The Status of Generator Tuning from Run2

Stephen Mrenna

Computing Division
Generators and Detector Simulation Group
Fermilab

HERALHC 03/13/07

Work with: G. Choudalakis, R. Culbertson, C. Henderson, B. Knuteson

Thanks to: R. Field, MLM

Motivations for Comparing Data to Monte Carlo

1. Validation
    - Check if Data and MC consistent in a control region to extrapolate into the signal region
2. Developing Corrections
    - Make the Data and MC agree in a control region to use for different physics
3. Tuning
    - Test the basic physics and fit the phenomenological parameters inside the event generators

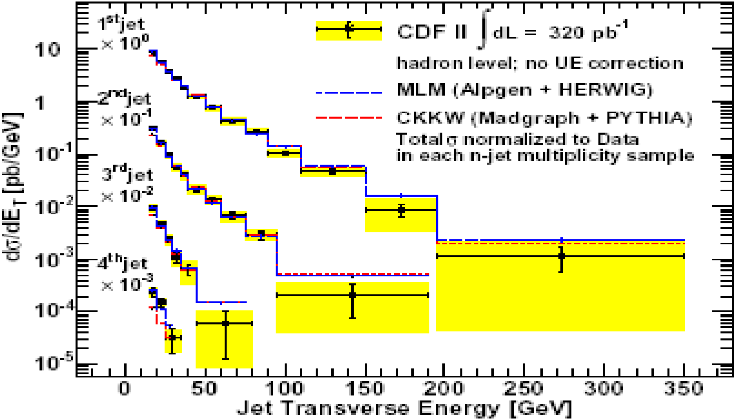
Experiments do mainly 1. and 2.

There are Exceptions

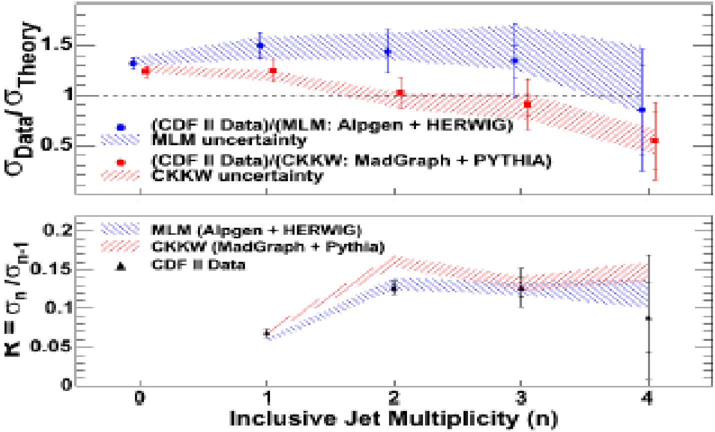
- Pythia UE Tunes
  - R. Field (CDF) + students + ATLAS people
- `Vista` analysis of *all* high-p_T data
  - Knuteson-MIT group + Culbertson (CDF) + SM
- Isolated other cases
  - DØ dijet correlations, a few others

Example of Validation

- Method: Compare different predictions of W+multijet events
- Goal: Estimate the systematic errors on the $t\bar{t}$ cross section measurement from theory



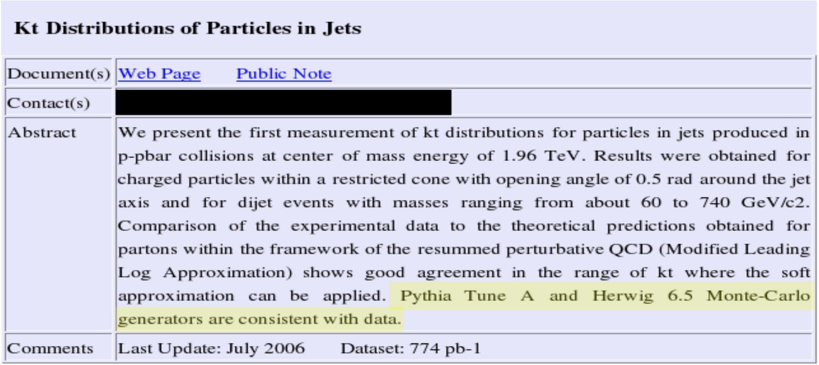
[CDF Prelim: Cooper,Huston, Messina, Waters]



[CDF Prelim: See upcoming talk!]

