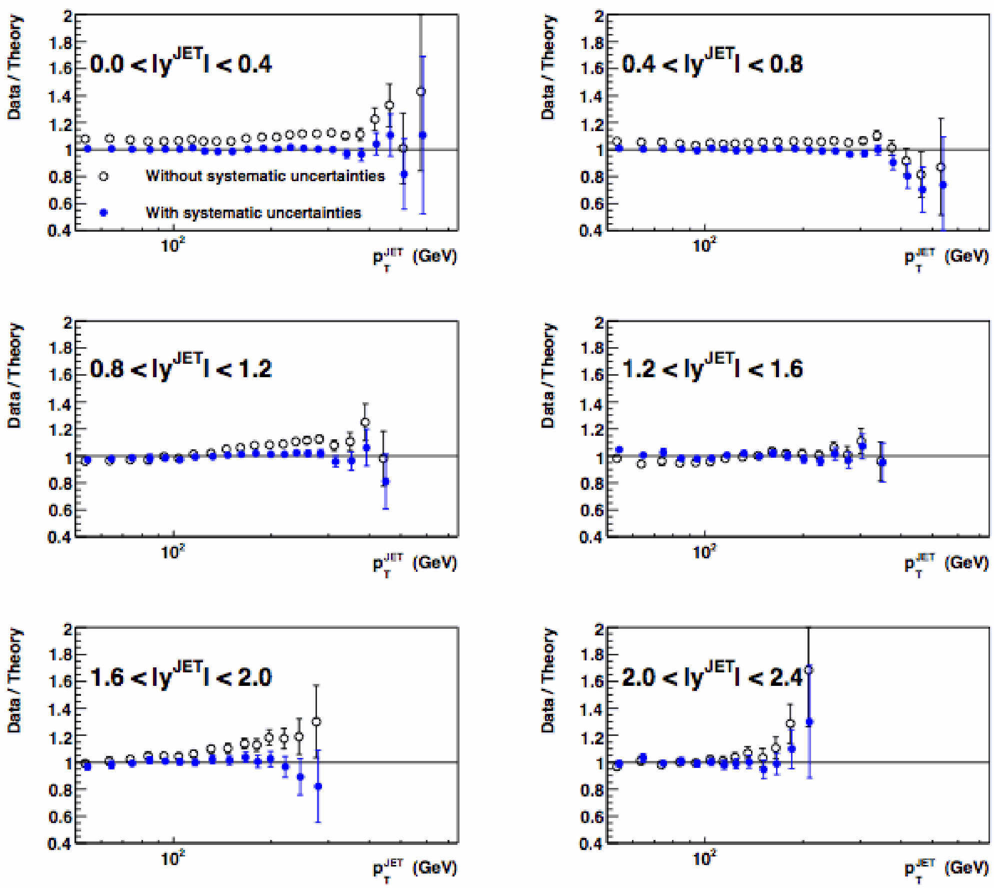
- DØ and CDF data are compatible with the latest MSTW08 PDF central fit; improved agreement with extrapolation from new HERA data
- Both favor softer (and more natural) high-x gluon PDF than Run I data

  ⇒ useful for physics predictions at the LHC

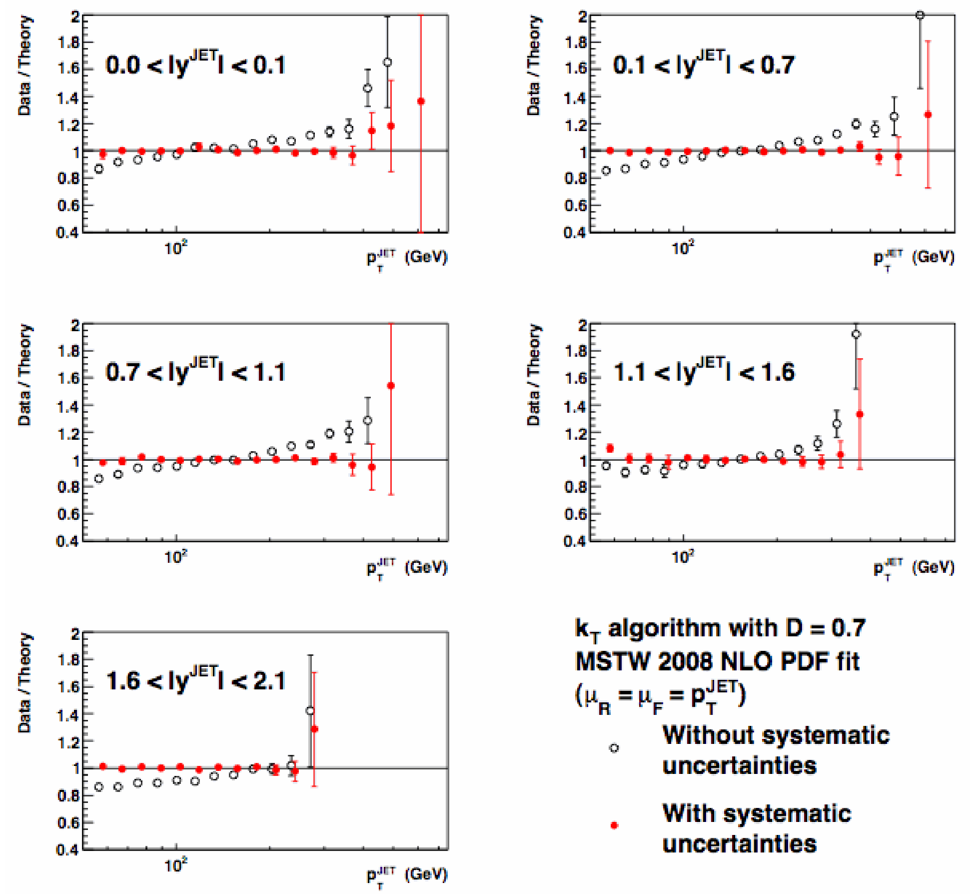
arXiv:0901.0002

DØ Run II inclusive jet data (cone, R = 0.7)

MSTW 2008 NLO PDF fit ($\mu_R = \mu_F = p_T^{JET}$), χ^2 = 114 for 110 pts.



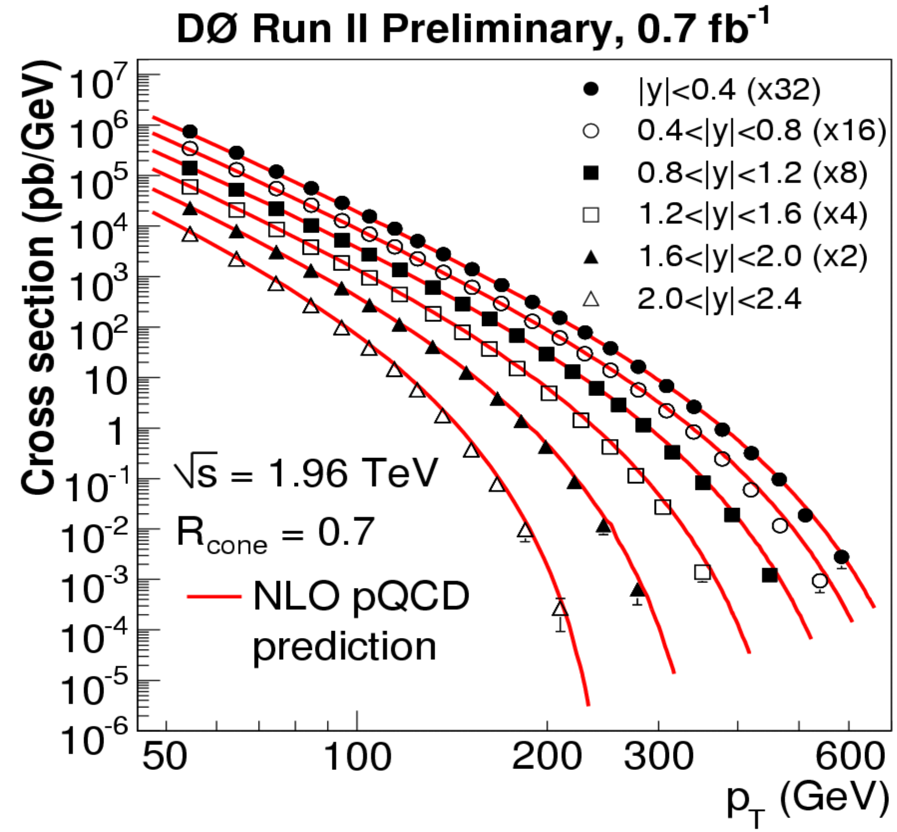
CDF Run II inclusive jet data, χ^2 = 56 for 76 pts.



Acknowledgements

I would like to thank:

- My advisors Christophe Royn (Saclay) and Greg Snow (UNL)
- My supervisor Jorma Tuominiemi (HIP)
- My funding agencies:
  - Graduate School for Particle and Nuclear Physics in Finland
  - Finnish Cultural Foundation
  - Magnus Ehrnrooth Foundation

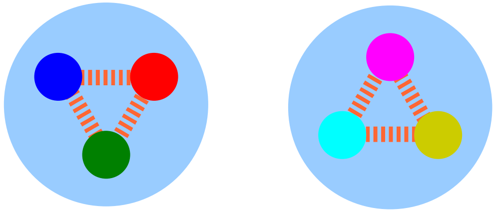


Fermilab Today result of the week on January 24th: www.fnal.gov/today (08-01-24)

Alexander Kupco, Inst. of Physics, Czech Republic
Christophe Royon, DAPNIA/SPP, Saclay, France
Mikko Voutilainen, UNL, Nebraska, HIP, Finland

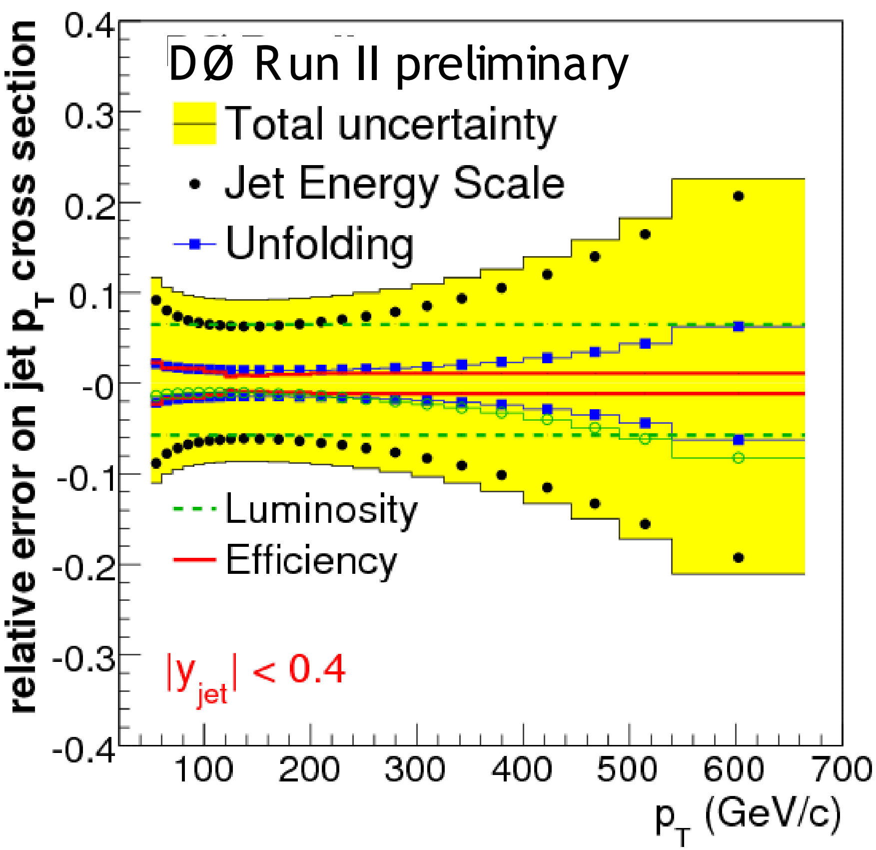
P.S. Tuula sends her greetings to everyone!

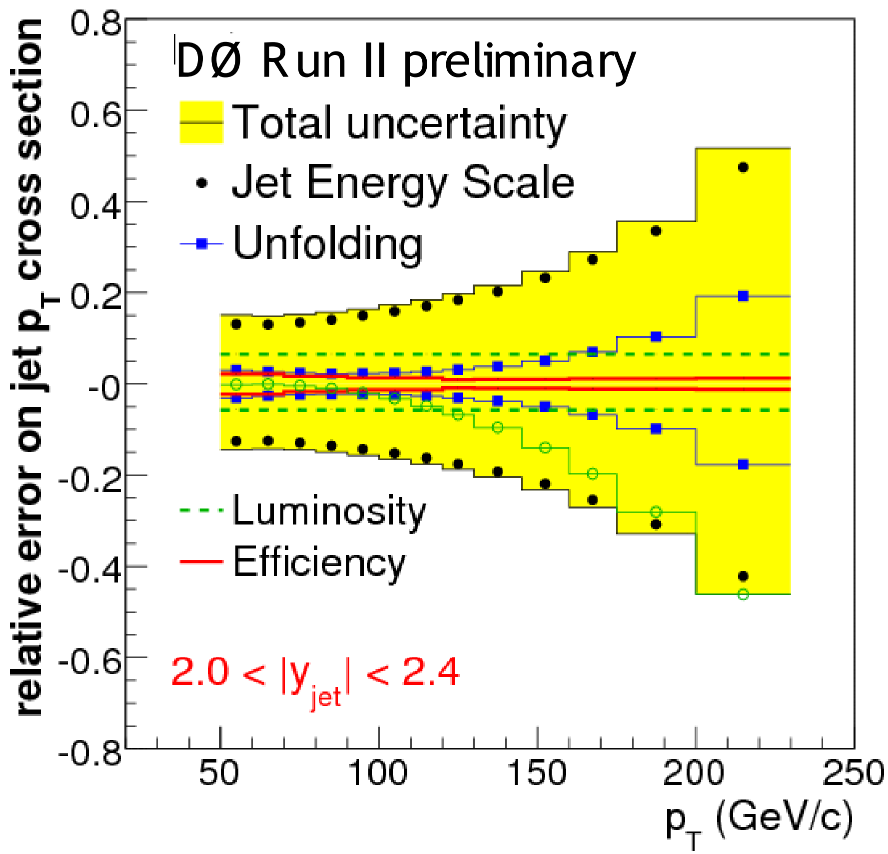
Back Up Slides



Components of uncertainty

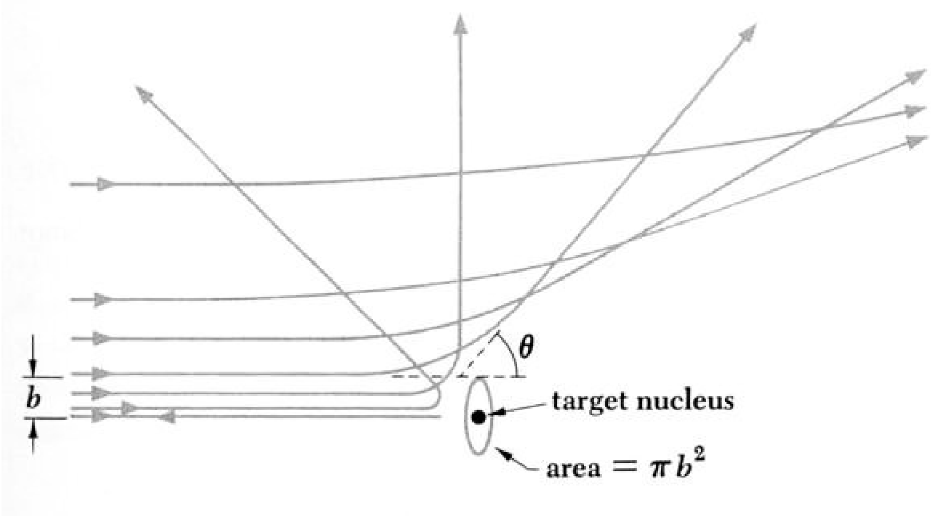
- Total uncertainty is dominated by the (much improved!) JES
- Unfolding (≈p_T resolution) uncertainty much smaller than JES
- Luminosity is a significant uncertainty at low p_T in CC
- Efficiency uncertainty negligible

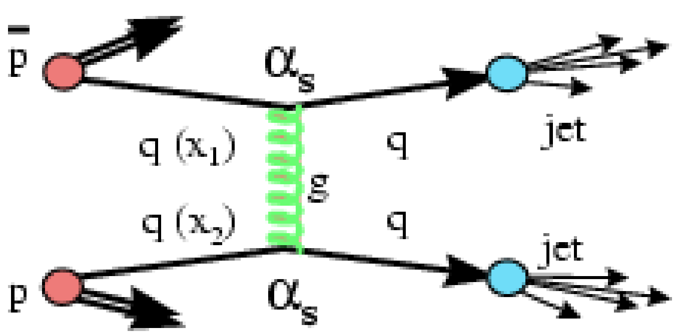


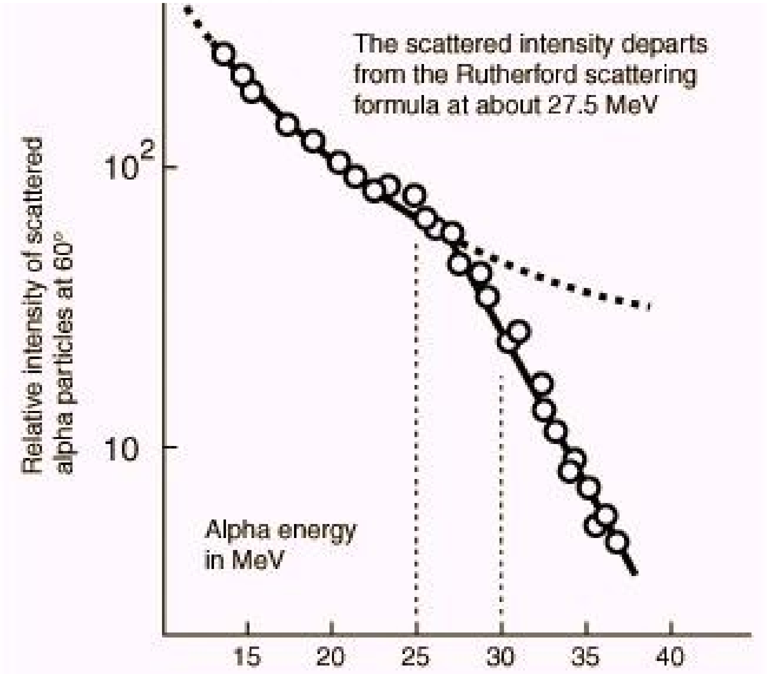


A Short History Of Scattering

- Rutherford shoots alpha particles to gold foil and finds atomic nucleus
- Higher energy probes start to break heavier nuclei
- Very high energy probes find parton substructure of the nuclei (quarks, gluons, virtual particles)
- Modern day deep inelastic scattering experiments probe the parton structure of the proton in detail and look for the unexpected

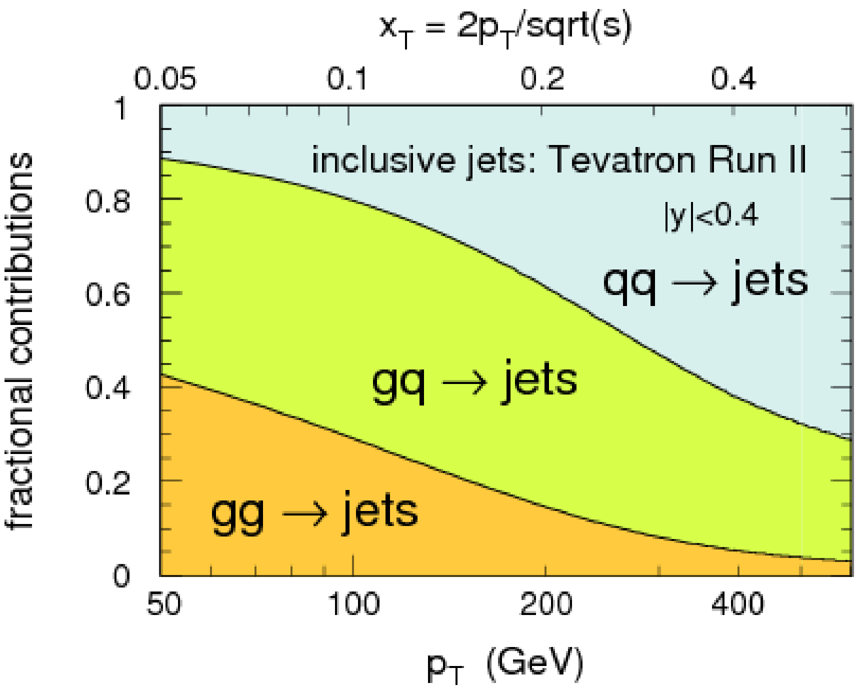




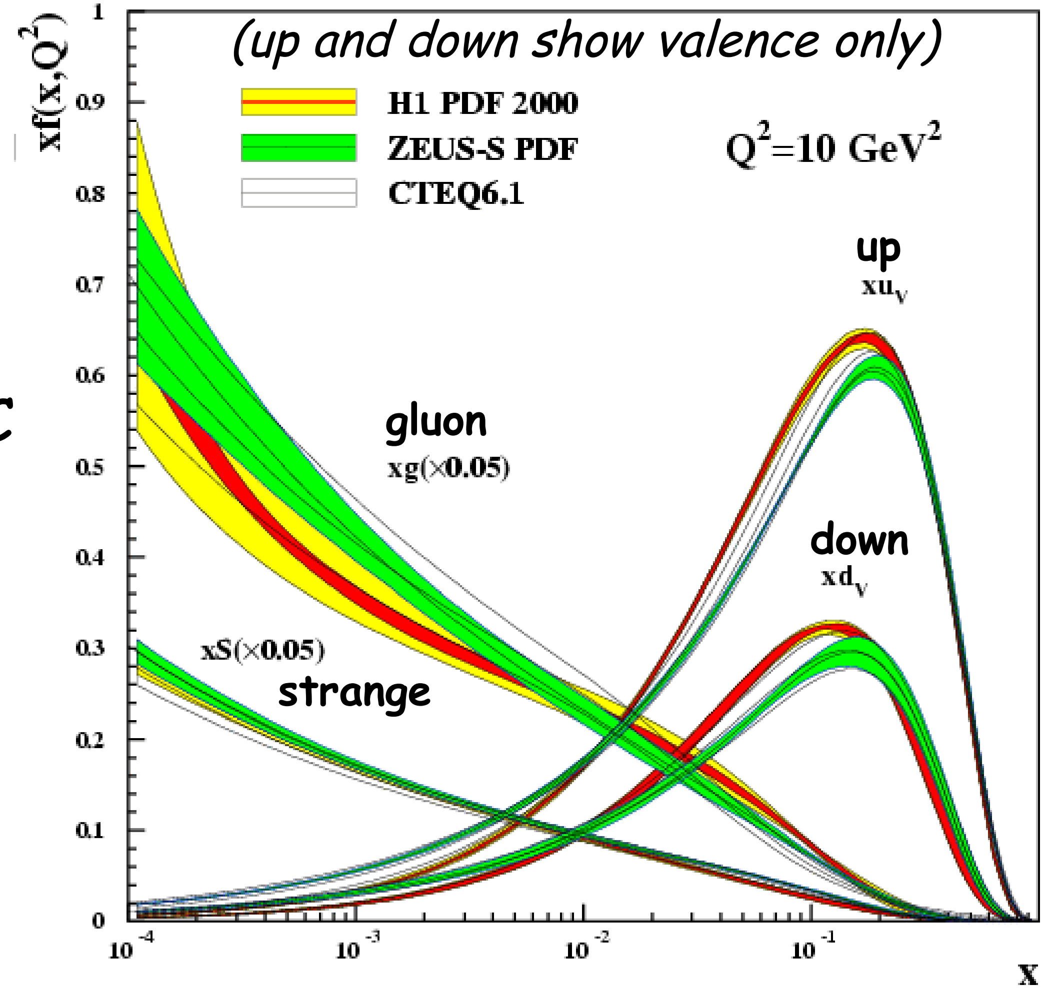


Parton distributions

- Inclusive jet cross section can constrain parton distribution functions (PDFs), especially the **gluon PDF at high *x***
- PDFs are needed *e.g.* to reliably calculate backgrounds at the LHC



Proton parton distribution functions

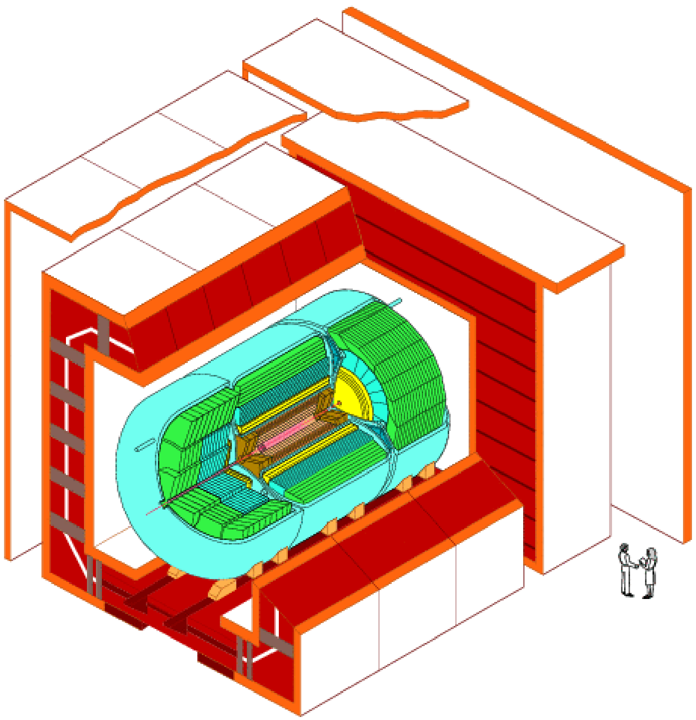


x: fraction of proton momentum carried by individual parton

f(x,Q²): probability of finding parton with momentum fraction x in interval dx

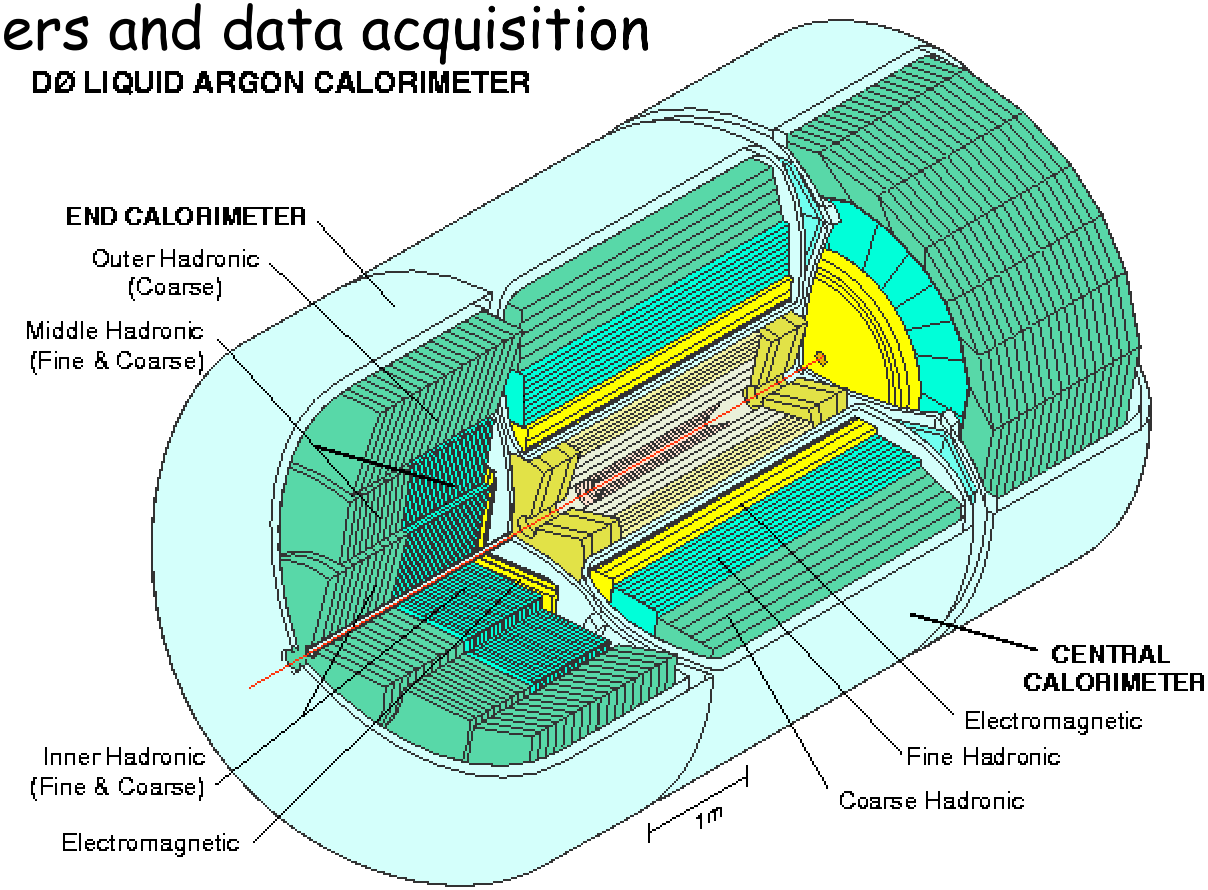
D0 experiment

- Main components: tracker, **electromagnetic calorimeter**, **hadronic calorimeter**, muon detectors
- Upgraded for Run II with new silicon and scintillating fiber trackers, 2 T solenoid magnet (for tracking), preshower detectors, and new electronics, triggers and data acquisition