- There are differences
- Not as important as other systematics for now
- Will be important in the near future

Another kind of validation



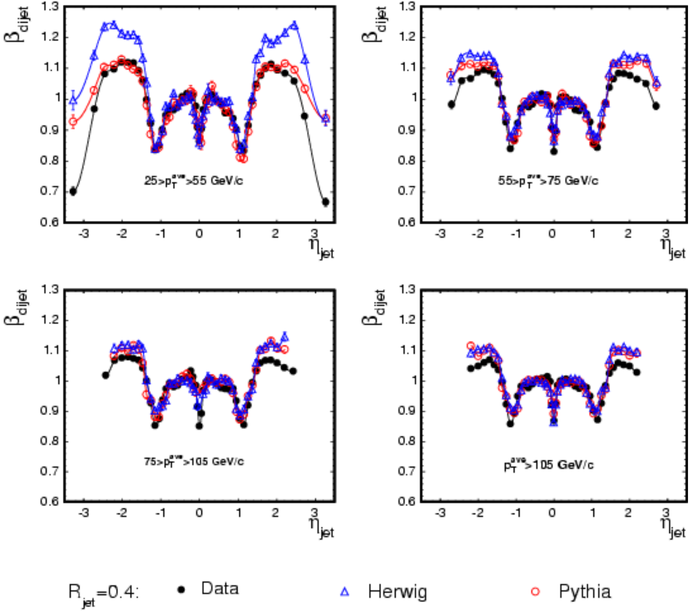
Kt Distributions of Particles in Jets

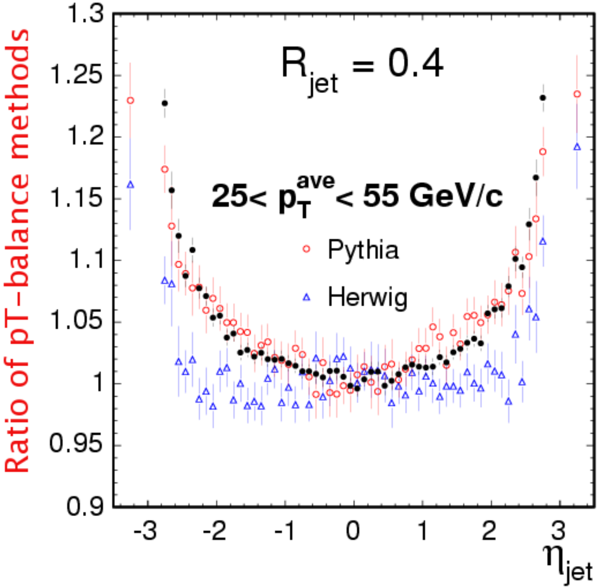
Document(s)	Web Page Public Note
Contact(s)	[redacted]
Abstract	We present the first measurement of kt distributions for particles in jets produced in p-pbar collisions at center of mass energy of 1.96 TeV. Results were obtained for charged particles within a restricted cone with opening angle of 0.5 rad around the jet axis and for dijet events with masses ranging from about 60 to 740 GeV/c2. Comparison of the experimental data to the theoretical predictions obtained for partons within the framework of the resummed perturbative QCD (Modified Leading Log Approximation) shows good agreement in the range of kt where the soft approximation can be applied. Pythia Tune A and Herwig 6.5 Monte-Carlo generators are consistent with data.
Comments	Last Update: July 2006 Dataset: 774 pb-1

Example of Developing a Correction Factor

- Method: Determine the correction factor between data and Monte Carlo for p_T balance in dijet or γ-jet events
- Goal: Determine an absolute energy calibration for jets to measure the top quark mass

pT(probe)/pT(trigger)