DØ Detector

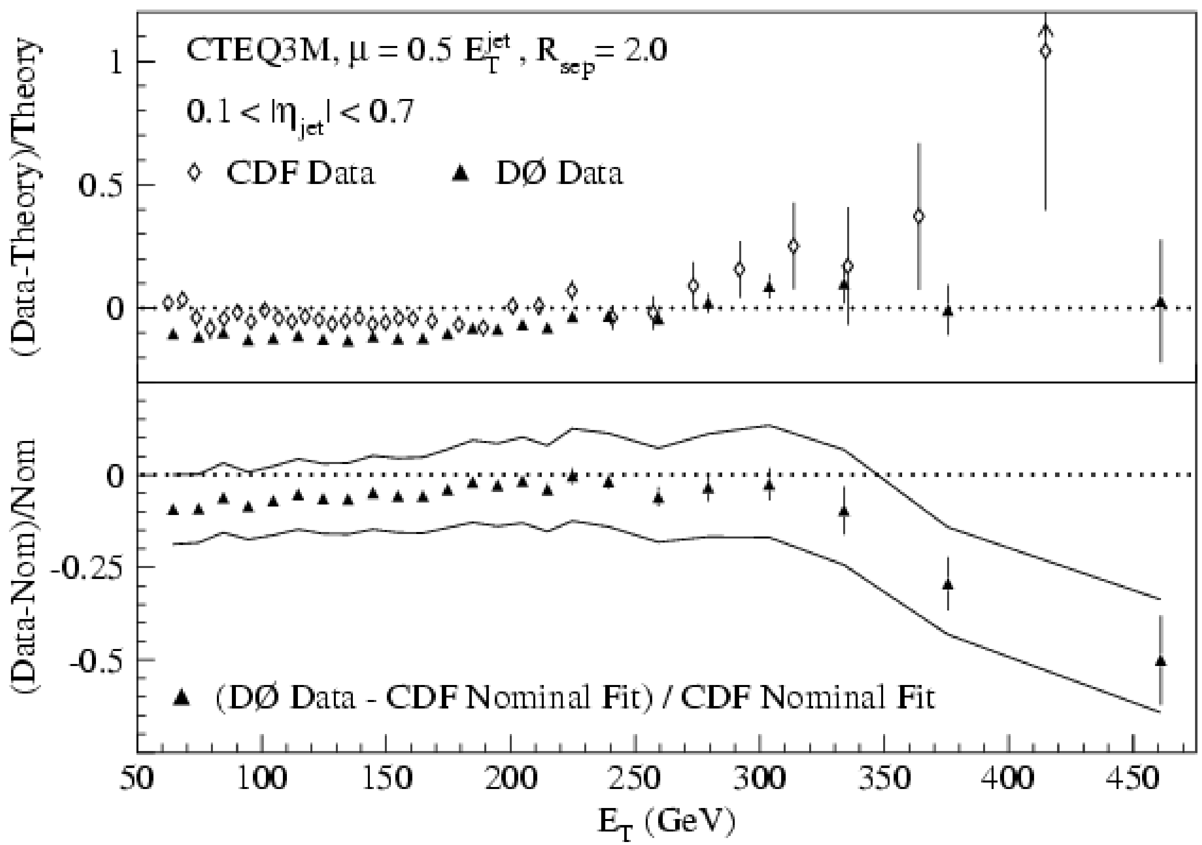
DØ LIQUID ARGON CALORIMETER



One of the more exciting stories

- The inclusive jet cross section measurement caused quite some excitement in Run I when CDF saw interesting features at high E_T

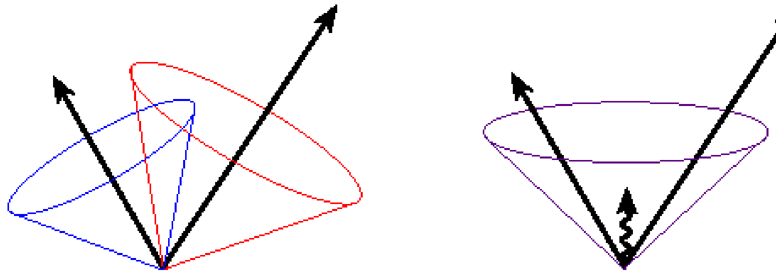
CDF to DØ comparison



- Both measurements were later shown to be explainable by increased gluon PDFs, renormalization scale in theory and cone algorithm; all within the allowed parameter range of QCD

Jet algorithm

- Detailed comparison to theory needs a precise definition of jet algorithm
- This measurement uses Run II Midpoint Cone with R_{cone} = 0.7



Run I Legacy Cone:

Draw a cone of fixed size in η–ϕ space around a seed

Compute jet axis from E_T-weighted mean and jet E_T from $\sum E_T$'s

Draw a new cone around the new jet axis and recalculate axis and new E_T

Iterate until stable

Algorithm is **sensitive to soft radiation**

Run II Midpoint Cone:

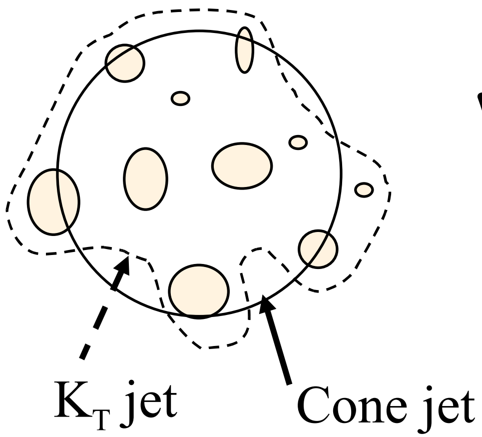
Use 4-vectors instead of E_T

Add additional midpoint seeds between pairs of close jets

Split/merge after stable protojets found

Improved infrared safety at NLO

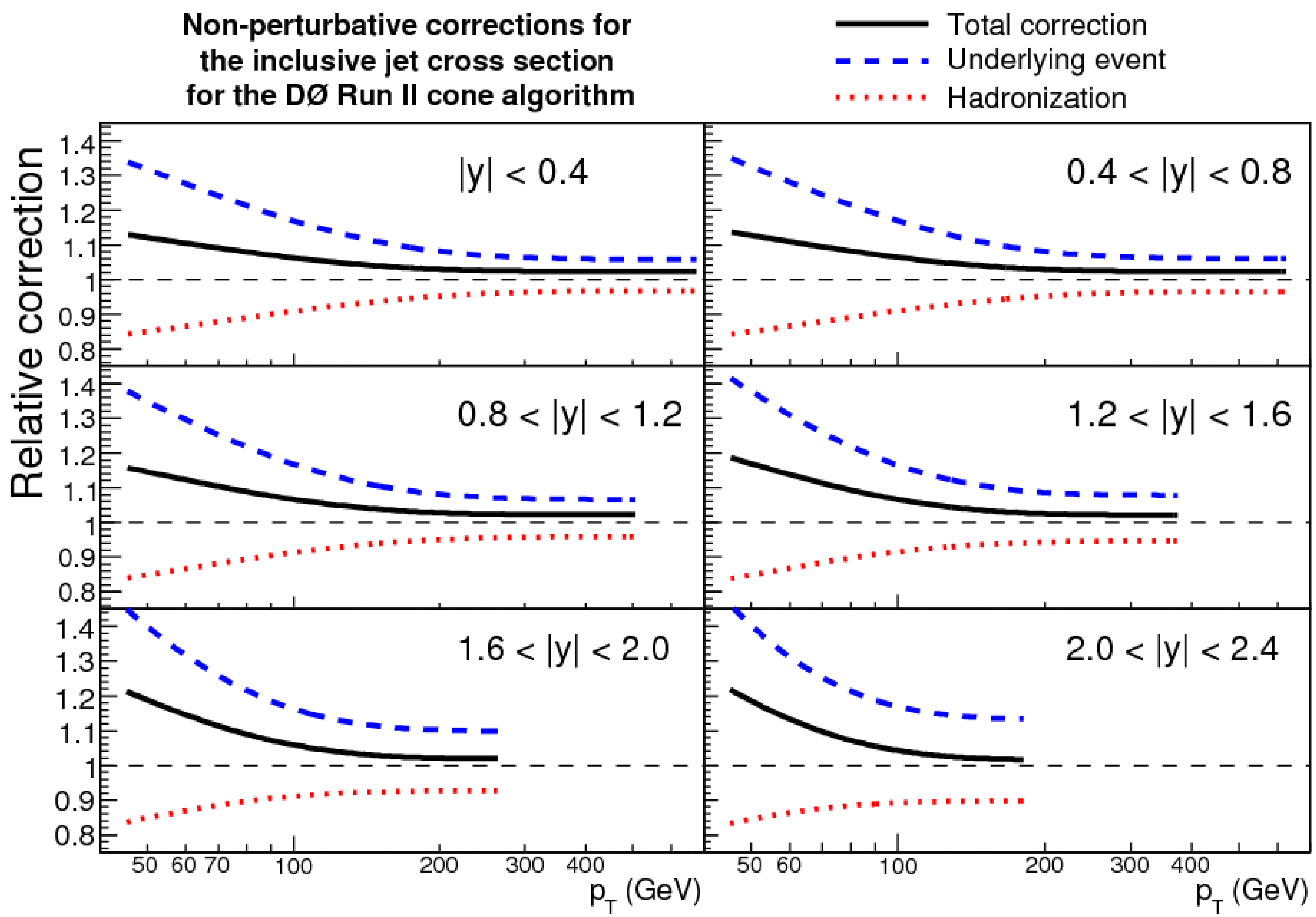
(D0 Run II/CDF MIDPOINT)



We characterize jets in terms of p_T and y

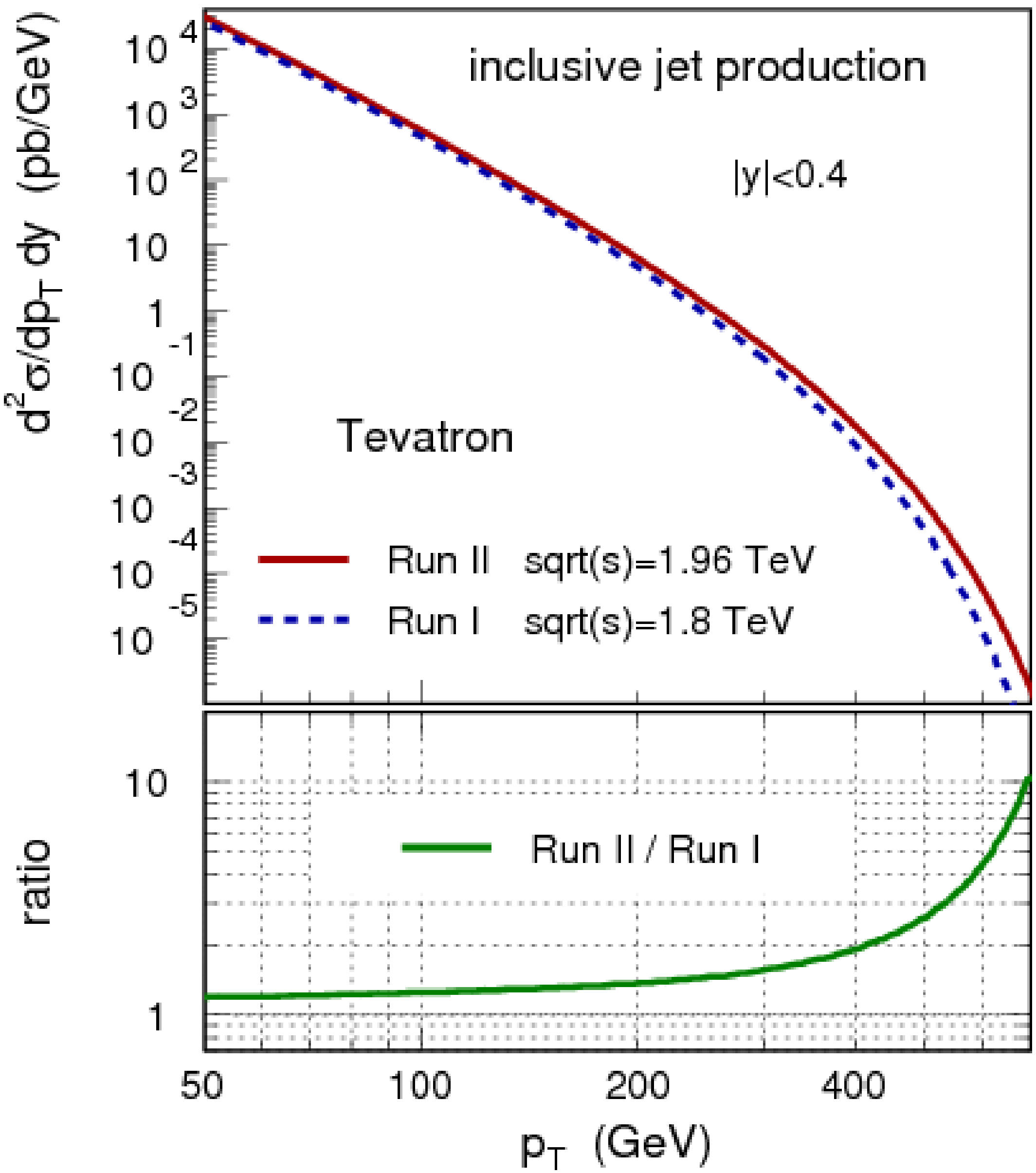
Non-perturbative corrections

- Hadronization and underlying event soft QCD effects and cannot be calculated with perturbation theory
- Pythia tune A used to calculate the non-perturbative corrections to theory



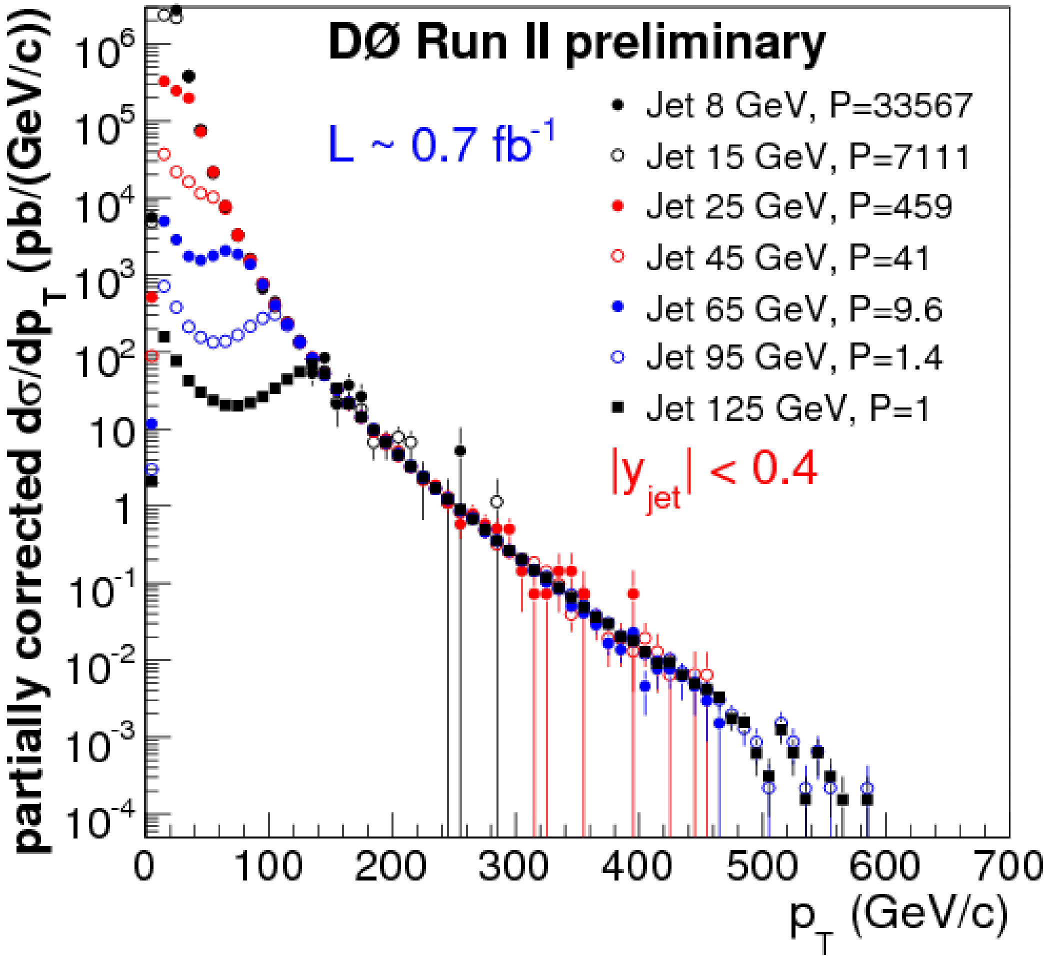
Run II advantage

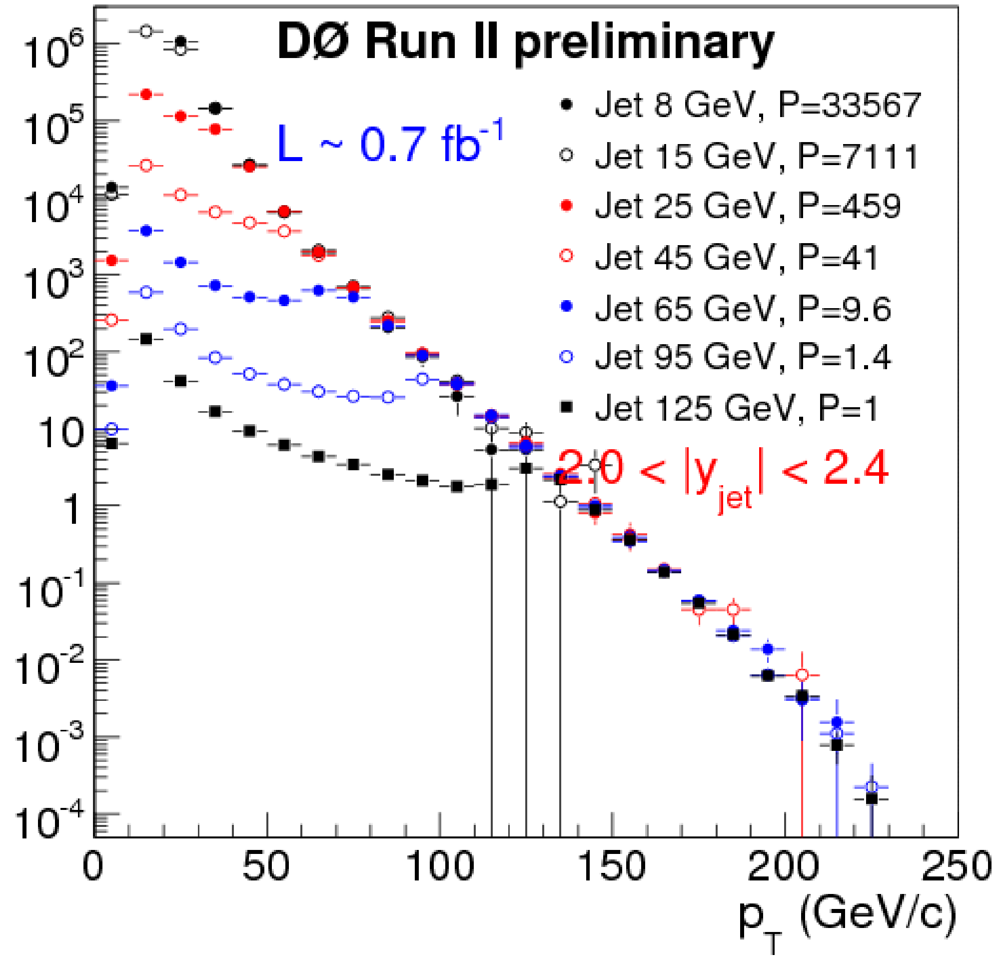
- Luminosity now ten times that of Run I ⇒ ×3 gluon PDF sensitivity
- Center-of-mass energy also 10% higher ⇒ three times higher cross section at p_T = 550 GeV
- Luminosity + cross section increase ⇒ ×5 quark substructure sensitivity



Triggers

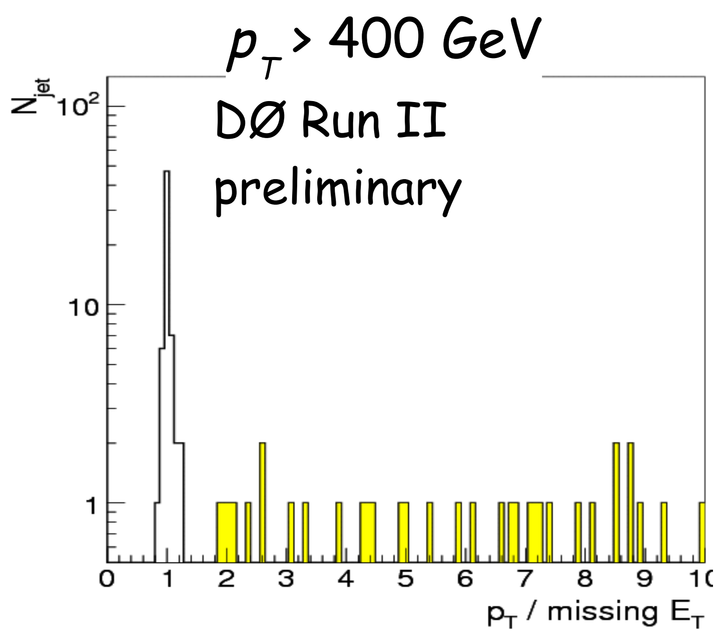
- Triggers fire on single jets above p_T threshold
- Measurement spans eight orders of magnitude in six rapidity regions
- Full p_T spectrum combined from seven different triggers