$\vec{p}_T \cdot \vec{\not{E}}_T$ vs. $p_{T1} - p_{T2}$

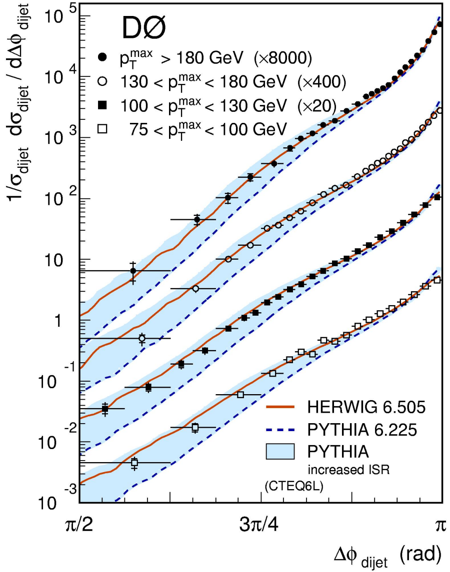
hep-ex/0510047, NIM

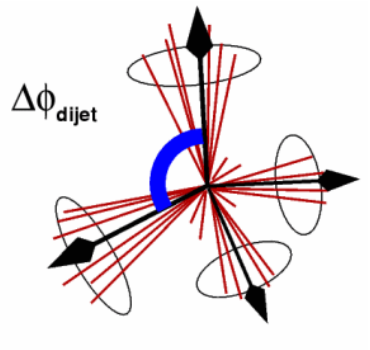
> *Since this behavior is only found in the dijet samples, we do not consider* `HERWIG` *dijet samples for the determination of the η-dependent corrections or their systematic uncertainties. In γ-jet, Z-jet or $t\bar{t}$ events no such problems are seen. At this moment we do not have any explanation for the differences. It could be due to initial or final state radiation, due to the underlying event modeling or many other effects, and it will be studied again in [the] ... future ...*

History of Tunes A-D

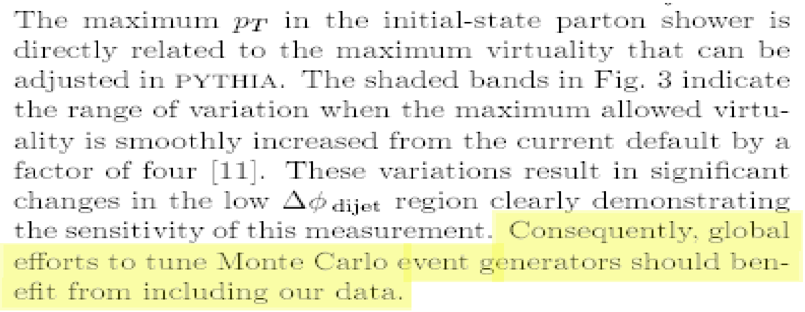
1. Theorist [R. Field] joins CDF
2. He can look at charged tracks, because he can't screw that up
3. Uses this as a vehicle to study UE
4. Finds the "best" tune is one that enhances ISR and has decreased MI interaction
5. Side note: I express concerns about best fit, but cannot quantify the size of an effect. Concerns are dismissed.
6. Tune A is almost exclusively adopted by the experiments and used for LHC extrapolations (with caveats)
7. Lots of good physics ensues

DØ Dijet Azimuthal Correlation



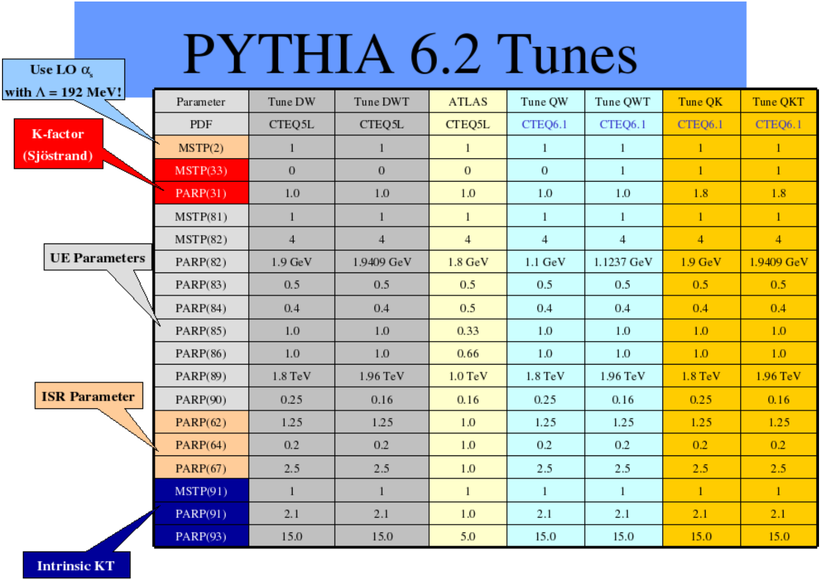


PRL 94, 221801 (2005)



The maximum p_T in the initial-state parton shower is directly related to the maximum virtuality that can be adjusted in PYTHIA. The shaded bands in Fig. 3 indicate the range of variation when the maximum allowed virtuality is smoothly increased from the current default by a factor of four [11]. These variations result in significant changes in the low $\Delta\phi_{\mathrm{dijet}}$ region clearly demonstrating the sensitivity of this measurement. Consequently, global efforts to tune Monte Carlo event generators should benefit from including our data.