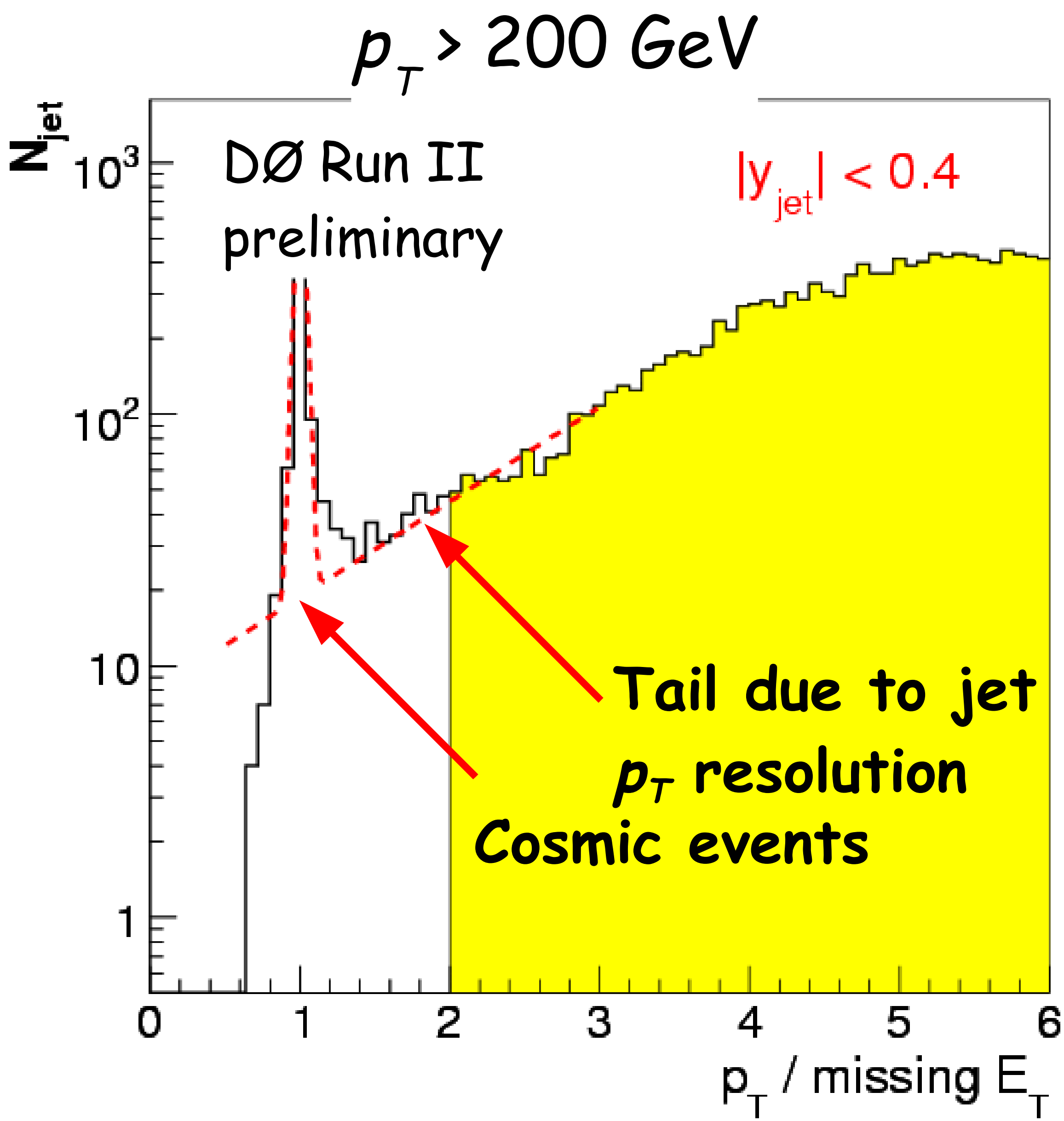




Cosmic background

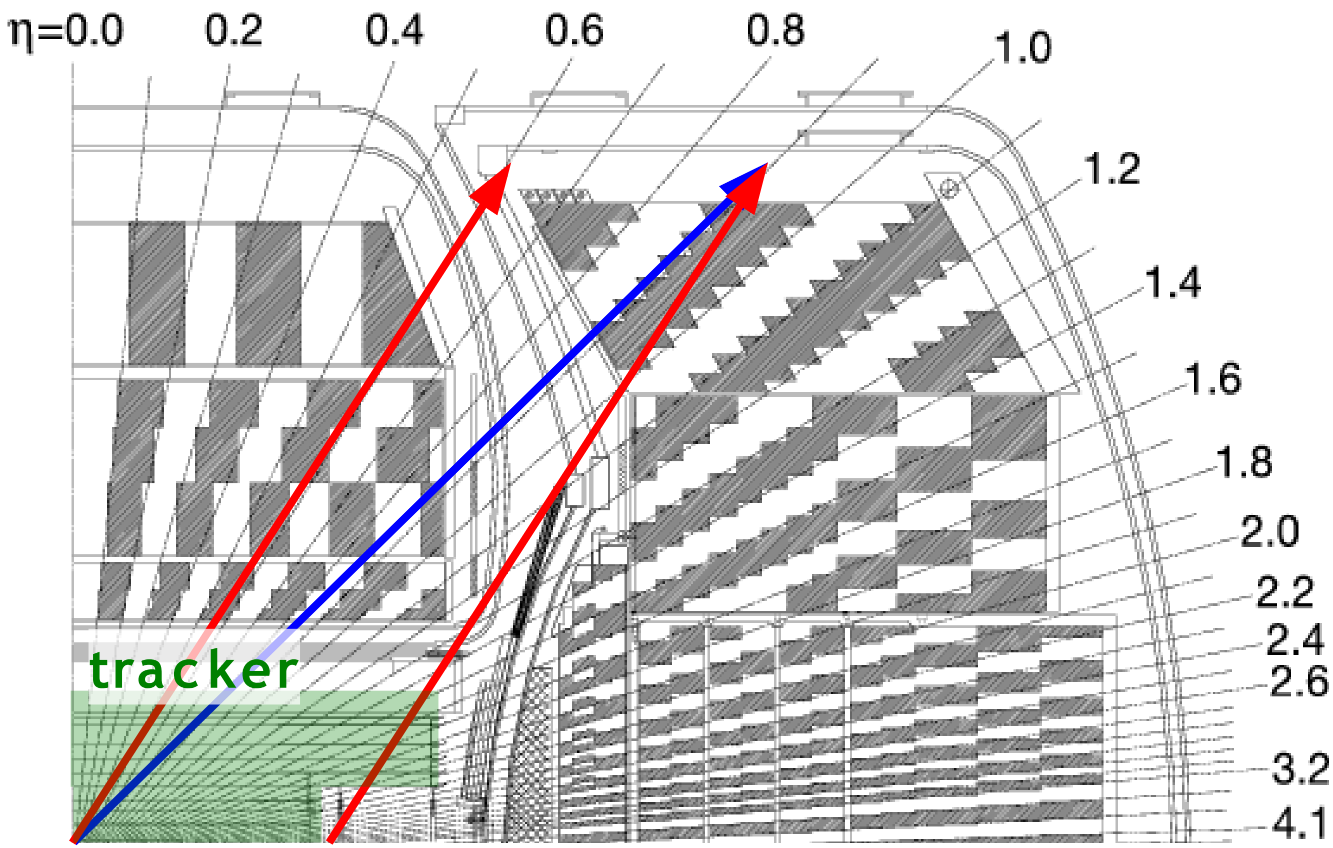
- Cosmic peak at p_T / MET ~ 1 comparable to inclusive jet cross section at p_T > 400 GeV
- Missing-E_T cut is important to remove cosmic background:
  - high rejection (100%)
  - low inefficiency (<0.5%)





Vertex cut

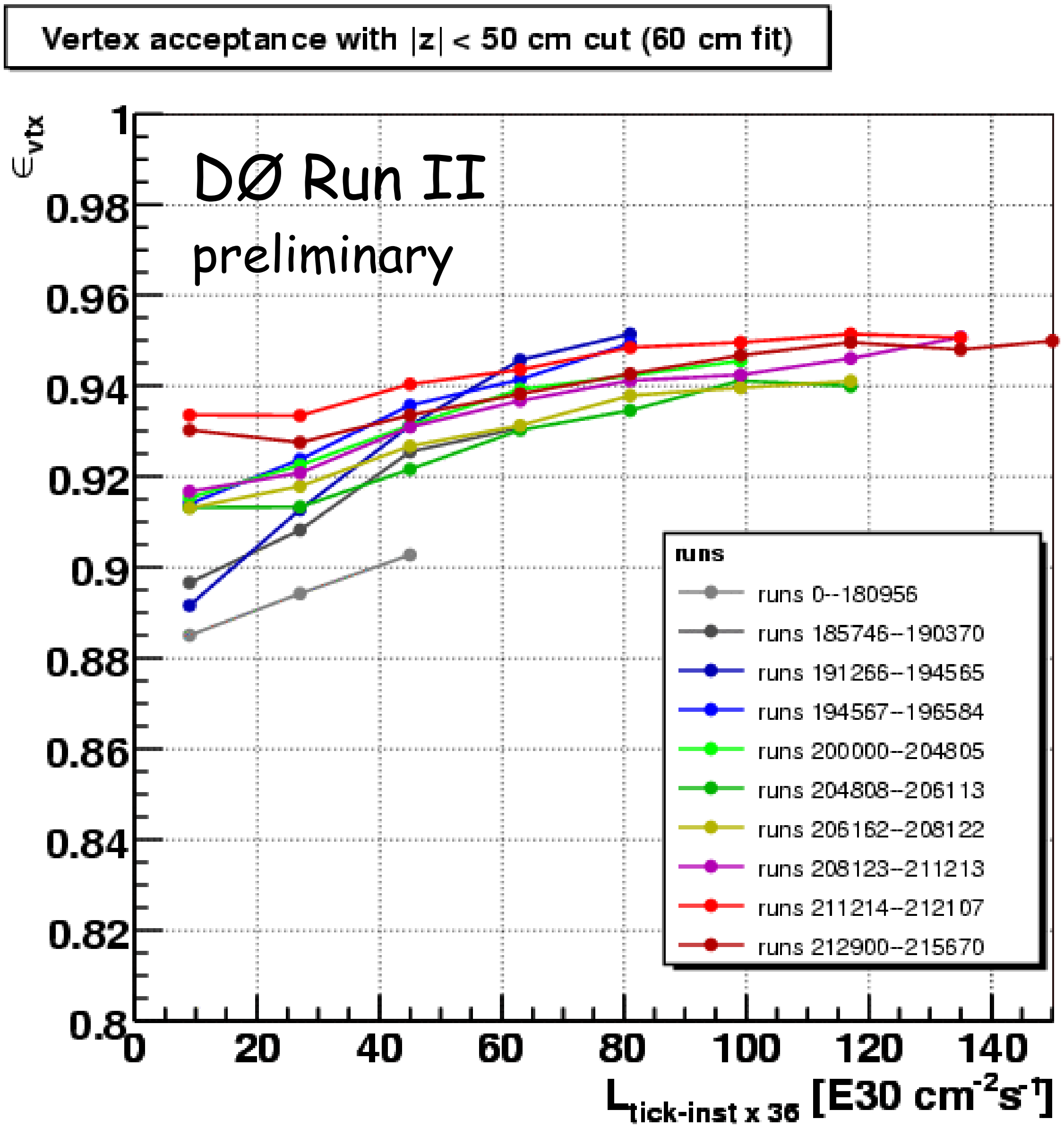
- Interaction vertex position is required to be within $|z_{vtx}| < 50$ cm of the calorimeter to improve jet p_T resolution
- Jets at large z_{vtx} can hit the calorimeter at a weird angle and at worse miss most of the calorimeter



- Vertex is needed for p_T reconstruction (E from the calorimeter, p_T with the vertex)
- Tracking efficiency quickly degrades beyond $|z_{vtx}|$=40—50 cm

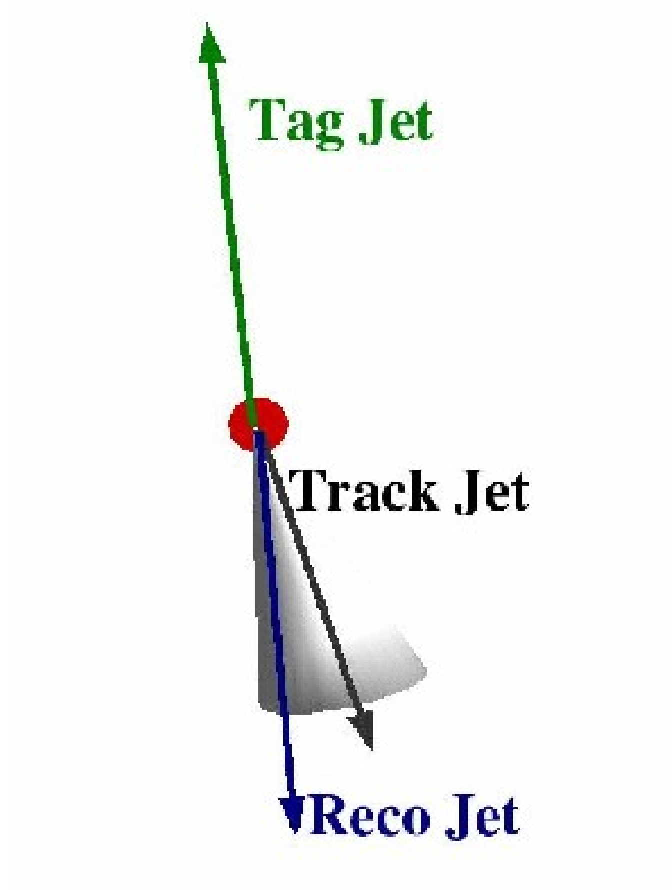
Vertex efficiency

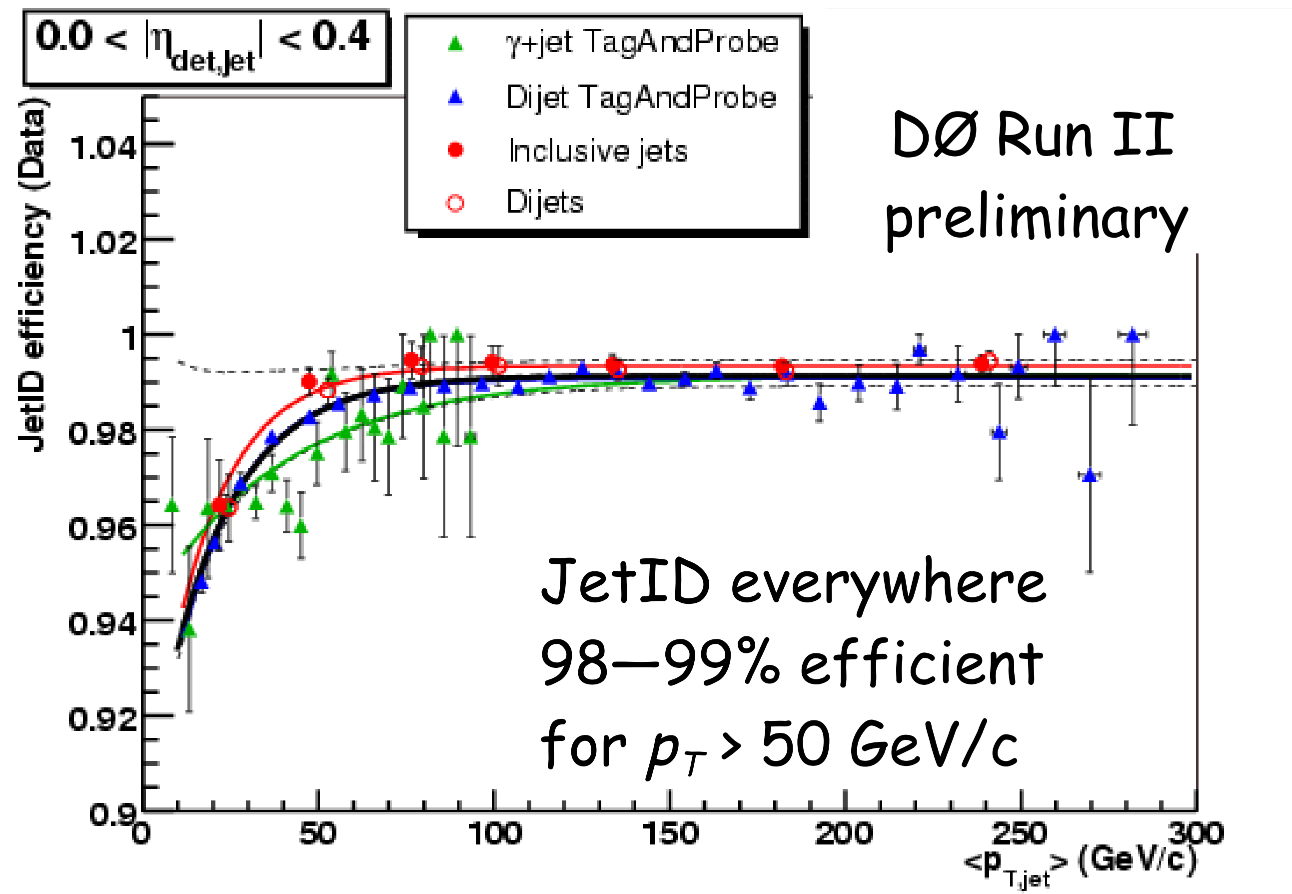
- Vertex cut efficiency is calculated from the longitudinal beam shape
- Time and luminosity dependence
  - Beam parameters (β^*) changing in epochs
  - Beam heating with time in store (luminosity ~ 1 / time)
- Average inefficiency 7.0±0.5%
- Leading inefficiency, others much smaller



JetID inefficiency

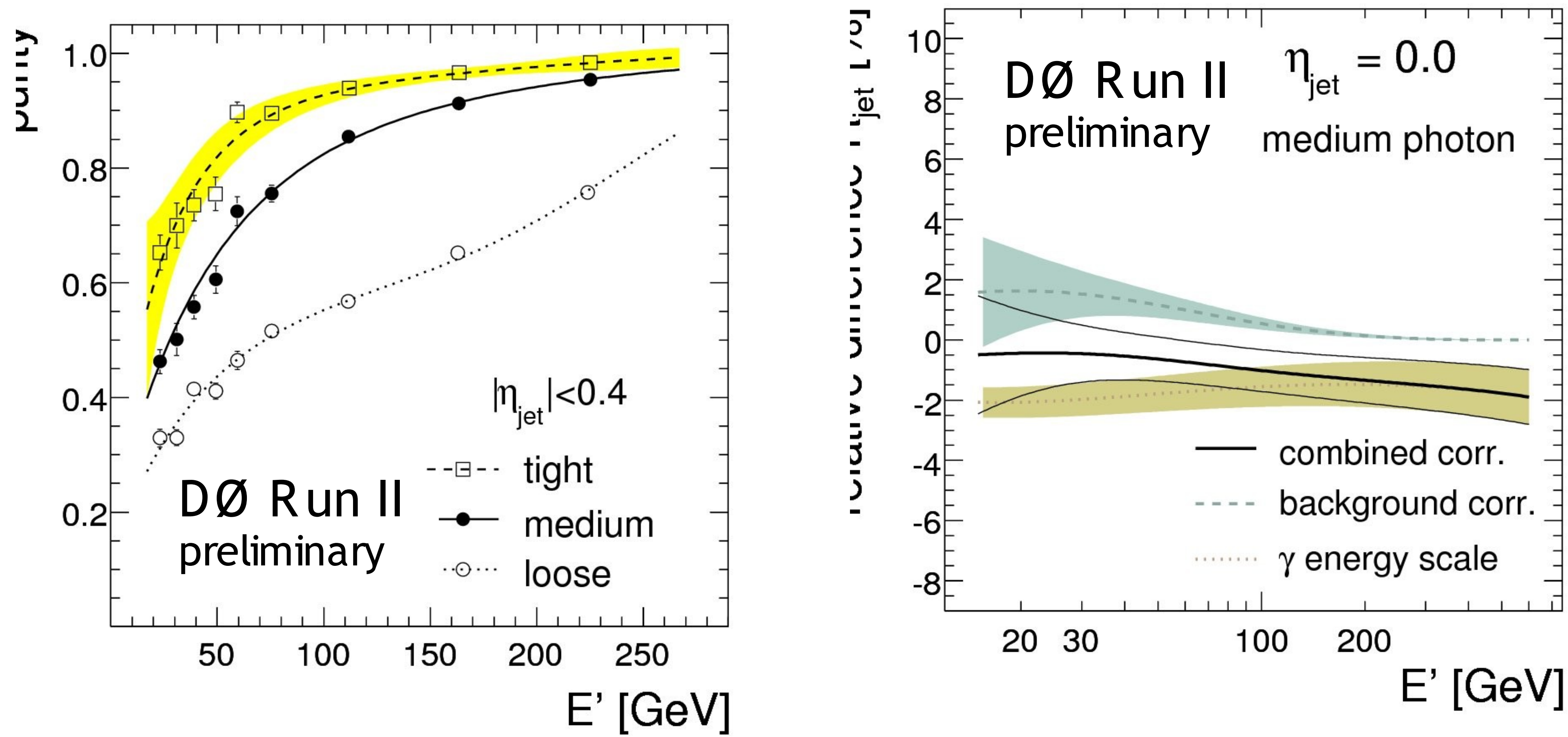
- JetID efficiency determined with the tag-and-probe method:
  - Tag is a good jet (or a photon) and an opposite track jet ⇒ good event
  - Probe is a reconstructed jet close to the track jet
- Cross checks with different samples and direct cut fraction





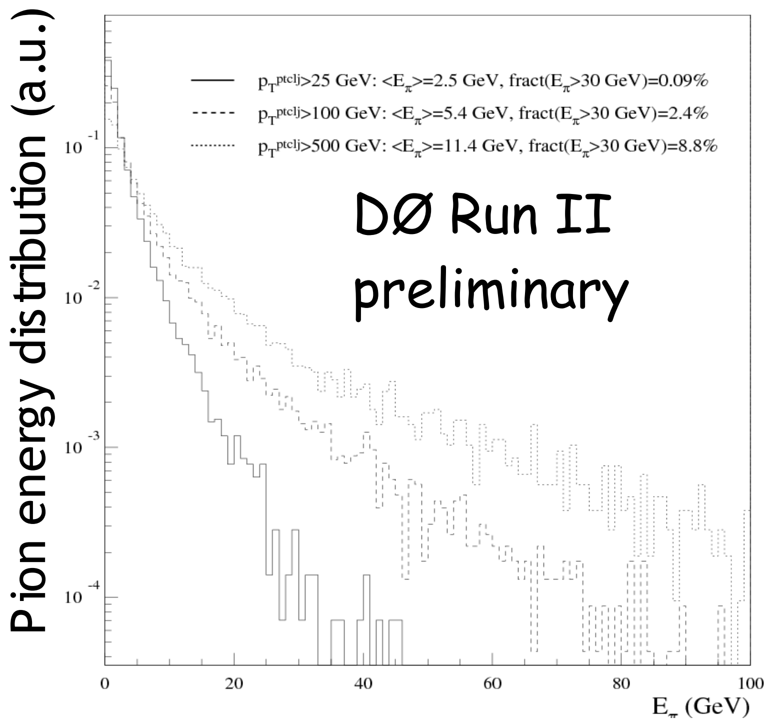
EM-jet background

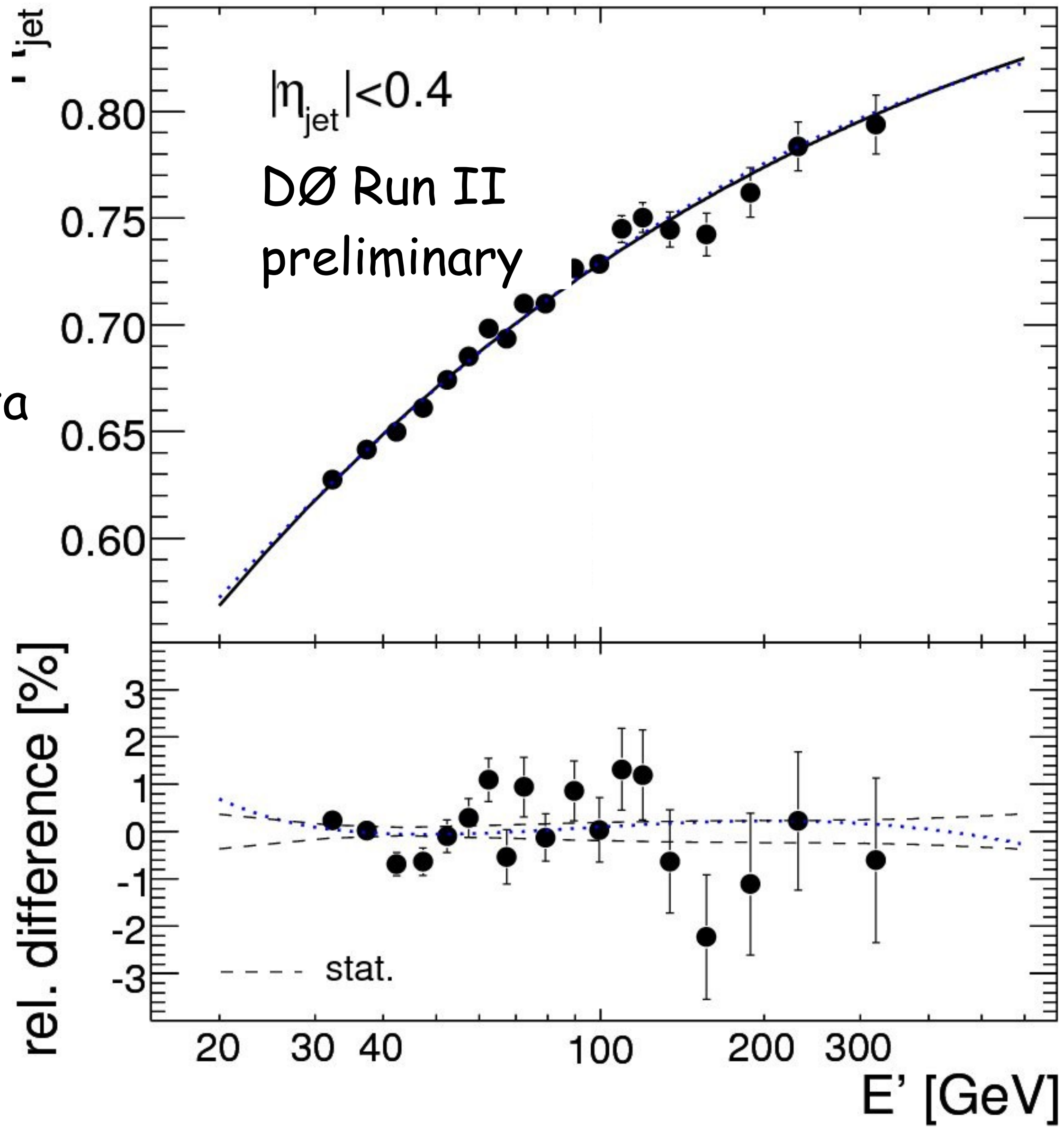
- Even with the tightest photon ID, γ+jet sample has significant EM-jet (leading $\pi^0 \rightarrow \gamma+\gamma$) background (dijet cross section $\times$1000)
- To reduce systematics, derived purity and energy scale for EM-jets, which are considered as part of the calibration sample ⇒ (γ/EM-jet)+jet sample



JES: CC response fit

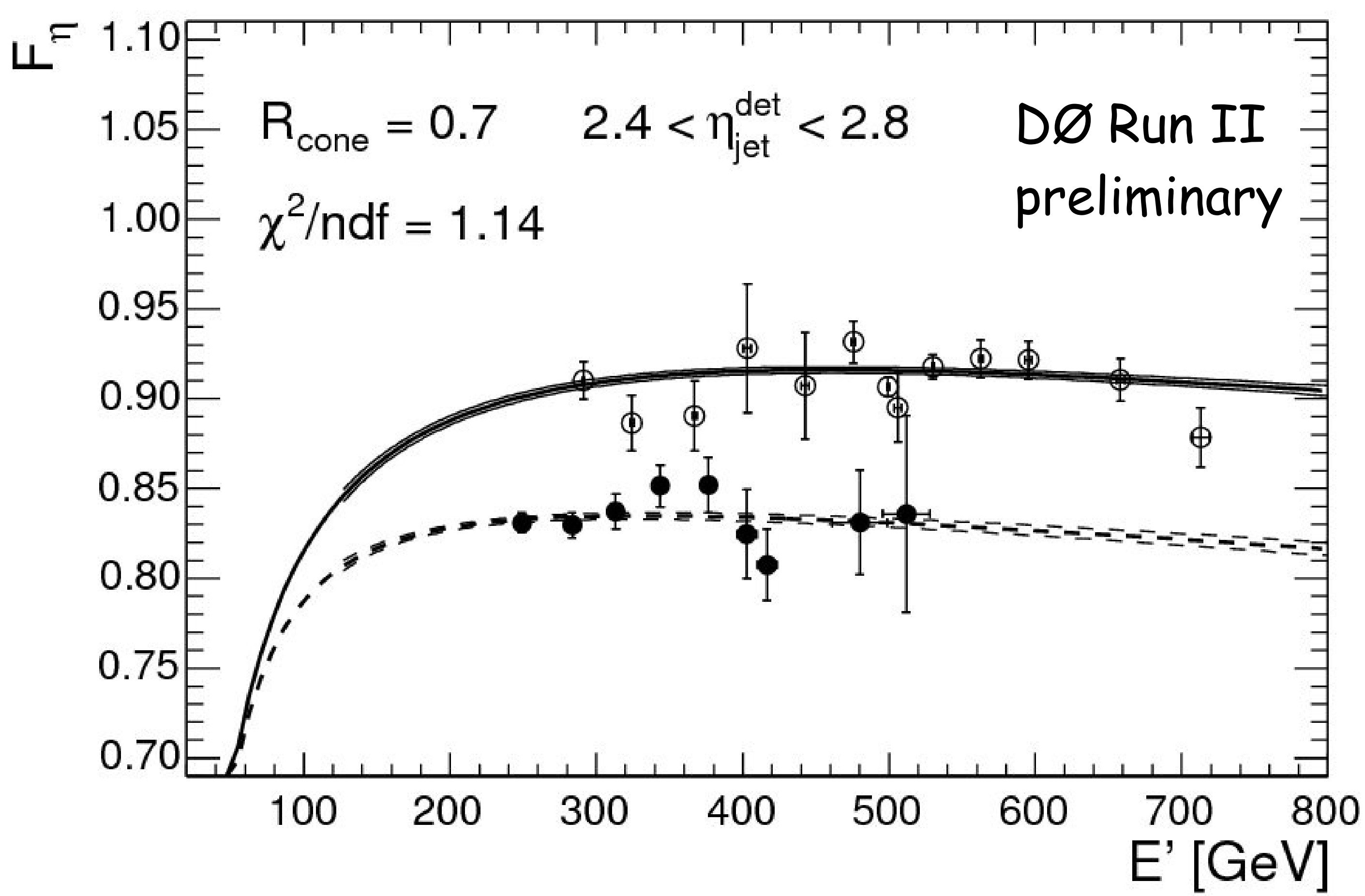
- Over 2% extrapolation uncertainty reduced by scaling single pion response in MC to γ+jet data
- Predict high p_T jet response by fitting low p_T pion response
- Agreement with isolated track data





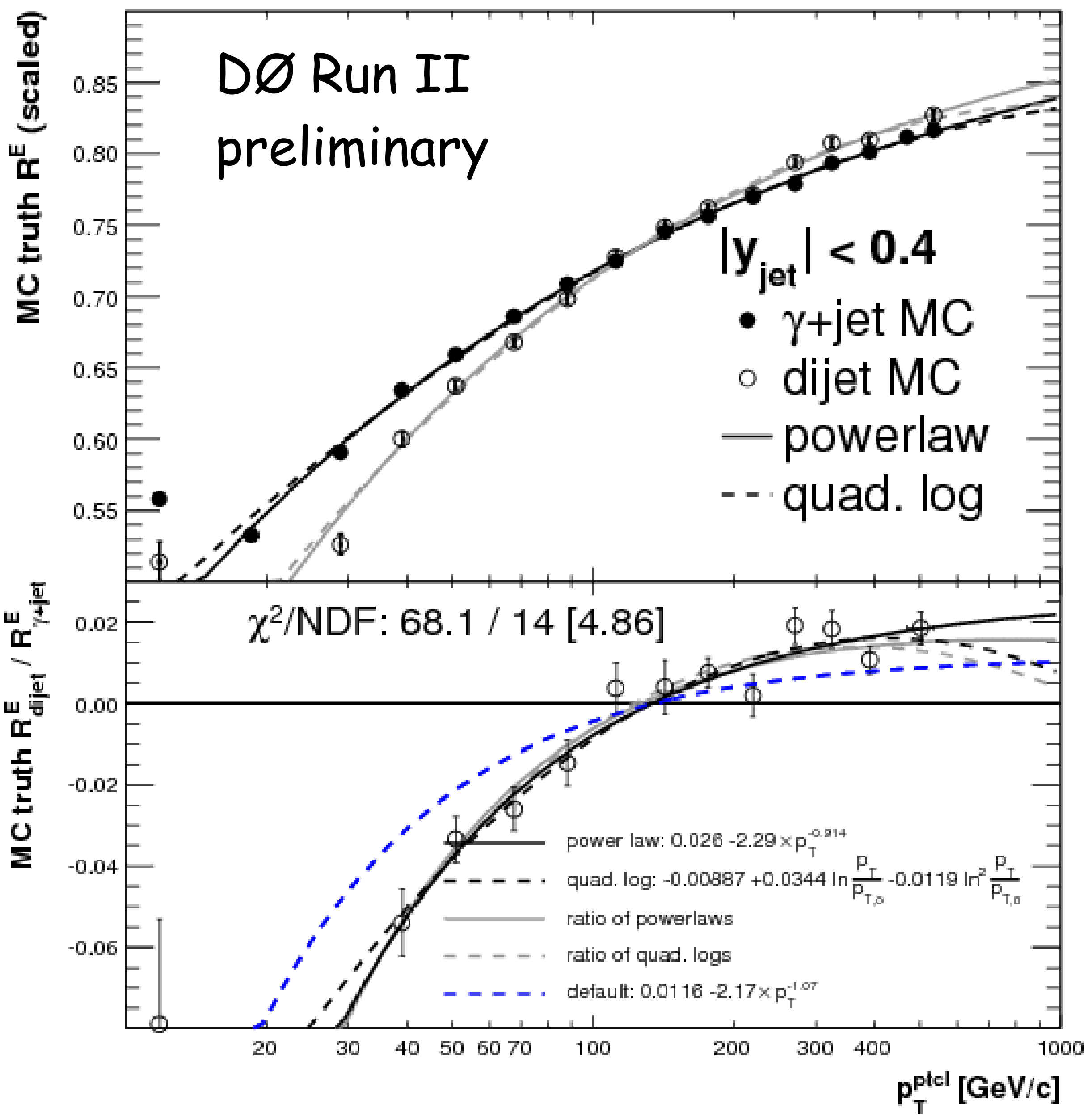
Eta-intercalibration

- Response η-dependence calibrated with respect to central jets and photons
- Dijets increase statistics at high p_T in the forward region compared to γ+jets
- Simultaneous fit to dijet and γ+jet samples taking into account sample differences
- Resolution bias for central jet in dijets explicitly corrected for and calibrated using central jet pairs