Status of UE Tunes



PYTHIA 6.2 Tunes

Use LO α_s with Λ = 192 MeV!

K-factor (Sjöstrand)

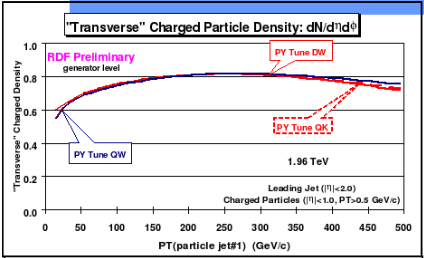
UE Parameters

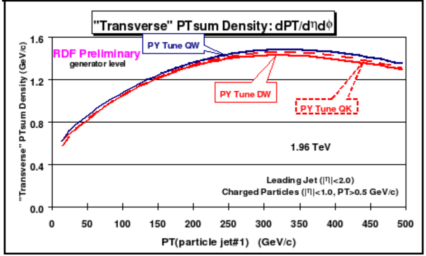
ISR Parameter

Intrinsic KT

Parameter	Tune DW	Tune DWT	ATLAS	Tune QW	Tune QWT	Tune QK	Tune QKT
PDF	CTEQ5L	CTEQ5L	CTEQ5L	CTEQ6.1	CTEQ6.1	CTEQ6.1	CTEQ6.1
MSTP(2)	1	1	1	1	1	1	1
MSTP(33)	0	0	0	0	1	1	1
PARP(31)	1.0	1.0	1.0	1.0	1.0	1.8	1.8
MSTP(81)	1	1	1	1	1	1	1
MSTP(82)	4	4	4	4	4	4	4
PARP(82)	1.9 GeV	1.9409 GeV	1.8 GeV	1.1 GeV	1.1237 GeV	1.9 GeV	1.9409 GeV
PARP(83)	0.5	0.5	0.5	0.5	0.5	0.5	0.5
PARP(84)	0.4	0.4	0.5	0.4	0.4	0.4	0.4
PARP(85)	1.0	1.0	0.33	1.0	1.0	1.0	1.0
PARP(86)	1.0	1.0	0.66	1.0	1.0	1.0	1.0
PARP(89)	1.8 TeV	1.96 TeV	1.0 TeV	1.8 TeV	1.96 TeV	1.8 TeV	1.96 TeV
PARP(90)	0.25	0.16	0.16	0.25	0.16	0.25	0.16
PARP(62)	1.25	1.25	1.0	1.25	1.25	1.25	1.25
PARP(64)	0.2	0.2	1.0	0.2	0.2	0.2	0.2
PARP(67)	2.5	2.5	1.0	2.5	2.5	2.5	2.5
MSTP(91)	1	1	1	1	1	1	1
PARP(91)	2.1	2.1	1.0	2.1	2.1	2.1	2.1
PARP(93)	15.0	15.0	5.0	15.0	15.0	15.0	15.0

PYTHIA 6.2 Tunes





	1.96 TeV		14 TeV	
	P_{T0}(MPI) GeV	σ(MPI) mb	P_{T0}(MPI) GeV	σ(MPI) mb
Tune DW	1.9409	351.7	3.1730	549.2
Tune DWT	1.9409	351.7	2.6091	829.1
ATLAS	2.0	324.5	2.7457	768.0
Tune QW	1.1237	296.5	1.8370	568.7
Tune QK	1.9409	259.5	3.1730	422.0
Tune QKT	1.9409	259.5	2.6091	588.0

- ➔ Remember the p_T cut-off, P_{T0}, of the MPI cross section is energy dependent and given by $P_{T0}(E_{cm})$ = PARP(82)×$(E_{cm}/E_0)^{\varepsilon}$ with ε = PARP(90) and E_0 = PARP(89).
- ➔ Average charged particle density and PTsum density in the "transverse" region ($p_T > 0.5$ GeV/c, $|\eta| < 1$) versus P_T(jet#1) at 1.96 TeV for PY Tune DW, Tune QW, and Tune QK.