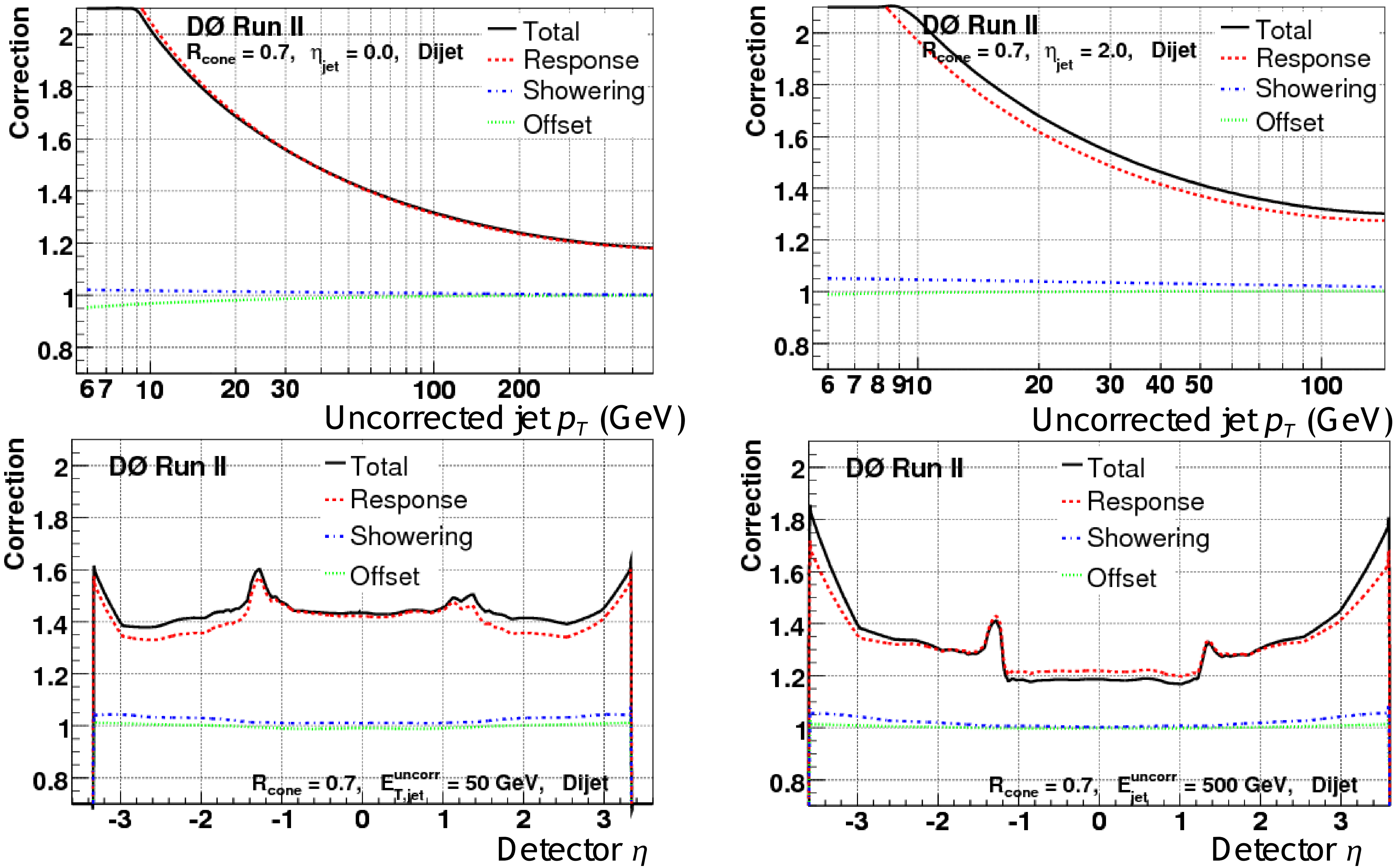


JES: Dijet corrections

- MC with **single pion response scaled to data** is used to derive the ratio of dijet and γ+jet responses in CC (-4% at 50 GeV, +2% at 400 GeV)
- Showering and bias corrections also rederived for dijets using tuned MC
- η-intercalibration for dijets directly from data
- Additional corrections for E / p_T difference and rapidity bias ⇒ four-momentum calibration

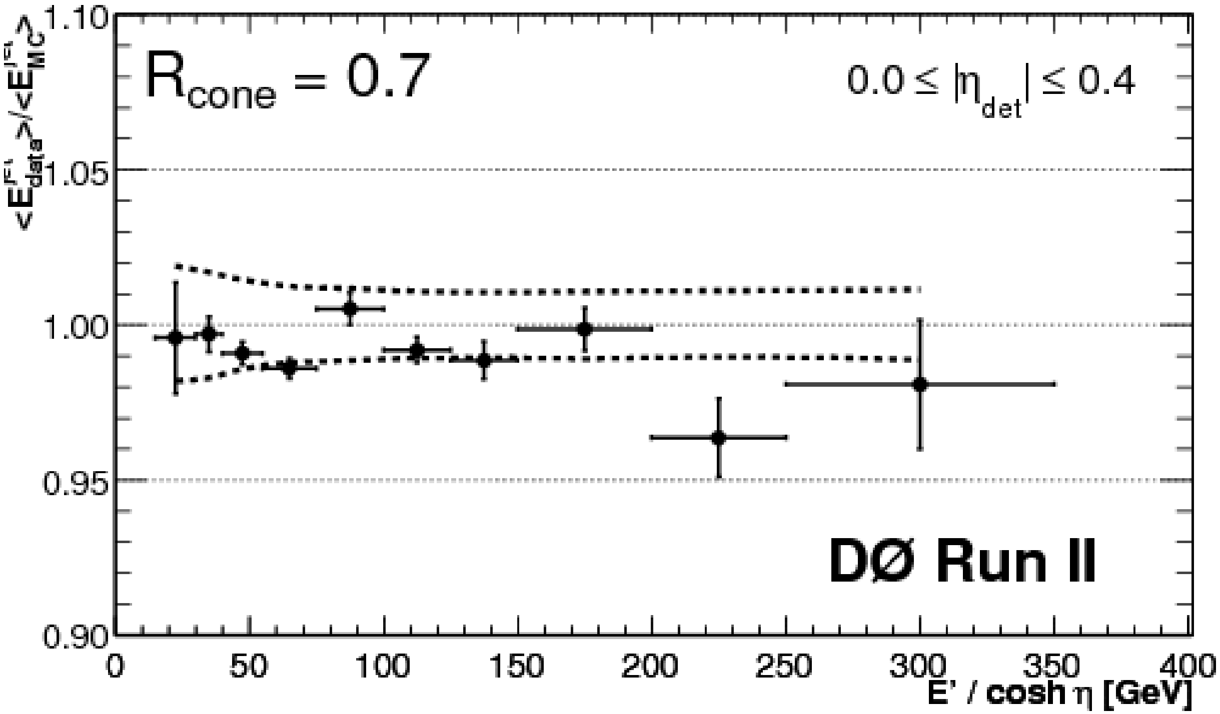


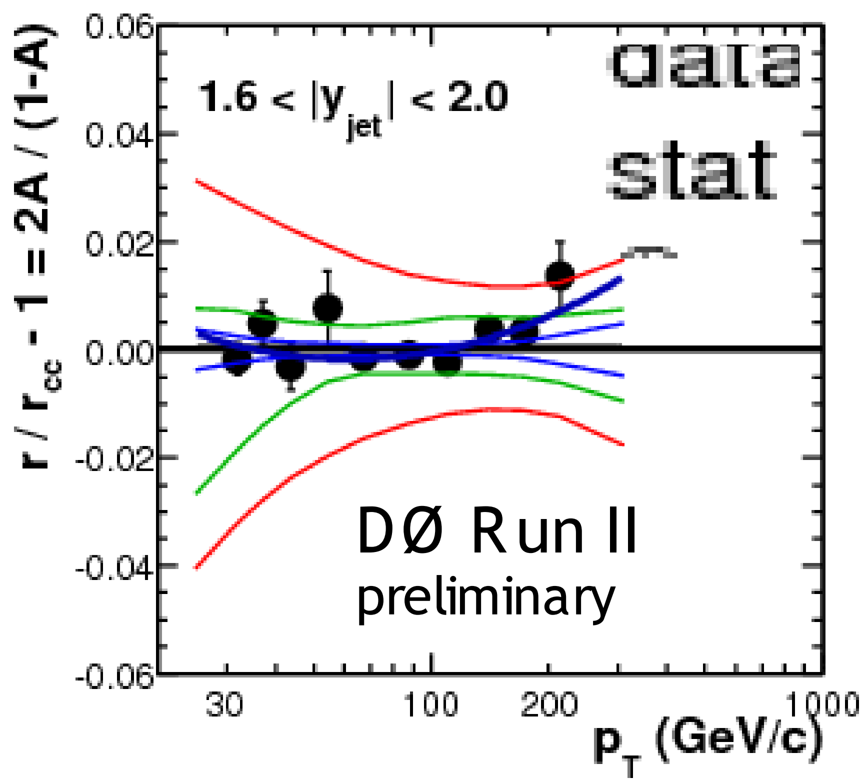
Dijet JES corrections



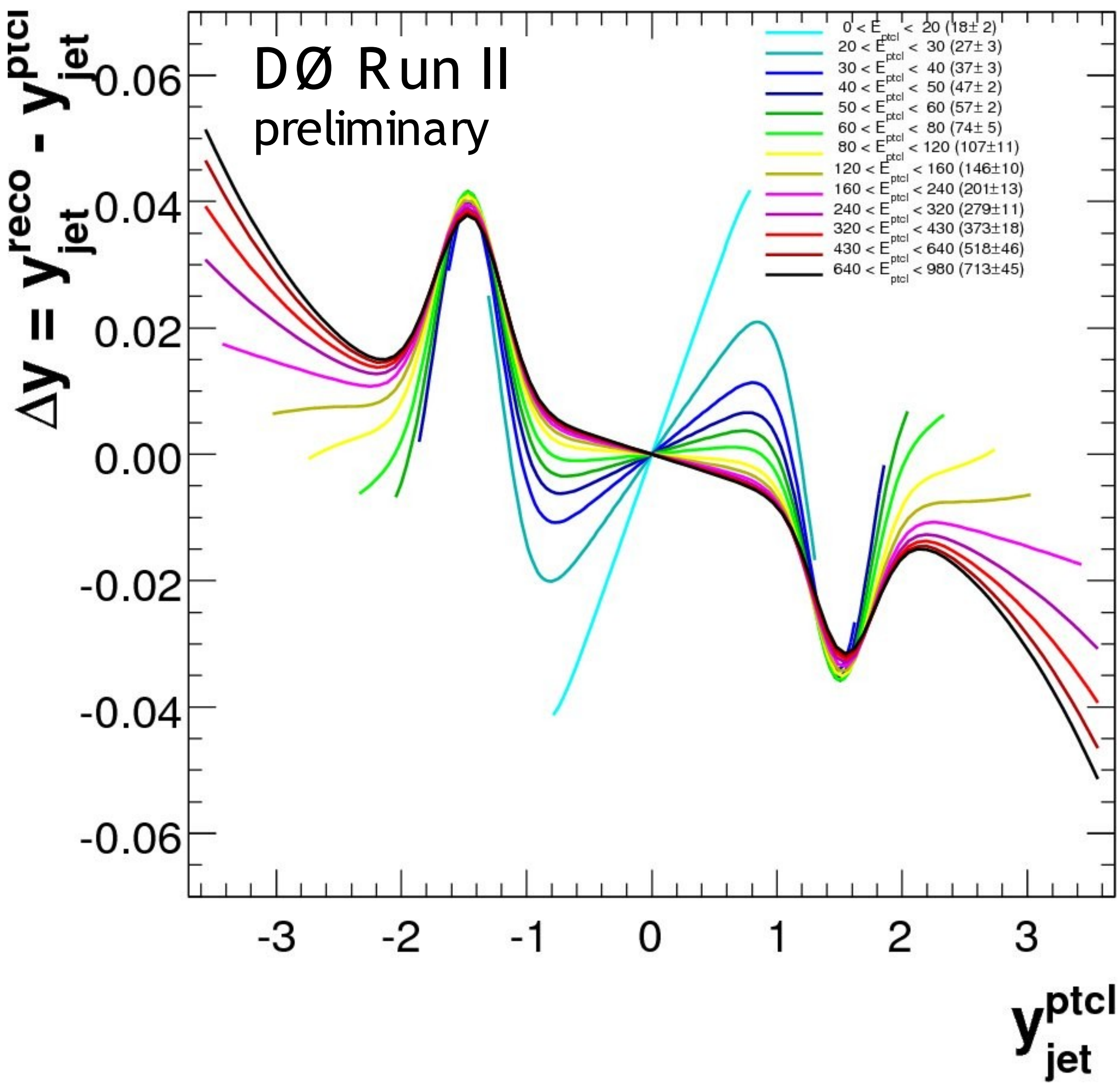
JES: Closure tests

- γ+jet closure tests consistency of JES corrections for absolute scale in CC
- Dijet closure tests the consistency of forward JES relative to CC
- Closure calculated from dijet asymmetry $A = (p_{T,fwd} - p_{T,cc}) / (p_{T,fwd} + p_{T,cc})$
- Explicit correction for residual resolution bias





JES: Rapidity bias



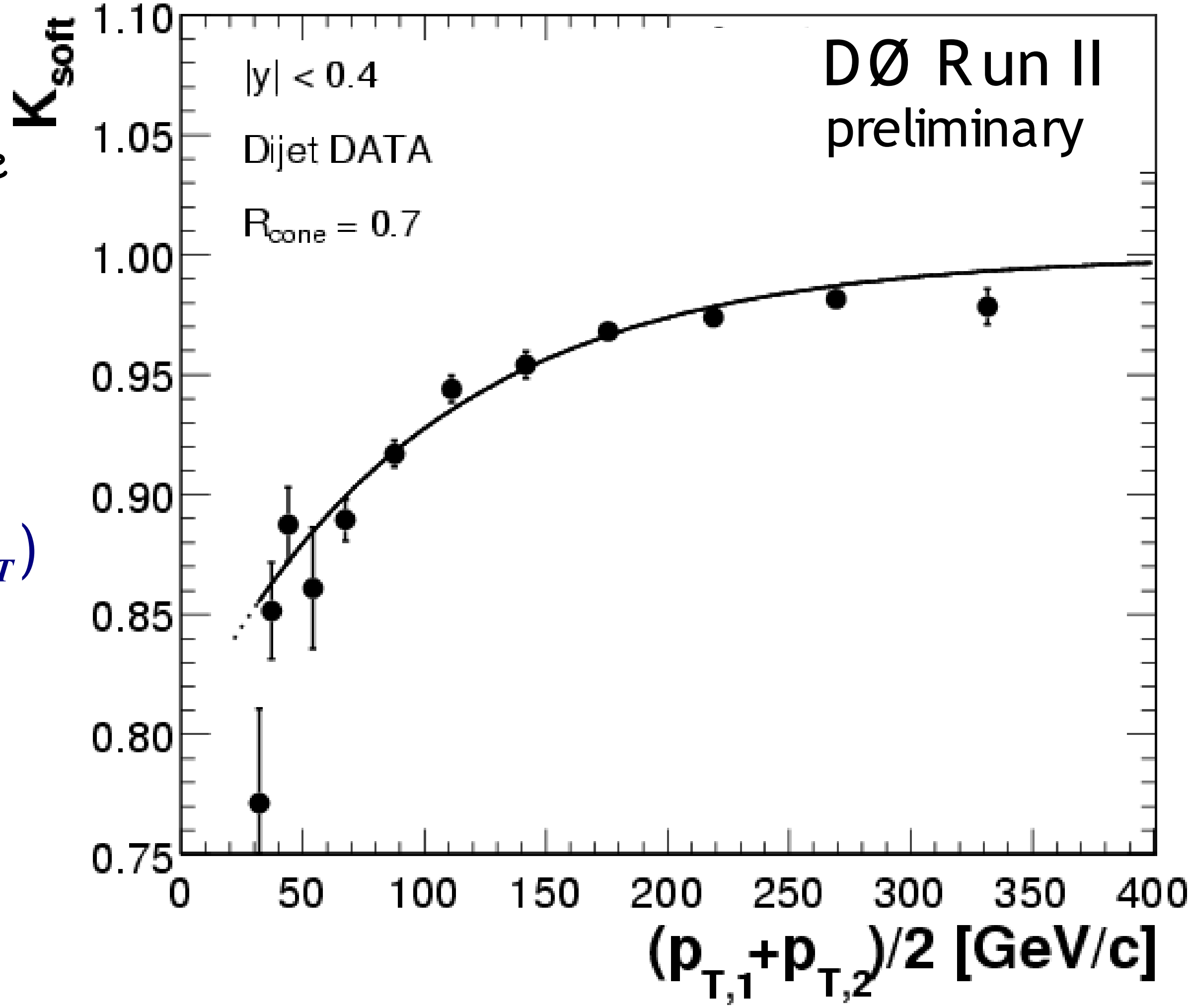
- Small detail: correction to p_T is much more important
- Jets are biased in rapidity on average toward the center of the calorimeter
- At most (in ICR), bias little less than half a cell width

Soft radiation correction

- Soft radiation estimated by increasing reconstruction threshold $p_{T,soft}^{cut}$ and the bias
- Extrapolation to $p_{T,soft}^{cut} \rightarrow 0$ gives the correction
- Soft radiation correction vanishes asymptotically at high p_T:

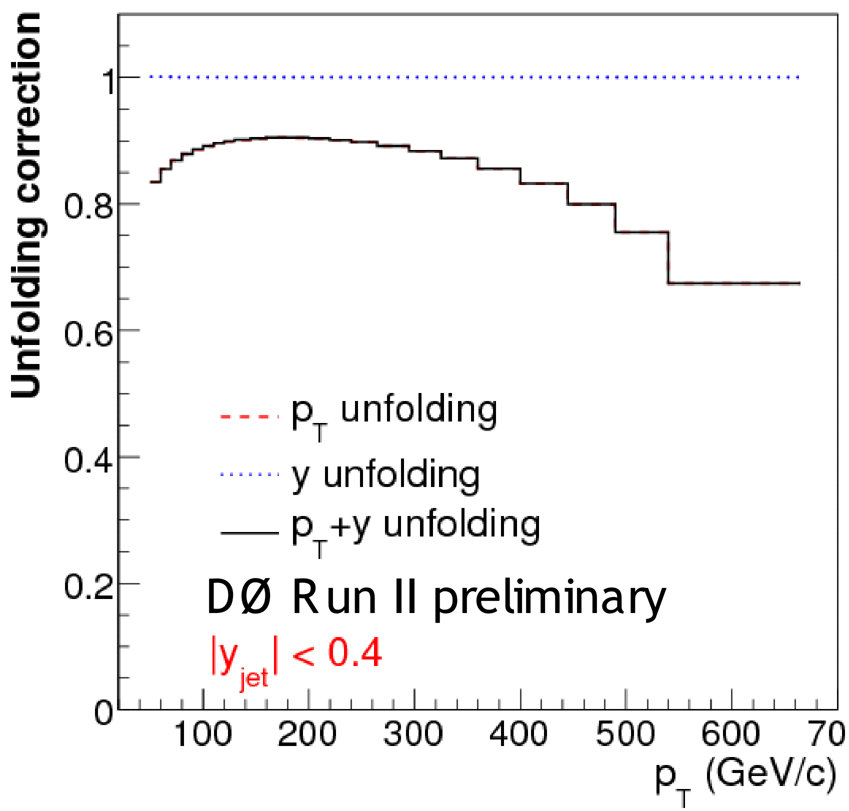
$$k_{soft}(p_T) = 1 - \exp(-a_0 - a_1 p_T)$$

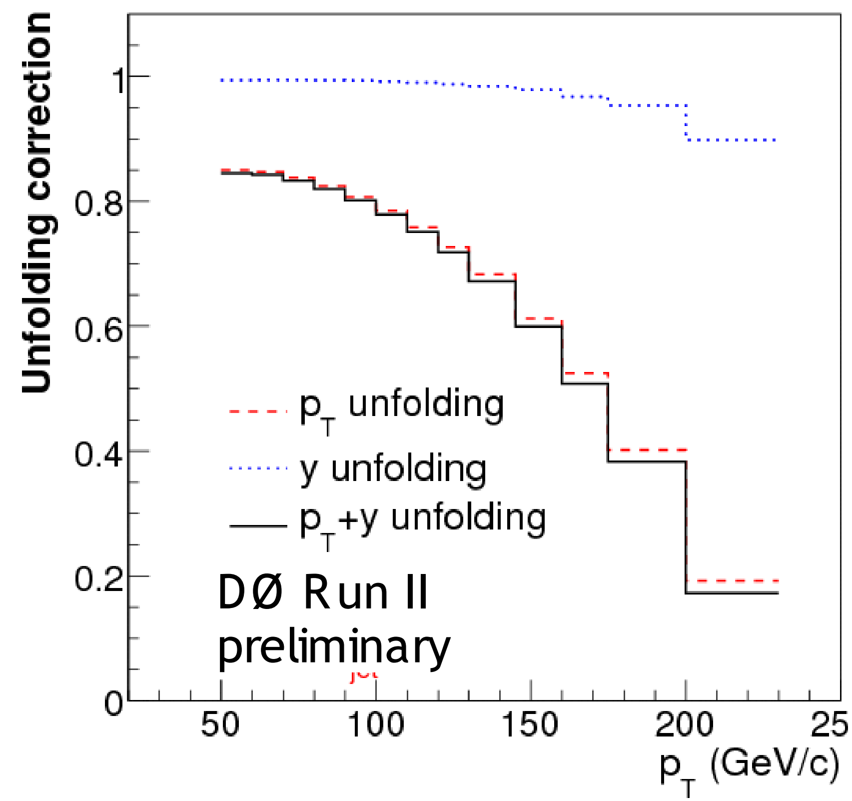
- Particle level imbalance from asymmetry in pure particle level MC after soft radiation correction
- Small correction, <10% everywhere



Jet p_T unfolding

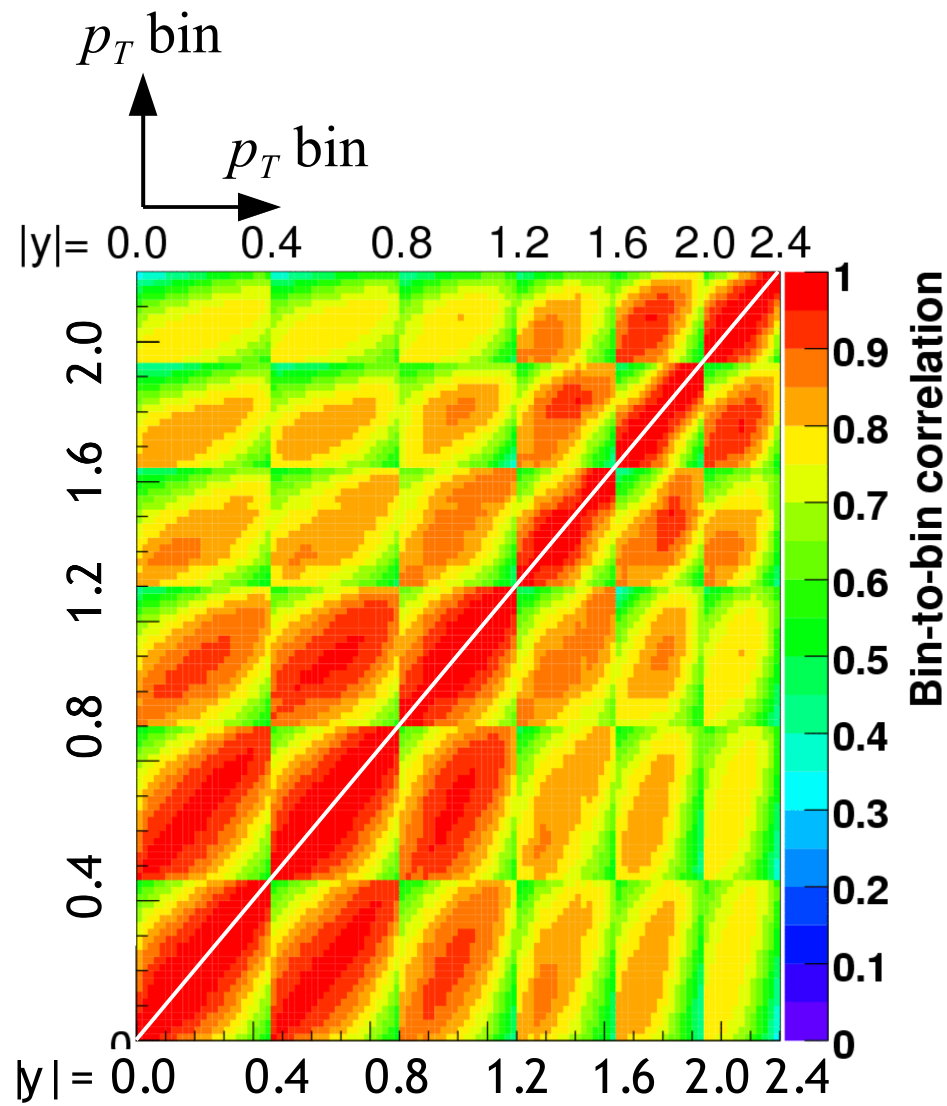
- Observed cross section is higher than true because more events migrate from high (and low) $p_{T,\mathrm{ptcl}}$ into a given bin of measured p_T than migrate out of the bin due to jet p_T resolution ⇒ net increase
- Model the true cross section (ansatz method) and smear it (⇒resolution!) to obtain the observed cross section and then iteratively fit this to data





Uncertainty correlations

- The uncertainty correlations are provided in the format CTEQ uses: set of independent variations (sources) describing how points move together
- Average bin-to-bin correlation of about **80% with RMS of 10%**
- Using the correlation information in the global PDF fit should **further reduce the effective uncertainty** in the measurement



Uncertainty correlations

- Leading sources are from JES:
  - **EM energy scale** ($Z \to e^+e^-$ calibration)
  - **Photon energy scale** (MC description of e / γ response, material budget)
  - **High p_T extrapolation** (fragmentation in Pythia/Herwig, PDFs)
  - **Rapidity decorrelation** (uncertainty in η-dependence)
  - **Detector showering** (goodness of template fits)
- Only five highest out of 23 correlated systematics shown

PRL 101, 062001 (2008)