Comments

- RF Tuning project is undermanned
- “tuning” process is not algorithmic
- Many refits spurred by `Vista`

History of `Vista`

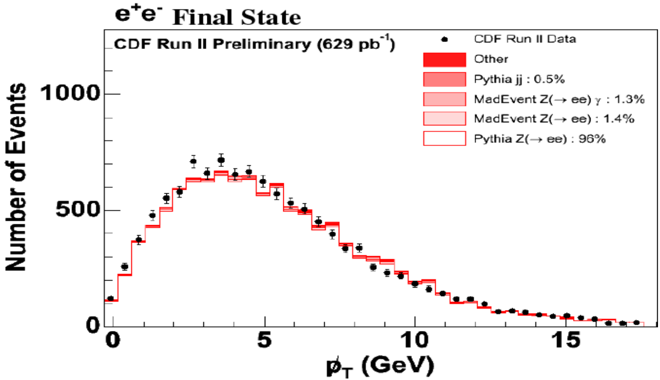
A global comparison of Standard Model predictions to the high p_T data

1. Tools developed by Knuteson (MIT) + collaborators to test consistency of Standard Model predictions vs. data
    - The endgame is to find deviations that "'cannot"' be explained by Standard Model, but that is a "'long-term"' goal
    - more details in later talk
2. I became part of the "team" as discrepancies developed and matrix element-improved events were needed
3. We have found [first] and solved [first] a number of problems and served as a catalyst for RF tunes

Case History

1. "UE" Tune consistent with p_T of the Z
2. Spike in $d\phi(j,j) = \pi$
3. Large $dR(j_2, j_3)$ in 3-jet events

Large Intrinsic k_T

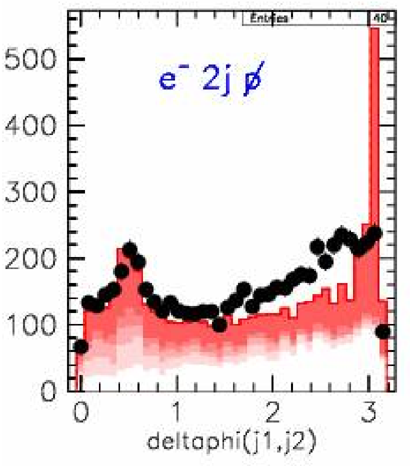


Catalysis for "-W"[1] tunes

---

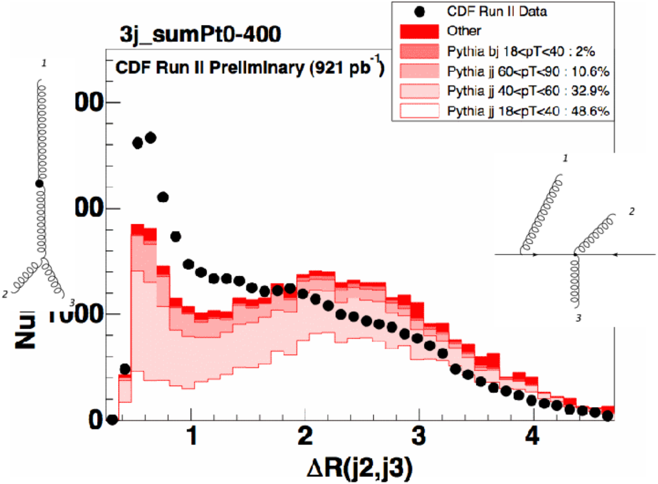
[1] W=Willis Sakumoto

High-p_T is sensitive to UE



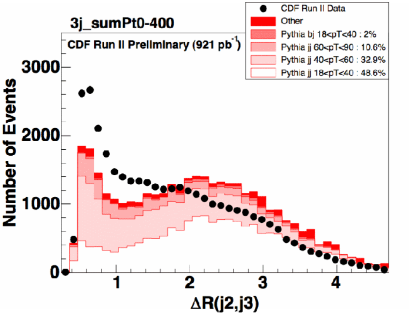
Allow FSR for multiple parton interactions

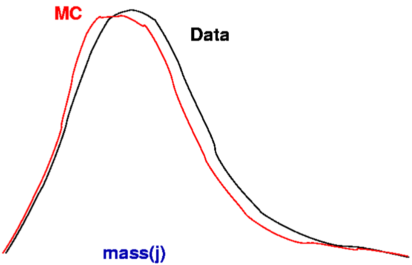
Tune A gives too much ISR

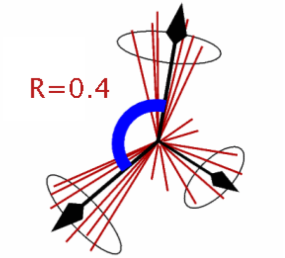


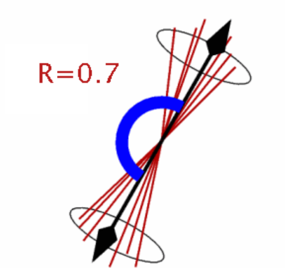
Don't increase starting scale for ISR

Case in Point: dR(j2,j3) and minMass(j)





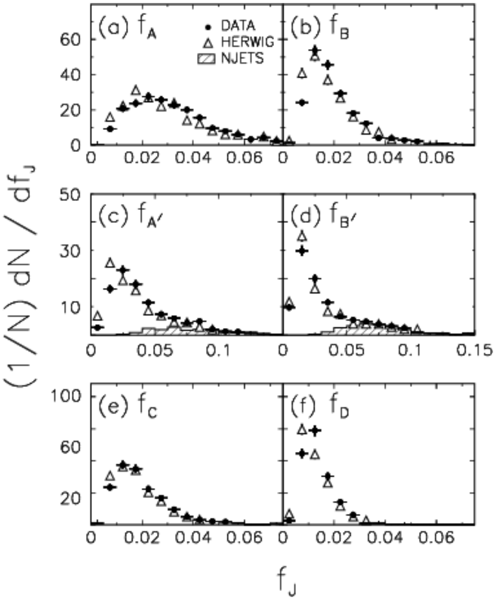




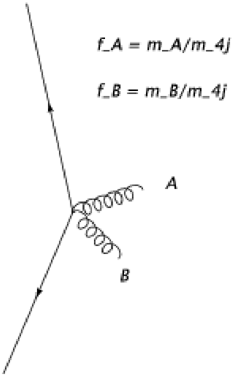
How is this possible?

Steve Geer led an effort to study multijets in Run I

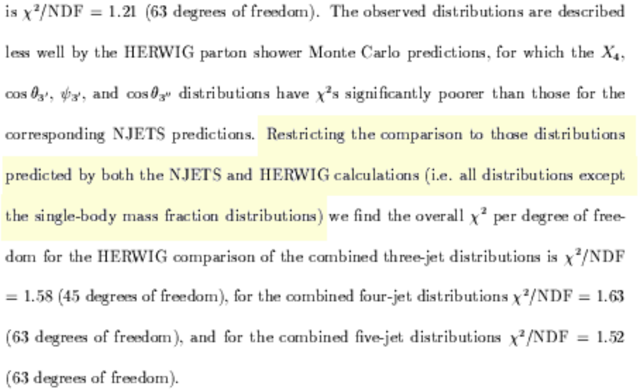
- Nice analysis of 3-, 4-, and 5-jet production
- Comparison of Herwig and simple models to the data
- Some notable discrepancies (f_i = mass fractions)
- These are not hidden in the text
- Main discrepancies dropped at the end when quoting overall goodness-of-fit



hep-ex/9605004, CDF-RunI



Kinematic set-up

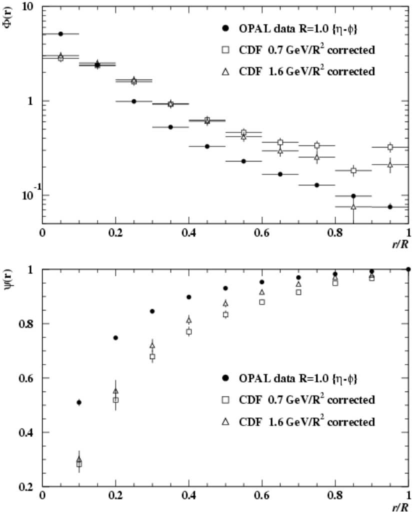


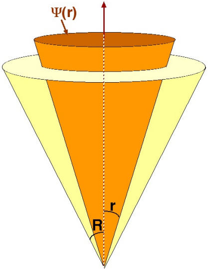
is χ^2/NDF = 1.21 (63 degrees of freedom). The observed distributions are described less well by the HERWIG parton shower Monte Carlo predictions, for which the X_4, $\cos\theta_{3'}$, $\psi_{3'}$, and $\cos\theta_{3''}$ distributions have χ^2s significantly poorer than those for the corresponding NJETS predictions. Restricting the comparison to those distributions predicted by both the NJETS and HERWIG calculations (i.e. all distributions except the single-body mass fraction distributions) we find the overall χ^2 per degree of freedom for the HERWIG comparison of the combined three-jet distributions is χ^2/NDF = 1.58 (45 degrees of freedom), for the combined four-jet distributions χ^2/NDF = 1.63 (63 degrees of freedom), and for the combined five-jet distributions χ^2/NDF = 1.52 (63 degrees of freedom).

f_i removed from the overall fit

no `NJETS` prediction for small f_i

OPAL cone jet studies, Z.Phys.C63:197-212,1994



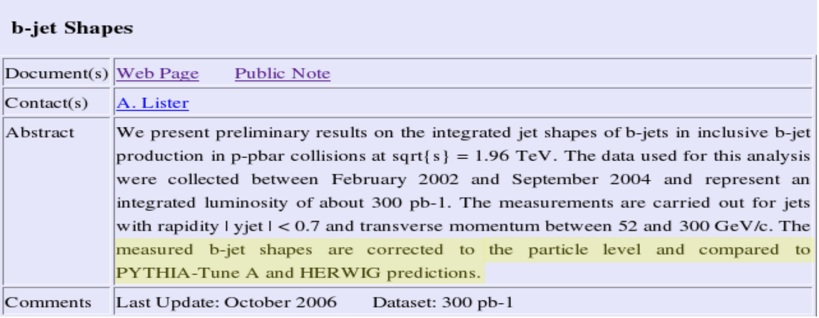


- LEP Jets .NE. TeV Jets
- Attributed to either UE or gluon jets
- Implies TeV Jets fatter!
- Would be useful to have access to the Z pole data

Notes

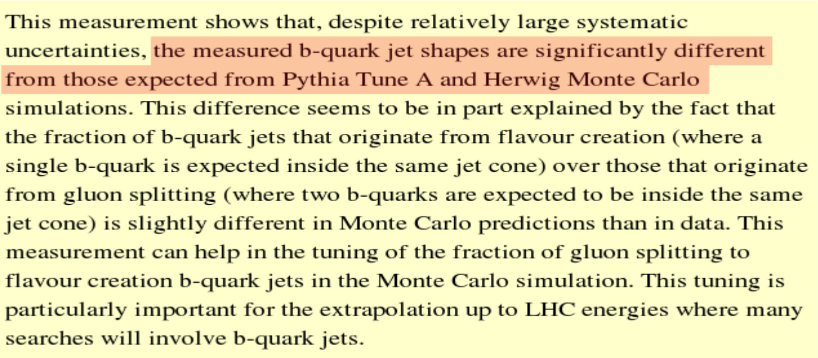
- Problem is rather “universal”
- High statistics
- Doesn’t seem to depend on jet definition
- Doesn’t seem to depend on detector
- Doesn’t seem to depend on generator
- Reproducible in orthogonal analyses
- We are [I think] converging on a solution(s)

Seen by others and in other final states



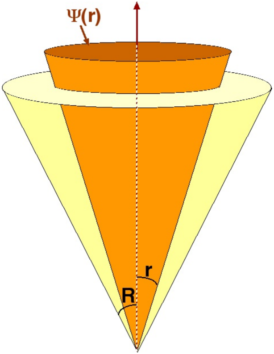
b-jet Shapes

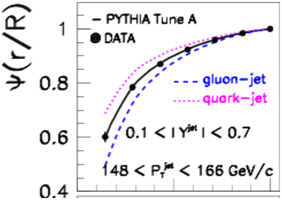
Document(s)	Web Page Public Note
Contact(s)	A. Lister
Abstract	We present preliminary results on the integrated jet shapes of b-jets in inclusive b-jet production in p-pbar collisions at sqrt{s} = 1.96 TeV. The data used for this analysis were collected between February 2002 and September 2004 and represent an integrated luminosity of about 300 pb-1. The measurements are carried out for jets with rapidity | yjet | < 0.7 and transverse momentum between 52 and 300 GeV/c. The measured b-jet shapes are corrected to the particle level and compared to PYTHIA-Tune A and HERWIG predictions.
Comments	Last Update: October 2006 Dataset: 300 pb-1

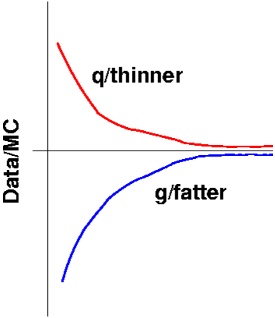


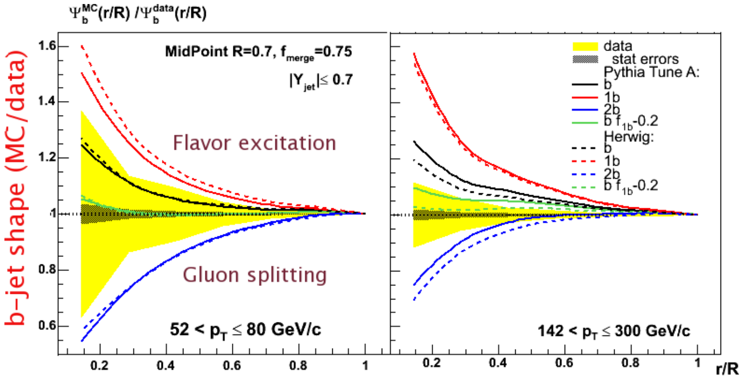
This measurement shows that, despite relatively large systematic uncertainties, the measured b-quark jet shapes are significantly different from those expected from Pythia Tune A and Herwig Monte Carlo simulations. This difference seems to be in part explained by the fact that the fraction of b-quark jets that originate from flavour creation (where a single b-quark is expected inside the same jet cone) over those that originate from gluon splitting (where two b-quarks are expected to be inside the same jet cone) is slightly different in Monte Carlo predictions than in data. This measurement can help in the tuning of the fraction of gluon splitting to flavour creation b-quark jets in the Monte Carlo simulation. This tuning is particularly important for the extrapolation up to LHC energies where many searches will involve b-quark jets.

Jet Shapes

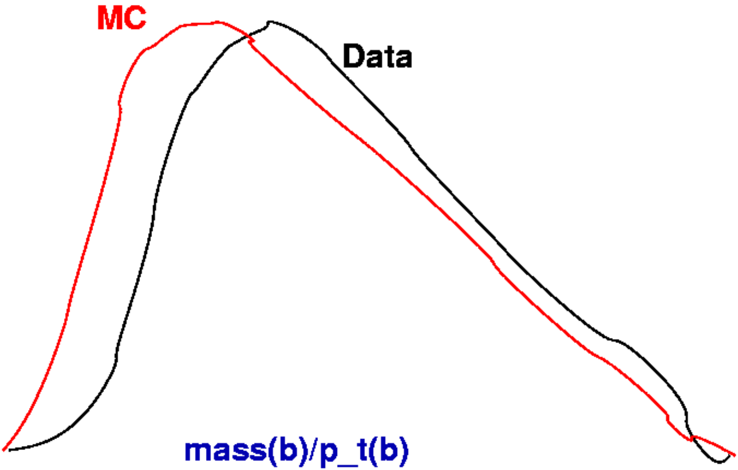

Ψ(r)
r
R


Ψ(r/R)
1
0.8
0.6
0.4
– PYTHIA Tune A
● DATA
- - - gluon–jet
...... quark–jet
0.1 < | Y^jet | < 0.7
148 < P_T^jet < 166 GeV/c


Data/MC
q/thinner
g/fatter



- Even after correcting $g \to b\bar{b}$, jet shapes differ
- MC/data > 1 means jets are thinner



Vista cartoon

The Importance of Automation

- Quick testing and feedback
- Ability to look at many channels at once
- Want/need methods to quickly reweight "old" Monte Carlo and converge on an answer

Want possibility to test hypotheses overnight or as soon as reasonably possible

Conclusions I Draw

- Only a handful of people doing this
- Should be done early on in debugging the experiment
- Bring different people together from the start

  A guiding force is needed to keep the project on track, or information is lost

- We should be prepared at the LHC start-up

  Getting ready for the 900 GeV run